THE UNIVERSITY OF HONG KONG

ICMA'97
April 28-30, 1997
Hong Kong

INTERNATIONAL CONFERENCE ON MANUFACTURING AUTOMATION

PROCEEDINGS
Volume One

EDITORS: S.T. TAN, T.N. WONG and I. GIBSON

THE UNIVERSITY OF HONG KONG
LIBRARIES

This book was received in accordance with the Books Registration Ordinance Section 4

PROCEEDINGS

OF

THE INTERNATIONAL CONFERENCE

ON

MANUFACTURING AUTOMATION

ICMA'97

April 28-30, 1997
Hong Kong

Editors: S.T. Tan,T.N. Wong and I. Gibson
The University of Hong Kong

VOLUME ONE

Organized by

Department of Mechanical Engineering, The University of Hong Kong

Sponsored by

K.C. Wong Education Foundation

Leeport (Holdings) Ltd.

Department of Manufacturing Engineering,
City University of Hong Kong

Department of Mechanical & Automation Engineering,
The Chinese University of Hong Kong

Department of Mechanical Engineering,
The University of Hong Kong

William M.W. Mong Engineering Research Fund,
Faculty of Engineering, The University of Hong Kong

Endorsed by the Rapid Prototyping Association of the Society of Manufacturing Engineers

For the complete set of two volumes, ISBN 962-85138-1-8
For volume 1, ISBN 962-85138-2-6
For volume 2, ISBN 962-85138-3-4

Foreword

I am delighted to welcome you to the International Conference on Manufacturing Automation (ICMA '97) organized by the Department of Mechanical Engineering, The University of Hong Kong.

In the same spirit as the first ICMA held in 1992, this conference aims to bring together leading experts in the area of manufacturing automation with a view to providing a forum for exchange of ideas and research findings, and to fostering a better link between industry and academic research. The programme contains 4 keynote lectures and a collection of some 180 contributions from over 25 countries, covering latest advances in a wide range of topics of manufacturing automation I hope that your expectations of the Conference will be fulfilled and wish you a fruitful and enjoyable week.

The Organizing Committee wish to acknowledge the very generous support from the sponsors of this event. Thanks are due to all members of the International Scientific Committee and keynote speakers for helping to make this event a rich and truly international one. The assistance of all session chairmen, student helpers and secretaries is gratefully acknowledged. The encouragement and facilities provided by the Head of the Department of Mechanical Engineering, The University of Hong Kong are most appreciated.

S .T. Tan
Chairman
Organizing Committee

April 1997

Committees

LOCAL ORGANIZING COMMITTEE

S.T. Tan (Chairman)	(HKU)
I. Gibson (Deputy Chairman)	(HKU)
T.N. Wong (Deputy Chairman)	(HKU)
K.W. Chan	(HKU)
Y.H. Chen	(HKU)
S.H. Choi	(HKU)
K.C. Hui	(CUHK)
C.L. Li	(CityUHK)
W.S. Sze	(HKU)
W. Wang	(HKU)
K.M. Yu	(HKPU)

INTERNATIONAL SCIENTIFIC COMMITTEE

A. Bowyer	(UK)
K. Case	(UK)
Y.K. Chan	(HK)
T.C. Chang	(USA)
N.N.Z. Gindy	(UK)
F.J.A.M. van Houten	(Netherlands)
T.L. Kunii	(Japan)
Y.T. Lee	(Singapore)
J.D. Liang	(Canada)
G.C.I. Lin	(Australia)
H. Marcus	(USA)
T. Medland	(UK)
A. Middleditch	(UK)
B. Mieritz	(Denmark)
T. Nakagawa	(Japan)
A.Y.C. Nee	(Singapore)
Q.S. Peng	(China)
T. Pintat	(Germany)
A.A.G. Requicha	(USA)
H.B. Voelcker	(USA)
J.Y. Wang	(Singapore)
Y.G. Wang	(China)
T. Wohlers	(USA)
J.R. Woodwark	(UK)
Y. Yan	(China)
M. Yeung	(Canada)
M.F. Yuen	(HK)
S. Zhang	(China)
J. Zhou	(China)

TABLE OF CONTENTS

Keynote Lectures
Worldwide Trends in Rapid Prototyping 1
Terry T. Wohlers
Responsiveness of Machining Environment 9
N.N. Gindy and S.M. Saad
The 21st Century Manufacturing Automation and the Social Impact 21
Tosiyasu L. Kunii
The Current State of Affairs in Dimensional Tolerancing: 1997a 30
H.B. Voelcker

CAD (Session 1A)
Application of a collaborative CAD on interactive remote coediting 44
Y.C. Kao and Grier C.I. Lin
Intelligent packing design based on assembly model 50
G.P. Huang, J. Zhou, Q. Yu, C.H. Wang and M.L. Yuan
Computer aided engineering system for cylindrical gearing 56
Rungfang Li, Tengjiao Lin, Chao Lin, Qian Tang and Bingkui Chen
Parametric geometry modelling based on an extended hypergraph 60
Xiangxu Meng, Xueqing Li, Gongbin, Jiaye Wang and Shenquan Liu
Isometric transformation between 2D and 3D surfaces 66
Z. Wu and M.M.F. Yuen

Computer Integrated Manufacturing (Session 1B)
Aspects of the design of a multiprocessor system for a flexible manufacturing system 72
H. Abachi and R. Ibrahim
Integration of machines and software packages to achieve computer integrated manufacturing 77
Henry Lau, Felix T.S. Chan and Ralph W.L. Ip
An efficient approach to manufacturing communications in a CIM environment 83
Xinquan Ding, Guanqun Gu and Jieyi Wu
CAPP systems in CIM and traditional production environment 89
N.K. Huang, N. Cui, Z.M. Zhang, J.X. Xu and Q.J. Fan
A preliminary study of a fuzzy clustering and assignment problem-based cell formation algorithm 95
S. Susanto, R.D. Kennedy and J.W.H. Price

Rapid Prototyping and Tooling (Session 1C)
How to improve the accuracy of prototype built by the laminated object manufacturing system 105
Bin Liu, Yuejia Xiao, Ming Han and Shuhuai Huang
Study on multifunctional rapid prototyping manufacturing (M-RPMS) system 111
Yongnian Yan, Guang Zeng, Wei Zhang, Haibin Guo, Liangwei Wu and Renji Zhang
Design of SLS (selective laser sintering) in M-RPMS (multi-functional rapid prototype manufacturing system) 117
Renji Zhang, Guanghua Sui, Haibin Guo and Yongnian Yan
Early stage rapid feedback design system based on rapid prototyped technology 122
Qingping Lu, Li Bao, Jianguo Wang, Jun Zhu and Yongnian Yan

Robotic Based Manufacturing (Session 1D)
An on-line calibration methodology for robot relative positioning inaccuracy 128
Tien-Fu Lu and Grier C.I. Lin
Force control using virtual trajectory generation 134
T. Fielding and J.K. Mills
Payload alignment control using task moment feedback for multi-robot assembly operation 140
M. Laliberte and J.K. Mills

Shape glass cutting direct drive robot 146
Hui-tang Chen, Ping Jiang and Yue-juan Wang

CAD (Session 2A)
STEP based semantic product model 152
Yan Luo, Xinjian Zhou and Ming Mo
Assembly modelling of product information based on self-organization 158
Jianrong Tan, Xiuting Wei and Chao Huang
A universal geometric interface implemented with structured language 164
Shixin Xu, Hegao Cai and Yongjie Wang

Computer Integrated Manufacturing (Session 2B)
Build integrated CAD/CAM system based on software reverse engineering 170
Yiping Lu, Jianping Lin and Ji Zhou
Design and development of material management system in CIMS of iron & steel complex 176
Zhiwen Tang and Dingwei Wang
Monitoring behavioral evolution for manual assembly system using object petri nets in CIM environment 182
F. Wu, W.H. Liu and T.Y. Xiao

Process Planning (Session 2C)
A taxonomy for interactive and blackboard based CAPP 188
G. Van Zeir, J-P. Kruth and J. Detand
Smart drilling - a fuzzy decision system for planning drilling operations 196
Matthew Jia Ma and R. Du
Tolerancing, tool adjustment and machining parameter selection problems in multi-pass turning operations 202
M. Liang, M. Zuo and T. Yeap

Geometric Modelling (Session 2D)
Three-valued calculi in the problem of conversion from CSG to boundary models 208
A.G. Gorelik
G^1 refinement of data for rapid prototyping 214
D.J. Walton and M.Yeung
The boolean operation of polygon with holes based on edge recognition 220
Yunxing Wu, Mike Du and Shiping Liu

CAD (Session 3A)
Implication of design for inspection 225
L. Wang and G.C.I. Lin
A Neural network approach to parametric design 231
Y.H. Chen
Element method for 2-D shape transformation 237
K.C. Hui and Y.D. Li
Interacting with a virtually elastic object 243
K.C. Hui and M.C. Ma
2D boundary recovery based on medial axis transform 249
Yujin Hu, M.M.F. Yuen and Ji Zhou

Computer Integrated Manufacturing (Session 3B)
System of cost control and analysis decision support under CIMS circumstance 255
Baodong Xu, Bin Gao, Jiazhong Chang and Bo Xu

Product data management system 260
Hongyuan Xue, Ji Zhou and Yifang Zhong
Prototyping tool for flexible manufacturing systems 266
Dalila Boudebous
Reference models as an approach for the construction of flexible production planning and control systems 272
W. Dangelmaier, A. Kuhn and T. Langemann
A task-centred methodology to support an integrated and open computer aided manufacturing systems design environment 278
B. Wu and R.S. Hull

Rapid Prototyping and Tooling (Session 3C)
Study of rapid intelligent tooling system based on RPM Technology 284
Zhaohui Du, Renji Zhang and Yongnian Yan
The application of hot melt adhesive to laminated object manufacturing 290
Xianglin Zhang, Jin Huang, Yan Li and Shuhuai Huang
Internal company use of rapid prototyping and 3D-digitizing in the engineering design process 294
C. Birke, C. Beyer and K.-H.-Grote
A computation efficient slicing algorithm for rapid prototyping 300
S.H. Choi
Rapid product development using a RP-Centered interactive design process 306
Hsien-Jung Wu

Robotic Based Manufacturing (Session 3D)
A path planning algorithm for robot manipulators in FMS environment 311
Yulin Ma, Wei Wang and Yili Fu
Model and sensor based control of industrial robots 317
Ole Madsen and Hans Holm
A formal method to specify at an abstract level robotics-based manufacturing systems 323
J.M. Condom and K. Ouriachi
A sophisticated assembly planning system for flexible robot-based manufacturing 329
H. Mosemann, F. Rohrdanz and F.M. Wahl
State space modelling of a class of discrete event systems for robotic manufacturing control 335
Hans Holm

CAD (Session 4A)
Function and behaviour description in intelligent mechanical CAD system 341
C.L. Li, S.T. Tan and K.W. Chan
Dimension-driven parameterized design of free form shapes 347
Kai-Ming Yu and Wei-Ming Zhu
Petri net aided tolerancing for prismatic components 353
P.Eshwaraiah, A.L. Sarat Babu, K.Ravi Raju and N.K. Tewari
Automatic mesh generation for simplification of finite element analysis 360
Junko Fukuda
Designing at a distance 366
A.J. Medland
A tolerance model with surface irregularity consideration 374
C.K. Lui, W.S. Sze and S.T. Tan

Computer Integrated Manufacturing (Session 4B)
On the assemblability of a product 383
Hung-Yao Hsu and Grier C.I. Lin

Software prototyping and object-oriented specification as a methodology for customer-oriented development of software systems - an example from the automotive industry 389
K. Mertins, R. Albrecht, S. Beck, O. Bahns and B. La Pierre
Improvement of the planning quality in highly networked processing industry production by means of simulaton and cooperative local control stations 395
J. Pirron, J. Arnold and O. Reisch
An alternative concept to MRPII for mass customization based on the object-oriented paradigm 401
Claus Rautenstrauch
Simulaton for evaluation of scheduling rules in flexible manufacturing systems 407
F.T.S. Chan and R.W.L. Ip
Shop-Floor scheduling and data capturing system for autonomous manufacturing island 414
F.W. Lam, Y.J. Ma, W.B. Lee and S. Zhang

Rapid Prototyping and Tooling (Session 4C)
Rapid prototyping of golf clubs 420
R. Jones and S.R. Mitchell
An approach on decision support for selection of process chains of the tool and die making industry under special considerstion of rapid prototyping technologies 426
H. Durr ahd U. Kaschka
Feature based optimization of part orientation for rapid prototyping 434
W.M. Micky Ng and S.T. Tan
Rapid prototyping using robot welding - process description 443
Fernando Ribeiro and John Norrish
An extended slicing method for rapid prototyping 449
Kwan H. Lee and Il-sang Yoo
Rapid prototyping on a large scale 455
I. Gibson and G. Mensing

Feature Based Technology (Session 4D)
Research of feature conversion and representation of product process information model 461
Jian Gao, Detao Zheng, Ping Zhang, Guowang Zhu, Jian Sun and Xiankui Wang
Research on customer-oriented CAPP system modelling technique of prismatic parts based on feature 467
Juhong Gao and Yanshen Xu
The structure of feature layers and its application in CAD/CAM integration 474
Qun Guo, Yifang Zhong, Li Wan and Ji Zhou
3D objects feature match based on relational graph 480
Xinhua Wang, Chuanhue Huo, Yuguang Li, Deren Li and Peizhi Huang
The feature modelling for a sculptured object 486
C.K. Au and M.M.F. Yuen
A method for recognizing feature interactions 492
C. Zhang, K.W. Chan and Y.H. Chen

CAD (Session 5A)
Standardization of basic design sequences of heat-exchanger for power plant 504
Hideyuki Yamaguchi, Zhihua Zhang, Isao Nagasawa and Masanobu Umeda
Implementation of design for assembly: assembly-centric modelling and digital mock-up system 510
Yoong Ho Jung
Integrated CAM design and manufacturing system 517
L. Tang
Tolerance design and analysis using the tolerance regions 523
Jyun-Ming Chen and Chi-Hui Chen

An approach to automate the bending sequence of progressive dies 529
C.C. Chan and C.H. Lai

Computer Integrated Manufacturing (Session 5B)
Simulation approach in planning and design of manufacturing cells - a case study 535
F.T.S. Chan and R.W.L. Ip
The application of genetic algorithm in the planning of FMS system 542
K.W. Keung, W.H. Ip, Y. Li and D.W. Wang
An FMS process planning system based on coloured petri nets 548
T.N. Wong, K.H. Ng and M.C.K. Kwok
Dynamic mixed dispatching of FMS 554
A. Krishna Kishore and N. Ramesh Babu

Process Planning (Session 5C)
Develop silicone steel sheet cutting patterns with filled-in blanks 560
Yaodong Cui, Lian Zhu, Aiming Jin and Ming Cao
An integrated system of CAPP and job shop scheduling 565
Chao Deng, Peigen Li and Bin Luo
Study on NC automatically programming for cartridge valve blocks 571
Lin Xie
Generation of alternative process plans by net model 577
Min-Hyoung Kang, Ji-Hyung Park and Myon-Woong Park
Tools for assembly in a virtual environment 583
R.G. Dewar, J.M. Ritchie, I.D. Carpenter and J.E.L. Simmons

Feature Based Technology (Session 5D)
Exact recognition of compound feature by feature adjacency matrix elimination 588
Yong Yu and Tie-Chang Zhang
Towards adaptable product development 594
Y. Garden, J.P. Jung, Y. Lanuel, S. Leinen, B. Martin, C. Minich, E. Perrin, C. Poinsignon and I. Stemart
Automatic selection of die structure based on product features of deep-drawn components 600
W.M. Sing and K.P. Rao
Identification of grain growth in wood machining using artificial neural networks 606
V. Karri

Author/Session/Page Index A1-A5

WORLDWIDE TRENDS IN RAPID PROTOTYPING

Terry T. Wohlers
Wohlers Associates, Inc.
Fort Collins, Colorado USA

ABSTRACT

Companies are changing their fundamental approach to design, and thus becoming leaner and producing products faster. Rapid prototyping (RP) is playing a part by helping them make design improvements earlier in the process when changes are less expensive. This paper focuses on advances in materials used for RP, tooling applications such as silicon rubber tooling, and bureaus that offer RP and complimentary services. It also discusses a new class of RP for concept modeling, research and educational efforts, government programs, market growth, and the future.

KEYWORDS

Rapid prototyping, tooling, CAD/CAM, freeform fabrication.

1. INTRODUCTION

An impressive number of industries have discovered RP and its benefits. Automobile manufacturers and suppliers use it to prototype instrument panels, rear view mirror housings, exhaust manifolds, and tire tread designs. Aerospace companies use it to produce throttle body parts and other mechanical subassemblies that go into aircraft. Manufacturers of business machines use it to prototype copier and fax machine parts. Computer manufacturers use RP for many of the plastic injection molded parts that go into keyboards, modems, and the computers themselves.

Style and the feel of a product plays an important role in the design process, so cellular phone manufacturers, such as Motorola, model new designs with RP. Manufacturers of toys, such as Hasbro, take advantage of RP to model molded plastic pieces that go into almost every toy imaginable. Makers of medical supplies, devices, and instrumentation put RP to work when designing everything from blood containers to diagnostic equipment. Government labs, the US Department of Defense, NASA, and other agencies use RP to help them design weapon systems, satellites, ships, tanks, and equipment of all types. Anything that you can model in three dimensions, you can create with RP.

2. ADVANCES IN MATERIALS

When StereoLithography (SL) from 3D Systems was first made available in 1988, most of the parts it produced were used as visual aids. 3D Systems admits that early RP systems did not produce accurate parts, limiting the application of the technology. Due to poor mechanical properties of the material, the parts could not withstand the kinds of tests that most prototypes must endure. Today, the materials used to make RP parts are much better. Parts are stronger and accuracy is superior. Strength and accuracy vary widely from part to part and system to system, but often they are good enough for many prototyping and testing applications.

SL users can today choose among a range of epoxy and acrylate resins. These materials are not necessarily the strongest available, but they build some of the most impressive-looking parts. Fused Deposition Modeling (FDM) from Stratasys uses ABS plastic and investment casting wax to build parts. With its strong mechanical properties, ABS has become popular among FDM users. Selective Laser Sintering (SLS) from DTM offers the widest selection of materials. Users can choose from polycarbonate, nylon, glass-filled nylon, and a proprietary pattern material that DTM calls TrueForm. The nylon materials work well for functional prototypes, while TrueForm is being used as patterns for investment casting and soft tooling applications. With SLS, you can also sinter parts using a plastic-coated sand and steel. The sand is being used to produce sand molds and cores for sand casting, while the steel is being developed and used to produce prototype injection mold inserts.

Laminated Object Manufacturing (LOM) from Helisys builds parts using layers of paper, so the final part resembles wood. Because its laser beam contacts only the periphery of each layer, the process works well for large, chunky parts that contain a lot of volume. The process also produces parts using plastic sheet material, although the plastic is relatively new. Solid Ground Curing (SGC) from Cubital produces parts in acrylate and epoxy resin and can make many parts at once. Using patented technology invented at MIT, Direct Shell Production Casting (DSPC) from Soligen produces ceramic shells for investment casting of metal parts.

Model Maker (MM) from Sanders and Ballistic Particle Manufacturing (BPM) from BPM Technology use proprietary wax materials. One of the strengths of the MM system is the production of small, intricate parts with a fine surface finish. While most RP systems produce layers in the 0.08 - 0.25 mm range, the MM can build parts with layers as thin as 0.013 mm. Thin layers produce parts with superior surface finishes and improves dimensional accuracy, although it takes longer to build parts using thin layers.

Kira Corp. (Japan) and Kinergy (Singapore) produce wood-like objects using paper lamination processes similar LOM. Japanese companies CMET, Denken, D-MEC, Teijin Seiki, Meiko, and Ushio, and German companies EOS and Fockele & Schwarze, each manufacture stereolithography systems that produce parts from light-sensitive photopolymers. EOS also offers a system based on laser sintering that competes with DTM in Europe. The machine, called EOSINT, uses polystyrene, nylon, sand, and metal to produce parts.

3. TOOLING APPLICATIONS

As the technology and materials have improved over the years, so has the value of RP. More durable materials and good part accuracy, coupled with improved surface finish, has made it possible to use the parts as patterns for prototype tooling. Usually, if the work requires more than a few copies of the same part, it makes economic sense to produce and finish one part and then use it to produce a mold – a practice that is in widespread use today. Using this approach, you can mold parts in a material that more closely emulates the material used in the final product, making them more suitable for testing. At many companies, more than half of all RP parts are used as patterns for molds.

One of the most popular approaches is RTV (room temperature vulcanization) silicon rubber tooling. With this type of tooling, an RP pattern is positioned in a mold box and encased in silicon rubber. After removing trapped air in a vacuum chamber and curing the silicon, the RP pattern is removed by cutting the rubber with a sharp instrument to create a two-part mold. The mold is then used to cast parts in wide variety of urethane materials. Compared to producing duplicate parts using the RP equipment, the RTV process is relatively inexpensive. Plus, it can duplicate impressive detail. Silicon rubber permits you to flex the mold and remove parts with moderate undercuts that otherwise get locked into the mold.

Another prototype tooling method that is growing in popularity is epoxy tooling. Because epoxy is tougher than rubber, you can expect better wear and longer mold life. You can usually get from 25 to 50 pieces from a rubber mold, but epoxy enables you to get from 100 to more than 1,000 pieces, depending on the parts you are molding. However, an epoxy mold can cost twice as much and it takes longer to make. To extend mold life, mold makers create a composite mold material by adding a filler material such as aluminum to the epoxy resin.

If your prototype molding needs are even more demanding, consider spray metal tooling using an RP part as the pattern. Spray metal is a process in which you spray hot metal particles onto the surface of a pattern. The spray metal mold is backed with epoxy resin to give the mold shell strength and rigidity. This process usually costs more than epoxy molds and the production of the mold can take longer.

Other methods of prototype tooling include cast metal molds using a combination of techniques. One method is to use investment castings to produce the core and cavity. You can potentially speed this process using 3D Systems' QuickCast method or Soligen's DSPC. Still, it usually requires that you machine the surfaces of the mold. Due to cost, lead time, and the resources needed to make it all happen, investment cast tooling has not become popular. Another cast tooling method involves rubber and plaster materials. You begin with RP masters of the core and cavity, then you go to rubber, then plaster, and finally to metal. This process is also quite involved, so not many are using it.

Additional methods include metal vapor deposition and nickel ceramic composite methods. A third is to use an RP part as an electrode for electrical discharge machining (EDM). These methods show varying degrees of promise and are not in widespread use today.

Machined aluminum using CNC milling machines is the most popular prototype tooling method. Machined aluminum molds can yield as many as 100,000 molded pieces, but these molds can be the most expensive to produce and often require weeks of lead time. For this reason, DTM and EOS have developed and are offering alternative methods to producing machined prototype injection molds. Both take advantage of their sintering equipment. DTM calls its process RapidTool which takes advantage of its RapidSteel material. Using its Sinterstation 2000 and 2500 machines, users can produce molds that are nearly as strong and durable as machined aluminum.

4. SERVICE BUREAUS

CAD solid modeling, rapid prototyping, and tooling requires special skills, methods, and machines, and not every company wants to develop these capabilities in-house. As a result, as many as 250 service bureaus (SBs) worldwide are offering solid modeling, RP, and rapid tooling services. These companies have emerged over the years as the demand for RP has grown. The way it works is like this: You provide the SB with STL, IGES, or CAD data, or drawings and sketches, and they will send you a quote. The work can vary from one part to hundreds. Turn around time varies accordingly.

My experience has been that if you need one or two RP parts from an SB, you can usually get it within a week. Delivery depends on the SB's backlog. If the work involves soft tooling, such as RTV silicon rubber, and many parts, set aside two to three weeks. Other factors that affect turn-around time include the size and quantity of parts, the benching and finishing of the pattern(s), and the design data that you provide. If you provide a clean STL file, the SB could be running the part the same or next day. If you provide data that requires CAD design time, it could be several days, even weeks, before the SB produces the RP part. In this case, the CAD work could cost more, much more, than the RP part.

SBs have developed as an important part of the RP industry. They provide thousands of models and prototype parts to countless organizations around the world. Also, SBs serve as training wheels for companies who are new to RP. Sometimes these companies develop RP capabilities in-house; others choose to use SBs over the long term. And some companies use both in-house RP resources and SBs, particularly in overload situations. In any case, SBs educate and provide a low-risk way of introducing a company to what the technology has to offer. Many of them offer niche services, such as spray metal tooling or laser digitizing. These services help separate themselves from their competition.

If you consider annual revenues from the RP industry, SBs represent a significant portion. In 1995, SBs were credited with generating an estimated 46% of all money spent on RP around the world, according to a report published by Wohlers Associates, Inc. Furthermore, SBs grew at a rate of about 43% from 1994 to 1995. New SBs are emerging as established ones expand their number of employees, machines, and capacity. Meanwhile, "mega-SBs" are developing. Examples include Japanese-owned ARRK Creative Network with its nine locations in the US and Compression Inc. with its six US locations. Compression, for example, employs 240 people and occupies 12,077 square meters of office and lab space. The company owns and operates 15 RP systems representing four technologies, 36 CNC machining centers, 15 injection molding machines, 45 seats of Pro/Engineer, and 65 workstations.

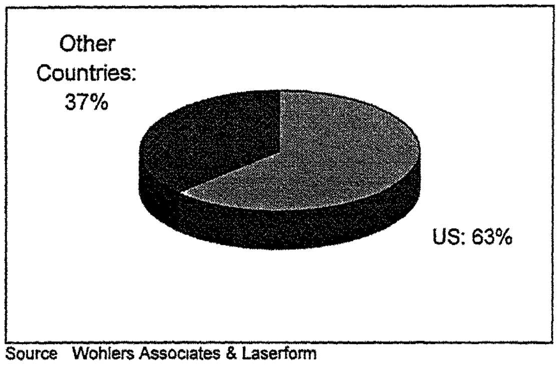


Fig. 1. Service bureaus generated an estimated US$135.5 million in RP-related revenues in 1995. About $86 million of it was generated in the US.

In recent months, Plynetics and Prototype Express – two of the most established and recognized SBs in the US – merged to form yet another mega-SB. Combined, the new company employs 130 people and operates 20 RP machines, seven CNC machining centers, and six injection molding machines. The new company has a total of 5,667 square meters of space at three sites. Most recently, Laserform – another popular and well-established SB – was acquired by Plynetics Express, the new name of the merged companies.

Advantages to this large SB business model are a sharing of people, machines, and advertising dollars. The challenge will be to maintain the kind of personalized service that customers have come to expect from small SBs. Most SBs employ 10-20 people.

5. CONCEPT MODELERS

Another developing area that's receiving a lot of attention is the idea of using RP for concept modeling and early design review. From the beginning, RP has been used for this purpose, but until recently, system manufacturers have not developed and priced machines specifically for this market. Most RP systems are relatively expensive to buy, operate and maintain, require special expertise, and are located at a central site. And often designers must wait a week, sometimes longer, to get a part. By that time, the design may have changed. For these reasons, high-end RP has had limited appeal for this purpose among the thousands of engineering groups around the world.

Three US companies are trying to address the opportunity by offering machines less capable than more expensive RP equipment, but good enough for quick models. The machines, essentially 3D printers, sit close to the users of the CAD solid modeling systems. The challenge is to develop a machine that is as easy and safe to operate as copiers, printers, and fax machines, and is quiet and fast. Also, it should not take up more space than most other office equipment and it must carry a price tag that engineering groups can swallow.

Stratasys, 3D Systems, and BPM Technology have made it clear that they are going after this market in an aggressive way. Stratasys is offering Genisys, a 3D printer that extrudes a reasonably strong polyester compound using a technique similar to the company's FDM products. The origins of Genisys came from IBM's Watson Research Center, an RP development that Stratasys later acquired from IBM. At US$55,500, the 102 kg Genisys product builds models that fit inside an 203 mm cube using 0.36 mm layers. The cost to build a small Genisys model is in the $25 to $50 range.

3D Systems developed Actua 2100, a 3D printer priced at $60,000 that takes advantage of ink jet technology. Using 96 jets arranged in a linear array, the 318 kg Actua machine prints 0.114 mm layers of a thermoplastic material that resembles the properties of a hard wax. Maximum part size is 254 x 203 x 203 mm. Detail and surface finish are impressive. BPM Technology has introduced its $35,000 Personal Modeler 2100 in a 125 kg package. Using a single ink jet orifice, the machine deposits 0.076 mm layers of a wax-like material. Maximum part size is 254 x 203 x 152 mm.

You could include a fourth product in the 3D printing class of machines. Sanders Prototype offers a small RP unit called Model Maker that operates quietly and safely in an office environment. It measures 521 x 444 x 584 mm and weighs 41 kg, making it the smallest and lightest RP machine available. Most people are using it to produce very small, but highly accurate, wax patterns for secondary processes such as investment casting. It's an excellent machine for this purpose, but most users will agree that concept modeling is not it's strength. Plus, it is not being positioned or sold as a concept modeler, although its small size and $65,000 price tag causes it to overlap into this area.

6. RESEARCH AND EDUCATION

Many universities and private and government research labs are exploring new RP techniques and tooling applications and several have made an impression. Milwaukee School of Engineering and Georgia Tech (both in the US) are on track to offering a top programs nationally. In the UK, the University of Nottingham, with its student projects, sponsored research, and service organization, rivals the best academic programs available anywhere. And University of Tokyo's Professor Takeo Nakagawa continues to push the limits, with 50 of 100 developments (several in RP) leading to industrial applications and commercialization.

Some of the most astonishing work in RP is at a high school. Bergen County Technical School's Academy for the Advancement of Science and Technology has been running RP equipment for years. They include RP as an important part of student projects – assignments that involve every facet of the product development process. The students form small companies that do market research, R&D, engineering, prototyping, testing, simulation, tooling, and production. Through impressive advertising and sales campaigns, the kids interface with customers and recoup their costs.

Many other institutions have excellent reputations for their work in RP. Among them are Carnegie Mellon, Clemson, MIT, Stanford, University of Dayton, University of Louisville, and the University of Texas at Austin. Most of them are researching or developing one or more aspects of the technology, while others have integrated RP into their instruction.

7. GOVERNMENT PROGRAMS

Many US government agencies are funding RP research and development programs. In August 1995, as many as 45 RP-related projects were being sponsored by the National Science Foundation (NSF), the Defense Advanced Research Projects Agency (DARPA), the Department of Energy (DOE), and the National Institute of Standards (NIST). The total amount of money being spent on these multi-year programs exceeded $45 million.

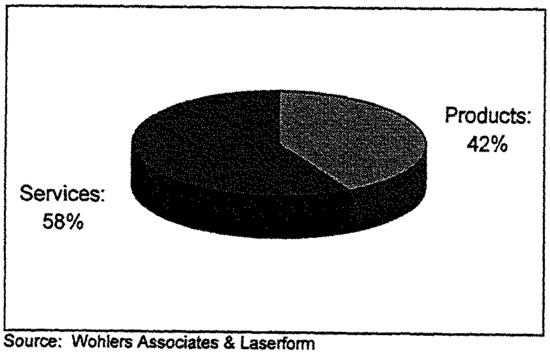
In Japan, the Ministry of International Trade and Industry (MITI) is spending 800 million yen (about $8 million) over a 4-year period. The work began in 1994. Its purpose is to stimulate RP system sales by improving resins, hardware (such as galvanometer scanning technology), software, and applications. The Japanese Ministry of Education and other groups in Japan are also sponsoring programs that they hope will accelerate the growth of RP in Japan.

In Europe, many application-oriented projects have been funded over the years. The funds for many of these projects have come from European Community (EC) Directorates such as Brite EuRam, Esprit, and Eureka. They have funded projects such as Computer Aided Rapid Prototyping (CARP), European Action on Rapid Prototyping (EARP), INSTANTCAM, and many others worth millions of dollars. Most of the work has been focused on applying and fine-tuning existing technology, such as stereolithography from 3D Systems and EOS.

With its 46 research institutes, the Fraunhofer Society (Germany) has been one of the few European organizations to develop new RP system technology. Its work is focused on the direct sintering of metals using high power lasers, as well as other unique approaches to rapid prototyping. Sandia National Labs and Los Alamos National Labs, both in the US, are also working on the sintering of metals using lasers.

8. GROWTH

The worldwide market for RP has grown steadily over the past several years. Since its beginning in 1988, the market has grown on average by 58% per year. How does this compare to the early days of related industries that are now established and respected? From 1963 to 1973, the machine tool market grew by an average of 8.5% annually, according to Julius Dorfman of CIMdata. From 1970 to 1981, the numerical control (NC) market grew by an average of 22% per year, Dorfman estimates. This illustrates the impressive nine-year growth of an industry that skeptics believed may not develop.

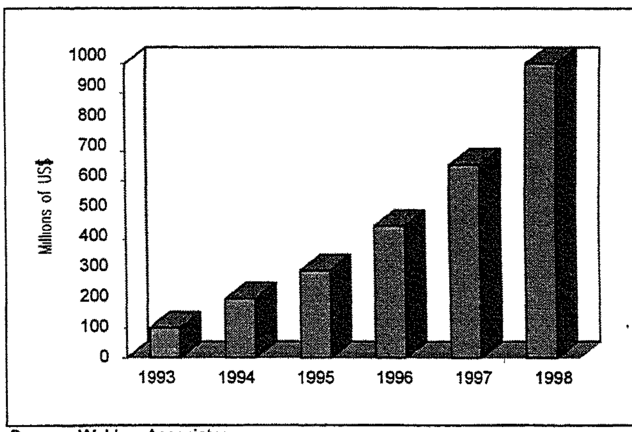


Source: Wohlers Associates & Laserform

Fig. 2. The primary market, consisting of RP products and services, was an estimated US$295.1 million in 1995. The "services" segment was estimated at $170.3 million.

The primary RP market grew by 49% to an estimated $295.1 million in 1995. The primary market consists of products and services affiliated with the sale and support of machines and materials and the production of RP parts at service bureaus. The secondary market grew by a 112% to $176.1 million. This market segment includes the secondary tooling, plastic duplicates, and metal castings created from RP patterns.

If RP continues to grow as it did in 1993, 1994, and 1995, it could reach the $1 billion milestone by 1998. This does not take into account the total economic impact from the technology, a figure that is difficult to calculate. How does one measure the dollar impact of a good design versus a poor one? If a substandard product generates $100 million, would a refined version of it generate twice that amount? And if RP helps speed the product to market weeks or months before a major competitor, how much more revenue does this mean for a company?



Source: Wohlers Associates

Fig. 3. The RP market could reach the $1 billion milestone by 1998 if the industry continues to grow as it did in the 1993 - 1995 timeframe.

What's known is that many organizations are buying multiple RP systems to keep up with demand. The process has moved out of its experimental phase at many companies and is now an important part of everyday business. Companies not using the technology are taking note. They are discovering that if they don't modernize their approach to product modeling and prototyping, it may be difficult for them to compete in the future.

9. THE FUTURE

Future uses of RP are promising. In the field of medicine, surgeons will favor RP models over x-ray images for planning critical operations such as craniofacial surgery. The process begins with CT or MRI scans. By taking many scans at small increments, medical professionals produce the cross section data needed to drive the RP process. Already, many models of human body parts have been built in the US, Europe, and Japan. However, the process is expensive and, in most cases, insurance companies do not yet cover it.

As the price of RP models become reasonably inexpensive, architects will use it to produce proposed building designs, enabling customers to more easily visualize how the final building will look. Scaled models of proposed high rise buildings must undergo wind tunnel testing to determine how the air turbulence will affect it and surrounding buildings. RP will help reduce this cost and speed the process of producing these models.

Artists will combine the flexibility of computer modeling with RP to produce exotic sculptures. With CAD solid modeling dipping under the $500 mark, schools will take advantage of low-cost RP to teach important concepts in science, math, industrial technology, manufacturing, engineering, biology, anatomy, medicine, and other subjects. Individuals will take advantage of it a home to experiment with new ideas and inventions that before were not possible.

10. SUMMARY

Those experienced with RP acknowledge that it saves significant time in new product development. It helps them avoid expensive delays caused by tooling problems – mistakes that are now being caught and corrected much earlier in the process. RP enables them to more easily and accurately communicate the proposed design to groups inside and outside the company. All of this early input dramatically increases the chances of getting it right the first time.

RP is being driven by the need to design new products faster than ever before. What use to take a year or two must be done in months. And as product life cycles shrink, engineers' workloads increase. Essentially, they must do more with less time. Companies must employ new ways of developing products. Tools such as solid modeling and RP, integrated properly into a concurrent engineering environment, can make the difference between profit and loss.

RESPONSIVENESS OF MACHINING ENVIRONMENTS

N. N. Gindy and S. M. Saad

Department of Manufacturing Engineering and Operations Management,

The University of Nottingham,

Nottingham NG7 2RD,

England.

ABSTRACT

In this paper manufacturing responsiveness is related to the ability of manufacturing systems to utilise its existing resources to make a rapid and balance response to the predictable and unpredictable changes. Better understanding of the inherent (hidden) flexibility that exist within a manufacturing system can therefore lead to significant improvement in system performance and responsiveness. In the reported research a conceptual framework for representing the capabilities of machine tools and machining facilities using generic capabilities units termed "Resource Elements" is presented as well as a mathematical basis of calculating the manufacturing system flexibility using the Resource Elements. Simulations are used to examine manufacturing system performance and compare Resource Element-based scheduling with conventional machine-based approaches. The results show that significant improvements in system performance and the system's ability to cope with disturbances can be achieved if manufacturing facilities are represented and scheduled based on the Resource Elements concept.

KEYWORDS

Manufacturing Responsiveness, Resource Elements, Manufacturing System Flexibility, Scheduling, Simulation.

1. INTRODUCTION

Some of the significant factors influencing the market for manufactured products are already apparent: global competition; shortened product life cycle; increasing requirements for quality; increasing need for product customisation; faster paced advances in increasingly complex technology and rapidly expanding options in materials and processes. To achieve success in such an environment, it is not sufficient anymore for manufacturing companies to focus on being competitive in terms of product cost and quality but must also achieve the highest possible performance in an unpredictable environment.

Manufacturing Responsiveness relates to the ability of manufacturing systems to make a rapid and balanced response to the predictable and unpredictable changes that characterise today's manufacturing environments. The development of appropriate measures and methods of assessment for the various facets and attributes of manufacturing responsiveness is an important step towards being able to optimise the utilisation of available system resources to improve performance and responsiveness. The focus of the work reported in this paper is on improving the responsiveness of "machining environments". Emphasis is placed upon achieving a better understanding of the inherent (hidden) flexibility that exist within a manufacturing system in order to improve: (1) The utilisation of available system resources to improve operational performance; (2) The ability of a manufacturing system to cope with internal and external disturbances under tight due date targets.

The reported work is part of ongoing research on the design of next generation manufacturing systems. Underpinning the work, a new methodology has been developed for representing the processing requirements of components at the process planning stage and the capabilities of manufacturing resources based upon generic capability units termed resource elements (RE's).

In this paper the conceptual framework for the research is briefly outlined and new measures for system flexibility are defined in section 2. In section 3 the experimental methodology and results are presented. Section 4 gives the main conclusions of the work to date.

2. CONCEPTUAL FRAMEWORK

The basic concept underpinning the research is shown in Figure 1. The assumption is that rigid, static and hierarchical manufacturing systems will be replaced by systems exhibiting great adaptability to rapid change which are able to produce low volume low cost products. Next generation manufacturing systems will be made up of relatively simple, distributed, autonomous but co-operating sub-systems or resource groupings organised as virtual factories.

A manufacturing facility is treated as an objective driven logical grouping of distributed resources which can equally be termed "cell" or "factory" where its resources not have to be in physical proximity. The only real constraint on the size of a cell is the limits of exercising proper control over the resources contained within it and achieving its objectives. The process of cell formation is dynamic, objective driven, self-optimising and self-organising. Resources are normally combined to carry out a task or perform a function and cell formation is based upon collaborative strategy for the utilisation of sufficient resources to achieve objectives. A variety of generic capability units (resource elements) are associated with sub-machine, machine, cell and factory levels to provide appropriate levels of detail for different applications. Resource elements describe the exclusive and overlapping capabilities of the available resources and hence indicate the similarity and uniqueness (scarcity and abundance) of manufacturing resources. This information is used extensively during process planning, modelling, simulation and resource allocation to the cell. The cell/factory performance is assessed and measured related to the multi-objectives it has to achieve. An example application of the resource elements in machine shop environment is shown in section 2.1.

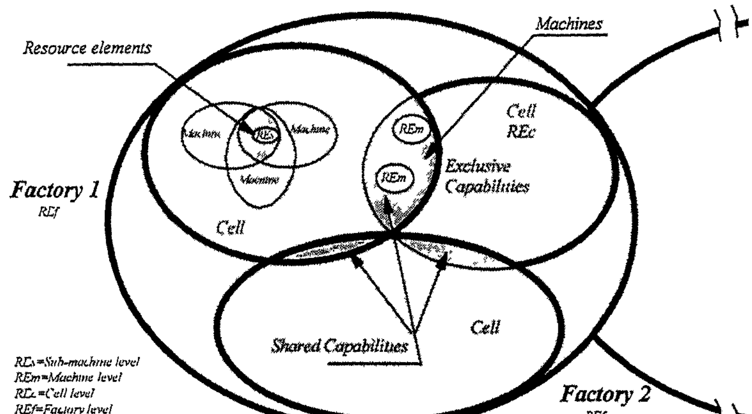


Fig. 1 : Conceptual framework

2.1 Resource Elements in Machining Environments

In a machining facility Resource Elements (RE's) are defined as facility specific capability units, which capture information relating to the distribution (commonality and uniqueness) of form generating schema among the available machine tools. A Form Generating Schema (FGS) is a technologically meaningful combination of a cutting tool of specific geometry, a set of relative motions between a part and the cutting tool, and the typical levels of technological output (surface finish, tolerances etc.) associated with using that combination of tool and relative motions. The available machine tools in a manufacturing system can be described using a set of RE's where each RE represents a collection of form generating schema such that the exclusive and the shared capability boundaries between all the available machine tools comprised in a manufacturing facility are uniquely identified. Figure 2 shows a diagrammatic representation of a machining facility based on its RE content. Full explanation of the RE concept is beyond the scope of this paper, for more details refer to Gindy *et al.* [1].

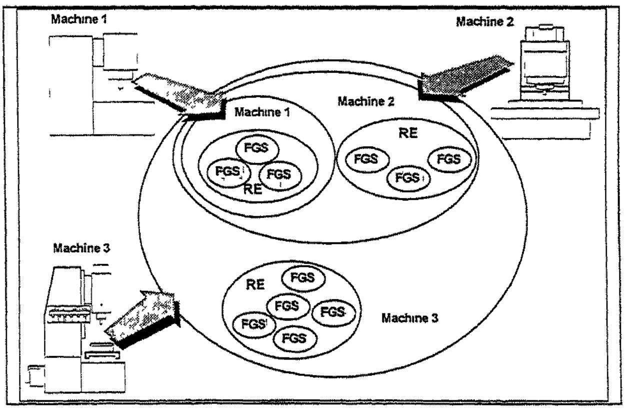


Fig. 2 : Representation of machining facility using REs

2.2. Machine Flexibility

Machine flexibility as defined in Sethi and Sethi [2] relates to the various types of machining operations that a machine tool can perform without requiring a prohibitive effort in switching from one operation to another. The measures proposed by Brill and Mandelbaum [3] are based upon the relative importance of the operations a machine executes and the efficiencies of machine tools in performing such operations. Son and Park [4] measure "machine" flexibility in terms of the capability of the machine to add value to raw materials. Das and Nagendra [5] measure "machine" flexibility as the sum of the efficiencies by which it performs different operations. In a similar treatment to Das and Nagendra's [5], the efficiency of machine tools based upon resource elements distribution can be calculated in terms of its set-up and processing time required for component manufacture. The overall "machine" flexibility in a machining facility is computed as:

$$Machf = \frac{1}{n}\sum_{k=1}^{n}\frac{1}{\sum_{j=1}^{m\in RE_k} RE_{k,j}}\sum_{j=1}^{m\in RE_k} E_{RE_{k,j}} \qquad (1)$$

where:

m = number of machines,
j = 1,2,........m,
n = number of different resource elements,
k = 1,2,.........n,
Re_{kj} = resource element k on machine j,
$E_{REk,j}$ = efficiency of resource element k on machine j,
$P_{REk,j}$ = processing time required by resource element k on machine j,

$S_{REk,j}$ = set up time required by resource element k on machine j,

The efficiency of $RE_{k,j}$ to perform the required machining becomes:

$$E_{RE_{k,j}} = \frac{(S_{RE_{k,j}})_{\min}}{S_{RE_{k,j}}} \times \frac{(P_{RE_{k,j}})_{\min}}{P_{RE_{k,j}}}$$

2.3 Load Flexibility

"Load" flexibility relates to the variation in the distribution machining tasks amongst the resources available in a manufacturing facility. A manufacturing facility which exhibit smooth and even load distribution is considered more efficient in utilising the inherent flexibility of its resources than a manufacturing environment in which resources are unevenly loaded or show bottle-necked resources (100% utilisation). Load flexibility therefore, can have a significant impact on the responsiveness of manufacturing systems in terms of its ability to cope with changes in production volumes and variety of manufactured components (demand patterns).

Load flexibility is represented as a relationship between the availability of an RE_k, required for a machining task, and the variation in the utilisation of machine tools in which RE_k is available. Machine utilisation is defined as the proportion of time that a machine is busy doing a useful task (Saad [6]). The variation in machine utilisation is measured utilising the standard deviation of machine utilisation from its mean value. Load flexibility is measured as:

$$Loadf = (1 - \frac{\sum_{k=1}^{n} U_{RE_k}}{n}) \; (1 - \sigma_{U_j}) \qquad (2)$$

where:

U_{REk} = utilisation of RE_k, (k=1, 2, ...n)
s_{uj} = standard deviation of machine utilisation. (j=1, 2, ...m)

2.4 Routing Flexibility

Routing flexibility of a manufacturing system as defined in Sethi and Sethi [2] is the ability of a manufacturing system to produce a part by alternate routes utilising the available system resources. Similar definition has been used by Das and Nagendra [5], Falkner [7], Buzacott [8] and Carter [9]. Other authors such as Buzacott [8], Browne et al. [10] and Upton and Barash [11] emphasise the ability of the system to reroute parts in case of machine breakdown. In this work routing flexibility is considered to depend not only on the number of routes available to produce a product, but also on an assessment of the availability of each alternate route and its efficiency to produce a product (similar to Zahran et al. [12]). Routing flexibility of a manufacturing system to produce a certain number of different products is given as:

$$Routf = \frac{1}{D}\sum_{d=1}^{D}\frac{1}{t_d}\sum_{i=1}^{t_{id}^{k}}\frac{1}{\sum_{j=1}^{m\in RE_k} RE_{k,j}}\sum_{j=1}^{m\in RE_k} E_{RE_{k,j}}(1-U_{RE_k}) \tag{3}$$

where:

D = number of different products,
d =1,2,................., D,
t_d = number of operations required for product d,
t_{id}^{k} = operation i for product d using resource element k.

From equation 3, it can be seen that the number of repeated resource elements, the availability of resources and the efficiency of the machine tool in carrying out the required operations are all factors that influence routing flexibility in a manufacturing system and hence its responsiveness.

3. EXPERIMENTAL INVESTIGATION

3.1 System Descriptions and Simulation Model

Figure 3 shows an overview of the prototype system for a dynamic on-line scheduling environment used in this research. Simulation models of a manufacturing facility which can be operated as a collection of machine tools or as a collection of resource elements have been built and are used to compare the conventional machine-based scheduling approach with RE-based scheduling under a wide variety of operating conditions (dispatching rules and due date assignment rules).

The simulation models used for this study are based upon simplified models of the machining facilities belonging to a large industrial company. The machine shop has 22 machine tools that perform a wide variety of machining operations e.g. turning, milling, drilling etc. Parts and materials are transported in the system by a combination of forklift trucks and cranes. Each part entering the system is routed through the machine shop according to its processing requirements. Two discrete simulation models were developed to represent the machining facility. The first is a conventional machine-based representation in which parts are dispatched to machine tools according to their machine-based routes. In the second model, the same machining facility is represented as a collection of resource elements operated as a manufacturing cell.

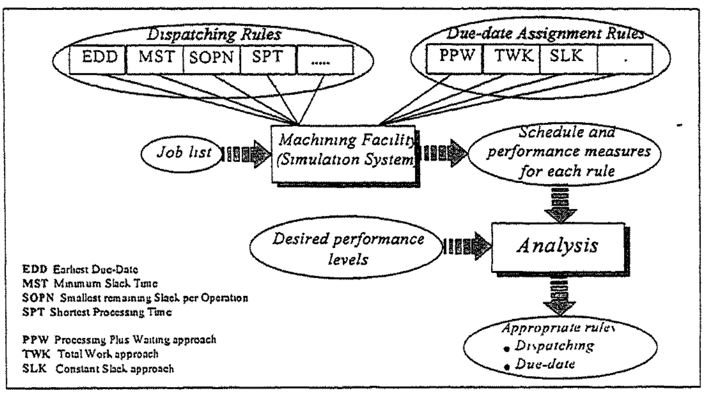


Fig. 3 : Scheduling environment

3.2 Dispatching Rules

Three, commonly used, dispatching rules were investigated in this study. Two of the selected rules use due-date information during scheduling and the third is based upon processing time information alone as detailed in Table 1.

Name	*Symbol*	*Description*
Earliest Due-Date	EDD	Select the job which has earliest due-date: $x_{i,j} = D_i$
Minimum Slack Time	MST	Gives priority to the minimum slack time : $x_{i,j} = D_i - t + \left(P_i - \sum_{l=0}^{j} p_{il}\right)$
Smallest remaining Slack per OperatioN	SOPN	Gives priority to the minimum remaining slack per operation: $x_{i,j} = \frac{D_i - t}{M_i - m_i(t)}$
Shortest Processing Time	SPT	Select the job that has smallest operation time: $x_{i,j} = p_{i,j}$

Table 1. Machine dispatching rules.

where:

I = job index,
J = operation index,
D_i = the due date of job *i*,
$p_{i,j}$ = the processing time required for the *j-th* operation of job *i*,
$x_{i,j}$ = priority index of the *i-th* job at the *j-th* operation,
T = present time,
P_i = total processing time of job *i*,
M_i = total number of operations of job *i*,
$m_i(t)$ = total number of operations of job *i* already done at time *t*.

3.3 Due-date Assignment Approach

Due-date assignment is an important factor influencing the performance and profitability of manufacturing companies. It normally serves as a basis for production planning and control. Several studies are reported in the literature which attempt to assess the applicability of specific due date assignment rules under specific operating conditions. For example, Kanet and Christy [13] compared two methods, TWK and PPW (see table 2) for setting manufacturing lead time in a general job shop when early shipment of completed jobs is not allowed. They reported that TWK appears to be the superior method when the performance criterion is mean tardiness. In practice, however, due-dates are normally decided based upon: 1) Negotiation with customers, 2) Market forces (competition) and 3) The constraints placed on the production control system due to job priorities, ability to deal with disturbance, capacity constraints, etc. Table 2 contains the three commonly used due-date assignment approaches used in this study.

Name	*Symbol*	*Description*
Processing Plus Waiting approach	PPW	Set job's allowance equal to the sum of its total processing time plus an allowance for non-productive inter-operation activities such as waiting in queue, waiting for transfer and transfer: $PPW_i = P_i + \frac{\bar{P}}{\bar{M}}(k-1) \times M_i$
Total WorK approach	TWK	Set job's allowance equal to the sum of its total processing time plus an allowance proportion to the total work content: $TWK_i = k \times P_i$
Constant SLacK approach	SLK	Set job's allowance equal to the sum of its total processing time plus a constant slack calculated based on the mean processing time of all jobs in the system: $SLK_i = P_i + \bar{P}(k-1)$

Table 2. Due-date assignment approaches

(where: k = due-date tightness, $\bar{P}$ = mean processing time and $\bar{M}$ = mean number of operations)

3.3 Manufacturing Responsiveness

In this work Manufacturing Responsiveness is used as a general term to refer to the ability of a manufacturing system to make a rapid and balanced response to the predictable and unpredictable demands of the manufacturing environment. It is argued that the root to improving the responsiveness lies in maximising the utilisation of the inherent flexibility of its available resources in order to: (1) achieve the "best" possible operational performance in terms of meeting performance targets while coping with unpredictable internal and external disturbances and (2) operate the manufacturing system such that the allowances added to product processing time are minimised (tightest possible due dates). Setting due dates "too tight" normally lead to deterioration in system performance measures (increasing mean tardiness and average flow time etc.). A responsive manufacturing system however, is a system that can be operated such that its objectives are achieved at the tightest due dates without compromising the confidence in being able meet the set due dates (minimising tardiness). There are several facets and possible measures of manufacturing responsiveness that are currently under investigation in this research. The ability to set tight due dates however, is one of the obvious manifestations of a responsive manufacturing system. In the reported work the level of due date tightness k refers to the allowances added to the processing times of components which leads to zero mean tardiness performance in the manufacturing system. K is taken as a simple measure for comparative assessment of the responsiveness of manufacturing systems. An increase in the value of K indicates a less system responsiveness.

4. SIMULATION RESULTS AND ANALYSIS

The purpose of the simulation experiments conducted is to provide performance comparison for a machining facility scheduled conventionally (machine-based) and the same facility scheduled based upon resource elements under a wide variety of dispatching and due date assignment rules and to compare the responsiveness of both systems. Two parameters were considered in the experimental design, dispatching rules (see Table 1) and due-date assignment rules (see Table 2) at a variety of due date tightness levels. Two performance indicators, mean tardiness (MT) and average flow time (AFT) were used to assess operational system performance. The system's ability to cope with internal and external disturbances was also examined under uniform and skewed load distributions at various machine breakdown rates.

4.1 Operational Performance

As can be seen from the Figure 4 and 5, for all the dispatching rules and the due-dates approaches investigated, the *RE*-based system achieved a superior performance in comparison with the machine-based system. Moreover, *RE*-based system is much less sensitive to the selection of dispatching rules and due-date assignment approaches. Mean tardiness level is considerably higher in machine-based compared with the than *RE*-based system and the performance of the machine-based system is very sensitive to the selection of dispatching rules and due-date assignment approaches. Tardiness is inversely proportional to due-date tightness and therefore, the mean tardiness performance of the machine-based system is lower when the TWK approach is employed in combination with EDD or SPT as the dispatching rules. This observation confirms the earlier research results from Kanet and Christy [13]. This is not the case in the resource element-based system where PPW gives better results in combination with EDD and SPT. Due to space restriction, the detailed results will be presented at the conference (also refer to Gindy and Saad [14]).

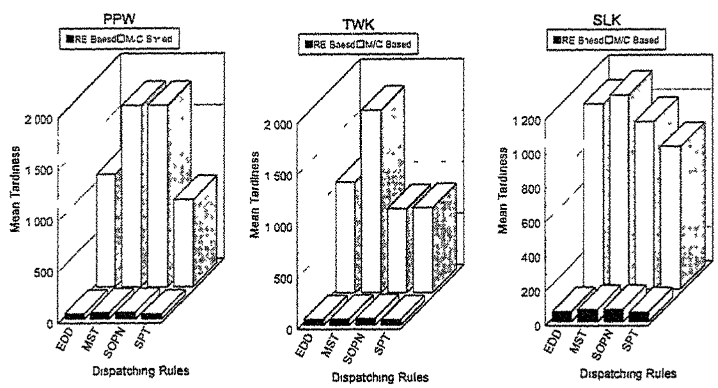


Fig. 4 : Mean tardiness (RE-based Vs M/C-based) at due-date tightness k=2

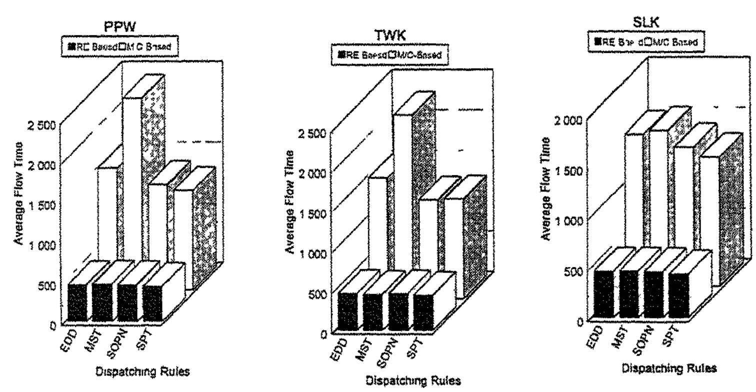


Fig. 5 : Average flow time (RE-based Vs M/C-based) at due-date tightness k=2

4.2 Resource Utilisation

As can be seen from Figure 6, the average utilisation of system resources under investigation in the RE-based system is higher (40%) than the machine-based system (22%) In the machine-based system inspite of low average utilisation of system resources, some key resources are utilised 100% which indicates some bottle-necked resources (machine 13). In the RE-based system however no resources are utilised more than 82%.

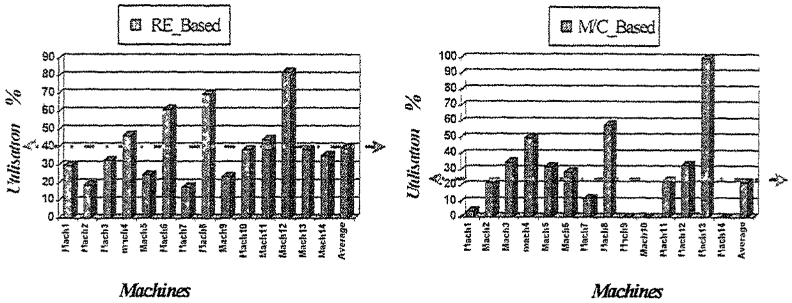


Fig. 6 : Load distribution in case of RE-based and machine-based systems

4.3 Demand Pattern

Changes in demand patterns is considered as one of the external disturbances that can influence manufacturing system performance. Figure (7) shows the relationship between average flow time taken as a performance measure and the effect of changing the demand on the manufacturing facility from uniform to skewed load distribution. It can be observed from the figure that the average flow time for components in the system is lower in the RE-based system compared with the machine-based scheduled facility and that the RE-based system is much more capable of coping with changes in demand pattern.

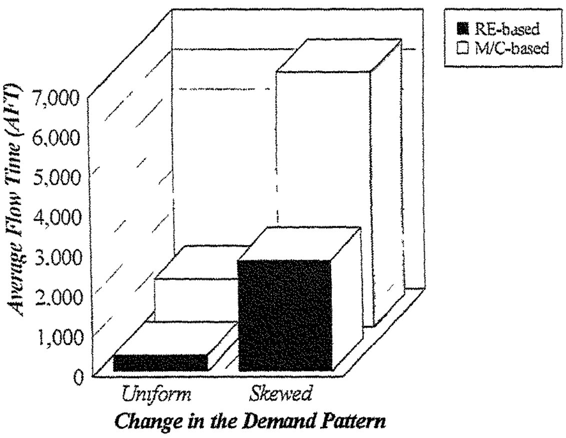
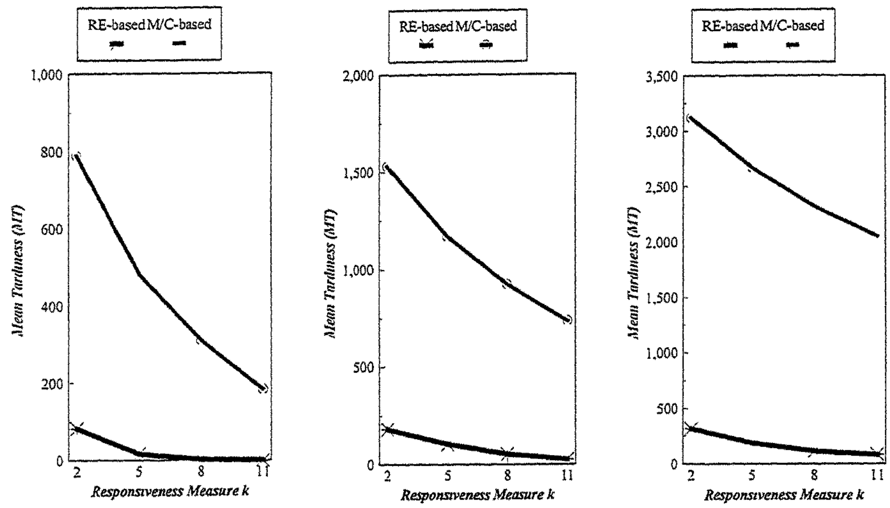


Fig. 7 : The effect of the demand pattern on the average flow time

4.4 Machine Breakdown

The system ability to cope with internal disturbances such as machine breakdown was also examined under a wide variety of operating conditions. Figure 8 shows the relationship between responsiveness measure k and the system's mean tardiness (MT) at various machine breakdown rates. As expected, system responsiveness deteriorates as the machine break down rate increases. However, in the RE-based system the rate of deterioration in system responsiveness is lower and hence it is much more able to cope with disturbances.



a) No break down *b) 5 % FR* *c) 10 % FR*

Fig. 8 : Mean tardiness performance of the two proposed systems (at PU=60%, Uniform load)

4.5 Manufacturing Responsiveness

Figure 9 shows system performance (mean tardiness) at various levels of due date tightness (K). For the RE-based system the point at which the mean tardiness approximates zero is at K= 5 for all due date assignment approaches used. For the machine-based system the same value for K (at which mean tardiness approximates zero) varies depending on the due date assignment approach used and is significantly higher (k=15.19 for TWK, k=15.7 for SLK and k=18.79 for PPW) compared with the RE-based system. As stated in section 3.3, the K value that achieves zero mean tardiness is taken as a simple comparative measure for the responsiveness of manufacturing systems. It can be seen therefore, that responsiveness improves significantly if manufacturing facilities are represented and scheduled based on the resource elements concept. For the same levels of performance an RE-scheduled system allows much tighter assignment of due dates without compromising the confidence in meeting these due dates.

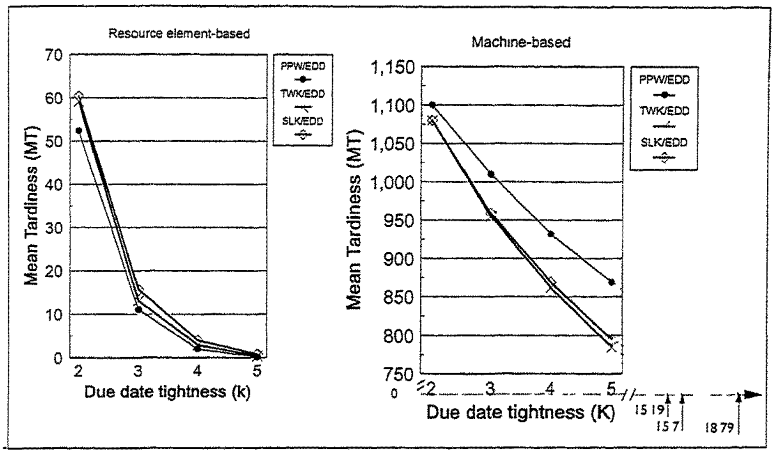


Fig. 9 : System responsiveness

5. CONCLUSIONS

Considerable experimental work has been carried out to investigate and compare manufacturing system performance achieved using resource element-based scheduling and conventional machine-based approaches. The comparison is performed using a variety of due-date assignment approaches, dispatching rules and due-date tightness levels. The results clearly indicate that significant improvements in system performance measures and responsiveness can be achieved when a machine shop is represented as a set of generic capability units. The research results to date can have far reaching implications for the operation of manufacturing systems. It shows that adopting the resource elements concept when representing and scheduling machining facilities can lead to significant improvements in system performance, achieve higher responsiveness in terms of setting and meeting tight due dates and improves the system's ability in dealing with internal disturbances (e g. machine breakdowns) and external disturbances (e.g. changes in demand patterns).

6. ACKNOWLEDGEMENT

The reported research is supported by the UK EPSRC research grant ref.: GR/J 90022. Their financial contribution is gratefully acknowledged.

7 REFERENCES

1. Gindy N. N. Z., Ratchev T. M., and Case K., 'Component Grouping for Cell Formation Using Resource Elements', International Journal of Production Research, Vol. 34, No. 3, pp. 729-759, 1995.
2. Sethi A. K, and Sethi S. P., 'Flexibility in Manufacturing A Survey', The international Journal of Flexible Manufacturing Systems, Vol. 2, pp. 289-328, 1990.
3. Brill, P. H., and Mandelbaum, M., 'On Measures of Flexibility in Manufacturing Systems', International Journal of Production Research, Vol. 27, No. 5, pp. 747-756, 1989.
4. Son Y. K., and Park C. S., 'Economic Measure of Productivity, Quality and Flexibility in Advanced Manufacturing Systems', Journal of Manufacturing Systems, Vol. 6, No.3, pp. 193-207, 1987.
5. Das S. K. and Nagendra P. (1993) 'Investigation into the Impact of Flexibility on Manufacturing Performance', International Journal of Production Research, Vol. 31, No. 10, pp. 2337-2354,

1993.
6. Saad, S. M., Design and analysis of a flexible hybrid assembly system. PhD thesis, Department of Manufacturing Engineering and Operations Management, University of Nottingham, England, 1994.
7. Falkner, C. H., 'Flexibility in manufacturing plants', Proceedings of the Second ORSA/TIMS Conference on Flexible Manufacturing Systems: Operations Research Models and Applications, edited by K. E. Stecke and R. Suri, pp. 95-106, Elsevier Science Publishers B.V., Amsterdam, 1986.
8. Buzacott J. A., 'The Fundamentals of Flexibility in Manufacturing' Proceedings of the First Int. Conference (Brighton, UK), Elsevier, pp. 13-22, North Holland, Amsterdam, 1982.
9 Carter M. F , 'Designing Flexibility Into Automated Manufacturing Systems', Proceedings of the Second ORSA/TIMS Conference on Flexible Manufacturing Systems: Operations Research Models and Applications, edited by K. E. Stecke and R. Suri, pp. 107-118, Elsevier Science Publishers B.V., Amsterdam, 1986.
10. Brown J., Dubois D., Rathmill K., Sethi S. P., and Stecke K. E., 'Classification of Flexible Manufacturing Systems', The FMS Magazine, pp. 114-117, 1984.
11. Upton, D. M., and Barash, M. M., 'A grammatical approach to routing flexibility in large manufacturing systems', Journal of Manufacturing Systems, Vol. 7, pp. 209-221, 1988.
12. Zahran, I. M., Elmaghraby, A. S., and Shalaby M. A., 'Evaluaion of flexibility in manufacturing systems', IEEE International conference on systems, MAN, and Cybernetics, Los Angeles, CA, USA, 1990.
13. Kanet J J.and Christy D. P., 'Manufacturing Systems with Forbidden Early Shipment: Implications for setting manufacturing Lead Times', International Journal of Production Research, Vol. 27, No. 5, pp 783-792, 1989.
14. Gindy N. N., and Saad S. M., 'Factory Model Implementation, Experimental Investigation of Resource Element-Based and Machine-Based Scheduling', Technical Report No. 2-2 EPSRC/GR-J90022, 1995.

THE 21ST CENTURY MANUFACTURING AUTOMATION AND THE SOCIAL IMPACT

Tosiyasu L. Kunii
Laboratory of Digital Art and Technology
1-25-21-602 Hongo
Bunkyo-ku, Tokyo 113, JAPAN
tlk@mbp.com

ABSTRACT

Manufacturing automation has lead the global social changes going across the three centuries since its birth in Yorkshire. The 21st century is expecting drastic advances in manufacturing automation with its social impact in several orders of magnitude greater than the predecessors. This research clarifies its nature and direction.

KEYWORDS.

industrial revolution, mechatronics, flexible manufacturing, process graph

1. THE NATURE OF THE 21ST CENTURY INDUSTRIAL REVOLUTION

Let us begin to understand the nature of the society we live. It is usually called a modern society. It is generally a type of society whose advances are based on the power of the totality of all the members and, conversely, all the members are benefited by the social advances, sometimes social evolution and sometimes social revolution. The prevailing of such a society on the earth was initiated by manufacturing automation born in Yorkshire two and half centuries ago. The move named the industrial revolution, spread out to the continental Europe and the United States within a century, and has continued on. The automation brought the industrial revolution was based on mechanics. We can view it as the automation of manual manufacturing by muscle power, or the muscle power automation in short. The muscles are controlled by the nerves. For more sophisticated automation, therefore, nerve control automation was required. Its realization by mechatronics, a word coined in Japan, meaning the electronic control of mechanics, took place around half century ago in Japan, caused the second industrial revolution and has propagated worldwide within a few decades. The mechatronics is now the most common and basic technology in manufacturing automation.

What, then, will be the forthcoming manufacturing automation toward the next century? Since the nerves are commanded by the brain, the answer is clear and definite. It is brain automation. The brain performs intellectual and emotional functions. Brain automation is, therefore, intellectual and emotional automation. The only intellectually functioning automated machine is a computer. A computer functions basically as a logical automated machine. All intellectual functions have been converted to logical functions inside a computer automatically. An emotionally functioning computer has been tried to be realized in many ways. More comfort is one of the expectations of emotional automation. An artistically sophisticated quality product including a multimedia content product is an example of a type of product by emotional automation. Artistic sophistication is a result of expert level performance. We have identified an expert technique as a singularity in a martial art after several years of research [1]. We expect the result to be used to control a group of multi joint robots to perform expert level artistic work. The emotional brain has been studied for long [2]. We also have been studying emotional facial expressions, automated capturing of them in particular, for years as the first step of realizing direct emotional communication systems[3-5]. Artificial intelligence is the general area dedicated to understanding and artificially realizing of intelligence. Our concern is rather focused, technological and social in its nature. From a manufacturing automation view point, what we matter is the fact that computer commanded mechatronics will be the core technology of the third industrial revolution in the 21st century.

What called flexible manufacturing (FM), small quantity multiple type production automation and on demand manufacturing for electronic commerce, all belong to simple applications of the technology.

2. A NEW AUTOMATED COMPONENT SHAPE IDENTIFICATION METHOD FOR FLEXIBLE MANUFACTURING

Flexible manufacturing means a type of automated manufacturing where different products are manufactured usually in relatively small quantity on demand through an identical manufacturing line. This requires an efficient method to automatically identify varieties of components of different products. The identification is usually by the component shapes. The components flow through a manufacturing line, and for easy set up of the components on the manufacturing line, the components are placed not necessarily in fixed orientations or locations. The identification of the shapes of thus arbitrarily oriented located and oriented components poses a difficulty in finding an efficient method.

An efficient method of orientation- and location- independent identification is realized only by finding a small set of orientation- and location- independent invariants of shapes. As such, we can utilize the curvature of shapes as shown previously [6]. Given the curvature of a shape, we can characterize it by characteristic points such as peaks, pits and passes as a small set of orientation- and location- independent invariants. This means a drastic increase in the efficiency of the shape identification. For all the components, we store the sequences of characteristic points. Then we can identify any components flowing through the manufacturing line by comparing small sets of symbols, namely the sequences of characteristic points. Further, we can usually reduce the identification problem of this type to the template matching problem of circular codes consisting of characteristic points symbols.

As a simple example we showed a case of the identification of animals from their silhouettes in Figures 1-6.

The generalization of the example to higher dimensional cases are under study.

3. EFFICIENCY ENHANCEMENT OF MANUFACTURABILITY TEST

Before trying to manufacturing any products we have to impose manufacturability test to it. The test is usually an extremely complex process. Decreasing the complexity is essential for efficiency increase. A very simple but generally applicable principle is shown here to automate the test by a computer.

The object to be manufactured has a shape. The shape consists of a solid body or bodies, and a surface or surfaces. The manufacturing is done by accessing the surface as

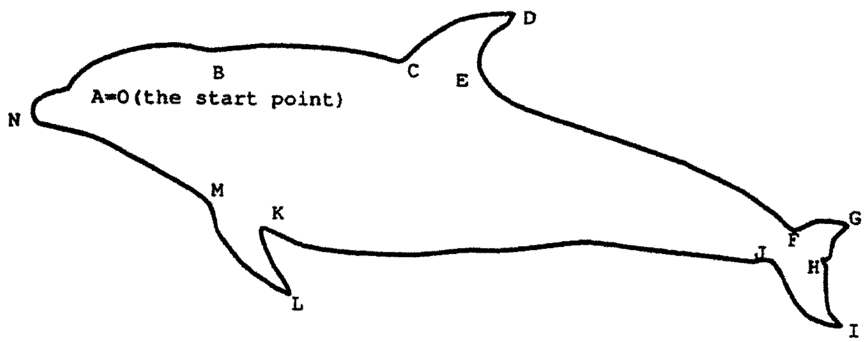


Fig. 1 : The outline curve of a dolphin silhouette.

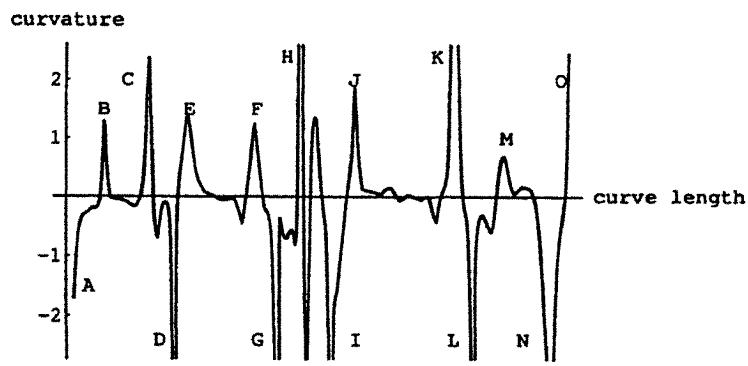


Fig. 2 : Curvature along the curve.

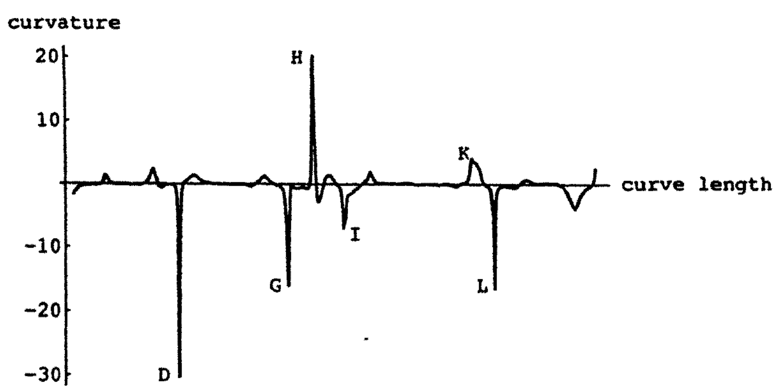
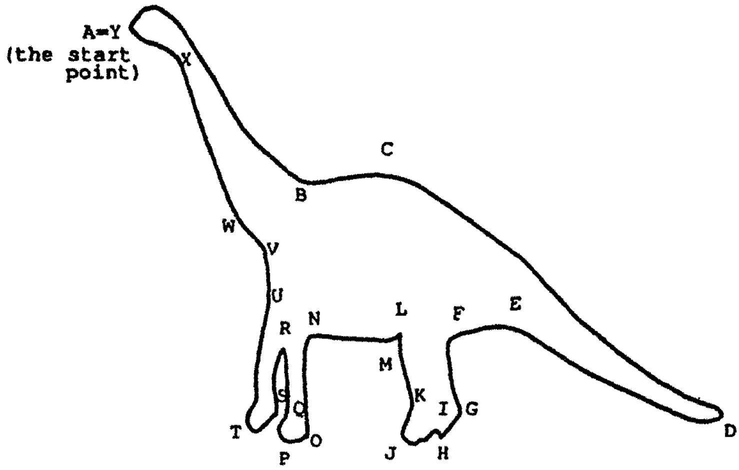


Fig. 3 : Curvature along the curve.



The outline curve of a dinosaur.

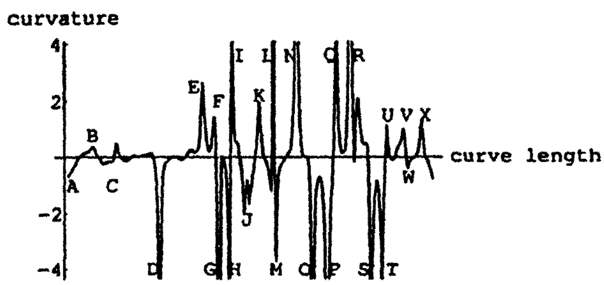


Fig. 5 : Curvature along the curve.

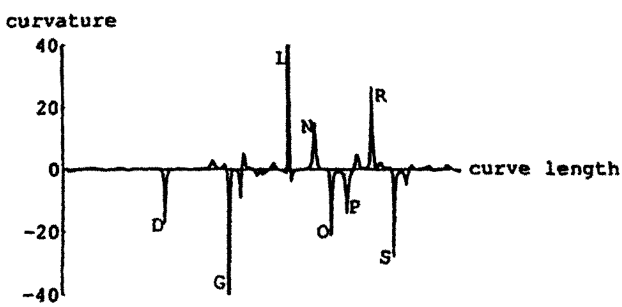


Fig. 6 : Curvature along the curve.

needed. Therefore, for the object to be manufacturable, the surface has to be accessible. We denote this by

manufacturability $\subset$ accessibility.

Hence,

manufacturability $\subset$ accessibility $\subset$ visibility.

By using this relationship, we can decrease the manufacturability test cases. First, we conduct the visibility test. Since the non visibility means non accessibility, we need to conduct the accessibility test on the visible cases only. Next, since the non accessibility means non manufacturability, we need to conduct the manufacturability test on the accessible cases only.

4. ASSEMBLY EFFICIENCY ENHANCEMENT

Given a product consisting of n components, the assemblability test complexity can be drastically decreased in the orders of magnitude. We found this in 1991[7,8], and the conclusion is summarized as follows.

The worst case complexity of the assemblability test is in the order of 2^n if the test is to be conducted against the n components. Instead, the worst case complexity of the assemblability test is decreased to the order of n^2 if the test is conducted by disassembling the assembly in an assembling way.

Just to see the meaning of the decrease of the complexity, let us take a case where n=50 and the testing time of a pair of components is 1 micro second. 2^n is 30 years while n^2 is only 3 milli seconds.

Further efficiency increase is attained by imposing the previous rule to the assemblability test :

assemblability $\subset$ accessibility $\subset$ visibility.

5. MANUFACTURING AUTOMATION DRIVEN SOCIAL STRUCTURE CHANGES

Manufacturing automation has changed the social structure drastically. Manufacturing automation has made it possible for increasing number of social commodities to be mass produced. Mass production has brought in mass sales. The social power structure change has taken place nationally and internationally by making it possible for all people, irrespective of their current social status, to join mass production and mass sales activities. Both the national and international powers have been governed by the assets created by manufacturing automation. The flows of the assets as the material and money flows have been basically controlling the world power structure. Information has been governing the control[9,10]. As explained earlier, the third industrial revolution is through the automation of the information governed control as well as the automation of processes.

6. THE LEVELS OF FACTORY MANUFACTURING EFFICIENCY INCREASE

Given a factory, let us see that there are 3 levels of incremental manufacturing efficiency increase methods at the increasing costs. Usually, a factory consists of manufacturing machines that perform manufacturing processes, laid out on the floor.

The least cost and the least efficiency increase method is realized as a material flow tuning method that tunes the material flow only. When more efficiency increase is desired, the layout tuning method is applied as the second least cost effective method that changes the floor layout of the manufacturing machines while the material flow is tuned. Further efficiency increase is achieved by the machine updates while the floor layout and the material flow are tuned; for simplicity we call the method a machine update method. This is the most expensive method.

7. CORPORATE REENGINEERING, BUSINESS RESTRUCTURING, AND MANUFACTURING AUTOMATION

Corporate reengineering and business restructuring are, as a matter of fact, essentially identical to the efficiency increase methods at a factory level. They all belong to a process graph usually represented as a DAG (directed acyclic graph) as shown in Fig. 7. Under a title "Beyond Reengineering" Michael Hammer proposes the "process centered organization"[11]. This means manufacturing automation, once understood well, becomes a universal discipline going across organizational restructuring, business automation and financing automation. Through the advances of computer mediated global communication, further global automation will take place in the 21st century at almost all the scenes of our life. Then, the DAG representation will serve as a crucial tool to understand the complexity of global automation.

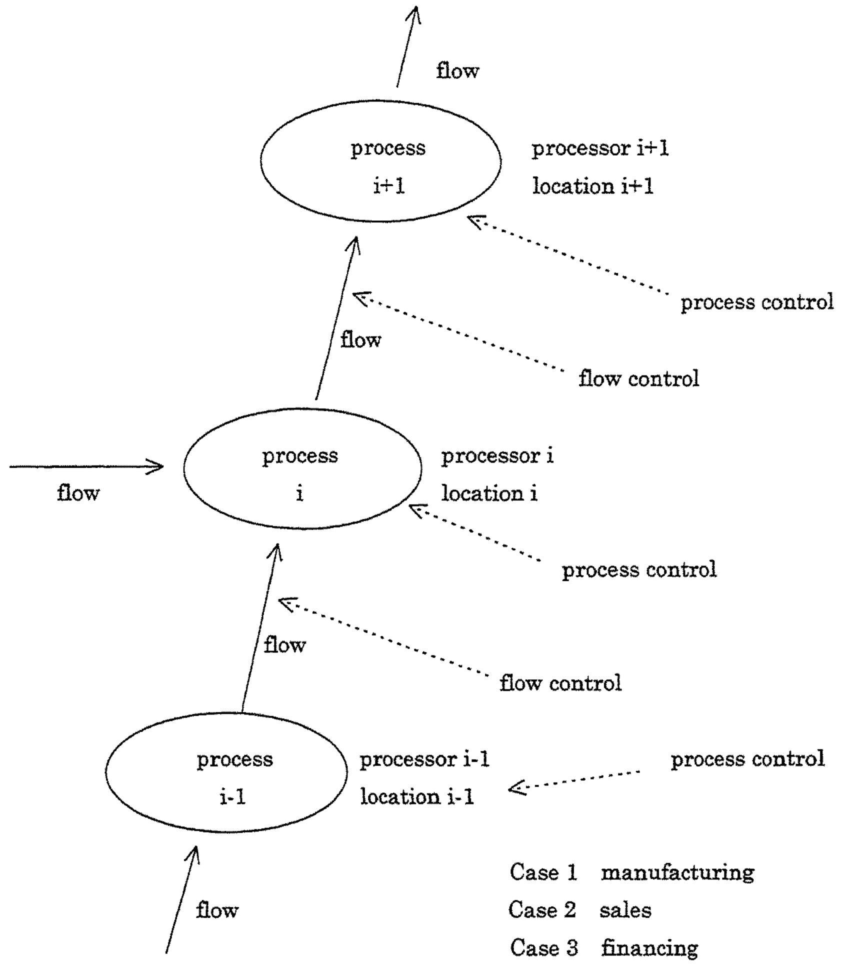
By extending the DAG representation to a hierarchical graph representation [12] incorporating power sets, we can decrease the complexity at the view level because of the capability of a hierarchical graph to present process graphs at the abstraction level or the aggregation level matching the given view level.

8. CONCLUSIONS

Towards the truly global society in the 21st century, a series of methods presented so far in this paper aim at global automation that can be generalized in an integrated way to automate the entire processes taking place in the modern society, such as automated manufacturing linked with electronic commerce that represents automated commerce, also with automated financing and even with automated financial trading. The third industrial revolution to come in the 21st century is, thus, not just extremely global but also unimaginably deep in changing the social structure. Intensive researches are essential to understand, to prepare for and participate in it. We manufacturing automation specialists are, as I pointed out through a process graph as a DAG, qualified to undertake the responsibility of the heavy integration of manufacturing, commerce and finance. The complexity of the whole problem of integrated manufacturing automation as described here requires the advanced engineering approach to deal with the total problem in a systematic and consistent way.

9. ACKNOWLEGMENTS

This research has been carried out for a few decades. Many people has given me various opportunities to look into the sites of factory automation such as Ricoh office machine manufacturing factories and Amada automated machine pavilions. I am grateful to Professor S. T. Tan of the University of Hong Kong for giving me an opportunity to think over the whole automation problems by inviting me to prepare this paper. MOVE Consortium and the University of Aizu have granted the partial support to this research while the author was affiliated with the University of Aizu. The Shorinji Martial Art Headquarters, Mr. Hiroshi Matsuda and Mr. Shinya Miura in particular, has contributed significantly to the understanding



process ::= case 1 manufacturing process
case 2 sales process
case 3 financing process

flow ::= case 1 manufacturing material flow
case 2 merchandise flow (e.g. logistics)
case 3 financial flow

Fig. 7 : A process graph

of human expert techniques. The assemblability discrimination method was developed through the partial support of Ricoh Software Research Center directed by Dr Hideko S. Kunii.

10. REFERENCES

1. Kunii, Tosiyasu L., Tsuchida, Yukinobu, Arai, Yasuhiro, Matsuda, Hiroshi, Shirahama, Masahiro and Miura, Shinya, "A Model of Hands and Arms Based on Manifold Mappings," in Communicating with Virtual Worlds (Proceedings of CG International '93), June 21-25, Lausanne, Switzerland, Thalmann, Nadia Magnenat and Thalmann, Daniel (eds.) pp.381-398, Springer-Verlag, Tokyo, 1993.
2. LeDoux, Joseph, "The Emotional Brain: The Mysterious Underpinning of Emotional Life", Simon and Schuster, New York, 1996.
3. Saji, Hitoshi, Hioki, Hirohisa, Shinagawa, Yoshihisa, Yoshida, Kensyu and Kunii, Tosiyasu L., "Extraction of 3D shapes from the moving human face using lighting switch photometry", in Creating and Animating the Virtual World (Proc. Computer Animation '92), Nadia Magnenat-Thalmann and Daniel Thalmann, (eds), pp. 69-86, Springer-Verlag, Tokyo, 1992.
4. Saji, Hitoshi, Shinagawa, Yoshihisa, Takahashi, Shigeo, Hioki, Hirohisa and Kunii, Tosiyasu L., "Measuring Three-Dimensional Shapes of Human Faces by Incorporating Stereo Vision with Photometry Using Blending Functions", in Fundamentals of Computer Graphics (Proc. of Pacific Graphics '94), (August 26-29, 1994, Beijing, China), Chen, Jiannan, Thalmann, Nadia Magnenat, Tang, Zesheng and Thalmann, Daniel (eds.), pp. 3-18, World Scientific, 1994; also available in Proceedings of the 4th International Conference on Computer-Aided Drafting, Design and Manufacturing Technology (CADDM'94), Tang, Rongxi (chief editor) pp.3-10, International Academic Publishers, 1994.
5. Saji, Hitoshi, Shinagawa, Yoshihisa, Kunii, Tosiyasu L., Hioki, Hirohisa, Hara, Kazuhiro, Asada, Noriaki and Yasumoto, Masato, "Characterization of Object Shapes by Singular Points: with an Application to Feature Extraction of Human Facial Expressions", in Visualization and Intelligent Design in Engineering and Architecture (Proc. VIDEA'93), pp.29-43, Computational Mechanics Publications and Elsevier Science Publishers, 1993.
6. Kunii,Tosiyasu L. and Maeda, Takao, "On the Silhouette Cartoon Animation", in Proceedings of Computer Animation '96, (June 3-4, 1996, Geneva, Switzerland), Thalmann, Nadia Magnenat and Thalmann, Daniel (eds.), pp.110-117, IEEE Computer Society Press.
7. Kunii, Tosiyasu L., Noma, Tsukasa, and Lee, Kyu-Jae, "Assemblability Discriminating Method and Assembling Sequence Generating Method", United States Patent No. 5058026, October, 1991.

8. Kunii, Tosiyasu L., Noma, Tsukasa and Lee, Kyu-Jae, "SYDEM: A New Approach to Computer-Aided Design of Assemblies and Assemblability Testing", in Visual Computing: Integrating Computer Graphics with Computer Vision (Proc. CG International '92), Kunii, Tosiyasu L. (ed), pp. 469-479, Springer-Verlag, Tokyo, 1992.
9. Kunii, Tosiyasu L., Pax Japonica (in Japanese), President Co., Ltd., Tokyo October, 1988.
10. Kunii, Tosiyasu L., "Creating a New World inside Computers -- Methods and Implications --", in ASCILITE '89 (Proc of the Seventh Annual Conference of the Australian Society for Computers in Learning in Tertiary Education), G. Bishop and J. Baker (eds.), pp. 28-51, (Gold Coast, Australia, December 11-13, 1989); also available as Technical Report 89-034, Dept. of Information Science, Faculty of Science, The University of Tokyo.
11. Hammer, Michael, Beyond Reengineering: How the Process-Centered Organization Is Changing Our Work and Our Lives, Harper-Collins, New York, 1996.
12. Kunii, Tosiyasu L. and Harada ,Minoru, "SID: A System for Interactive Design ", in Proceedings of National Computer Conference 1980, AFIPS Conference Proceedings, Vol. 49, pp. 33-40, AFIPS Press, Arlington, Virginia, 1980.

THE CURRENT STATE OF AFFAIRS IN DIMENSIONAL TOLERANCING: 1997a

Herbert B. Voelcker
The Sibley School of Mechanical & Aerospace Engineering
CORNELL UNIVERSITY
Ithaca, NY, USA

ABSTRACT

Man-made artifacts are inherently imprecise. Designers control variability by specifying allowable limits on variations through dimensional and material-property *tolerances*. Manufacturing engineers honor these limits through judicious selection of manufacturing processes, control of these processes, and inspection procedures that test conformance to design tolerances.

Major strides have been made over the past thirty years in developing mathematical foundations and practical techniques for modeling nominal (ideal-form) mechanical parts and assemblies, but – surprisingly – there has been almost no development of effective mathematical theory and formal techniques for handling form *variations*. Tolerancing continues to be handled through a collection of largely manual, *ad hoc* techniques that evolved from two centuries of machine-shop and drafting-room practice. This will change soon, however, because computerization of design and manufacturing is driving us into a new era of mathematically based tolerancing.

This paper summarizes the evolution of mechanical tolerancing practices, the general character of tolerances as they are currently understood, the current state of tolerancing technologies, and the recent surge of activity aimed at rationalizing and 'mathematizing' form tolerancing. The paper concludes with two examples that illustrate current research frontiers.

KEYWORDS

parametric | geometric | statistical tolerance, dimensional variability | tolerance | metrology

1. INTRODUCTION

Manufacturing and assembly processes are inherently imprecise, and thus the artifacts they yield vary in form, material properties, and performance. Mechanisms for accomodating and controlling variability are woven through the entire production system. *Tolerances* are the primary variation-control tool in design. Control in manufacturing is exercised through process selection, process control, and inspection procedures – all aimed at meeting the tolerances set in design. This paper is focused mainly on spatial-form and -relationship tolerances. These are commonly called 'dimensional tolerances'; they influence material-property and performance tolerances in often subtle ways.

Major strides have been made over the past thirty years in developing mathematically sound techniques for modeling nominal (ideal-form) mechanical parts and assemblies, and modern CAD systems exploit some of these techniques heavily. Surprisingly, there has been almost no development of effective mathematical theory and formal techniques for handling form *variations*. Tolerancing continues to be handled through a collection of largely manual, *ad hoc* techniques that evolved from two centuries of machine-shop and drafting-room practice. However, in the 1980s computerization of design and manufacturing exposed serious flaws in this edifice of codified practice, and this triggered a burgeoning scramble to

- 'mathematize' current tolerancing practices insofar as possible [Walker 94];
- build an underlying mathematical theory of tolerancing that (1) reflects the 'physics' of the effects of form variations, (2) is mathematically consistent [Requicha 93], and (3) is consistent with current practices, because industry has large investments in product designs, worker training, production machinery, and operating procedures that are keyed to current practices [Voelcker 93];
- evolve appropriate custodial communities, notably
  - a research community to continue development of tolerancing fundamentals and techniques,
  - teaching communities to educate engineers and train technicians, and
  - links-to-practice communities of technology vendors and standardization authorities, to promote and disseminate rational, standard methods on a worldwide basis [Bennich 94].

We stand today at an unusually interesting turning point in technical history, because an important clump of technologies is poised to move over the next 10–20 years from a base of codified practice to a new base grounded in science and mathematics, with minimization of industrial disruption as a primary constraint.

This paper summarizes the evolution of form tolerancing and its surrounding culture, the general character of tolerances as they are currently understood, and the state of the major tolerancing technologies in use today. The paper concludes with two examples – one centered on statistical tolerancing, the other on assembly modeling – that illustrate current research frontiers. An appendix summarizes a recent major reorganization of the ISO committees responsible for international tolerancing and metrology standards.

2. HISTORY & CULTURE

Figure 2-1 summarizes many of the key advances in symbolic form specification[1], variation control, and measurement technology. The left column denotes major eras in the evolution of design and manufacturing: the era of mechanization, in which guided cutters and powered tools replaced human dexterity and muscle, the era of product proliferation, in which ever more precise processes and part interchangeability replaced small-lot production by artisans with 'mass production' based on highly specialized machines and workers, and finally, the current era of automation, in which human sensory and reasoning skills and specialized machines are being replaced with versatile automata.

ERAS		*SYMBOLIC FORM SPECIFICATION*		*VARIATION CONTROL & MEASUREMENT*
Mechanization		*Sketching*		
Guided cutters Powered tools	1637	Algebraic Geometry } R. Descartes	< 1800	{ Proportional dividers Crude absolute scales
↕	1795	Descriptive Geometry } G. Monge	1750-1800	} Vernier & micrometer principles
		⇒ ***Drafting***	1820	Functional gages
Proliferation				⇒ ***Interchangeability***
Precision Interchangeability Mass Production		Dimensioned Drawings ←	1850	Precise shop-floor measurement
		→	1900	Toleranced drawings
↕	1930's	Drafting Standards	1940	Geometric tolerances
	1950	Computers; NC Machines	1950	***Process independence***
	1960	C-aided drafting	1960	Statistical tolerances
Automation	1970	Wireframe CAD	1970's	CMM's
	1980	Solid modeling; IGES	1980's	'Metrology Crisis'
	1990	'Feature' modeling; STEP Assembly modeling	1990's	'Mathematization' of tolerances

Figure 2-1: Historical highlights.

Some points of special significance[2] ...

- Sketching stretches back to antiquity. Engineering drafting is attributable to Gaspard Monge's treatise on descriptive geometry, which traces its formal lineage to Descarte's algebraic geometry.
- Interchangeability – that is, making parts uniform enough to enable products to be assembled without selective fitting – was sought actively in the early 1800s and attained gradually through *gaging technology*. (A gage – or gauge – is a mechanical artifact that simulates a mating-part feature at a worst-case condition. A simple example: a precisely sized plug that does or does not fit into a

[1] There are three major approaches to form specification: physical modeling, procedural methods (specifying an artifact through a recipe for producing it), and symbolic methods – predominantly drafting and computer modeling. Symbolic methods are superior to the others in almost every respect.

[2] See [Booker 79], [Hounshell 84], and their references for more complete accounts.

controlled hole. Gages yield binary 'good', 'bad' data rather than continuous measures of deviations from a target value.)

- Drawings carried no dimensions until about 1850, when mass production of accurate scales, calipers, and micrometers made precise shop-floor measurement affordable. Tolerances began to appear on dimensions in drawings in the late 1800s.
- The rise of 'out-sourcing' (external procurement, by competitive bidding) in the 1900s required that drawings – the primary mechanism for part and assembly specification – be standardized. This process began in several nations between the two World Wars.
- Procurement problems in World War II revealed ambiguities in dimensional-limit tolerances (called 'parametric tolerances' below) and stimulated the development of *geometric* tolerances. These same problems led to the important principle of *process independence* (discussed below).
- Industrial use of Coordinate Measuring Machines (CMMs) for part inspection proliferated in the 1980s, but early CMM algorithms produced results different from those obtained with traditional methods (gages, conventional 'two-point' manual measurements). This divergence triggered a 'Metrology Crisis' that is being resolved only gradually [Hook 94].
- The 'Metrology Crisis' triggered the current drive to rationalize and 'mathematize' tolerancing and metrology.

3. THE ROLE AND CHARACTER OF TOLERANCES

This section summarizes the role and some general characteristics of dimensional tolerances. It is aimed mainly at readers who know little or nothing about tolerancing.

3.1 Scale and Precision

Tolerances are used to control variations in form and relation, where 'variation' is best viewed as a *relative* phenomenon that depends on (1) how precisely one can *measure* a variation relative to a part or product, and (2) the *effects* of a variation – for example, what changes in (relative) efficiency it engenders.

Figure 3-1a shows that mechanical artifacts span an unusually broad range of dimensional 'sizes', from micron-level sensors and actuators to kilometer-scale strip mills and super-tankers. The size range is about 10^9 when expressed as a ratio [Swyt 92]. The ratio of smallest observable unit on the left to largest size on the right is at least 10^{10} ... far too large to be spanned by any known manufacturing or measurement technology.

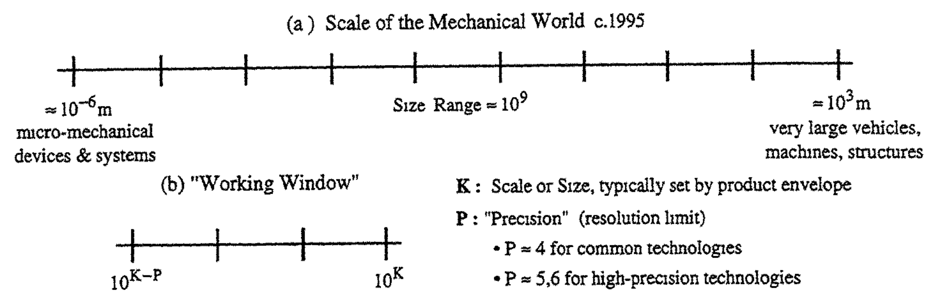


Figure 3-1: Scale and precision.

Particular product families – computer disk drives, say, or marine diesel engines – span smaller ranges, typically of order 10^4 to 10^6. Each family can be associated with a 'dimensional working window' – see Figure 3-1b – defined by a scale (or size) parameter K which sites the window in an absolute scale, and a 'precision parameter' P which sets the size of the window in terms of smallest resolvable units. P is a measure of the relative precision of a product family or manufacturing process. The historical growth of P is one measure of progress in engineering.[3]

[3] In the 1700s conventional goods were built to about .01 (P=2) precision, but the construction of early steam engines in that era forced the level to .001. Today most high-quality conventional goods lie on a precision

In summary: one usually sees toleranced dimensions in the form 10.00 ± .01, but it is often useful to represent them in the form 10.00 (1 ± 10^{-3}), where '10' sets the scale and 10^{-3} sets the precision. Further, complex mechanical systems should be thought of as 'hierarchies of precision windows'. Thus the 100-meter envelope of a jet airliner may be precise to a few centimeters, while subsystems within the aircraft may be precise to a few micrometers.

3.2 Tolerancing for Assembly

Tolerances are generally bilateral (set upper and lower limits of some kind), with one limit usually set by assembly requirements – addressed here through an example – and the other set by functional requirements (addressed in Section 3.3). The example we shall use is based on the electrical bus bar shown in Figure 3-2a,b. The bus bar's nominal cross-sectional dimensions, H=10 and W=5, are to be toleranced.

The assembly constraint – that the bar always fit into its hangers – is satisfied by matching worst-case conditions. The critical state for the hangers, shown in Fig. 3-2c, occurs when the hole is at its lower dimensional limits, 10.10 and 5.05. This is called the Maximum Material Condition (MMC), because the material surrounding the hole is 'locally maximal'. The corresponding MMC state for the bar (maximum material 'in' the bar) occurs when its dimensions are at their upper limits. Values for the upper limits – 0.10 and 0.05 for H and W respectively – follow from equating the two MMC states.

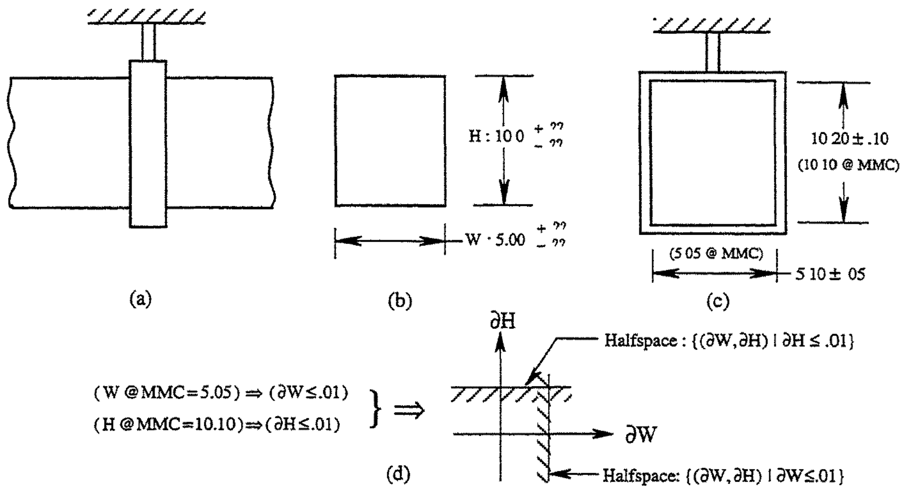


Figure 3-2: Tolerancing for assembly using MMC criteria.

These results are shown in Fig. 3-2d as halfspaces that define by intersection, in (∂W, ∂H) space, a semi-infinite *dimensional conformance zone* for the bus bar. The parameters ∂W, ∂H are defined as follows. Let

$$H = H_0 + \Delta H \ , \quad W = W_0 + \Delta W \qquad \text{(3-1a,b)}$$

represent actual values of the H and W dimensions, where ΔH, ΔW are deviations from the nominal values $H_0 = 10$ and $W_0 = 5$. Then

$$H = H_0(1 + \tfrac{\Delta H}{H_0}) = H_0(1 + \partial H), \quad \text{or} \quad \partial H = \tfrac{\Delta H}{H_0} = \tfrac{H}{H_0} - 1 \ , \qquad \text{(3-2a,b)}$$

and similarly for W and ∂W. Thus ∂H, ∂W are 'fractional deviations' from their nominal values. Each is upper-bounded by .01, as shown in Fig. 3-2d. *Any point that lies in the doubly hatched region – the dimensional conformance zone – represents a bus bar that will always assemble with its hangers.*

---

plateau of about 10^{-5} which often can be associated, in highly simplified terms, with the thermal expansion coefficients of typical engineering materials such as steel.

3.3 Tolerancing for Function

To set lower tolerance limits for the bus bar in Fig. 3-2, we need additional constraints and appeal first to 'functionalism'. The bus bar is to conduct electricity with low resistive loss. The nominal resistance per unit length, r_0, is given by

$$r_0 = K \cdot R_0 / A_0 \qquad (3\text{-}3)$$

where R_0 is the nominal resistance per unit volume (a material property), $A_0 = H_0 \cdot W_0$ is the nominal cross-sectional area, and K is a constant. The actual values of each parameter – r, R, A – vary about their nominal values r_0, R_0, A_0. If the variations are small,

$$\partial r \approx \partial R - \partial A \approx \partial R - (\partial H + \partial W) \qquad (3\text{-}4)$$

where the '∂' variates are fractional deviations analogous to ∂H, ∂W above.

Low values of r are desirable (low resistive losses) whereas high values are not, and so ∂r must be bounded from above. For illustrative purposes, assume that $\partial r \leq .05$ (5% maximum deviation above nominal) and $|\partial R| \leq .03$ (±3% maximal variations in bulk resistivity: this reflects what we are prepared to pay for raw material). Then

$$\partial A \approx \partial H + \partial W \geq -.02 \ , \qquad (3\text{-}5)$$

which establishes the diagonal third halfspace in Fig. 3-3a and creates a finite dimensional conformance zone.

The conformance zone in Fig. 3-3a is the most permissive that will meet both the assembly and functional constraints, but it requires that the H and W dimensions be measured for inspection 'simultaneously'. If one requires that the H and W dimensions be *independently* inspectable, then the conformance zone must be rectangular. Fig's 3-3b,c,d provide examples of three such zones with their associated tolerance values. Additional constraints, such as the relative cost of controlling H vs. controlling W, are needed to optimize the aspect ratio of the rectangular zone. The resulting lower tolerance limits define what are called, for obvious reasons, Least Material Condition (LMC) states for the bus bar.

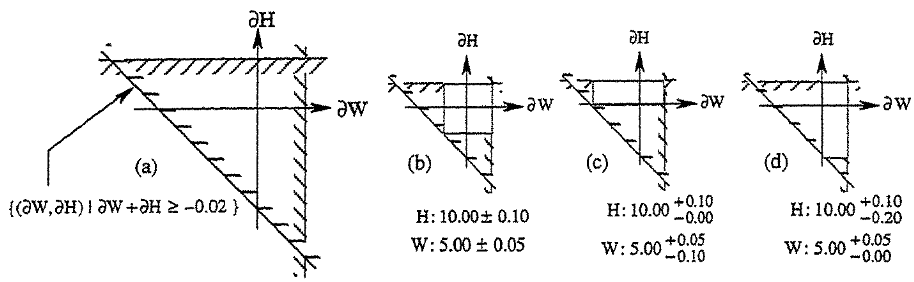


Figure 3-3: Tolerancing for function with coupled (a) and independent (b-d) LMC conditions.

3.4 Remarks on the Relative Importance of Assembly vs. Functional Tolerancing

The bus bar example illustrates the bilateral nature of most tolerances. One set of limits is determined by assembly requirements, using various kinds of MMC-like 'fitting together' criteria, while another set is determined by functional arguments. The bus bar example suggests the two sets are of commensurate criticality because the tolerance *values* are commensurate, but this is seldom the case in practice.

Tolerancing for assembly, rather than for function, has been the predominant concern in most product designs for at least half a century. Most products simply work better, and are easier to assemble, if they are built from precise components, and the cost of precision has decreased almost continuously for two centuries. Higher precision means tighter tolerances, and thus smaller margins for 'fit-together' allowances ... and for approximations and errors in assembly tolerancing analyses.

Further, tolerancing for assembly is primarily a *geometric* exercise, whereas tolerancing for function involves various kinds of energy exchanges – elastic or plastic deformation, heat or fluid flow, electrical phenomena, and so forth. The sensitivity of energy exchange phenomena to form variations – which is what functional tolerancing seeks to control – is not well understood, but two observations help to explain why functional tolerancing is usually less central than assembly tolerancing.

Note first that energy exchanges almost always involve material properties (eleastic moduli, bulk resistivity, and so forth), and these cannot be controlled to the same levels of precision as geometric properties. In the bus bar example: it is technically feasible to tighten the hanger tolerances which dominate assembly by 10^3, but it is not feasible to tighten the tolerance on bulk resistivity by 10^3.

Secondly, the intrinsic characteristics of some (perhaps many) phenomena 'smooth' the effects of variations in form geometry. Heat conduction, which is a diffusion process described by Laplace's Equation, is a good example.: small changes in boundary geometry simply don't have much effect on the temperature distribution over the interior of a thermal conductor unless the conductor is geometrically 'thin', e.g. a sheet or plate. However, there may be important exceptions to this observation. With boundary *layer* phenomena, for example, functional tolerancing may dominate assembly concerns. We simply don't understand these issues at present.

3.5 Accumulation, Process Independence, and Conformance

Most mechanical parts contain several or many toleranced features, and feature tolerances often 'interact' and *accumulate*. The analysis and management of tolerance accumulation ('stackup', in trade jargon) is *the* central challenge in design for manufacturing. It is illustrated and discussed in the examples in Section 6, which show that there are more open questions than known answers.

In the past, parts were often specified by local 'process callouts' – for example, "Bore Ø3.250 x 1.500 deep", or "Finish-mill then grind ... " in lieu of explicit tolerances. Implicit tolerancing through process specification can work well within vertically integrated companies, but it complicates outsourcing, and it caused enough problems in World War II military procurement to result in its prohibition in postwar American tolerance standards. Thus for the past several decades the dictum, called *process independence,* has been, "Define the result you want in geometric terms, via tolerances, not how to attain the result with particular processes." It's worth noting that there are compelling formal reasons as well as practical reasons for process independence. One such: how can equivalence be established between two definitions, defined through different sets of processes, for the same part?

Finally: tolerances should be inspectable by finite sequences of physically realizable measurements, or equivalent gaging operations. This may seem self-evident, but there are many traps and subtleties in assessing actual parts for conformance to tolerance specifications ... too many, and too subtle, to address in this paper. But to clinch the point: we show below that parametric tolerances – the primary tolerancing mechanism for almost a century – are not rigorously inspectable. They have been largely replaced over the past 25 years with geometric tolerances.

4. PARAMETRIC AND GEOMETRIC TOLERANCES

The examples above are based on *parametric* tolerances. The central concept is simple and intuitively appealing: (1) regard the dimensions engineers use to define nominal parts as parameters whose numeric values may vary in actual parts, and (2) specify allowable variations for each parameter, usually with upper and lower limits as in the examples above, or by statistical means as discussed below.

Intrinsic ambiguities in parametric tolerancing began to be recognized in the early and mid-1900s, as industry moved to progressively higher levels of precision. Two of these are shown in Figure 4-1.

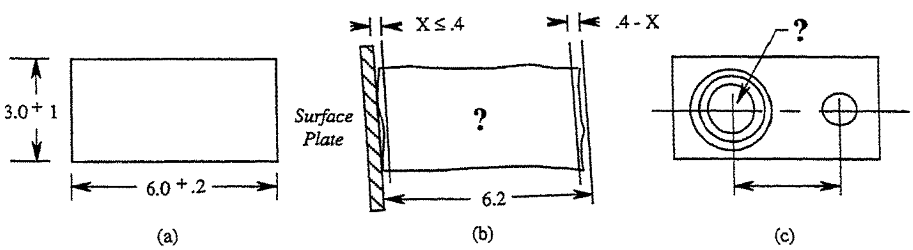


Figure 4-1: Ambiguities in parametric dimensioning and tolerancing.

The specification in Fig. 4-1a is clear enough if actual parts produced to the specification have precisely planar and parallel faces, 'square corners', and so forth, but real parts exhibit none of these characteristics. Thus, how should the 6.0 ± .2 dimensional specification be verified? Fig. 4-1b shows one possible technique, but there are many others ... and each may yield different results. Fig. 4-1c shows a seemingly harmless control dimension that is ambiguous in several subtle ways. For example: how is the center of the stepped hole to be determined? (Three nominally concentric cylindrical features cannot be exactly concentric: which governs?) These are serious ambiguities in the context of assembly design, and they stimulated the development of *geometric* tolerancing in the 1940s and '50s.

Geometric tolerancing (Geometric Dimensioning and Tolerancing, or 'GD&T') is based on three central notions.

- Conformance to a geometric tolerance requires that a surface feature, or an attribute of a feature (e.g. the axis of a hole), lie within a prescribed spatial zone. Note that this is a true *geometric* criterion, whereas conformance to a parametric tolerance is inherently numeric.
- A geometric tolerance usually controls explicitly only one specified property of a feature, such as form (flatness, cylindricity) or position. (However, subtle interactions between different tolerances on the same feature can complicate matters considerably.)
- Some containment zones (e.g. for form) can be positioned freely in space, whereas others (e.g. for position) are located on parts through reference features called *datums*.

The use of containment zones deals directly with *imperfect form*, and is one of the hallmarks of gometric tolerancing.[4] Figure 4-2 shows the zone associated with a surface profile tolerance (symbol ⌓): it is a spatial 'slab' of thickness 't' whose centerplane is located 'd' units above the *datum* (reference) feature A. Dimension 'd' is boxed in the specification to indicate that it is 'Basic', i.e. defines the 'True Position' of the zone with respect to datum A. Dimension 'd' is inspected by using measurement methods that are typically 10X more precise than the magnitude being checked for conformance.

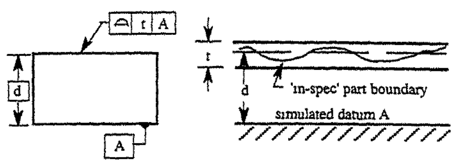


Figure 4-2: A surface profile tolerance.

While GD&T addresses the weaknesses of parametric tolerancing, it has its own weaknesses. These stem in part from its definitional media – national and international standards, in which potentially precise concepts and criteria are defined informally, using prose, graphics, and special-case examples. These informal definitions are now being replaced with mathematical definitions, as noted later.

Finally, parametric tolerancing will not be abandoned entirely, and in fact is enjoying a minor resurgence in some circles. It is useful when form errors can be ignored safely, as in most studies of functional tolerancing, and it continues to have considerable intuitive appeal.

5. THE CURRENT STATE OF AFFAIRS

This section summarizes the author's view of the current state of tolerancing technology and the major streams of activity aimed at improving it. *An important caveat: virtually everything that follows is directed at tolerancing for assembly.* Relatively little is known about tolerancing for function, and at present there is not much research directed at functional tolerancing.

5.1 Technology Summary

Table 5-1 presents the summary. The table's organization reflects the following slightly more formal interpretation of 'tolerancing scheme' than used heretofore.

A tolerancing scheme is a triple, (variability rep-scheme, criteria, composition rules), whose members provide

- *means (a representationn scheme) for representing spatial variability in a part or assembly,*
- *means (criteria) for placing limits on variability through said representations, and*
- *rules for composing (representing the interaction of) variabilities and criteria.*

---

[4] Imperfect form can be handled in other ways – for example, by replacing imperfect planar and cylindrical features with fitted versions of perfect features. This approach is popular in some European and ISO circles.

The table covers two representation schemes, Parametric and Geometric, and two criteria, Worst Case (WC) and Statistical (ST). Notes on the schemes and criteria follow the table.

	REPRESENTATION OF VARIABILITY	
⇓ CRITERIA ⇓	PARAMETRIC ('Dimensions')	GEOMETRIC (Zones, Datums, etc.)
WC (Worst Case)	Informal name: P/WC Origin: < 1900 • Few and simple standards • Primary rule is WC stack-up • Declining industrial use	Informal name: G/WC Origin: ≈ 1940 • Elaborate national, international standards • Local (size, zone) interaction rules; no established chaining (stackup) rules • Widespread industrial use
ST (Statistical)	Informal name: P/ST Origin: ≈ 1955 • Codified company practices; no established standards; recent ISO activity • Primary rule is *rss* stack-up • Growing industrial interest and use	Informal name: G/ST Origin: ≈ 1990 • No established standards; recent ISO activity • No established rules; major open issues • Industrial interest

Table 5-1: Summary of contemporary tolerancing schemes.

Notes

1) Parametric representations are simply conventional dimensions whose actual values may vary; these, with a WC criterion, are the Parametric Tolerances of Section 4. Vectorial representations [Henzold 94] are a late-1980s variant popular in some European circles.

2) Geometric schemes use the representational machinery of Geometric Tolerances (Section 4): zones, datums, and so forth. GD&T, as defined in the standards, carries WC criteria, but there is growing interest in statistical interpretations (the 'G/ST' scheme).

3a) WC criteria are used to guarantee assembleability. For a given nominal part: every actual part that conforms to WC tolerance specifications on the nominal part is a member of a 'WC-conforming population'. Each nominal part in a nominal assembly has as associated WC-conforming population. Parts chosen arbitrarily from (appropriate) WC-conforming populations are guaranteed to assemble, and any member of a WC-conforming population may be replaced with any other member with no loss of assembleability. (The latter property is 'interchangeability'.)

3b) WC criteria are applied to parametric representations ('dimensions') through limits that define strict numeric containment intervals for the actual values of dimensions. The primary (and essentially only) composition rule covers WC 'stack-up' over linear dimensions chains; it is illustrated in the examples of Section 6.

3c) WC criteria for geometric representations appear either as 'sizes' of containment zones, or as limits for features-of-size that require careful interpretation. The composition rules cover the local interaction of feature-of-size limits and zone limits (yielding 'Virtual Condition', 'Resultant Condition'). There are no commonly accepted rules for chaining WC zone-based tolerances.

4a) ST criteria are used to reduce the cost of parts in complex assemblies. The essential idea: if variabilities in actual parts are viewed as random fluctuations that cluster near nominal values, then greater variability than would be allowed under WC doctrine can be accepted in each part *if* some small percentage of assembly failures can be tolerated ... because the greater variabilities will 'average-out most of the time'. ST criteria are applied to populations of actual parts, called ST-conforming populations, never to individual parts. Thus, parts drawn from (appropriate) ST-conforming populations will assemble successfully with a predictable probability, A member of an ST-conforming population that assembles successfully may be replaced with another member of the same population with a predictable (generally small) probability of assembly failure.

4b) Arguments for using ST criteria with parametric representations appeared in industry about 40 years ago (e.g. [Brooks 56]), and effective industrial use of ST criteria has grown since then, but past and current industrial applications are based on methods that are often proprietary and idiosyncratic. There are no agreed, standard methods for applying ST criteria to parametric representations, but a new ISO working party is attempting to devise some. The example in Section 6.1 suggests why there are no standard methods.

4c) There are no standard methods for applying ST criteria to geometric representations, and until recently there has been little interest in doing so – thus the G/ST tolerancing scheme in Table 5-1 is more goal than reality. [Braun 97] presents experimental data that illustrate some of the subtleties in this area.

In summary: the evolution of tolerancing technology has been based on WC criteria, used first with parametric representations and then with geometric representations. The GD&T defined in current national and international standards uses WC criteria and geometric representations, and is the predominant tolerancing technology in industrial use today. There is growing interest in ST criteria, especially with geometric representations, but there are no standard methods and it is not obvious that we know enough about either geometric representations or ST criteria to devise good standard methods in the near future.

5.2 Forward Motion: the Main Streams of Activity

Activities aimed at improving tolerancing technology have grown in breadth, depth, and scale over the past decade, and are likely to continue growing for (at least) another decade. They can be summarized at present in terms of four streams of activity.

Stream 1: Maintenance and mathematization of current standards

In America the traditional custodial committees for standards – Y14.5 (Dimensioning & Tolerancing), B89 (Dimensional Metrology), and B46 (Surface Phenomena) – continue to seek evolutionary improvements, and are assessing the implications of the recent major ISO reorganization summarized in the Appendix. Two new committees, Y14.5.1 and B89.3.2, were formed to respond to the 'Metrology Crisis' of the mid-1980s. The first produced in 1994 a 'mathematical companion' [ASME 94b] for the new edition of the main American standard [ASME 94a]; this provides definitions in algebraic geometry for the concepts defined in the main standard through special-case examples cast in prose and graphics. A similar document for metrology is promised in the near future. In Europe, standards maintenance is being reshaped by the more ambitious goals noted in Steam 3.

Stream 2: Fixing problems and filling gaps

Mathematization of the American standard disclosed serious weaknesses in some traditional concepts – e.g. the definition of 'size', properties of datum systems – that are being attacked individually in the research community. [Suresh 94] provides an example of this work. Conceptual extensions are also being sought, notably in the area of statistical tolerancing. [Srinivasan 97a] summarizes some of the latter work.

Stream 3: Comprehensive rationalization

A project aimed at a major rationalization of ISO standards for tolerancing, metrology, and surface phenomena is being organized in 1997. The primary goals are to draw tolerancing and conformance metrology together, and to place both on uniform and moderately formal mathematical foundations that are compatible with emerging CAD standards – notably STEP [Laurance 94]. If this project succeeds – and success if not assured, because the goals are very ambitious given our present incomplete understanding of many issues – a new wave of standards can be expected in AD2000–2010.

Sream 4: Research on the physics of applications

Surprisingly little is known, in terms of traditional mathematical models of physical reality, about the 'physics' of variation control in assembly and in design-for-function. The example in Section 6.2 illustrates this claim. When viewed in this light, the foregoing activities can be viewed as efforts to provide increasingly powerful and rigorous solutions for ill-defined problems. This is not to say that the work in Streams 1-3 is 'improper'; it is both proper and necessary, because industry always needs the best technology and standards that can be provided with whatever knowledge is available. However, we are unlikely to devise truly *appropriate* (eventually 'ultimate') technologies and standards until we understand the physics of variation control much fully than we do at present.

6. TWO EXAMPLES OF 'FOREFRONT' RESEARCH

The examples below illustrate some established techniques and summarize some current research issues.

6.1 What is a Statistical Tolerance?

Figure 6-1 reviews the classical WC and ST stackup rules. The simple part in 6-1a carries two specifications: WC limit tolerances, and independent normal (Gaussian) random-variate statistics using the notation N(<mean>,<variance>). Note that the distance between faces f_1,f_2 inherits its nominal value and variability from the controlled dimensions. Fig. 6-1b shows that the WC variability grows linearly, whereas the statistical variability in Fig. 6-1c grows on an *rss* (root-sum-of-squares) basis ... because the mean and variance of a linear combination of independent random variates are respectively the algebraic and absolute sums of the component means and variances. This difference in accumulation rates can have profound effects, as the next example will show.

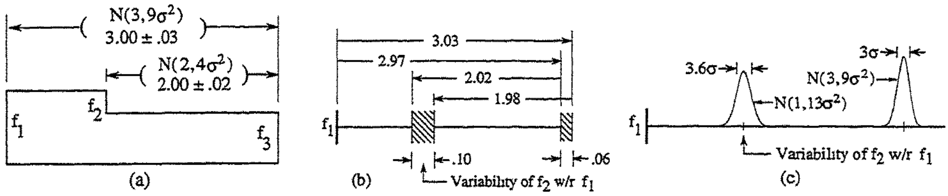


Figure 6-1: Worst-case (WC) and statistical (ST) stackup.

The main constraint on the disk brake assembly shown in Figure 6-2a is that gap G must be open (non-negative), so that the brake will assemble. If readers run the stackup rules to find values for the G-row in the parameter table (Fig. 6-2b), they will find that

- on a WC basis, G is 1.00±1.00 and thus 100% assembly is assured;
- on a ST basis, the probability of assembly *failure* is about .0014, even though 18% of each of the component parts cannot meet the WC part tolerances! This shows the power of statistical averaging.

Figure 6-2a: Disk brake assembly.

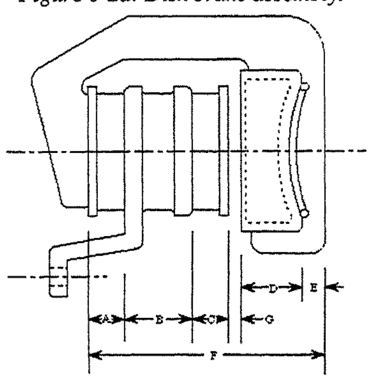


Figure 6-2b: Disk brake parameters.

Dimension	Nominal Value	Limit Tolerances	Standard Deviation
A	15.0	±0.2	.150
B	30.0	±0.3	.225
C	15.0	±0.2	.150
D	50.0	±0.1	.075
E	11.0	±0.1	.075
F	122.0	±0.1	.075
G			

Now to the main issue: how should statistical tolerance specifications for each of the brake's component parts be written? Figure 6-3 shows some possibilities, using the A-part as an example.

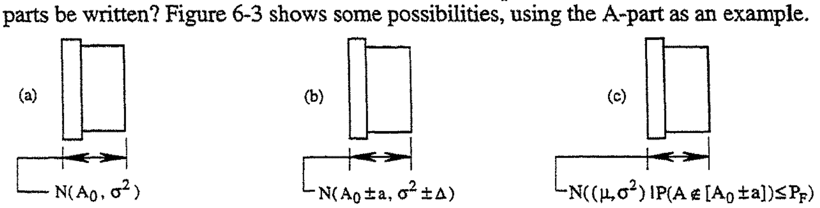


Figure 6-3: Candidate definitions for a statistical tolerance specification for part 'A'.

- The 'a' proposal (Fig. 6-3a) mimics Fig. 6-1a: it requires that the population of actual values of the controlled dimension exhibit normal statistics with mean A_0 (the nominal value – 15.0 in Fig. 6-2b) and a specified variance. But this is clearly illegal, because it requires *exact* parameter values.

- The 'b' proposal puts independent tolerances on the distribution parameters, but this is naive. An efficient tolerancing scheme would allow tradeoffs: if the mean-shift is large, the variance should be small, and conversely.
- The 'c' proposal offers a mechanism for making such tradeoffs: allow any values for (μ, σ^2)that will keep the actual A-values in a specified interval [A_o±a] with probability 1–P_F. This is a Constant Failure Rate (CFR) criterion, which is related to industrial practices based on the 'process capability indices' C_p, C_{pk}, C_c [Braun 97, Voelcker 97]. Other related candidate methods are discussed in [Srinivasan 97a] and [O'Connor 97].

  ... Those are some of the easier issues. Here are two more difficult questions.

- All of the foregoing proposals for statistical part-level specification *specify* normal population statistics ... apparently precisely normal statistics. Is this 'legal'? Is it sensible? (Probably neither.)
- Whatever mechanisms are finally chosen for part-level specification almost certainly must accomodate parameter variations and some level of distribution diversity in the population statistics – but these properties invalidate most of the assumptions underlying our earlier assembly analysis. Thus we have a higher level question: what models, methods, and data are needed to predict statistical assembleability when 'legal' part-level specifications are required? Some steps toward possible answers to this question are reported in [O'Connor 97] and [Srinivasan 97b].

  In summary: what is a statistical tolereance?

6.2 Tolerancing a Hinge – An Exercise in Assembly Modeling

The problem discussed here – tolerancing a hinge – appears deceptively simple, but in fact it contains interesting subtleties. It has proven to be seminal in terms of posing issues and eliciting results of promising generality.

Figure 6-4 shows an unrolled two-sector hinge. It can be regarded as a one-dimensional tolerancing problem if such issues as form and sector-side parallelism are ignored. Will the hinge assemble with the tolerances shown? The conventional answer is found through worst-case stackup analysis. Thus, mate faces a_1 and b_1 and calculate the WC variational zones (as in Fig. 6-1b) for the faces in the remaining pairs (a_2, b_2), (a_3, b_3), and (a_4, b_4). If the zones in any pair overlap, there is potential interference and assembly cannot be guaranteed with (a_1, b_1) mated ... so repeat the exercise with a different pair of faces mated. The result: for the tolerances shown in Fig. 6-4, each of the four possible matings shows potential interference, and assembly cannot be guaranteed with any prescribed mating.

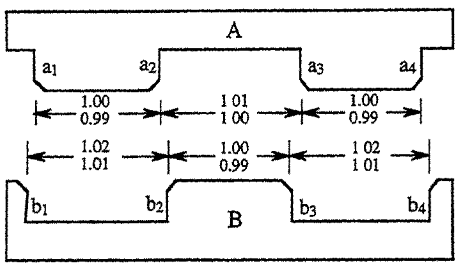


Figure 6-4: An unrolled two-sector hinge.

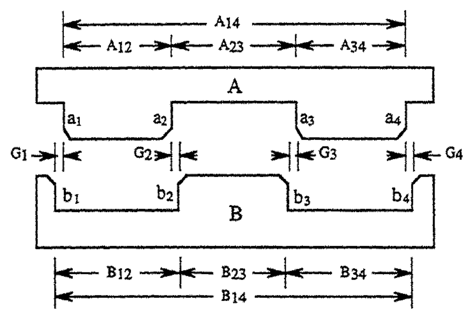
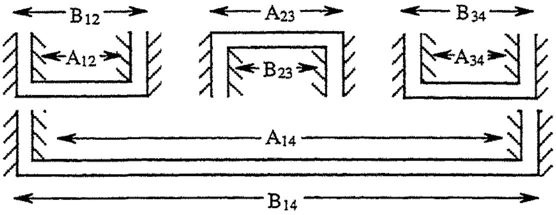


Figure 6-5: Gap analysis.

The first significant research result [Parratt 94]: assembly *can* be guaranteed with the given values if a specific mating is not required – that is, if the assembly can 'float'. This can be shown by writing equations for the gaps G_1 ... G_4 in Figure 6-5 in terms of the sector dimensions A_{12}, A_{23}, ... These gaps must not be negative, and from this fact the assembly conditions shown graphically and symbolically in Figure 6-6 can be derived. In the graphic, the sector faces are grouped into three mating pairs (e.g. A_{12}, B_{12}) of 'ordinary' features of size (as that term is used in geometric tolerancing), plus one pair (A_{14}, B_{14}) of 'generalized' features of size. All pairs must mate as shown. Readers can verify assembleability by confirming that the hinge as toleranced in Fig. 6-4 meets the conditions shown in Fig. 6-6.



Necessary & Sufficient Conditions for Assembly

Max $A_{12} \leq$ Min B_{12}

Min $A_{23} \geq$ Max B_{23}

Max $A_{34} \leq$ Min B_{34}

Max $A_{14} \leq$ Min B_{14}

Figure 6-6: Conditions for floating assembly.

Two important points: first, in floating assemblies specified face matings (a_1 to b_1, say) cannot be guaranteed. If a specified mating is essential to the operation of the hinge, then stackup analysis is an adequate analytical procedure (for hinges, but not for all specified-mating problems), and tolerances must be tightened to eliminate any zone overlaps. The second point: both the stackup and floating-assembly procedures can be generalized to hinges having an arbitrary number of sectors. The maximum number of conditions that must be checked in a stackup analysis with a specified mating grows linearly with the number of sectors, whereas the number of floating assembly conditions grows with the square of the number of sectors.

The second significant research result [Parratt 94]: some tolerance patterns ('topologies') yield a worst-case part called the *Maximum Material Part (MMP);* its distinctive characteristic is that all (ordinary *and* generalized) features of size are at their maximum materia conditions – largest separation for external features of size (e.g. A_{12}, A_{14} in Fig. 6-6), smallest separation for internal features of size (e.g. A_{23}, B_{14} in 6-6). For example: the chain tolerancing scheme used in Figure 6-7a (and in Fig. 6-4) does not have a unique worst-case part, whereas the scheme in Fig. 6-7b does; it is attained with the values shown, where (for example) $B_{12}-b_{12}$ denotes the lower limit of $B_{12}\pm b_{12}$. Means for testing dimension trees for MMP-ness are discussed in [Parratt 94].

The Maximum Material Part concept is important because it enables a WC-conforming population (sometimes called a 'variational class') to be represented with a *single* worst-case member – the MMP – that contains, in the set-theoretic sense of containment (and modulo rigid motions), all other members of the population. This should simplify assembly design considerably. As a concept, MMP is a major generalization of the Maximum Material Condition (MMC) that applies to features of size (not whole parts) in geometric tolerancing.

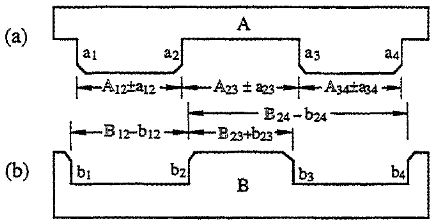


Figure 6-7: Dimension schemes that do (b) and do not (a) have a distinct MMP.

Parratt developed the notions of Generalized Feature of Size (GFS) and MMP only for one-dimensional parts and some special ('orthogonal') two- and three-dimensional parts. Subsequent research by others implies that the GFS concept may not extend to 'general' two- and three-dimensional parts, but the prognosis for generalization of the MMP concept, which is ostensibly more powerful, are brighter. Some early results are reported in [Robinson 97].

Stepping back from the details of the hinge problem: assembly analysis of the disk brake in Section 6.1 involved a single constraint (that a gap be non-negative), whereas the hinge involves multiple constraints (that multiple gaps, as in Fig. 6-5, be non-negative). Is the hinge a simple generalization of the brake, or is it 'structurally different'? Does the hinge represent a distinct class of assembly problems? If so, what are some other distinctive classes, and how many are there? We know very little about assemblies in abstract (e.g. topological) terms, and until we can say more, we will continue to design by special-case analysis, and our attempts to devise better tolerancing technologies will be correspondingly myopic.

7. THE FUTURE

The near-term future – the next five years, say – is reasonably predictable: the streams of work summarized in Section 5.2 will go forward. What happens in the subsequent five-year period will be influenced, if not determined, by progress in Stream-3 activities.

Progress in one of the main lines of Stream-3 work – collaborative rationalization of tolerance and CAD standards – is almost inevitable and overdue. Progress in ISO's major tolerancing and metrology rationalization project is harder to predict.

In the longer term, we can certainly expect continuing evolutionary improvements in tolerancing methods and their application. Whether we can do significantly better than that will depend strongly, in the author's opinion, on how much progress is made in Stream 4.

8. ACKNOWLEDGEMENTS

The views and opinions in this paper are solely those of the author. He has learned much in recent years from discussions with Greg Hetland, Ted Hopp, Alan Jones, Mark Nasson, Alvin Neumann, Vijay Srinivasan, and others. The author is especially indebted to his students: Peter Braun, Ed Morse, Steffen Parratt, Dean Robinson, and K. Suresh – they are great teachers. All of the foregoing rank as indirect contributors to this paper, although none bear any responsibility for what is actually said here.

Research in tolerancing at Cornell University has been supported by the AT&T Bell Laboratories, the Brown & Sharpe Manufacturing Company, the Ford Motor Company, the National Research Council, the National Science Foundation, Sandia National Laboratories, and the Alfred P. Sloan Foundation.

9. APPENDIX: THE 1996 ISO REORGANIZATION

ISO – the International Organization for Standardization – was founded in 1946. Its Central Secretariat is in Geneva, Switzerland, and more than 100 countries participate in its work. Over 10,000 ISO Standards covering a great variety of products and technologies have been published since its founding.

Organizationally, ISO is relatively broad and shallow. The Central Secretariat handles matters of common concern (setting standards for standards, national memberships, and so forth). A large number of Technical Committees (TCs) covering topical areas operate under the Central Secretariat. TCs usually charter Special Committees (SCs), and SCs usually establish Working Groups (WGs).

ISO activities irelevant to this paper were covered until very recently by the following three long-standing Technical Committees.

- TC 10 – Technical drawings, product definition and related documentation, with Dimensioning and Tolerancing handled by TC 10/SC 5.
- TC 3 – A mixture of topics including CMMs, statistical tolerancing, vectorial tolerancing, and several others.
- TC 57 – Surface phenomena and, again, a mixture of other topics.

These TCs worked over the decades at a stately pace and coordinated their efforts only loosely. In the early 1990s they chartered a special 'lateral study committee' called the Joint Harmonization Group (JHG) to investigate instances of overlap and inconsistency in the technical areas covered by the TCs, to identify significant gaps in standards coverage, and to recommend remedial action to the parent TCs. The JHG was led by a small group of Danes headed by Dr. Per Bennich, who had very clear and, at the time, radical ideas about how the JHG's mandate should be pursued. Succinctly, he envisioned a 'master matrix' with technical goals on one axis and levels of implementation technology on the other axis [Bennich 94],. He set the JHG's goals as identifying gaps and conflicts in ISO standards coverage, with the matrix as a primary working tool, and recommending remedial action. Suffice it to say that the JHG worked vigorously, found numerous gaps and conflicts, proposed ambitious remedial action plans to its parents, and started to pursue some of the plans 'informally'.

In 1996 s surprising thing happend: the three parent TCs were dissolved (at their recommendation), and the JHG was reconstituted as a single new Technical Committee – TC 213 : "Dimensional and geometrical product specifications and verification" – to cover and rationalize the very broad domain of dimensional and surface tolerancing and metrology, including practices and implementation technologies. TC 213 has retained the JHG's vigorous style and leadership, and now has the authority to prosecute ambitious projects which it lacked as a study group. [ISO 95] describes the JHG's (now TC 213's) 'global view' of its technical world.

10. REFERENCES

[ASME 94a] ASME (American Society of Mechanical Engineers), *Dimensioning and Tolerancing*, ASME Standard Y14.5M-1994, New York, 1994.

[ASME 94b] ASME (American Society of Mechanical Engineers), *Mathematical Definition of Dimensioning and Tolerancing Principles*, ASME/ANSI Standard Y14.5.1M-94, New York, 1994.

[Bennich 94] Bennich, P., "Chains of standards – A new concept in GPS standards", *ASME Manufacturing Review*, vol. 7, no. 1, pp. 29-38, 1994.

[Booker 79] Booker, P. J., A History of Engineering Drawing, Northgate Publishing Ltd, London, 1979.

[Braun 97] P. R. Braun, E. P. Morse, and H. B. Voelcker, "Research in statistical tolerancing: Examples of intrinsic non-normalities, and their effects", *Proc. 5th CIRP Seminar on Computer Aided Tolerancing*, Ed. H. ElMaraghy, University of Toronto, Toronto, 1997 <in press>.

[Brooks 56] Brooks, K. A. , "Statistical dimensioning", Technical Report TR 107.011.409, International Business Machines, Inc., Endicott, 1956.

[Henzold 94] Henzold, G., "Comparison of Vectorial Tolerancing and Conventional Tolerancing", *Proc. 1993 International Forum on Domensional Tolerancing and Metrology*, V. Srinivasan and H. B. Voelcker, Ed., CRTD-27, pp. 147-160, American Society of Mechanical Engineers, New York, 1994.

[Hook 93] Hook, R., "Interaction of dimensioning, tolerancing, and metrology", *Proc. 1993 International Forum on Domensional Tolerancing and Metrology*, V. Srinivasan and H. B. Voelcker, Ed., CRTD-27, pp. 1-4, American Society of Mechanical Engineers, New York, 1994.

[Hounshell 84] Hounshell, D. A., From the American System to Mass Production 1800-1932: The Development of Manufacturing Technology in the United States, Johns Hopkins University Press, Baltimore, 1984.

[ISO 95] ISO (International Organization for Standardization), "Geometrical Product Specifications (GPS) – Masterplan", ISO/TR 14638:1995(E), Genève, 1995.

[Laurance 94] Laurance, N., "A high level view of STEP – A formal specification of the information content of a product design", *ASME Manufacturing Review*, vol. 7, no. 1, pp. 39-46, 1994.

[O'Connor 97] O'Connor, M. A. and Srinivasan, V., "Composing distribution function zones for statistical tolerance analysis", *Proc. 5th CIRP Seminar on Computer Aided Tolerancing*, Ed. H. ElMaraghy, University of Toronto, Toronto, 1997 <in press>.

[Parratt 94] S. W. Parratt, "A theory of one-dimensional tolerancing for assembly", Ph.D dissertation; Technical Report CPA94-2, Sibley School of Mechanical & Aerospace Engineering, Cornell University, Ithaca, 1994.

[Requicha 93] Requicha, A. A. G., "Mathematical definition of tolerance specifications", *ASME Manufacturing Review*, vol. 6, no. 4, pp. 269-274, 1993.

[Robinson 97] Robinson, D. M., "Geometric tolerancing for assembly with Maximum Material Parts", *Proc. 5th CIRP Seminar on Computer Aided Tolerancing*, Ed. H. ElMaraghy, University of Toronto, Toronto, 1997 <in press>.

[Srinivasan 97a] Srinivasan, V., "ISO debates statistical tolerancing", *Proc. 5th CIRP Seminar on Computer Aided Tolerancing*, Ed. H. ElMaraghy, University of Toronto, Toronto, 1997 <in press>.

[Srinivasan 97b] Srinivasan, V., O'Connor, M. A., and Scholz, F. W., "Techniques for composing a class of statistical tolerance zones", in Advanced Tolerancing Techniques, Ed. H. C. Zhang, J. Wiley, New York, 1997 <in press>.

[Suresh 94] K. Suresh and H. B. Voelcker, "New challenges in dimensional metrology: A case study based on 'Size'", *ASME Manufacturing Review*, vol. 7, no. 4, pp. 291-303, 1994.

[Swyt 92] Swyt, D., "Challenges to NIST in dimensional metrology: The impact of tightening tolerances in the U. S. discrete-part manufacturing industry", NISTIR 4757, National Institute of Standards and Technology, Gaithersburg, 1992.

[Voelcker 93] Voelcker, H. B., "A current perspective on tolerancing and metrology", *ASME Manufacturing Review*, vol. 6, no. 4, pp. 2258-268, 1993.

[Voelcker 97] Voelcker, H. & Morse, E., "De-mystifying C_p and C_{pk}", *mfg.*, vol. 4, no. 1, 1997 <in press>.

[Walker 94] Walker, R. K. and Srinivasan, V., "Creation and evolution of the ASME Y14.5.1M standard", *ASME Manufacturing Review*, vol. 7, no. 1, pp. 16-23, 1994.

APPLICATION OF A COLLABORATIVE CAD ON INTERACTIVE REMOTE COEDITING

Y. C. Kao and Grier C. I. Lin
Centre for Advanced Manufacturing Research
University of South Australia
Australia

ABSTRACT

This paper presents the development of a collaborative CAD system to interactively coedit CAD geometry and its application. Traditional single-location CAD/CAM system provides normally the interaction between a user and the user's computer to increase the individual productivity which, by its nature, is a single-location application and does not support the necessary communications channel between team designers. However, the trend towards global manufacturing has apparently made a collaborative CAD system necessary for geographically distributed designers to work as a team. A collaborative CAD system is a multi-location application which can facilitate the communication channel enabling team designers to co-edit the CAD geometry interactively and concurrently.

KEYWORDS

Collaborative CAD, Remote editing, Computer-Supported Collaborative Work

1. INTRODUCTION

The development of CAD can be traced back to the Automatic Programming Tools (APT) project[1] started in the 1950s at the Massachusetts Institute of Technology (MIT). APT was not, by its nature, designed for interactive operation. The CAD technology was designed as an interactive application for a single-user and therefore a CAD user can only communicate with a computer's central processing unit (CPU). The commercial CAD packages became available in the early 1970s, nearly at the same time (1969) that the United States Department of Defence (DoD) launched its Advanced Research Projects Agency network[2] (ARPANET); but the typical configuration of a CAD system was still designed as a standalone system. The Internet came into existence as a natural development resulted from the ARPANET, and the National Science Foundation's network[3] (NSFNET) in 1980s. The CAD users still could not communicate synchronously on editing and discussing the CAD geometry among themselves through CAD/CAM computer software, even though the CPU was a time-shared mainframe system linking several terminals running the CAD software. With local area network (LAN) becoming popular, CAD systems began to be implemented on a network server, allowing users to access from any computer station networked to the server. However, interaction among CAD users was still much the same as before except that more advanced communication services were increasingly used to shorten the time span for necessary communication.

Advances in computing technology have led to a common practice of CAD collaboration assisted by using collaborative tools to establish a computer-supported cooperative working environment. Examples are: text-based messages such as text e-mail, text talk and facsimile; audio-based messages such as voice mail, telephone and Internet phone; two-dimensional drawings such as blueprint drawing by facsimile and mail (by express delivery); asynchronous computer support such as multimedia e-mail, file transfer (FTP) and world-wide web (WWW) services; and synchronous computer support including shared applications, multimedia audio, whiteboard and video, 3D geometry viewers, and desktop conferencing.

2. OVERVIEW OF THE COLLABORATIVE CAD SYSTEM

The collaborative CAD system (*COCAD*) is based on UNIX interprocess communication (IPC) and a connection-oriented client and server model underlying the Transmission Control Protocol/Internet Protocol (TCP/IP) for synchronous network data transaction. The TCP/IP architecture[4] accommodates the communication technologies in its subnetwork level. The utilisation of TCP/IP architecture is transparent to developers and users at the application level.[5] The networking modules are developed by C programming language integrating and linking the Extended Grapl-IV[6] programming language of a commercial CAD/CAM system Anvil-5000. The developed system can facilitate two users to coedit three-dimensional non-uniform rational B-spline (NURB) curves and surfaces which means the NURB geometry edited by one designer is reflected almost immediately on the other designer's screen through network. Both Internet and narrow band Integrated Services Digital Network (ISDN) were adopted to implement the *COCAD* system in linking Sun workstations running CAD/CAM applications. The *COCAD* system is designed to provide an environment allowing geographically dispersed team users to communicate and to dynamically revise CAD geometry.

Network communication is the key in extending a traditional single-location CAD system to a multi-location application. The limitation of the single-location CAD is that it does not support the communication channel between the geographically separated users for necessary communication at the same time. Network function is applied to develop the *COCAD* system for CAD collaboration and within the *COCAD* system to support data communication among related application programs.

2.1 The Configuration of *COCAD*

The system configuration of *COCAD* depends upon the adopted CAD/CAM system. The development of *COCAD* is based on the application programming interface (API) provided by the CAD/CAM system in order to access and manipulate CAD geometry data. Generally the system structure of *COCAD* has two types. The first type is to construct *COCAD* totally based on the API of a CAD/CAM system if the API provides the user with the networking libraries. The other is to develop external networking modules for network communication and to integrate with the API functions. The *COCAD* system is the second type because the API of the ANVIL-5000 CAD/CAM does not support network socket connection. Therefore the system configuration of COCAD, as shown in Figure 1, consists of a global server *gserv*, a global geometry server *ggserv*, a local server *camrs* and a local client *camrl*.

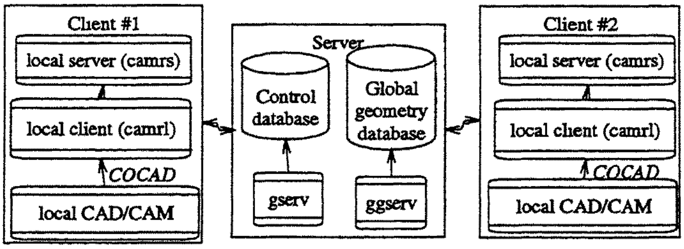


Figure 1 System configuration of *COCAD*

The development of the *COCAD* system uses the XCALL subroutine provided by GRAPL-IV in ANVIL-5000, linking to an external program (EP). The XCALL functions enable a GRAPL-IV program to exchange data with and run an EP which is external to the ANVIL-5000. Both GRAPL-IV program and the EP are spawned through XCALL, run in parallel, and the amount of data to be passed between these two processes is variable, so handshaking is needed. Data is exchanged between GRAPL-IV and an external program, through a shared memory. When a GRAPL-IV program invokes an XCALL command, the data are passed and used as arguments. Constants are passed in one way to the external program while all the variables, of any length, are sent in entirety to the external program, and must be sent back in entirety. The shared memory buffer (SMB) is used for the data exchange. Shared memory is an algorithm for two or more processes to share a memory buffer. A process means a program that is running under certain operating system. The external program is in charge of the socket communication which links the CAD/CAM system to the other computer platform through a network. The role of the GRAPL-IV program and the EP is just like the client-server in a standalone computer.

2.2 The Relation between *COCAD* and CAD/CAM Software

There are two interfaces among different CAD/CAM systems. One is direct database conversion, and the other is through neutral file format. Although neutral file data exchange, for example, Initial Graphics Exchange Specifications (IGES) and Data eXchange File (DXF), has been popularly adopted, there are still many dedicated modules developed for direct data exchange. Neutral file exchange contains only general CAD data and there is no guarantee of 100 per cent data exchange, even though the exchange is for the same CAD/CAM software to output and input the same data. Both of these provide a channel for the data exchange as a whole or on selected CAD geometry. If there is only one curve geometry out of thousands to be edited, either the whole CAD data or the complete curve data has to be exchanged again. The other user must manually delete the original curve before or after the new curve data is received. This results from the fact that usually the pre- and post-processor of data exchange do not automatically replace the original geometry with the new one. Therefore special information has to be added to enable the *COCAD* system to revise CAD geometry on the other side. The relationship between the *COCAD* system and the adopted CAD/CAM system is shown in Figure 2. One of the important *COCAD* data formats is the identification number (entity name), of the added geometric entity for each geometry to be traced if it is necessary.

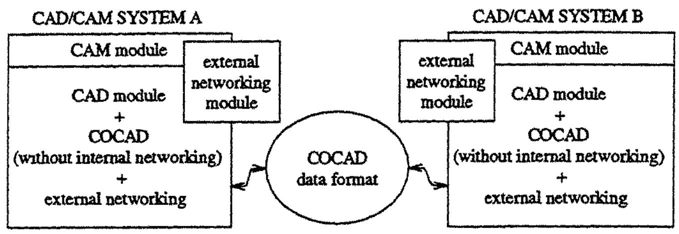


Figure 2 Relation of *COCAD* and CAD/CAM

2.3 Data Communication Algorithm

The data communication in the *COCAD* system is based on the communication strategy of a connection-oriented socket-based interprocess communication and several local disk files. The local server *camrs* and the local client *camrl* are running on the same computer platform. "*camrl*" is actuated by *COCAD* through XCALL, a GRAPL-IV subroutine, to link with *camrs*. Only one user (the creator) can edit the CAD geometry at any one time. On the creator's side, *camrl* contacts with *camrs* through a predefined TCP port. "*camrl*" stores the new data, that is output from *COCAD*, to a common disk file if *camrl* has the right to write to the common disk file. Once *camrs* is contacted by *camrl*, *camrs* forks itself to handle the communication with *camrl*. "*camrs*" then contacts with the global geometry server *ggserv* and updates the data through a connection-oriented socket connection.

If the communication between *camrl* and *camrs* is on the observer's side, *camrl* will contact *camrs* through a predefined TCP port. Once *camrs* is contacted by *camrl*, *camrs* forks itself to handle the communication with *camrl*; and *camrl* stays in a waiting status until receiving *camrs's* response. "*Camrs*" then contacts *ggserv* to obtain or to wait for the updated data through a connection-oriented socket connection. Once *camrs* has the privilege to write onto the common data file, it revises the common data file by the updated data from *ggserv* and return the access right to *camrl*. "*camrl*" is ended and the control goes back to *COCAD* to manipulate the updated data if *camrl* has the response back from *camrs*. Once *COCAD* finishes the manipulation of the updated data, *COCAD* returns the privilege of writing onto the common data file back to *camrs*. Another client process *camrl* would be XCALLed to link to *camrs* for the next observation. The loop for the observation would not be ended until either the *creator* drops the control or explicitly asks the *observer* to pause the observation.

2.4 Data Format for *COCAD* Communication

The idea behind the design of the data format for the communication in *COCAD* is to reduce the amount of data that is necessary for the transmission through network to shorten the networking time. This is based on the assumption that the major part of the computer calculation is operated locally, that is to say the CAD/CAM workstation is powerful enough for efficient calculation. A sample data file for the communication of the CAD geometry is shown in Table 1. The first record in Table 1 is the 'Owner ID' where the IP address is used to distinguish the user who creates the geometric entity from the other. The second record is the 'Command Name' which could be EDIT, INSERT, DELETE, UNDELETE or other keywords. The third record is the 'Geometry Type' definition. The 'Geometry Type' could be NURBI, for interpolated NURB curve, NURBC, for closed NURB curve, or SURCM, for curve-meshed surface, etc. The fourth record is the 'Geometry Name' (entity name) which is unique to a specific geometry. The 'Geometry Name' is designed to keep track of the geometric entity because Anvil-5000 uses sequence number and other information to keep track of geometric entity. The sequence number is assigned by Anvil-5000 and cannot be changed by the user. The records from 'Data records start' to 'Data records end' store the data that is requested by the amalgamation of both 'Command Name' and 'Geometry Type'. For example, if the 'Command Name' is 'INSERT' and the 'Geometry Type' is 'NURBI', the fifth record in the table would be the total number of points. These data points define the interpolated NURB curve followed by the Cartesian coordinates 'X Y Z' of the points set. If the 'Command Name' is 'EDIT' and the 'Geometry Type' is 'NURBI', the fifth record in the table is the 'Geometry Name' of the point. The new Cartesian coordinates of the point are followed to edit the interpolated NURB curve.

Owner ID	x.x.x.x	x.x.x.x	x.x.x.x	130.220.19.77	130.220.19.76	130.220.19.76
Command Name	INSERT	EDIT	INSERT	INSERT	EDIT	INSERT
Geometry Type	NURBI	NURBI	SURCM	NURBI	NURBI	SURCM
Geometry Name	integer	integer	integer	10	10	40
Data records start	40	5	No. of U curves	5	2	3
...	X1 Y1 Z1	X5 Y5 Z5	...	10.0 10.0 10.0	20.0 20.0 10.0	10
...	...	NULL	No. of V curves	20.0 20.0 25.0		11
Data records end	X40 Y40 Z40	NULL	...	30.0 30.0 30.0		16
				40.0 40.0 20.0		2
				50.0 50.0 10.0		12
						20
				Sample 1	Sample 2	Sample 3

Table 1 The communication data format definition

Three sample communication data are also shown in Table 1. Sample 1 shows a data message from the user that is on a host computer with IP address 130.220.19.77. This message asks for an insertion (INSERT) of an interpolated non-uniform rational B-spline (NURBI) which has the entity name '10' and consists of five points. These points are (10.0 10.0 10.0), (20.0 20.0 25.0), (30.0 30.0 30.0), (40.0 40.0 20.0), and (50.0 50.0 10.0). Sample 2 is a message from host 130.220.19.76 which asks for editing (EDIT) an interpolated NURB spline that has the name '10'. The second data point of the interpolated NURB spline '10' is to be changed to (20.0 20.0 10.0). Sample 3 is a message from host 130.220.19.76 and asks for insertion of a curve-meshed surface (SURCM) which has the entity name '40'. This curve-meshed surface has three fixed (U) curves ('10', '11' and '16') and two variable (V) curves ('12' and '20'). These messages are sent to the other designer sequentially, based on the connection-oriented client and server model, during the collaborative co-editing session.

3. EXAMPLE

A coordinate measurement machine (CMM) with touch-triggered probe is usually used to digitise the surface data of a component to fulfil its reverse engineering task such as rapid prototyping. However, the CMM device is usually very expensive and is not located at every engineering department. Moreover, once the product geometry becomes more complicated and/or if the company is going to redesign or enhance the measured product, the measurement from CMM needs to be modified or revised. *COCAD* can be used as a tool between geographically distributed CAD users to access and edit the focused CAD geometry interactively and concurrently. The interactive session can help the engineering group to shorten the revision time by accessing the measurement result from the other department and make full use of the company wide or geographically distributed resources. An example is used to illustrate the application of the *COCAD* in coediting a turbine blade CAD geometry from a CMM as follows.[7]

The screen layout on the *creator's* side is shown in Figure 3. The upper-right window shows the *COCAD* system in Anvil-5000, which also displays the wireframe geometry of a turbine blade. The lower-left window is the message window of the global server *gserv*. The lower-middle window shows the global geometry server *ggserv*. The lower-right window represents the local server *camrs*. The middle-left window shows a portion of the built-in text 'talk' tool in Sun workstation. The text 'talk' was used for realtime communication of text in the collaborative turbine blade co-editing task. The telephone service and multi-media audio tools were also used for oral communication.

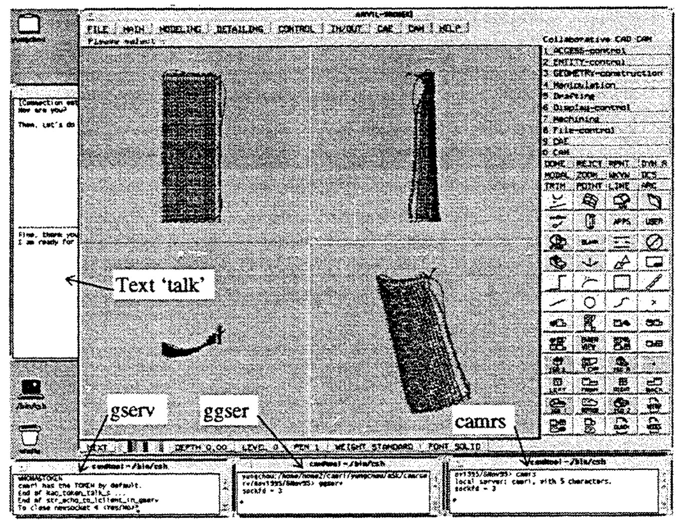


Figure 3 A collaborative CAD session of a turbine blade

4. DISCUSSION AND CONCLUSION

This paper has presented the algorithm in the development of a collaborative CAD system *COCAD*. *COCAD* has been successfully implemented under the Internet for an inter-continental link between the Tjing Ling Manufacturing Technology Research and Development Centre at National Cheng Kung University in Taiwan, and the Centre for Advanced Manufacturing Research at University of South Australia in Australia.

The *COCAD* system can be applied for a collaborative work and has extended a single-location CAD/CAM software to multi-location application; the designers at geographically dispersed locations can coedit CAD geometry data collaboratively at a distance.

External networking modules, as shown in Figure 2, are the key to CAD data communication in *COCAD*. The system configuration, as shown in Figure 1, needs modification to adapt the functionalities of different CAD/CAM application programs. The local client process *camrl* is not needed, if the Extended Grapl-IV language in Anvil-5000 provides networking subroutines, for the *COCAD* system to link local server *camrs*. This is because *camrl* is serving as the bridge between a CAD/CAM application and *camrs*. Moreover, if Extended Grapl-IV has subroutines to be used as a network server, the existence of *camrs* is not necessary. Both *camrl* and *camrs* are modules of the external networking module.

The limitation of an existing single-location CAD/CAM application was resolved for the extension to a multiple location application. A traditional single-location CAD/CAM software was adopted in exploring the necessary functions to construct a multi-location CAD/CAM application. UNIX interprocess communication, the connection-oriented server and client model and the network file system (NFS) were implemented as the basis for the data communication in *COCAD*. The geographically dispersed designers and engineers can work together, through the proposed configuration, as a team. The designers on one side of the world can work with the design engineers on the other side of the world simultaneously under a common design activity.

The capability of the *COCAD* system is limited by the traditional single-location CAD, although an exciting progress has been achieved by the proposed strategy for collaborative CAD/CAM application. The CAD/CAM system that was adopted did not support multi-user interface such as multiple mouse cursors. The interactive operation, that was provided by the GRAPL-IV functions in Anvil-5000 software, could not accept message from external programs. The GRAPL-IV program needs to communicate with the external program for the data and message exchange through the IPC. Therefore, the *observer* needs to observe the *creator's* drawing messages at all times, during the collaborative session.

A set of concurrent server and client program has been developed based on the connection-oriented socket-based model. The concurrent server can respond to the requests from the client immediately and therefore is used as the foundation for a synchronous CAD collaboration.

Network-based communication is now a global phenomenon. The data transmission speed depends heavily on the traffic load of the Internet, that is being shared by millions of users all over the world.[8] The data transmission through the Internet is not as reliable as through LAN. In this paper, a dedicated network, the narrow-band Integrated Services Digital Network (N-ISDN), has been implemented to establish a communication channel for reliable and synchronous network data transaction. The performance of the adopted N-ISDN has shown very promising result in data communication. The environment established in this paper can be applied to other software-based collaboration and served as the communication tools for education and training.

The prevailing trends of increasing business internationalisation and international competition have been well realised and the role of product development is becoming more important than ever before. Not only do they have to meet the consumers' needs at a lower cost and price-performance, higher quality and better after sale services, but must also be able to cope with the competition in a global environment. The previously far-reaching but precious international domain expertise and resources will be available through the networking functions such as the algorithm developed in this paper.

5. ACKNOWLEDGMENTS

The authors acknowledge and appreciate the Australian Federal Department of Employment, Education and Training's (DEET) support of this project under the Targeted Institutional Links (TIL) Program.

6. REFERENCES

[1] Rembold, U., Nnaji, B. O., and Storr, A., Computer Integrated Manufacturing and Engineering, Chapter 1 and 7, Addison-Wesley, 1993.

[2] Davison, John M., An Introduction to TCP/IP, pp. 1-6, Spring-Verlag New York Inc. 1988.

[3] Krol, E. And Hoffman, E., FYI on "What is the Internet?", [Online, accessed 6 Nov. 1996] ftp://nic.merit.edu/documents/fyi/fyi20.html, 1993.

[4] Santifaller, M., TCP/IP and NFS Internetworking in a UNIX Environment, Addison-Wesley. 1991.

[5] Comer, Douglas E., Internetworking with TCP/IP Volume I: Principles, Protocols and Architecture, Chaper 1, 2nd Edition, Prentice-Hall. International Edition. 1991.

[6] MCS, Reference manual of Extended GRAPL-IV, Manufacturing and Consulting Services Inc., 1993

[7] Kao, Y.C. and Lin, Grier C.I., "Implementation of a Collaborative CAD/CAM System", The Fourth International Conference on Automation Technology, Vol. 1, MIRL/ITRI, Hsinchu, Taiwan, R.O.C., July 8-11, 1996, pp. 163-169

[8] Ishida, Haruhisa and Landweber, Lawrence H., "Internetworking", *Communications of the ACM*, 36(8):28-30, August 1993.

INTELLIGENT PACKING DESIGN BASED ON ASSEMBLY MODEL

G.P.Huang, J.Zhou, Q.Yu, C.H.Wang and M.L.Yuan
CAD center
Huazhong University of Science and Technology
Wuhan, Hubei, P.R.China.

ABSTRACT

This paper is concerned with the research on automatic packing in a limited space with complex constraints. In this paper, product functional specifications and intents of designer are mapped into initial assembly design taking the form of assembly constraint information using the idea of top-down design. Firstly, the approach formulates a geometric assembly-based model. The geometric model can be produced by assembly design and directly used for mathematic model to reason and to analyze reasonable positions of components. Secondly, a mathematic model is set up according to the geometric model and product's functionality. In mathematic model, the spatial constraints are represented as equality constraint, inequality constraint and objective function. The result returns to be the input of geometric assembly-based model, so that packing design is realized by variational parameterization. In order to accelerate constraint propagation and solution during packing redesign, the technology of constraint reasoning based on knowledge is adopted.

KEYWORDS

Optimal Design, Assembly Design, Intelligent Design, Auto Design

1. INTRODUCTION

Packing design is the process of locating components into an available space while satisfying spatial relationships between components.In the old time packing design was proposed to solve the problems about building layout. Many two-dimension models have been built. Later packing design was developed in many other fields and a lot of research on three-dimension has been done. For instance, K.lee set up the model of reasonning out component location in assembly. D.Rochelean proposed interactive assembly model. L.S.Homen worked on assembly sequence's method. B.Freitag researched on control assembly system for automative manufacture. It is obvious that the process、sequence and path of assembly for manufacture were largely considered. The assembly design is not faced to product's function (DFF) but rather to assembly manufacture (DFA). In this paper, the methodology of the packing design faced to product's function is emphasized. The design process faced to function, geometric package model based on assembly and mathematic model of intelligent packing design are set up.

2. DESIGN PROCESS FACED TO FUNCTION

In functional space, component is a basic functional unit and subassembly is a kind of organic combination among these functional units, while assembly realizes the whole function composed of these combined functions. In physical space, assembly represents how component and subassembly organize in package design. Assembly is divided into coarse components by using the idea of top-down design and analyzed. One of the major requirements for a modeling system faced to function is a facility for conceptual, nongeometric modeling, which is mainly concerned with decomposing the product to be designed into subassemblies and their interfaces.

So the design process faced to function is as follows:

Firstly, to define the problem in terms of functional requirements(FRs). In this step, we establish FRs from product's needs. So the definitional step requires insight into the problem, and a knowledge base encompassing issues related to the problem. Using the initial stage of concept design for new auto as an example shown in Fig.1 illustrates this step.

Secondly, to decompose the product's function into subfunctions in lower lever step by step and set up their relationships. Fig.2 gives a "functional" view of interior package.

Thirdly, to describe the physical space for the divided functional modulars, i.e., to map from functional space to physical space. Fig.3 is the structure view, the mapping of Fig 2 in physical space.

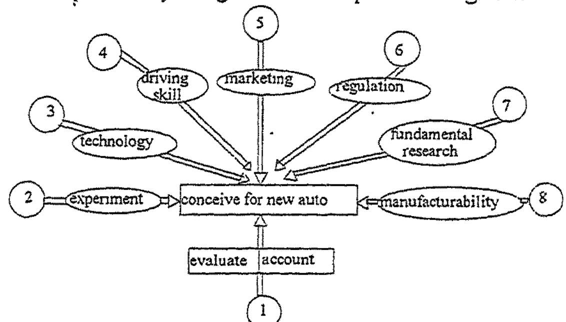


Fig.1: The initial stage of concept design for a new auto.

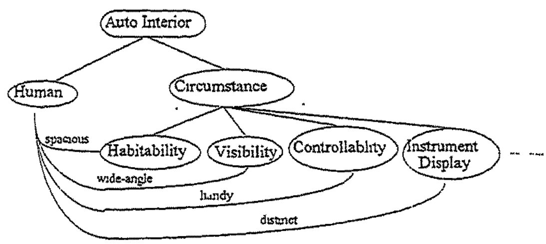


Fig.2. Functional view of interior package.

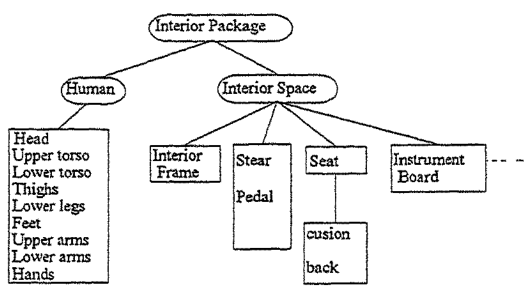


Fig.3: The structure view of Fig 2 .

FRs is analyzed according to the requirement of society and market. In the processes of concept design and assembly, design variables are defined and mapped from functional space to physical space and package entity is created to satisfy FRs. This symbiotic process of the "design" world, going from "FRs" to "concept design" to "package design" to "FRs" to "concept design," etc., can best be depicted by the design helix, shown in Fig.4.

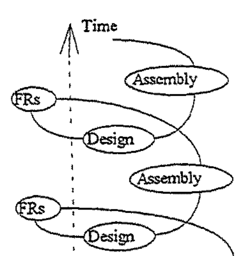


Fig.4: The design helix of the "design"world.

3. GEOMETRIC PACKAGE MODEL BASED ON ASSEMBLY

The initial geometric package model can be mainly dimensionless, concentrating instead on the geometric arrangement of the main components and their geometric interaction. Except for a few values critical for delivering the desired function of the design, the actual dimensions and coordinate values are unimportant in this step. The geometry mainly relies on geometric and transmitting relationships among components and subassemblies. We call the relationship constraint. Geometric relationships that present as equal and unequal equation are geometric element constraints of components. For an instance of a seat, seeing Fig 5(a), the relationships between back and cushion connected by a hinge is equal ones; but the relationship between a seat and its circumstance such as a steer is an unequal one. Transmitting relationship represents motion between components, such as rotation between back and cushion of a seat; that is, the angular α of the seat. These constraints mentioned above are essential informations in the process of spatial package. Fig 5(b) illustrates the constraints. Only when these constraints are satisfied, the locations of components are obtained. Fig. 6 is a directionless adjacent graph of Fig 5(b). In the up-down chain, the node of gra*_H represents components to be constrained. In the left-right chain, the node of gra*_N represents the components related to the constrained components. In gra*_N, each constraint is recorded and forms a chain of constraints, that is, mate_list.

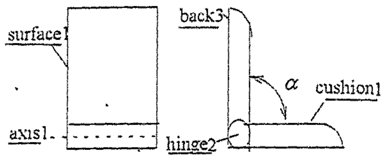


Fig. 5 (a): A seat made of cushion,back and hinge.

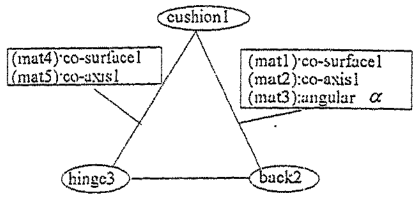


Fig. 5(b): The constraints exists in Fig 5(a).

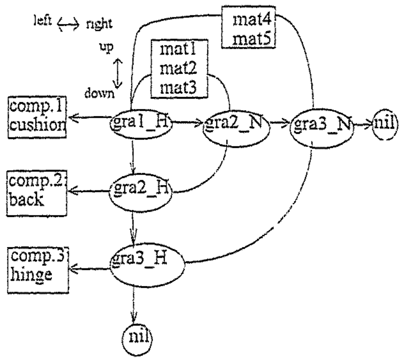


Fig. 6: The adjacent graph of Fig. 5 .

4. MATHEMATIC MODEL OF INTELLIGENT PACKING DESIGN

The rough geometric package model has been built by the design methodology mentioned above. However feature parameters that define shape of component and geometric parameters that define locations of components have not yet been correctly determined. How to determine these parameters requires special knowledge. At first, this paper adopted the structure of knowledge representation to organize the special knowledge as Fig.7.

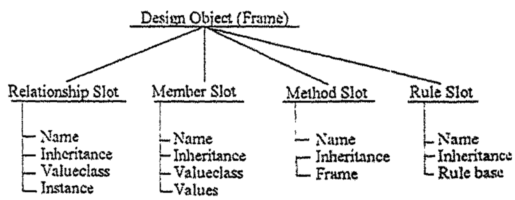


Fig. 7:The structure of knowledge representation.

Then, potential function is adopted to set up the objective function of packing design to determine geometric parameters in this paper.

Supposed that the number of entity is n, $P_{a_j}(u_a, v_a)$ is a reference point on the entity a_j, defined by parametric point (u_a, v_a) and Q_j is the known initial point. The objective function is represented by the form of potential energy as $E=\sum \omega_j (P_{a_j}(u_a,v_a) - Q_j)^2$; ω_j is a weight factor.

During interior package design, it is not easy to specify a function that exactly describes the whole human body potential field. We have divided the human body into a number of parts: head, neck, torso, thighs, lower legs, feet, upper arms, lower arms and hands. Each body part creates its own potential and the potential of the body is the union of the potential of all the parts. Given two potentials A and B, the potential $A \bigcup B$ (A union B) is defined to have the following potential function: Potential-($A \bigcup B$)= Potential-A + Potential-B. Hence, the function for the whole potential is: body-Potential= $\sum$ (body_part-Potential)$_i$.

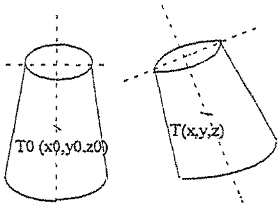


Fig. 8: The model of the human thigh.

Fig.8 gives an example of potential function for the thigh. We model a bounding volume of the thigh as a cone. The potential of the thigh generates an attraction of best posture for feasible position . The potential of the thigh is defined as

$$P(T(x,y,z))=(x-x_0)^2+(y-y_0)^2+(z-z_0)^2$$

where T_0 (x_0,y_0,z_0) is the best comfortable position of the gravity center of the thigh and T(x,y,z) is the feasible position of the gravity center of the thigh by geometric transformation.

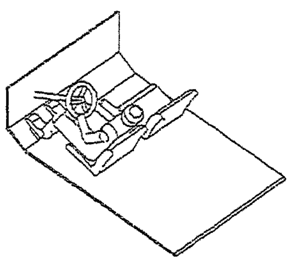


Fig.9: The illustration of interior package.

So far the paper utilizes a novel combination of assembly ,optimization and reasoning technique. Taking auto inner packing design as an instance (the figure shown in Fig.9), the positions of inner components such as floor,pedal,seat,steer,driver or passenger etc., can be reasoned and analyzed by computer. All programs are implemented by C++ language and knowledge base on the basis of graphic software GHCAD, optimal software OPB and intelligent software DEST, all of these are developed by our research group.

5. RESOLUTION OF CONSTRAINT NETWORK

Design variables and constraints constitute constraint network during concept and initial assembly design . In the network each set of constraint represents relevant function. During the process of mapping from functional space to design space and from design space to assembly design space, to input or modify function is to input or modify relevant variables. To add or delete function is to add or delete relevant constraint set. The components related to these changed components should be adjusted in constraint network to satisfy functional requirements. So the design strategy is required. From the whole point, design strategy propagates from variables and three-dimension geometry to entity element and feeds back from entity element to constraints. So the design strategy obeys the rules as follow:

Step 1. To satisfy design in inherent constraint sets as possible, otherwise to propagate to lower constraint sets step by step.

Step 2. To satisfy design in three-dimension geometry constraints in advance, that is, to attempt to adjust positions of components at first.

Step 3. To satisfy design in entity elements at last, because entity design is the most complicated.

The objective of constraint network resolution is to solve all design variables. When some disobeyed constraints happen, it is meaningful to point out these constraints and their relevant variables. So the initial variables and their affective degrees can be analyzed and modifying suggestion proposed.

Given :

{FRs} represents functional requirements of product.

{DPs} represents design parameters of components.

{PPs} represents geometric package parameters .

$[A]$ is the mapping matrix from functional space to design space.

$[B]$ is the mapping matrix from design space to assembly design space.

The design equation for design parameters may be written as

$$\{FRs\} = [A]\{DPs\} \quad (1)$$

The design equation for package geometric parameters may be written as

$$\{DPs\} = [B]\{PPs\} \quad (2)$$

The equations (1,2) can be combined into a single relationship linking FRs to PPs. The result is

$$\{FRs\} = [A][B]\{PPs\} = [C]\{PPs\} \quad (3)$$

where $[C]=[A][B]$.

For the situation of two-dimension, equation (3) may be represented as

$$\begin{Bmatrix} FR1 \\ FR2 \end{Bmatrix} = \begin{bmatrix} C11 & C12 \\ C21 & C22 \end{bmatrix} \begin{Bmatrix} PP1 \\ PP2 \end{Bmatrix} \quad (4)$$

when differential, the change in FRs due to a change in PPs may be written as

$$\begin{Bmatrix} \Delta FR1 \\ \Delta FR2 \end{Bmatrix} = \begin{bmatrix} C11 & C12 \\ C21 & C22 \end{bmatrix} \begin{Bmatrix} \Delta PP1 \\ \Delta PP2 \end{Bmatrix} \quad (5)$$

That is, the system behavior is such that the change in FRs, when the changes in PPs are given, may be written as

$$\Delta FR_i = \sum_{j=1}^{n} \frac{\partial FR_i}{\partial PP_j} \Delta PP_j \quad (i=1,2,\cdot\cdot\cdot,n) \quad (6)$$

We define $\frac{\partial FR_i}{\partial PP_j}$ as sensitiveness of initial geometric variables PP_j to FR_i. The causal relation graph Fig.10 represents between initial geometric variables PP_j and FR_i. If the sensitiveness is $\begin{cases} positive \\ negative \\ zero \end{cases}$, the initial variable has $\begin{cases} obverse \\ reverse \\ no \end{cases}$ effect on FR_i correspondently. And the bigger the value of sensitiveness, the more the effect. At the same time ,authors also notice that the analysis method of the sensitiveness is by partial differentiation with respect to initial variables at the present state of that point. So sensitiveness shows local characteristics. However it provides users with a good ground for evaluating results and redesigning.

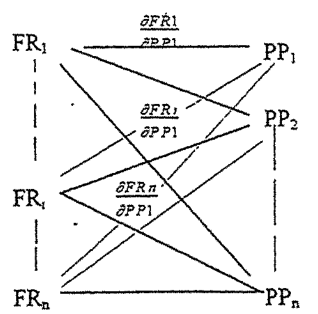


Fig.10: The causal relation graph.

6. CONCLUSION

This paper dwells on the problem of spatial packing. The positions of components are specified based on assembly relationships such as orient, mate, align and offset etc. between components . And the offset values are variational parameters. A mathematic model is set up according to the geometric model and product's functionality. Equality constraint is the requirement of particular spatial relationships between components being maintained exactly, while inequality constraint is that of being allowed in an limited space. Objective function is expressed as a form of energy: $F = \Sigma w_j (P^a_j (u_a, v_a) - Q)^2$. The objective function is minimized to exactly infer the spatial positions of components under the constraints. In the process, the relation among evaluating function, constraint equation and geometric variable constitutes a complex network graph. In order to accelerate constraint propagation and solution during packing redesign, the technology of constraint reasoning based on knowledge is adopted So the paper proposes a method of producing causal relationship graph . Every independent variable is partially differentiated to obtain variable sensitivity by the method of symbolic differential equation. The approach provides a good way to modify objective scheme.

7. REFERENCES

1. Rajneet Sodhi and Joshua U.Turner, Towards Modelling of Assemblies for Product Design, pp.85-96, Computer-Aided Design, Vol.26, No.2, Feb.1994.

2. J.J.Kim and D.C.Gossard, Reasoning on the Location of Components for Assembly Packaging, pp.402-407, Transactions of the ASME, Vol.113, Dec.1991.

3. Xinmin Zhao and Norman I Badler, Interactive Body Awareness, pp.861-866, Computer-Aided Design, Vol 26, No.12, Dec.1994.

4. Andrew Witkin , Kurt Fleischer and Alan Barr, Energy Constraints on Parameterized Models, pp.225-229, Computer Graphics, Vol.21, No.4, July 1987.

5. Suh, N. P., the Principle of Design, Oxford University Press, 1990.

6. M.Mäntylä, a Modeling System for Top-down Design of Assembled Products, IBM J.RES.DEVELOP., Vol.34, No.5, September 1990.

7. Kang Youshu, a Research of Theory and Method of Variable Assembly Design, Dissertation, Huazhong University of Science and Technology, China, June 1996.

8. Mao Quan, Research on Case Prototype Based Design Methodology and Design Support System, Dissertation. Huazhong University of Science and Technology, China, May 1995.

COMPUTER AIDED ENGINEERING SYSTEM FOR CYLINDRICAL GEARING

Li Runfang, Lin Tengjiao, Lin Chao, Tang Qian and Chen Bingkui
State Key Lab of Mechanical Transmission
Chongqing University
China (630044)

ABSTRACT

An integrated computer aided engineering system is developed based on the combination of gear geometry with numerical method in engineering. This CAE software system consists of several families, including geometric parameter estimation and optimization, strength calculation, solid modeling, finite element contact analysis, pre-processing and post-processing in FEM, gear tool design and computer aided drawing. By means of this useful tool, the automatic design, computer aided structure analysis, automatic mesh generation and automatic drawing for cylindrical gearing can be done conveniently and quickly in Rapid Products Development. This CAE system has been applied successfully to construction machinery and machine tools.

KEYWORDS

Computer Aided Design, Gear Transmission, Solid Modeling, Computer Graphics

1. INTRODUCTION

Gearing is the most significant basic component in mechanical transmission. The research of load-sharing and meshing performance of gearing is very important in improving load-carrying capacity, reducing impact and vibration, extending working life and ensuring transmission precision. The gear blank design, geometric parameters, technological parameters, tooth contact analysis and strain-stress analysis are relevant with computer applied technology. Computer applied software, including mechanical optimal design, 3-D solid modeling, finite element analysis, manufacturing process simulation, has also become the indispensable condition in modern products development. On the basis of our research results for years, the CAE software system in gear transmission has been developed successfully, which consists of three main families, i.e. cylindrical gear, straight bevel gear and spiral bevel gear. This software has been successfully applied to transmission system of automobile and construction machinery because of its high efficiency, reliable method, overall function and friendly interface. This paper will introduce the algorithm, function and analysis examples of computer aided engineering software system for cylindrical gear (CGTCAE).

2. FUNCTION AND OPERATION ENVIRONMENT OF SOFTWARE CGTCAE

GTCAE includes the following eight analysis modules: (1) optimal design and strength calculation of planetary gearing; (2) optimal design and strength calculation of plain external gearing; (3) 3-D solid modeling of gear; (4) automatic generation of tooth FEM model; (5) automatic generation of gear contact finite element analysis model; (6) tooth strain-stress analysis by FEM; (7) tooth elastoplastic contact finite element analysis; (8) FEM post-processing of tooth itself and meshing gears.

This software system was programmed by FORTRAN and C language. It could be operated under MS-DOS or Windows operating system and more than 8M bits inner memory is needed.

3. MATHEMATICAL MODEL OF OPTIMAL DESIGN

Parameters of cylindrical gearing could be classified into following categories: (1) application condition parameters (rated power, gear ratio, revolutions per minutes); (2) tooth profile parameters (addendum coefficient, bottom clearance); (3) the parameters relative to gear tools (radius of fillet); (4) gear material properties (hardness, limited stress); (5) the parameters relative to gear manufacture

(accuracy, surface roughness). The design variables which should be determined in gear optimization are gear modular, tooth numbers, helix angle, modification coefficient and tooth width.

In planetary gearing, assume that the minimal volume is considered as a optimal object, the criterion function can be expressed as following

$$f(x)=\frac{\pi}{4}x_2x_3^2(x_1^2+k(i_{ac}x_1)^2)\cos^2\alpha/\cos^2\alpha' \tag{1}$$

Reference pressure angle α and working pressure angle α' can be determined by the following equation

$$inv\alpha'=inv\alpha+\frac{2(x_4+x_5)}{x_1+i_{ac}x_1}tg\alpha \tag{2}$$

where x_1 — teeth number of sun gear; x_2 — face width; x_3 — gear module;
x_4 — modification coefficient of sun gear; x_5 — modification coefficient of planet gear;
k — number of planet gear; i_{ac} — the gear ratio of planet gear and sun gear.

Setting contact strength, bending strength, conterminous condition, concentric condition as a set of restrained condition, the SUMT procedure of restrained optimization is proposed while BFGS method is applied to search out the optimal solution in non-restrained optimization.

4. ELASTOPLASTIC CONTACT FINITE ELEMENT EQUATION

So far as elastoplastic material non-linear problem is concerned, the elastoplastic constitutive relationship can be expressed as following

$$d\sigma=D_{ep}d\varepsilon \tag{3}$$

where $d\sigma$ — stress change; D_{ep} — elastoplastic matrix; $d\varepsilon$ — strain change.

The iterative solving procedure of elastoplastic contact problem is based on mixed formula of finite element contact analysis combined with Newton-type iterative approach for non-linear discrete problem. The analysis procedure can be expressed by following equation

$$\left.\begin{array}{l}
\begin{bmatrix}F_a & F_c\\ Q_c & 0\end{bmatrix}\begin{Bmatrix}\Delta R_m^{(n)}\\ \Delta u_{em}^{(n)}\end{Bmatrix}=\begin{Bmatrix}-\Delta S_{pm}^{(n)}-\delta_{m-1}^{(n)}\\ \Delta P_{em}^{(n)}\end{Bmatrix}\\
R_m^{(n+1)}=R_m^{(n)}+\Delta R_m^{(n)}\\
\delta_{m-1}=S_{m-1}\\
S_m=F_a\Delta R_m^{(n)}+\delta_{m-1}-F_c\Delta u_{em}^{(n)}\\
\Delta u_m^{(n+1)}=\left[K_T(u_m^{(n)})\right]^{-1}(-\int_\Omega B^T\sigma_m^{(n)}d\Omega+\lambda_m P_0+R_m^{(n)}+\Delta R_m^{(n)})\\
u_m^{(n+1)}=u_m^{(n)}+\Delta u_m^{(n+1)}\qquad (m=1,2,\dots.M\qquad n=1,2,\dots.N)
\end{array}\right\} \tag{4}$$

where F_a — flexibility matrix;
F_c — transformation matrix relative to contact force;
Q_c — transformation matrix relative to displacement;
ΔR — vector of contact force increment;
Δu_e — vector of displacement on imaginary support point;
ΔS_p — relative displacement vector caused by applied load;
δ — initial gap vector; ΔP_e — vector of resulting applied load;
R — contact force vector; $K_T(u)$ — tangential stiffness matrix;
B — strain matrix; σ — stress vector;
P_0 — applied load vector; u — nodal displacement vector;
M — number of incremental loaded, N — number of plastic iterative procedure.

5. AUTOMATIC MESH GENERATION OF FEM IN GEARING

In the first place, a rotating coordinate system and a global coordinate system, which is linked together with pinion and gear respectively, is adopted to build up automatically finite element mesh of teeth in engagement at various positions. The generating procedure of tooth mesh can be described as following : (1) determine the location of contact line for a certain meshing gear, then calculate the pressure angle and coordinate of contact point at transverse plane. (2) regard the above contact point as dividing point, the addendum of pinion and gear is divided equally, then calculate the coordinate of all discrete point. Based on the discrete coordinate of addendum, the discrete coordinate of dedendum of pinion and gear can be obtained easily. (3) determine other nodes coordinate of tooth profile at transverse plane in terms of the fillet equation and dimension of gear bodies. (4) calculate the coordinate of other transverse planes in face width direction. (5) convert the nodal coordinate of pinion and gear into global coordinate.

Based on above-mentioned approach, a parametric pre-processing program is developed. Only inputting basic parameters of gears, all data and mesh for 3-D contact finite element analyzing of tooth can be obtained. Fig. 1 is the one gear-pair mesh of helical gear. Fig. 2 is the two gear-pair mesh of helical gear.

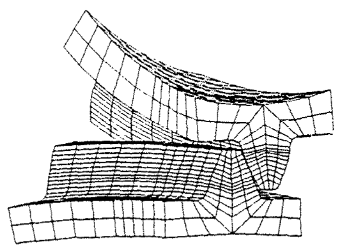


Fig. 1 : One gear-pair mesh of helical gear

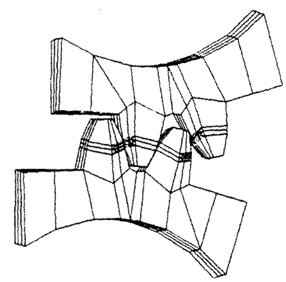


Fig. 2 : Two gear-pair mesh of helical gear

6. COMPUTATIONAL RESULTS

Fig. 3 is the solid modeling of various gears and spline shaft. While the solid modeling of optimized planetary gearing of wheeled reducer in loading machine is shown in Fig. 4.

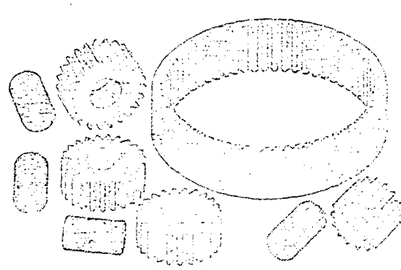


Fig. 3 : Solid modeling of gears and spline shaft

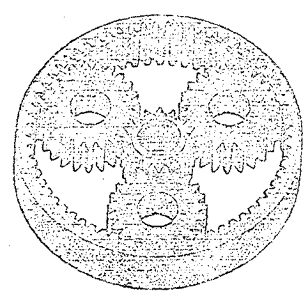


Fig. 4 : Solid modeling of planetary gearing

The surface displacement contour plot of helical gear is given in Fig. 5. Fig. 6 is the section stress contour plot of spur gear in 3-D finite element contact analysis. Fig. 7 (a) and (b) is the 3-D surface stress contour plot of pinion and gear respectively.

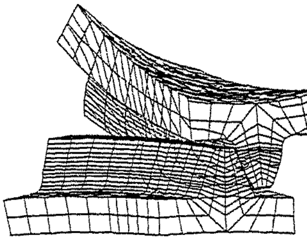


Fig. 5 : Surface displacement contour plot

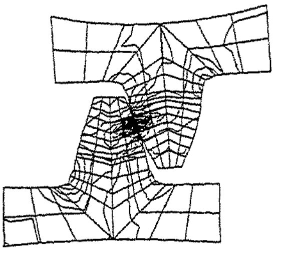
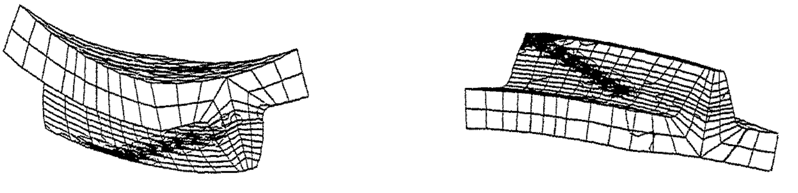


Fig. 6 : Section stress contour plot



(a) Pinion **(b) Gear**

Fig. 7 : Surface stress contour plot

7. CONCLUSIONS

The illustrative examples of cylindrical gear transmission applied to construction machinery and machine tools show that this procedure is valid and efficient. This integrated CAE software system is a useful tool for design and analysis automation.

8. REFERENCES

1 Yang Kaixiu Optimal Design of Planetary Gearing, Engineering Machinery, 1992.12. (in Chinese)

2 Li Runfang. 3-D Strain-stress Numerical Analysis and Experimental research of Meshing Gears, Chinese Journal of Mechanical Engineering, 1994, 30(2). (in Chinese)

3. J C.Zhu, R.F.Li, O.C.Zienkiewize. Adaptive Method for Engineering Analysis with Application to Contact Problem, Proc of 20th, MWMC, USA 1987.8.

PARAMETRIC GEOMETRY MODELING BASED ON AN EXTENDED HYPERGRAPH

Meng Xiangxu, Li Xueqing, Gongbin, Wang Jiaye and Liu Shenquan

CAD Laboratory
Institute of Computing Technology
Academia Sinica, P. R. China

Department of Computer Science
Shandong University
P. R. China

ABSTRACT

A new parametric geometry modeling based on an extended directed hypergraph is described in this paper. It can support the representation of dimensional constraints and geometric constraints. Moreover, the geometric structure changes and design constraints can also be represented. The parametric geometry modeling is finished while interactively constructing geometric objects. The constraint consistency is assured by construction steps. Overheadconstrained or underconstrained sketches can not take place. In this paper, The structures of the geometric modeling and the dynamic hidden line remove algorithms are given in details. The implementation suggests that the modeling is practicable and convenient for parametric geometry. It can be used to realize products oriented parametric designs and drafting.

KEYWORDS

Parametric Drafting, Geometric Constraints, Directed Hypergraph

1. INTRODUCTION

While algorithms for rendering models of mechanical engineering and other complex assemblies advance rapidly, methods for creating and manipulating such models are not keeping pace. For example, after looking at a rendering of a machine design, a designer may decide that certain dimensions of the machine should be changed. Although conceptually simple, this operation may require changing thousands of coordinates and coordinate transformations for various machine components, a time-consuming and error prone task. As a result of this difficulty, complex models are often not created until the design is nearly complete, thus limiting the impact of the computer on the design process.

Parametric designs offer a solution to this problem. Constraints automatically maintain desired spatial and dimensional relationships, so that a model can be manipulated using a relatively small number of parameters to focus on relevant design properties, while the constraint manager looks after the details. However few proposed methods of variational geometry are suited to complex drawings and assemblies. With numerical constraint solvers[1]-[3], the constraints are translated into a system of equations and are solved using an iterative method. When based on Newton iteration, such solvers require good initial values, so that the initial sketch must almost satisfy all the constraints already. Such solvers are quite general, and are capable of dealing with overconstrainted, consistent constraint problems. Nonlinear systems have an exponential number of solutions, but Newton iteration will find only one. Numerical solvers are therefore inappropriate when the initial sketch is only topologically correct.

Constructive constraint solvers[4]-[6] are based on the fact that most configurations in an engineering drawing are solvable by ruler, compass and protractor, or by using another, less classical repertoire of construction steps. In these methods, the constraints are satisfied constructively, by placing geometric elements in some order. This is more natural for the user, and it makes the approach suitable for interactively creating a sketch. In such solvers, it is difficult to handle the functional relationships between dimension variables that can express design intent very flexibly.

In this paper, A new parametric geometry modeling based on an extended directed acyclic hypergraph is proposed. A two-dimensional model is represented as a set of geometric elements and a set of constrains. With this work, the purposes and goals are as follows:

- The user should be given support when defining a geometric model. The constrains should be get automatically according to variable expressions and command parameters. The validity of model should be ensured by itself.
- The model should be capable of handling the functional relationships between dimension variables, because it can express design intent very flexibly. Simple functional relationships are the content of certain geometry theorems, such as the theorems of proportionality, and many other classical results. In general, functional relationships between dimension variables necessitate additional mathematical techniques and application oriented design knowledge. Moreover, design parameters should be supported to reduced geometric degrees of freedom (the numbers of dimension variables).
- The model should support topological information parameterizations of geometry. When the topological parameters were changed, the topological and structural information of geometry should be changed accordingly. For example, assume that the number of holes in a plane was a topological parameter. When the number was changed from four to six, the number of the holes drawn in the plane must be changed to six automatically.

2. THE GEOMETRIC MODELING BASED ON AN EXTENDED HYPERGRAPH

The core constraint model is based on an extended directed acyclic hypergraph (a graph where links are nary rather than binary). The *constrainter* is defined as follow in this paper. It is a constraint that locates a single item, called *the target item*, with respect to one or more other items, called *source items*. A *constrainter* serves a role similar to that of a formula in a speadsheet, determining the value of one entity based on the values of others. 'Value' has different meanings for different kinds of *constrainter*. The node of a hypergraph is an item that may be value (such as a geometric dimension, a topological parameter, a variable, a constant), character point and geometric element (such as a point, a line, a circle, a curve). *constrainters* are the nary links directed from the source items to the target. There are three kinds of nary. One is the constraints of variables that are represented by expressions, the other is the constraints of character points and the another is the constraints of geometric elements. All the items are linked by the constraints according its type. Before an item was added to a hypergraph, all of its source items must have been added to the hypergraph. All geometric elements are added to the hypergraph in an appropriate order.

2.1 Constrainters of Variables

The constrainters of variables is used to represent the constraint relations among dimensions of geometry. The *target item* is a value type. The *source items* may be value types, point types or geometric element types. The constrainters are represented by mathematics expressions.

Constraint expressions are defined as follows:

Expression is composed of operands and operators. It is similar to C language.

Operands can be constants, variables (a kind of node in a hypergraph) and a lot of functions.

Operators can be arithmetic operators, compare operators and Boolean operators.

The prototype implementation includes about 30 kinds of functions. It is mathematics functions (sine, cosine, tangent, square root and so on), geometric functions and others. The following is a partial listing of geometric functions.

Dist-point-line: Calculate the distance between a given point and a given line.

Dist-point-point: Calculate the distance between two points.

Angle-line-x: Calculate the angle between a given line and x axes.

Angle-line-line: Calculate the angle between two given lines.

2.2 Constrainters of Points

The constrainters of points is used to represent the point constraints relations with points, variables and geometric elements of geometry. The *target item* is a point type. The *source items* may be value types, point types or geometric element types. The constrainters are represented by the links and constraint types. There are about 40 kinds of constrainters of points in the prototype. The following is a partial listing of point constrainters.

Pnt-x-y: Locate a point by two values as its x, y coordinates.

Pnt-angle-length-point: Locate a point by the angle A and the length L with respect to the point P.

Pnt-line-line: Locate a point at the intersection of the two lines.

Pnt-line-circle-mark: Locate a point at the intersection of line and circle near the given reference point.

Pnt-circle-circle-mark: Locate a point at the intersection of circle A and circle B near the given reference point.

Pnt-outertanline-circle-circle-mark: Locate a point at the outer tangent point of circle A and circle B near the given reference point.

2.3 Constrainters of Geometric Elements

The constrainters of geometric elements is used to represent the element constraints relations with points, variables and geometric elements of geometry. The *target item* is an element type. The *source items* may be value types, point types or geometric element types. The constrainters are represented by the links and constraint types. There are about 20 kinds of constrainters of geometric elements in the prototype.

2.4 The Formal Representation of the Modeling

The geometric modeling based on an extended directed acyclic hypergraph (GMH) can be described as follows:

$PD=<\Sigma, C>$

PD is a hypergraph that represent a parametric drafting

Σ is a set of nodes of hypergraph, $\Sigma = G \cup P \cup D$,

where

$G=\{g_1, g_2, \cdots, g_l\}$ is a set of geometric element nodes that represent the geometric element, such as lines, circles, arcs, curves, dimensions and subgraphs.

$P=\{p_1, p_2, \cdots, p_m\}$ is a set of point nodes.

$D=\{d_1, d_2, \cdots, d_n\}$ is a set of value nodes that represent the dimensions, length, variables and any other constants.

C is a set of directed edges of hypergraph, $C=\{<v_{d1}, V_1>, <v_{d2}, V_2>, \cdots, <v_{dk}, V_k>\}$

where:

$v_{di} \in \Sigma$, $i=1 \cdots k$, Represent the target node of constraint

$V_i \subseteq \Sigma$, $i=1 \cdots k$, Represent the source nodes of constraints

If the in-degree of a node is zero, the node is a start point of the hypergraph. These nodes consist of the parameters of the engineering drawing. When the initial values of the nodes are given, an instance of the parametric model can be obtained. The node type can be value or point type.

Figure 1(a) show an example of a simple engineering drawing and Figure 1(b) its hypergraph representation. Assume the width of the rectangle is w, the height is h, the radius of fillet is r. Symbol w, h, r is variation dimensions.

The thick directed edge chains represent the calculation order of the geometric elements. According to the directed chain order, The correct results of constraint geometry can be obtained.

The thin directed edges represent the target nodes and source nodes of the constrainters. The tail of directed edges is point to the premise node of constrainters. The leading of directed

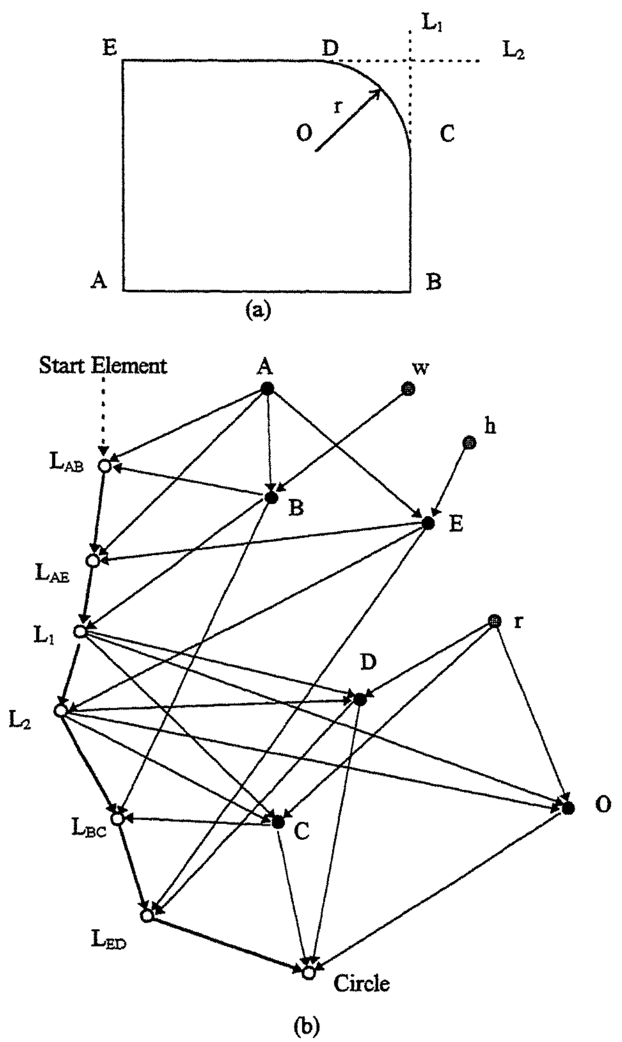


Figure 1

edge is point to the target node.

The node A, w, h, r is initial node. These are also the parameters of the drawing.

3. THE HIERARCHICAL MODEL AND DYNAMIC HIDDEN LINE REMOVE ALGORITHM

In the model GMH, a new kind of node—hyper-node is introduced. A hyper-node is a geometric element that is an instance of other GMH with the corresponding constrainters. When the hyper-node is processed, the values of the source nodes are transfer to the GMH and then calculate the GMH recursively. The hierarchical parametric geometric model is formed. The model is convenient to make engineering drawings and specially to produce the assembly of mechanical parts and building designs. To speed up the drawing works, we introduce another kind of geometric element—clip area, to achieve the dynamic hidden line remove for the instance of the model.

In the model, each geometric element belongs to a layer. The layer number is from 1 to 256. If the layer number of an element is bigger than the lay number of clip node, The element must be clipped with the clip area. Because the clip areas are changed with the parameters of the drawing, the clip areas need to be calculated first

Given a numerical value for each parametric constraint of a hypergraph, a variant is rapidly generated. Processing of an extended hypergraph is performed by a solver. The solver begins from the first geometric element and checks whether its source items were computed (if it was not computed, some source items would be computed first), then computes and outputs the geometric element with the clip area. In this way, All the elements are computed in an appropriate order.

The dynamic hidden line remove algorithm is divided into two steps. The first step is the creating process of dynamic clip area, And the second step is the hidden line remove for the geometric elements.

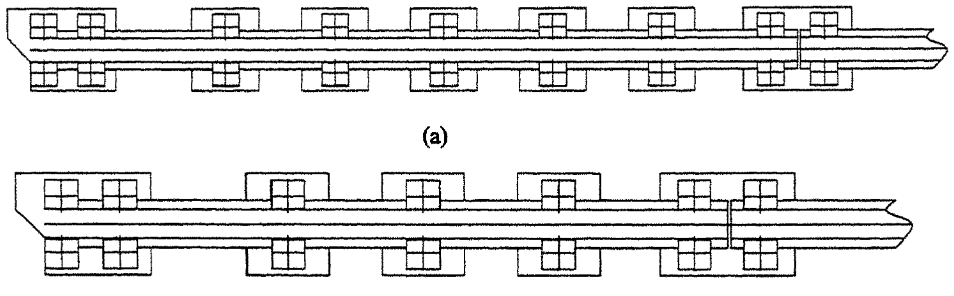
Step 1:
Initialize the clip area chains.
For each geometric element in the GMH Do
1. If it is a clip type node Then calculated the clip area and saves to the chain.
2. If it is a hyper-node Then create a new sub-layer clip area in the clip area chain and calculate all clip areas in the instance of nested GMH recursively.

Finally the hierarchical clip areas are created.

Step 2:
Put the current geometric into the clip stack.
For each geometric element in the stack.
For each nested GMH layer Do
    For each clip area in the layer Do
        clip the element with the clip area and put the visible part into the stack
Draw the visible part of the element

4. THE IMPLEMENTATION AND EXAMPLES

A prototype system PDPG based on the model has been developed by the researcher. It was written in C++ language. It can run on IBM PC compatible computer with DOS or WINDOWS 3.X operating system. It has been successfully used to create the parametric engineer drawings for a motor factory. Figure 2 shows the engineering drawings produced by the system. The long horizon line is clipped by the box element. The number of the box is a parameter. It changes from 5 (Figure 2 a) to 3 (Figure 2 b). Figure 3 shows a mechanical assembly drawing.



(a)

(b)

Figure 2

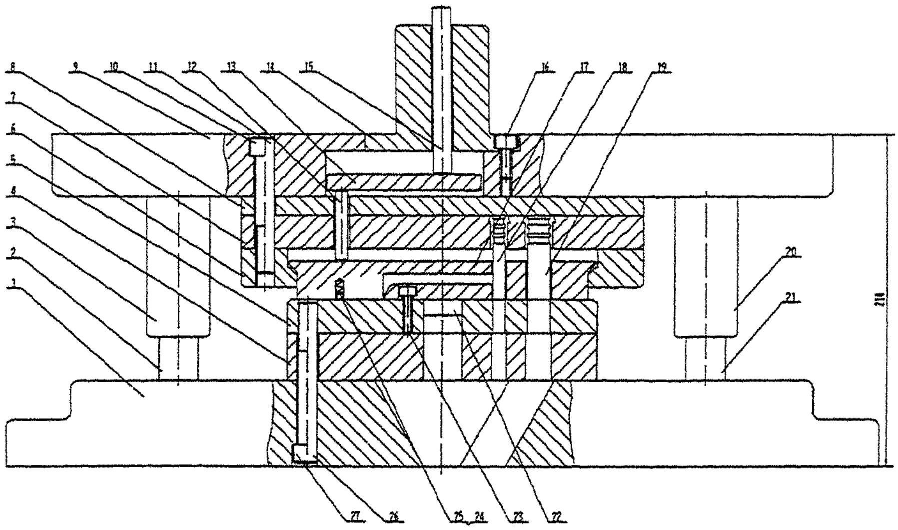


Figure 3

5. CONCLUSION

In this paper, A new parametric geometry modeling based on an extended directed acyclic hypergraph is described. It can support the representation of dimensional constraints and geometric constraints. Moreover, the geometric structure changes and design constraints can also be represented. The parametric geometry modeling is finished while interactively constructing geometric objects. The constraint consistency is ensured by construction steps. Overheadconstrained or underconstrained sketches can not take place. In this paper, The structures of the geometric modeling and the dynamic hidden line remove algorithms are given in details. The implementation suggests that the modeling is practical and convenient for parametric geometry. It can be used to realize products oriented parametric designs and drafting. The software has been used in several factories as a core system platform. The model will be extended to 3D case to support parametric solid modeling.

6. ACKNOWLEDGEMENTS

The project was supported by Shandong Natural Science Foundation. We gratefully acknowledge their support. The author also thanks all research fellows for many helpful discussions and suggestions.

7. REFERENCES

1. Lin, V. C. Gossard, D. C. Light, R. A., 'Variational Geometry In Computer Aided Design', Computer Graphics, Vol 15 No 3 (pp 171-177) 1981
2. Light R A. Gossard D C., 'Modification Of Geometric Models Through Variational Geometry', CAD, Vol 14 No 4 (pp 209-214) 1982
3. Aldefeld B., Variational Geometries Based On A Geometric-Reasoning Method, CAD, Vol 20 No 3 (pp 117-126) 1988
4. Watson S., Relational Geometry - A New Generation Of Two Dimensional CAD, Computer Aided Engineering Journal, 1988,8
5. Verroust.A, Schonek.F, and Roller.D, 'Rule-Oriented Method For Parameterized Computer-Aided Design', CAD, Vol 24 No 3 (pp) 1992
6. Meng xiangxu Wang jiaye, A New Generator For Parametric Drafting Program, CADDM,Vol 2 No 1, (pp) 1992

ISOMETRIC TRANSFORMATION BETWEEN 2D AND 3D SURFACES

Z. Wu and M. M. F. Yuen
Department of Mechanical Engineering
Hong Kong University of Science and Technology
Hong Kong

ABSTRACT

In this paper, an isometric mapping method for apparel design is proposed. The 3D garment surface and the 2D pattern surface were decomposed into a mesh with identical topological structure. Then the transformation operations of wrapping (2D to 3D mapping) and flattening (3D to 2D mapping) were conducted. The operations include three steps: meshing, mapping and skinning. In order to simplify the mapping processing, 3D curve and surface description was given in the form of directrix and metric matrices. The matrix forms represent the size and the shape of curves and surfaces explicitly allowing better control in the mapping process. This set of metric matrices forms the necessary information to conduct an isometric transformation from a 3D garment surface to 2D flat garment patterns. The mapping procedure is similar for transformation from 2D flat pattern to 3D garment with the addition provision that the characteristic points on the garment surface are prescribed.

KEYWORDS

Apparel Design, Isometric Mapping, Wrapping, Flattening, Feature-based

1. INTRODUCTION

Apparel manufacturing was traditionally conducted using 2D patterns to produce sample garments to be fitted on 3D mannequins. The fitness of the garment is largely dependent on the experience of the pattern maker. The process is usually time-consuming and inefficient. There is a need to apply the current Computer-Aided-Design (CAD) technology to assist the mapping between 3D garment surface and corresponding 2D clothing patterns, and thus enhance the efficiency of the industry.

In recent years, CAD/CAM technologies have been used in the apparel industry [1]. Hinds and McCartney's [2] [3] represented mannequin with bi-cubic B-spline surfaces. Garment panels were constructed around the static mannequin by geometric methods, and then are reduced to 2D cutting patterns. Thalmann et al [4] created a controlled and realistic animation of physically based flexible model using mathematical constraint methods. Another interesting approach by Kunii and Gotoda [5] incorporated both the kinetic and geometric properties of fabric for generating garment wrinkles and more realistic looking simulations were produced. Wozny et al [6] presented a mapping algorithm to fit a 2D woven-cloth model onto a 3D NURBS surface. In this method the 2D broad-cloth composite ply was deformed for areas in contact with the 3D surface.

In this paper, the authors present a method to describe the geometric features (shape and size) of both 2D clothing pattern and 3D garment surface. A new metric-based representation for curves and surfaces is proposed containing the directrix and metric information which is effective in the isometric mapping between 2D and 3D surfaces.

2. ISOMETRIC MAPPING FOR APPAREL DESIGN

The geometry of a clothing pattern is defined in two dimensional space. Sewing them together will form a garment in three dimensional space. However the shape of a garment surface is influenced by factors such as the characteristics of clothing material and the shape of mannequin. In order to simplify the problem, the proposed mapping between 2D clothing pattern and 3D garment surface is constrained to be metric consistent. The isometric mapping means that the metric measurement is preserved during the transformation while the 'shape' would be governed by some external constraints.

The garment surface generated by the mapping from the 2D pattern is assumed to be smooth with no folds nor wrinkles. To present the garment surface in its natural look, further transformation using physically based soft models is needed.

There are two kinds of isometric mapping in apparel design:

- **wrapping** - mapping 2D clothing pattern to 3D garment surface, and
- **flattening** - mapping 3D garment surface to 2D clothing pattern.

The 2D clothing pattern and the 3D garment surface are decomposed into a common topological mesh which will guarantee the topological consistency in both the wrapping and flattening processes.

The wrapping and flattening transformations each includes three operating steps: meshing, mapping and skinning. The mapping of free-form surfaces cannot be handled explicitly and a metric-based description of curves and surfaces is introduced to facilitate the isometric mapping process.

3. A METRIC-BASED DESCRIPTION OF FREE-FORM CURVES

Firstly, a 3D curve was subdivided to m1+m2+1 elements by a set of points $P_i\,(-m1 \le i \le m2)$ and $P_0\,(x_0\, y_0\, z_0)$ is called the base point. The symbols $l_i = \left| P_i - P_{i-i_s} \right|$ and $D_i = (P_i - P_{i_s}) / l_i$ denote respectively the chord length and direction of the ith section, where i_s is a function with the following characteristics:

$$i_s = \begin{cases} 1 & i \ge 0 \\ -1 & i < 0 \end{cases}.$$

Given two orthogonal unit vectors *Eu* and *Ev*, the third unit vector *Euv*= $Eu \times Ev$ could be obtained. *Eu*, *Ev* and *Euv* forms a 3D coordinate frame as shown in figure 1. The projections of *Di* on the three coordinate axes can be expressed as follows,

$$t_{ui} = Di \cdot Eu$$

$$t_{vi} = Di \cdot Ev$$

$$t_{uvi} = Di \cdot (Eu \times Ev)$$

Let β_i represent the angle between *Di* and the plane formed by *Eu* and *Ev*. Let α_i represent the angle between *Eu* and the projection of *Di* on the plane formed by *Eu*, *Ev*. From vector algebra,

$$\beta_i = arcsin(t_{uvi})$$

$$sin(\alpha_i) = t_{vi} / cos(\beta_i)$$

$$cos(\alpha_i) = t_{ui} / cos(\beta_i)$$

Where $0 \le \alpha_i < 2\pi$, $-\pi/2 \le \beta_i < \pi/2$.

As i varies from -m1 to m2, α_i, β_i and l_i will form three (m1+m2+1) vectors *[A]*, *[B] and [L]*. Define $\alpha_0 = x_0$, $\beta_0 = y_0$ and $l_0 = z_0$, we have

$$[A] = \begin{bmatrix} \alpha_{-m1} & \cdots & \alpha_{-1} & x_0 & \alpha_1 & \cdots & \alpha_{m2} \end{bmatrix}$$

$$[B] = \begin{bmatrix} \beta_{-m1} & \cdots & \beta_{-1} & y_0 & \beta_1 & \cdots & \beta_{m2} \end{bmatrix}$$

$$[L] = \begin{bmatrix} l_{-m1} & \cdots & l_{-1} & z_0 & l_1 & \cdots & l_{m2} \end{bmatrix}.$$

[A][B][L] form a metric-based representation of the 3D curve. While *[P]* can be expressed in x-coordinate, y-coordinate and z-coordinate by the three matrices *[X][Y][Z]*, the equivalence of the two forms is demonstrated at the end of this section.

Matrix *[L]* contains the metric information and describes the **size** or the length of the 3D curve. Matrices *[A]* and *[B]* contain the directrix information and describe the **shape** of the 3D curve. *[A][B][L]* are geometric parameters that are independent of the choice of the coordinate frame. Using *[A][B][L]* to describe a 3D discrete curve, the size and shape of the curve can be controlled explicitly. For example, if all $\beta_i = 0$ $(-m1 \le i < m2 \cap i \ne 0)$, the 3D curve will become a 2D curve in the plane formed by *Eu* and *Ev*; if $\beta_i = \beta_j\ (i \neq 0)$, $\alpha_i = \alpha_j\ (0 < i \leq m2)$ and $\alpha_i = p + \alpha_j\ (-m1 \leq i < 0)$, the curve will becomes a straight line. When *[A]* and *[B]* are unchanged while $[L'] = k[L]$, k is an arbitrarily scalar, then the curve will be scaled in size without change in direction. When *[L]* is unchanged while *[A]* and *[B]* changed to *[A']* and *[B']* respectively, then the curve will have the same length but takes a different shape.

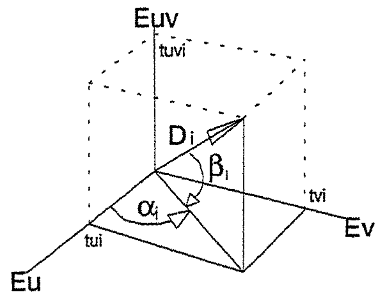


Figure 1: The coordinate frame of curve

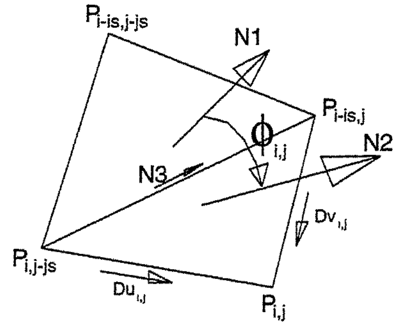


Figure 2: The coordinate frame of surface

4. A METRIC-BASED DESCRIPTION OF FREE-FORM SURFACES

Suppose a surface is subdivided into (m1+m2)×(n1+n2) quadrilateral elements by points $P_{i,j}\ (-m1 \leq i \leq m2,\ -n1 \leq j \leq n2)$ denoted by a (m1+m2+1)×(n1+n2+1) matrix *[P]*. $P_{0,0}\ (x_{0,0}, y_{0,0}, z_{0,0})$ is the base point of the surface, $P_{i,0}\ (-m1 \leq i \leq m2)$ form the base line in u direction, $P_{0,j}\ (-n1 \leq j \leq n2)$ form base line in v direction, i_s and j_s are two functions defined as,

$$i_s = \begin{cases} 1 & i \geq 0 \\ -1 & i < 0 \end{cases} \qquad j_s = \begin{cases} 1 & j \geq 0 \\ -1 & j < 0 \end{cases}.$$

Referring to figure 2, let $e_{i,j} = \left|P_{i,j} - P_{i,j-j_s}\right|$ and $g_{i,j} = \left|P_{i,j} - P_{i-i_s,j}\right|$, where $-m1 \leq i \leq m2\ (i \neq 0)$ and $-n1 \leq j \leq n2\ (j \neq 0)$. Defining two units vectors $Du_{i,j}$ and $Dv_{i,j}$ through $P_{i,j}$ as

$$Du_{i,j} = P_{i,j} - P_{i,j-j_s} / e_{i,j}$$
$$Dv_{i,j} = P_{i,j} - P_{i-i_s,j} / g_{i,j}$$

The unit normal vector of the plane constructed by $P_{i-i_s,j-j_s}$, $P_{i,j-j_s}$ and $P_{i-i_s,j}$ is given as

$$N1 = Dv_{i,j-j_s} \times Du_{i-i_s,j}$$

Similarly, the normal unit vector of the plane defined by $P_{i,j}$, $P_{i,j-j_s}$ and $P_{i-i_s,j}$ is given as

$$N2 = Dv_{i,j} \times Du_{i,j}$$

Let $N3 = (P_{i-i_s,j} - P_{i,j-j_s}) / \left|P_{i-i_s,j} - P_{i,j-j_s}\right|$ and let $\phi_{i,j}$ denotes the angle between *N1* and *N2*, then

$$\phi_{i,j} = arcsin(N1 \times N2) \quad (-\pi/2 \leq \phi_{i,j} \leq \pi/2)$$

When i varies from -m1 to m2 and j varies from -n1 to n2, $e_{i,j}$, $g_{i,j}$ and $\phi_{i,j}$ will form three (m1+m2+1)×(n1+n2+1) matrices *[E]*, *[G]* and *[Φ]*. Suppose that the base line in u direction is given by $[A]^u\ [B]^u\ [L]^u$, and the base line in v direction is given by $[A]^v\ [B]^v\ [L]^v$, the detailed information are:

$$\begin{cases} e_{i,0} = \alpha_i^u \\ g_{i,0} = \beta_i^u \\ \phi_{i,0} = l_i^u \end{cases}, \qquad \begin{cases} e_{0,j} = \alpha_j^v \\ g_{0,j} = \beta_j^v \\ \phi_{0,j} = l_j^v \end{cases} \qquad \text{and} \begin{cases} e_{0,0} = x_{0,0} \\ g_{0,0} = y_{0,0} \\ \phi_{0,0} = z_{0,0} \end{cases}.$$

The detailed information of [E], [G] and [Φ] are,

$$[E]=\begin{bmatrix} e_{-m1,-n1} & \cdots & e_{-1,-n1} & \alpha^{v}_{-n1} & e_{1,-n1} & \cdots & e_{m2,-n1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \iddots & \vdots \\ e_{-m1,-1} & \cdots & e_{-1,-1} & \alpha^{v}_{-1} & e_{1,-1} & \cdots & e_{m2,-1} \\ \alpha^{u}_{-m1} & \cdots & \alpha^{u}_{-1} & x_{0,0} & \alpha^{u}_{1} & \cdots & \alpha^{u}_{m2} \\ e_{-m1,1} & \cdots & e_{-1,1} & \alpha^{v}_{1} & e_{1,1} & \cdots & e_{m2,1} \\ \vdots & \iddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ e_{-m1,n2} & \cdots & e_{-1,n2} & \alpha^{v}_{n2} & e_{1,n2} & \cdots & e_{m2,n2} \end{bmatrix}$$

$$[G]=\begin{bmatrix} g_{-m1,-n1} & \cdots & g_{-1,-n1} & \beta^{v}_{-n1} & g_{1,-n1} & \cdots & g_{m2,-n1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \iddots & \vdots \\ g_{-m1,-1} & \cdots & g_{-1,-1} & \beta^{v}_{-1} & g_{1,-1} & \cdots & g_{m2,-1} \\ \beta^{u}_{-m1} & \cdots & \beta^{u}_{-1} & y_{0,0} & \beta^{u}_{1} & \cdots & \beta^{u}_{m2} \\ g_{-m1,1} & \cdots & g_{-1,1} & \beta^{v}_{1} & g_{1,1} & \cdots & g_{m2,1} \\ \vdots & \iddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ g_{-m1,n2} & \cdots & g_{-1,n2} & \beta^{v}_{n2} & g_{1,n2} & \cdots & g_{m2,n2} \end{bmatrix}$$

$$[\Phi]=\begin{bmatrix} \phi_{-m1,-n1} & \cdots & \phi_{-1,-n1} & l^{v}_{-n1} & \phi_{1,-n1} & \cdots & \phi_{m2,-n1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \iddots & \vdots \\ \phi_{-m1,-1} & \cdots & \phi_{-1,-1} & l^{v}_{-1} & \phi_{1,-1} & \cdots & \phi_{m2,-1} \\ l^{u}_{-m1} & \cdots & l^{u}_{-1} & z_{0,0} & l^{u}_{1} & \cdots & l^{u}_{m2} \\ \phi_{-m1,1} & \cdots & \phi_{-1,1} & l^{v}_{1} & \phi_{1,1} & \cdots & \phi_{m2,1} \\ \vdots & \iddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_{-m1,n2} & \cdots & \phi_{-1,n2} & l^{v}_{n2} & \phi_{1,n2} & \cdots & \phi_{m2,n2} \end{bmatrix}$$

The matrices *[E][G][Φ]* form a metric based representation of surface. Like the 3D curve case, *[P]* can be expressed by three (m1+m2+1)×(n1+n2+1) matrices of *[X] [Y] [Z]*.

[E] and *[G]* contain the metric information of the boundary curves and describe the **size** of the surface. *[Φ]* contain the directrix information and describes the **shape** of the surface. *[E][G][Φ]* are geometric invariant. Using *[E][G][Φ]* to describe a 3D discrete surface, the size and the shape of the surface are described explicitly. When $f_{i,j} = 0\ (i \neq 0 \cap j \neq 0)$, $P_{i-i_s,j-j_s}$, $P_{i,j-j_s}$, $P_{i-i_s,j}$ and $P_{i,j}$ will be in the same plane. When the two base lines in u and v directions are coplanar and $f_{i,j} = 0\ (-m2 \leq i \leq m1,\ -n2 \leq j \leq n1)$, all $P_{i,j}\ (-m2 \leq i \leq m1,\ -n1 \leq j \leq n2)$ will belong to the same plane. The new expression for the 3D surface also possesses similar scaling and bending functions as in the case of the 3D curve.

The matrices *[E]*, *[G]* and *[Φ]* provides a convenient description to support the transformation between 2D and 3D net meshes. Assuming the principle in 2D to 3D transformation and vice versa is to preserve all net edge lengths, the processes of flattening and wrapping can be considered as a special case of surface bending.

5. THE FLATTENING OPERATION IN APPAREL DESIGN

When a 3D net mesh flattens to a 2D net mesh, the following conditions should be satisfied,

$(e_{i,j})_{2D} = (e_{i,j})_{3D}$, $(g_{i,j})_{2D} = (g_{i,j})_{3D}$, $(f_{i,j})_{2D} = (0)\ (i \neq 0\ or\ j \neq 0)$

$(l_i^u)_{2D} = (l_i^u)_{3D}$, $(\beta_i^u)_{2D} = (0)$ $(-m1 \le i \le m2 \text{ and } i \ne 0)$

$(l_j^v)_{2D} = (l_j^v)_{3D}$, $(b_j^v)_{2D} = (0)$ $(-n1 \le j \le n2 \text{ and } j \ne 0)$

In this process, there are m1+m2+n1+n2 degrees of freedom needed to be defined, if the angular values of $(\alpha_i^u)_{2D}$, $(\alpha_j^v)_{2D}$ ($-m1 \le i \le m2$ and $i \ne 0$; $-n1 \le j \le n2$ and $j \ne 0$) are given, the 2D net mesh will be determined. Figure 3 shows an example illustrating the flattening of a 3D net mesh to a 2D net mesh.

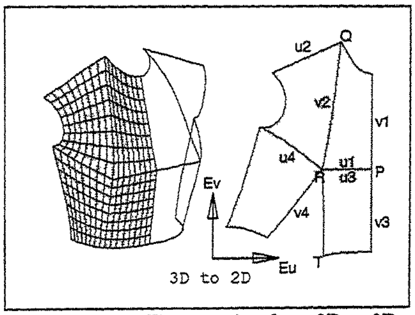


Figure 3: The mapping from 3D to 2D

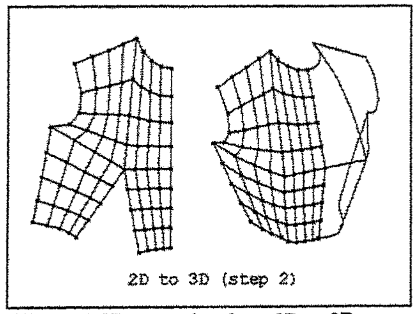


Figure 4: The mapping from 2D to 3D

6. THE WRAPPING OPERATION OF APPAREL DESIGN

When the 2D net information is given, the 3D net to be constructed should satisfy the following conditions,

$(e_{i,j})_{3D} = (e_{i,j})_{2D}$, $(g_{i,j})_{3D} = (g_{i,j})_{2D}$

where, $-m1 \le i \le m2 \cap i \ne 0, -n1 \le j \le n2 \cap j \ne 0$.

$(l_i^u)_{3D} = (l_i^u)_{2D}, (-m1 \le i \le m2 \cap i \ne 0)$

$(l_j^v)_{3D} = (l_j^v)_{3D}, (-n1 \le j \le n2 \cap j \ne 0)$

This process is more difficulty because $[\Phi]_{3D}$, $[A]_{3D}^u$, $[B]_{3D}^u$, $[A]_{3D}^v$, $[B]_{3D}^v$ need to be defined in a specific problem. The example below give a method for determining the above **shape** information of a 3D net.

A garment is designed to fit a specific mannequin, so it is inevitable that points on the garment surface will be related to corresponding points on the mannequin. Prescribing the mannequin model and garment model with a unified net mesh topology, and giving a suitable offset at each node point of mannequin will generate the garment surface. Using the metric-based transformation algorithm, the garment shape information $[\Phi]_{3D}$, $[A]_{3D}^u$, $[B]_{3D}^u$, $[A]_{3D}^v$, $[B]_{3D}^v$, will be obtained and a 3D garment surface net could be constructed. If the effect is not favorable, some adjustment could be attempted. Figure 4 is obtained by using the method mentioned above.

7. DISCUSSION

We know in differential geometry, the isometric mapping S→S' of a bi-parametric surface S(u,v) is only possible if the first fundamental form is preserved. That is E=E', F=F' and G=G', where:

$$E = \frac{\partial S}{\partial u} \cdot \frac{\partial S}{\partial u}, \quad F = \frac{\partial S}{\partial u} \cdot \frac{\partial S}{\partial v} \text{ and } G = \frac{\partial S}{\partial v} \cdot \frac{\partial S}{\partial v}.$$

E and G define the length information in u and v direction respectively and F defines the shear angle. Analytical differential geometry can perform isometric mapping to flatten developable surface. However, the garment surface is always constructed from free-form surface which do not has zero Gaussian curvature value for all of the points. A numerical procedure for isometric mapping for 3D appeal design could be the only viable choice. The proposed method only preserves the length in both u, v directions respectively and the included angle during the transformation and it would be considered as the next logical step to preserve the area of the patch. The area metric value in not included in the metric definition of the surface and only acts as a checking value in the current implementation.

Checking will be conducted to ensure that the 2D and the 3D metric values are within a specified error bound. If there is a significant mismatch in the 2D curve length and the corresponding 3D curve length in the process of 2D to 3D transformation or vice versa, a compensated method can be applied. For instance, if the 2D curve length is l_{arc} and the corresponding 3D curve length is l'_{arc}. Assume $l'_{arc}=kl_{arc}$, where k is a positive scale, a "scale" function can be imposed on the curve transformation by setting $[L]'' = \frac{1}{k}[L]'$. The overall length of the curve will be then compensated.

Similarly, if there is a significant mismatch in the 2D pattern area and 3D surface area in the process of 2D to 3D transformation or vice versa, a compensated method can be applied. Suppose the 2D pattern area is $a_{surface}$ and the corresponding 3D surface area is $a'_{surface}$. Assume that $a'_{surface}=k_u\, k_v\, a_{surface}$, where k_u and k_v are the adjustment values in u and v directions respectively, a "scale" function can be imposed on the surface transformation by setting

$$k_u = k_v = \sqrt{a'_{surface} / a_{surface}}$$

$$[E]' = \tfrac{1}{k_u}[E]' \text{ and } [G]'' = \tfrac{1}{k_v}[G]'.$$

In this way, the surface area will be preserved during the whole deformation of 2D to 3D.

8. CONCLUSION

In this paper, an isometric mapping between 2D garment pattern and 3D garment surface is proposed based on a new metric-based representation of curve and surface. In the new metric-based representation, the shape and size of the curve and the surface are presented in explicit form which will enable the effective flattening and wrapping of surfaces in garment design.

The transformation algorithms are stable and speedy. There is no accumulated error in the process 2D→3D→2D, that is, the original 2D pattern has no different with the last 2D pattern if the parameters remain unchanged.

The isometric mapping for apparel design described in this paper is based geometric mapping. In order to get some folds and wrinkles in the garment surface, a physical based deformable model would be needed. The relationship between feature of clothing material and the shape of wrinkle will be interesting subject and it needs researching further.

9. REFERENCES

1. H. Okabe, H. Imaoka, T. Tomiha and H. Niwaya, Three Dimensional Apparel CAD System, Computer Graphics, 26, 2, July 1992.

2. B.K. Hinds, J. McCartney and G. Woods, Pattern development for 3D surfaces, Computer Aided Design, volume 23, number 8, 1991.

3. B.K. Hinds, J. McCartney, C. Hadden and J. Diamond, 3D CAD for Garment Design, International Journal of Clothing Science and Technology, Vol. 4, No. 4, 1992.

4. M. Cargnan, Y. Yang, N.M. Thalmann, D. Thalmann, Dress Animated Synthetic Actors with Complex Deformable Clothes, Computer Graphics, 26, 2, July 1992.

5. T.L. Kunii and H. Gotoda, Modeling and Animation of Garment Wrinkle Formation Processes, Computer Animation'90, Switzerland, 1990.

6. M. Aono, D.E. Breen and M.J. Wozny, Fitting a woven-cloth model to a curved surface: mapping algorithms, Computer-Aided Design, Volume 26, Number 4, April 1994.

ASPECTS OF THE DESIGN OF A MULTIPROCESSOR SYSTEM FOR A FLEXIBLE MANUFACTURING SYSTEM

H. Abachi* and R. Ibrahim+
*** Electrical & Computer Systems Engineering**
+ Mechanical & Industrial Engineering
Monash University
Caulfield Campus
Victoria 3145 Australia

ABSTRACT

This paper outlines the aspects of architectural design of a multiprocessor system applicable to Flexible Manufacturing System (FMS). The implemented architecture is based on a message passing system in order to resolve the limitations of memory contention that naturally exists in the shared memory configuration. Incorporating multiprocessor system design in FMS architecture fulfils the system requirements and improves speed, accuracy, reliability and performance of the system. Fault-tolerance, flexibility and self-reconfiguribility of the overall system is further improved by incorporating Artificial Intelligence and Expert system in the body of the database management system.

KEYWORDS

Multiprocessor system design, performance, Artificial Intelligence, Manufacturing Message Specification and Flexible Manufacturing System.

1. INTRODUCTION

Rapid advancement of Very Large Scale Integration (VLSI) technology in recent years and decrease in computer cost has encouraged system designers to develop new advanced computer architectures in the area of distributed computer networks. This has contributed significantly to the development of many new concepts and design strategies for parallel processing applications in modern industrial Flexible Manufacturing System.

Generally it is evident that the area of multiprocessor system can be regarded as one of the engineering solutions to many computational problems that suffer from lack of speed, accuracy and performance. In addition the use of multiprocessor system can also improve other engineering areas such as reliability, fault-tolerance and flexibility. This paper demonstrates the design of a Flexible Manufacturing System, using parallel computer network configuration. This was implemented using a group of advanced computers equipped with tool handling capabilities in conjunction with automated materials handling systems. These systems have reconfiguribility characteristics, in order to facilitate manufacturing of a wide variety of parts that can be used in discrete batch production systems such as automotive, aerospace and machine tool industries. Their recent applications are further extended to include IC manufacturing and PCB assembly industries.

The design and implementation of a Flexible Manufacturing System which incorporates a multiprocessor application is complex and involves hardware and software system design. The design parameters for implementing this architecture include equipment selection, layout, material/tool handling system design, pallet and fixture design, control and communication software selection. In addition, selection and design of manufacturing database management system and communication network topology and protocols have received full attention during all stages of the system design implementation. The role of the communication software is to control the system activities by communicating with the PLC system. It also displays the control commands and shows the material movement on display.

2. PARALLEL PROCESSING ARCHITECTURE

Most industrial projects require fast, accurate and reliable processing power. A Flexible Manufacturing System is not excluded by these highly desirable industrial requirements. To fulfil this requirement one common approach is to implement a faster microprocessor or for some applications faster microcontroller. Alternatively use of parallel processing system could equally serve the same purpose. In order to satisfy the former criteria, system designers have adopted the use of microarchitecture as a design tool, or increase of clock frequency and implementation of superscalar and pipelining as an advanced system architecture. The latter approach could be fulfilled by implementing the most advanced parallel processing architectures which includes mesh, torus and hypercube with superhypercube architecture being the latest of this family.

Implementation of parallel processing could either be internal within the microprocessor itself or external through the architectural design of the overall system In this implemented system design the authors have adopted both approaches. This means by appropriate choice of the Pentium microprocessor as the master processor we have implemented parallel processing through pipelining and superscalar mechanism within the processor and externally by incorporating the multibus configuration into parallel processing architecture.

As the number of processors in any parallel processing system increases, undoubtedly the complexity of the system both from hardware as well as software point of view would increase. This in turn would enhance the possibility of both transient and permanent failures and pose serious reliability problems. One solution is to build redundant communication paths and provide spare units and introduce dynamic reconfiguration capability into network design. The performance and fault-tolerance of the system can further be improved if an Artificial Intelligence (AI)/expert system and Neural system are incorporated within database management of the system architecture. Hence the system would be capable of fault recovery, instead of permanent processor failures continually causing system failures.

3. SYSTEM CONFIGURATION

Generally Flexible Manufacturing System has three essential components which includes:CNC machine tools to process parts, a materials handling system to move parts, tools and an overall control and communication system to manage the FMS. The design and implementation of a FMS is a complex task and requires skills in system design (hardware and software) and product design.

The functions performed by the FMS computer control system can be grouped into six categories: control of each machine centre, distribution of control instructions to machines, production control, work handling system monitoring, tool control, and system performance monitoring and reporting. These computer operations can be accomplished by any of several slave processors under direct supervision of the master computer. In general the implemented configuration consists of three levels, The top level shown in the schematic diagram consists of the master processor which is connected to the second level comprising the slave processors. Third level consists of the subsidiary units (machine control units) which includes the robots, CNC Lathe and CNC mill.

4. ARCHITECTURAL DESIGN CRITERIA

One of the simplest multiprocessor system architectures is the common bus which has been used extensively due to its simplicity. However it incurs communication bottlenecks in large networks, and is vulnerable to faults due to the use of global shared memory. Various schemes have been employed such as multiported memory and cross-bar switch architectures to address these limitations. Their implementation usually results in considerable increase in cost, complexity and serve limitations in scalability.

In an attempt to resolve the limitations of memory contention and reliability inherent in shared memory architectures, message passing systems were developed. These systems provide point-to-point communication links between processors, and try to avoid use of the shared memory. This can be achieved when local or private memory is distributed at each processing node, and sharing information takes the form of routing a message from a source node to a destination node. These message passing systems are commonly used in the parallel processing system design. One of the architectures that is suitably designed and used in the area of Flexible Manufacturing System environment is Multibus Parallel Processing System based on the Master-Slave configuration as shown in Fig. 1.

In the first level of the implemented architecture the master processor is a Pentium computer with 166 Mhz clock frequency, 32 MB RAM, 2.5 GB hard disk and MODEM facility. In the second level there are four slave processors each being a 486 microcomputer with 32-bit CPU, 32 megabyte main memory which is a combination of semiconductor memory, E2PROM and RAM, and a 320 megabyte hard disk. These processors are individually connected to a colour display, input unit and floppy disk unit, cassette unit, printer and other peripheral units such as plotter and scanner if needed.

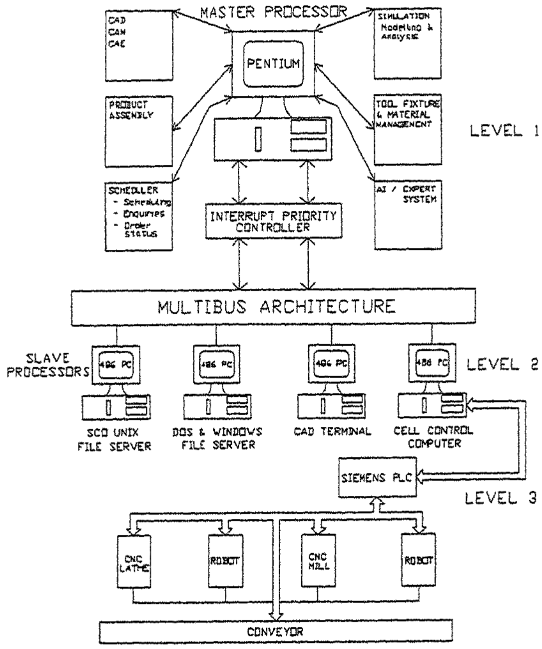


Fig. 1 : Multibus Parallel Processing System for a FMS Configuration

One of the slave processors performs as cell control computer and controls the PLC. The PLC has a central role and it is linked with all stations (NC machine, robot, conveyor) via parallel digital inputs/outputs and linked with the cell controller via RS 232 serial communication link. The designed FMS includes CNC Lathe, milling centre, and robots. If the driving capability of the overall system permits it can be further interfaced to additional automated sub-units. The FMS control system receives manufacturing instructions and necessary information via the Manufacturing Message Specification (MMS) variable access services [1, 2]. The MMS defines building blocks for generation of an abstract of a manufacturing device. It also specifies services which facilitate work with these building blocks. The individual subsidiary units including NC, PLC are tasked and supervised by the supervisory management program of the FMS control system via the MMS [3-6] and AI applications. This involves a protocol stack conversion between the NC system and the FMS control system. Thus the control systems have access to a file server which is responsible for the central data organisation.

5. OPERATION OF THE IMPLEMENTED ARCHITECTURE

One of the main applications of the parallel processing system is to increase the computational power of the system compared to the uniprocessor system. To achieve this, the implemented configuration is outlined. In this architecture the bidirectional nature of the direct communication links between the master processor and slave processors, and consequently the subsidiary functional units of the overall system, allows receiving and transmitting of information in a real time environment. In addition, due to the multiuser and multitasking nature of the hardware and software of the master and slave processors fast and reliable access to all subsidiary units can take place.

The operation of the implemented architecture can best be explained in terms of dividing the main manufacturing task into independent sub-tasks and loading them into the main memory of the master processor. The master processor then allocates each sub-task into the memory buffer of each slave processor. From this point onwards each subsidiary unit functions independently and the result of each sub-task is transferred to the master processor via slave processor. Upon completion of each sub-task subsidiary units send an interrupt request signal to the slave processor and that is forwarded to the master processor for allocation of a new sub-task for processing purpose. This process of allocating sub-tasks will continue until the task is completed. Upon existence of simultaneous interrupt and their generation by the subsidiary units an interrupt priority controller will allocate priority to the slave processors and they receive service in time

Within the FMS while subsidiary units are performing their tasks, the master processor can perform other functions and responsibilities. These include:

- CAD/CAM (design, drafting, NC tool path)
- Product Assembly (feasibility check for product)
- Schedular (scheduling, enquires, order status.)
- Simulation (modeling & analysis)
- Tool fixture & material management (availability and status check).

During this period the main functions of the slave processors include but are not limited to :

- SCD Unix (file server, DOS and Windows),
- File server, CAD terminal and cell controller which mainly involves the PLC controller.

For a given task, depending on the machine status, the PLC has to make a decision based on the information provided by the master processor on how to start and co-ordinate the whole system in a correct sequence of operation and feed back information relating to the status of the whole system for future reference to the master processor.

6. CONCLUSIONS

Advanced microprocessors and microcontrollers play a dominant role in the development of many new concepts and design strategies for modern industrial systems. This contribution is significant in the area of FMS.

The implemented architecture provides improvements to the production efficiency within manufacturing engineering. Among the most important features of the implemented configuration, are increasing the speed of operation, accuracy, system automation, flexibility and cost reduction which has significant importance in a manufacturing engineering environment.

In order to further improve the system performance, some Artificial Intelligence (AI) and expert system can be incorporated within the master processor. This option will provide self testing and random diagnostic tests and the capability of self reconfiguribility to the system. Taking all these factors into account, one can argue that the area of parallel processing has well and truly left the domain of the pure research laboratory and is finding its way into a wide range of applications in the manufacturing and industrial sectors. This rapidly growing acceptance of parallel processing in the industrial environment is based on the fact that parallel processing can improve computational power compared with a uniprocessor system.

It is generally accepted that in order to improve hardware and software system performance in the manufacturing and industrial environment, implementation and continuous upgrading of all parallel computer based equipment must take place.

7. REFERENCES

[1] Halsall, H., Data Communications. Computer Networks and Open Systems, Addison-Wesley Publishing Company, 1992.

[2] ISO 7498, Information Processing Systems Open Systems Interconnection, Basic Reference Model, 1984.

[3] ISO 9506-1, Information Processing Systems Open Systems Interconnection Manufacturing Message Specification, (Part I) Service Definition, 1984.

[4] ISO 9506-2, Information Processing Systems Open Systems Interconnection Manufacturing Message Specification, (Part 2) Protocol Specification, 1984.

[5] ISO DIS 95-6-3, Information Processing Systems Open Systems Interconnection Manufacturing Message Specification, (Part 3) Robot Specific Message System, 1984.

[6] ISO DP 9506-4, Information Processing System Open Systems Interconnection Manufacturing Message Specification, (Part 4) Numerical Control Message Specification, 1984.

INTEGRATION OF MACHINES AND SOFTWARE PACKAGES TO ACHIEVE COMPUTER INTEGRATED MANUFACTURING

[1]Henry Lau, [2]Felix T.S. Chan, [3]Ralph W.L. Ip
[1]Advanced Technology Education Centre, Australia
[2]Dept. of Ind. & Manuf. Systems Engg., University of Hong Kong
[3]Dept. of Manuf. Engg., City University of Hong Kong

ABSTRACT

The success of Computer Integrated Manufacturing (CIM) relies on an efficient information flow system. The information flow system must be able to provide accurate, updated and efficient bidirectional flow of information among the computer systems of a company. In real companies, machines from various vendors with various specifications and computer software running on various platforms such as Unix, DOS/Windows, Virtual Memory System (VMS) are used. To achieve CIM, techniques must be developed to link these machines and software packages for "seamless" data exchange. This paper attempts to examine the techniques and methodology to link "multi-vendor" machines and "multi-platform" packages together in a cost-effective way.

KEYWORDS

Computer Integrated Manufacturing, Information Exchange, Computer Systems

1. INTRODUCTION

There are three ways to achieve the linking of systems together for information exchange, namely turnkey systems, specific format translation and neutral data format (1,2,3). Turnkey systems are supplied by a number of big computer firms including IBM and Hewlett Packard but these systems are unaffordable by most medium to small size companies. Specific format translation is aimed at the specific link of computer systems by reformatting the data of the source system so that the destination system can "read" the data. Data integration using neutral data format is to integrate systems together by adopting a common or "system-independent" format so that various systems can "read" the data.

Real companies usually have a wide variety of multi-vendor computing equipment and integration of such equipment inevitably leads to a diversity in communication hardware and software. Consequently, traditional information exchange systems consist of specific links for the particular computer systems involved in the information exchange process. The links are specific in the sense that particular communication hardware and software are used for the linkage. Changes to both communication hardwares and softwares are necessary if any of the involved computer systems are modified or replaced by another system.

2. DATA EXCHANGE BETWEEN VARIOUS SUBSYSTEMS

In a previous paper related to the linking of CAD system to MRP-II system (1), the technique using format translation is described. The connection between AutoCAD and IMPCON (an MRP-II package) is a simple RS-series connection between two computer systems (4). However, the communication connection is more complex when more computer systems are involved in the data exchange process. In such a case, it is useful to divide the data into groups. As shown in Fig. 1, the manufacturing subsystems are divided into three groups namely Marketing, Design/Planning and Manufacturing. The Marketing group consists of two subsystems namely Sales/Marketing and Purchase. These two subsystems are involved in the data exchange with MRP-II system which belongs to the Design/Planning group. As indicated in the diagram, this group contains four manufacturing subsystems including CAD, MRP-II, Production Scheduling and Computer aided Process Planning (CAPP). Design and planning data is shared among the subsystems of this group as well as subsystems of Manufacturing group which consists of manufacturing subsystems such as CAM, Automatic Storage and Retrieval System (AS/RS), Automatic Testing Equipment (ATS) and Flexible Manufacturing System(FMS). The manufacturing of the product is processed in this group. This group receives the design and planning information from the Design/Planning group and it also transfers manufacturing data such as actual production and inventory data to the Design/Planning group. The effective exchange of information among these three groups is important to enhance productivity in the organisation.

3. INFORMATION FLOW SYSTEM FOR SPECIFIC LINKING OF SYSTEMS

For a simple file transfer between two systems, a RS-series connection is sufficient. However, in the case as shown in Fig. 1, RS-series connection is not enough. Computers running on various operating systems such as DOS/Windows, Unix and Macintosh can be networked for file transfer using well developed commericial networking products. In addition, security of files e.g. restriction of access of certain files by unauthorized users, can also be achieved. However, this networking of computer systems using commercial networking products such as Netware is inadequate for achieving the data exchange as shown in Fig. 1. Networking of computers ensures the transfer of files among computers based on some handshaking rules (protocols), it does not ensure the transfer of selected data e.g BOM, tool data from one system to the other. In this respect a specially designed Information Flow System , which ensures that the destination system can "read" the data from the source system, is required. In short, apart from "physical" communication connection between systems, an efficient Information Flow System must also be able to deliver "selected" data, "reformatted" the data to ensure "seamless" integration with the destination system. In the AutoCAD to IMPCON example (1), the AutoLISP programs are used for the purpose of "selecting" the BOM data and convert the format to suit the IMPCON system.

4. COMMUNICATION LINK COMPONENTS

A number of firms such as IBM, Microsoft, Novell and SUN Microsystems have developed networking products which can handle file transfer among various computing operating systems e.g. DOS/Windows, Unix, OS/2 (5,6). However, these networking products are not able to connect to CNC machines, robots and Automatic Guided Vehicles (AGVs) because most of these equipment do not use any of the common operating systems such as DOS/Windows or UNIX. Instead, manufacturers of this equipment e.g. Fanuc have developed their own simplified "data interpretation/management system" for interfacing the electronic parts of the equipment to the mechanical parts. However, it should be noted that some of the latest "hi-tech" manufacturing equipment are equipped with the common operating system such as UNIX and DOS/Windows. For example, the Selective Laser Sintering Machine Model 2000 runs on UNIX while Digitbotic Non-Contact Laser Scanning Machine and Laminated Object Manufacturing Machine Model LOMM 2030 run on Windows NT. For those machines which do not support a common Operating System (OS), specific communication links are necessary. These are described in the following context.

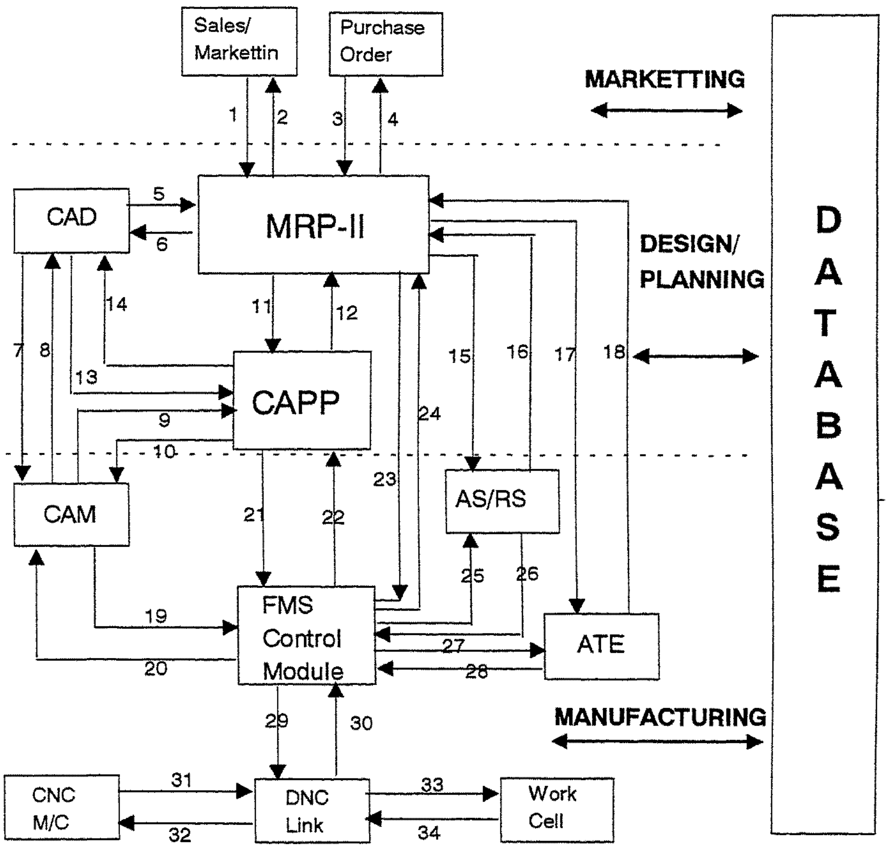


Figure 1 : Information Exchange among Manufacturing Systems

1. Master Schedule Data
2. Actual Prouction Data
3. Place Orders
4. Delivery Progress
5. Bill of Material
6. Inventory Records
7. Geometri Data of Design
8. Information Feedback
9. Feedback Status
10. Process Data
11. Product Information
12. Process Planning Summary
13. Geometric Data
14. Status Feedback
15. Lists of Parts Drawn
16. Inventory Situation
17. Retrieval of Data Analysis
18. Quality Results
19. Status Feedback
20. NC Programming Data
21. Process Operations Info.
22. Data Enquires
23. FMS Production Schedule
24. Orders Completed
25. Matl. Handling Monitoring
26. Feedback Data
27. Quality Checking
28. Results Feedback
29. Monitor M/C Sequence
30. Machine Status Report
31. Feedback Current Status
32 & 33. Download Program
34. Feedback Current Status

5. TYPES OF COMMUNICATION LINKS

The special communication links used in industry can be catorgorised into five types namely RS-series connection, Machine Interface Terminal, Secondary Adaptor, Host Adaptor and Network Interface Module. These five types are described as below.

(a) RS-series connection - This is the simplest type of connection between a computer and the computer system of a machine. In this connection, data can be sent from a computer to the Computer Numerical Control (CNC) machines through the RS-series connection. However, a program is required to enable the transmission of data e.g. a file with NC codes from the computer to the machines. In a Direct Numerical Control (DNC) link, a supervisory computer downloads data e.g. NC codes, to the various CNC machines by means of a RS-series connection.

(b) Machine Interface Terminal (MIT) - MIT helps accomplish a two way data communication links between the supervisory computer and other machines by means of serial (COM Port) or parallel (Printer Port). Suppliers of MIT also provide the communication software which runs on the supervisory computer in order to ensure data transfer to the connected equipment such as CNC machines and robot.

(c) Secondary Adaptor (SA) - This is a key communication link for the integration of the multi-vendor machines. The SA comprises a microprocessor which has a mixture of Read Only Memory (ROM) and Random Access Memory (RAM) according to the interface required. This adaptor is used in conjunction with a Local Area Network (LAN) in order to achieve an efficient performance (6).

(d) Host Adaptor (HA) - This is a general purpose control interface connecting the host computer and the local area network. The host adaptor directs data from the host computer to the various connected systems of the network. It also ensures data integrity in transmission of data in both directions.

(e) Network Interface Module (NIM) - This product allows Programmable Logic Controllers (PLCs) to be interfaced to the network. It is exclusively designed for linking PLC to the network. Its roles in the system are the same as the secondary adaptor.

Referring to Fig. 2, these different types of communication links can be used to connect different computerised manufacturing equipment such as robot, CNC machines, Automatic Guided Vehicle (AGV) etc to a supervisory computer. The supervisory computer normally runs under UNIX or DOS/Windows platforms and therefore it can be connected to the IBM, SUN, Microsoft or Novell networking products. Referring to Fig. 1, the FMS CONTROL MODULE may well be the supervisory computer as illustrated in Fig. 2 for monitoring the equipment of the Flexible Manufacturing System (FMS). With this approach, the manufacturing equipment such as robots, CNC machines, AGVs are integrated with other computer systems such as CAD, CAM, MRP-II indirectly using these special designed communication link products.

6. SUMMARY

The project related to the link-up of AutoCAD and IMPCON (1) for BOM transfer demonstrates the requirements for specific link among computer systems. Essentially, there are two prerequisites for specific link of computer systems. The first is the communication link hardware and software for the specific equipment and the second is related to accurate identification and reformat of data. An efficient information flow system must be able to ensure that required data is transferred and "accepted" by the destination system.

Most of the traditional computerised manufacturing equipment do not support the common OS such as UNIX and DOS/Windows because they have their own computer system which interfaces with the mechanical parts of the equipment. This means that the commercial networking products such as Novell's Netware cannot be used to connect them together. However, some industrial communication link components can be used to link these equipment to a computer which plays the role of monitoring the data communication process and therefore it is also called "supervisory computer". As these supervisory computers run on common OS, they can be connected to a network as provided by commercial computer firm. With this approach, the computerised manufacturing equipment can be connected "indirectly" to other computer systems such as CAD, CAM and MRP-II thus a complete CIM connection involving machines, softwares packages can be accomplished (7,8).

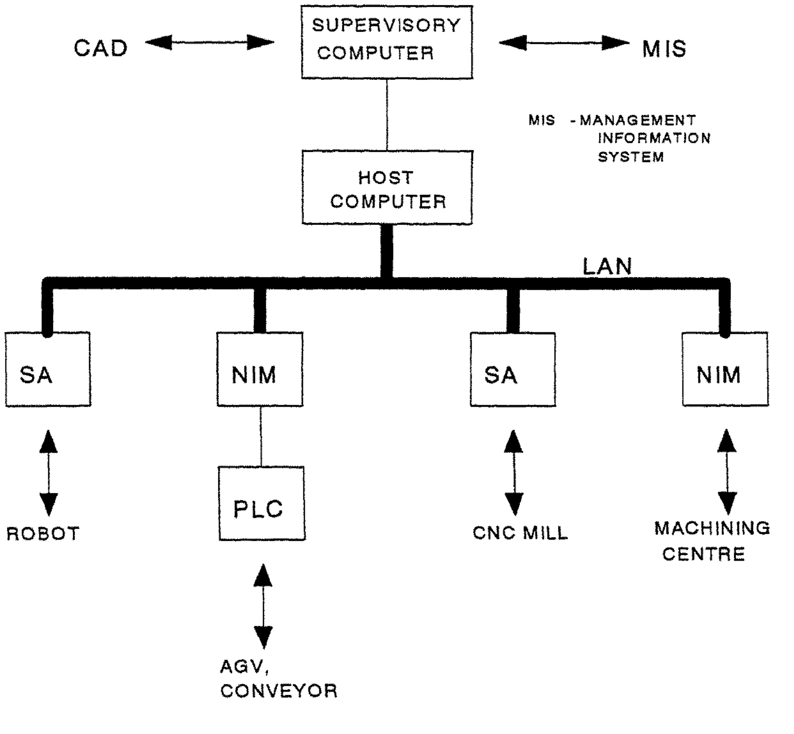


Figure 2 : SA, NIM Connected to LAN

7. REFERENCES

1. Zockel, M. & Lau, H., Development of a generic link between CAD and MRP-II systems, Pacific Conference on Manufacturing Proceeding, Vol 1, 512-518, 1990.
2. Lau, H. & Zockel, M., A Generic Non Standard-specific Rationalised Information Flow and Storage System for Computer-based Manufacturing, Proceedings for Industrial Automation '94 Conference in Singapore, 1994.
3. Lau, H. & Zockel, M., The Role of GNSIFS in Simultaneous Engineering, Proceedings of 3rd International Conference on Manufacturing Technology in Hong Kong, December, 1995.
4. Newton, M.J. & Hurst, K.S., CAD data communications -- an update, Computer-Aided Engineering Journal, February, 1988.
5. Nicoletti, G.M., Computer-integrated manufacturing (CIM), manufacturing automation protocol (MAP), and automated mass finishing systems, SME technical paper MS87-158, 1987.
6. Nicoletti,G.M., CIM: LAN ommunications, protocol standards, and real-time control, Automach Australia, Conference proceedings, May 26-29, 1986, Sydney, Australia, 1986.
7. Ranky, P.G., Computer Integrated Manufacturing, Prentice/Hall International, 1988.
8. Ranky, P.G., A Generic CIM System Modelling Methodology with Practical Examples. Pacific Conference on Manufacturing, Australia, 1990.

AN EFFICIENT APPROACH TO MANUFACTURING COMMUNICATIONS IN A CIM ENVIRONMENT *

Xinquan Ding , Guanqun Gu and Jieyi Wu
Department of Computer Science and Engineering
Southeast University
Nanjing 210096, P.R.China
e-mail : xding@seu.edu.cn

ABSTRACT

In this paper, networking for programmable devices and systems in the automated factory environment is discussed. First, the important features and principles of MMS are described. Then, a method of implementing the interconnection of MAP and BITBUS based on MMS is presented. After that, the hardware and software required to connect the robot to the network system are described in detail. Finally, a summary of the implementation of MMS messaging system in the heterogeneous networks is made in the last part of the paper.

KEYWORDS

Computer Networks, Network Interconnection, Manufacturing Message Specification, Virtual Manufacturing Device, Robot

1. INTRODUCTION

Today's factories are composed of programmable devices and automated systems from a variety of vendors, each with a different proprietary protocol. To enable these different machines to communicate with each other, specific interfaces must be designed to translate one vendor's protocol into another's. Integrating these islands of automation within a factory has proved to be a time-consuming and expensive process. To solve this problem, General Motors made a standardization effort which resulted in the advent of the manufacturing automation protocol (MAP). MAP is based on the ISO/OSI reference model. It is a seven-layer, token bus communication specification LAN suitable for the information integration in a factory automation environment. One of the contributions that MAP has made to the international standard effort is the manufacturing message specification(MMS). MMS is the key component of the MAP specification. It is an application layer standard, and the services it offers allow the user to manipulate a generic software model and to control manufacturing devices.

In the automated factory environment, computer networks are used to connect computers and computer-based devices together. Usually, single type network can not meet all communication needs in a factory. TOP or Ethernet in the office, MAP for the shop and Fieldbus to the end devices is the typical industrial network hierarchy. In this paper, we focus on the interconnection of MAP and Fieldbus. Among several Fieldbus protocols, BITBUS is a good choice because of its high-performance and low-cost. BITBUS network architecture may be divided into three layers : the physical layer (RS-485 transmission link), the data link layer (synchronous data link control , SDLC) and the application layer (remote access and control , RAC). BITBUS is much mature in technique and widely used in China to connect the low-level devices such as sensors, actuators and machine controllers, etc. However, BITBUS doesn't provide users with MMS messaging functions. In this project, we have built a set of MMS services in the application layer of BITBUS according to the ISO/IEC 9506 standard[1, 2]. In order to implement network interconnection, we have also designed a gateway and presented a method of MMS implementation in MAP network.

MMS was designed as a communication standard for messaging between various manufacturing machines[3, 4]. However, most computer-based devices existing in the shop floor are

* This work was supported by the National " 863 " Programme for Research and Development in High Technology under Grant 863-511-09-0144.

incapable of MMS messaging functions. In this paper, we also propose a method of adapting a typical manufacturing device — a computer-controlled robot to MMS compatible one. The remote communication and control of the robot are developed and demonstrated. The method will be useful in adapting other computer-controlled devices to the network system based on MMS.

The remainder of the paper is organized as follows : Section 2 outlines some important features and principles of MMS. Section 3 focuses on the interconnection of heterogeneous networks based on MMS. This section discusses the development of MMS protocol software in MAP network and BITBUS network respectively, and describes the design and implementation of MAP/BITBUS gateway in detail. Section 4 takes a robot as an object of study to demonstrate how a computer-based device could be integrated into the network system and manipulated remotely by using MMS. Finally, Section 5 concludes the paper with a summary.

2. AN OVERVIEW OF MMS

MMS makes use of an abstract object modeling technique in order to exactly describe the device model and the service procedure. In this technique, the objects, their attributes, as well as operations on them are described. The main object defined in MMS is the virtual manufacturing device (VMD). A VMD is an abstract representation of a specific set of resources and functionality at the real manufacturing devices. Each VMD contains exactly one executive function and zero or more program invocations, each of which depends on one or more domains. Besides the VMD object, MMS defines other objects such as the transaction object, the domain object, the program invocation object, etc. These objects are included in VMD and used for the abstract representation of VMD resources and MMS service procedure. All services offered by MMS can be regarded as operations on these objects.

MMS adopts a communication model called Client-Server[5].The Client plays a role of sending the service request to the Server. The Server responds to the request of the Client. It receives the request-PDU, takes corresponding actions to the service request and accordingly modifies the status of the VMD. At last, the Server sends back the response-PDU to the Client as an answer (see Fig.1).

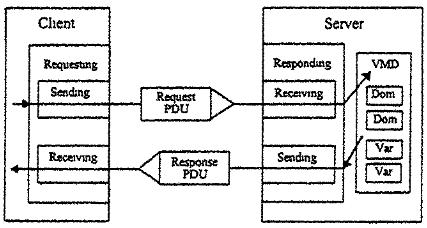


Fig.1: MMS Client-Server model

MMS protocol specification describes the protocol procedure of transmitting data and control information between application entities in the MMS context. It also describes the structure of the MMS PDU used to transmit data and control information. MMS services can be classified as two types : confirmed and unconfirmed. A confirmed service requires a response, and communication is complete only when a response is received. A confirmed service contains four primitives, i.e. request, indication, response and confirmation. An unconfirmed service is one that does not require a response. The application conveys the message and does not expect a reply or acknowledgment. Usually, the details of safe delivery of messages are addressed by the lower layer protocols[6]. An unconfirmed service contains only two primitives : request and indication.

3. IMPLEMENTATION OF NETWORK INTERCONNECTION

3.1 MMS Implementation in MAP Network

In this project, the MAP network we adopt is the product from Concord Communication Inc.

The application layer of the network provides MMS protocol software conforming to the ISO/IEC 9506 specification. As viewed from programming , the Concord MAPware MMS consists of a library of C functions providing the user with a high-level, real-time interface to the MMS application service element which is called a MAPware MMS interface[7]. The interface consists of two parts : the paired-primitive interface (PPI) and the virtual machine interface (VMI). PPI allows MMS user to take complete control over the data sent in MMS messages. VMI offers a higher level of functionality than PPI and provides automatic functions to take care of actions that are specific to the operating system such as reading and writing variables and files. Fig.2 shows the major component of MAPware MMS.

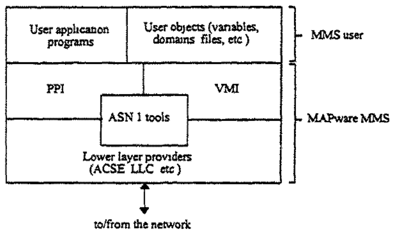


Fig.2: Major components of MAPware MMS

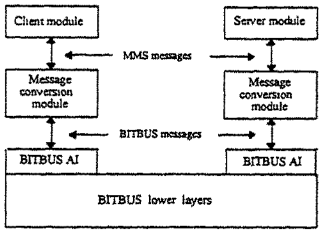


Fig.3: Module structure of MMS system in BITBUS

3.2 MMS Implementation in BITBUS Network

BITBUS is a high-speed serial bus providing transmission of short messages at low cost, with a master-slave network topology. BITBUS uses the RS-485 transmission link and synchronous data link control (SDLC) protocol which is used in data security and access rights of individual stations. BITBUS message protocol includes an order and replay mechanism, node and task address, as well as command and response status. It also provides for higher level commands with the remote access and control (RAC) capability.

Note that BITBUS lacks MMS messaging functionality in the application layer, we developed our own MMS system in BITBUS network, according to the ISO/IEC 9506 specification. Because the MMS implementation depends on the support of BITBUS lower layers, a BITBUS application interface (AI) which provides the MMS system with an external features of BITBUS must be designed. The functions of the application interface developed are : ① establish data buffers (including a sending buffer and a receiving buffer) ; ② set a hardware interrupt and receive information from the iPCX 344 board via FIFO ; ③ send information to the iPCX 344 board by means of polling ; ④ make use of the units reserved for users in the interrupt vector table, and set INT 64H software interrupt and communicate with MMS application program for information sending and receiving.

BITBUS network architecture may be divided into three parts, which correspond to the physical layer, the data link layer and the application layer of the ISO/OSI reference model. MMS system in BITBUS network consists of three modules (see Fig.3). The functions of each module are described as following : ① Client module : According to the types and parameters of the service request provided by the user, it fills in related operation-specific data structures to form corresponding MMS service request messages, and delivers them to the message conversion module. When receiving MMS confirmation messages from the message conversion module, the Client module is responsible for reporting the results of the service request. ② Server module : When receiving MMS service request messages, it analyzes the type of the service request and takes corresponding actions. Then, on the basis of service results, the server module fills in related operation-specific data structures to form MMS service response messages and delivers them to the message conversion module. ③ Message conversion module : It performs the conversion between MMS messages and BITBUS ones, and calls the interrupt service program set in the BITBUS application interface to implement the sending and receiving of BITBUS messages.

3.3 Design of the Gateway

There are mainly three kinds of devices used to interconnect local area networks (LANs) : bridges, routers and gateways. Bridges, which operate at the data link layer of the ISO/OSI reference model, do not have to perform protocol conversion. They simply look at the packet address to see where the packet is going. A router has more intelligent capabilities than a bridge because it can handle several levels of addressing. However, routers are protocol-dependent and can be only used to link LANs that have the identical protocol. A gateway operates at the highest level of the ISO/OSI reference model. It interconnects networks with different architectures by processing protocols to allow a node on one type of LANs to communicate with a node on another type of LANs.

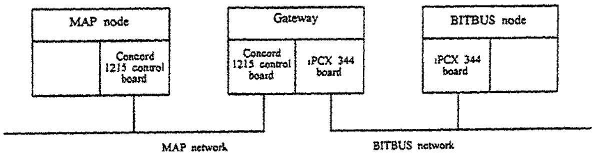


Fig.4: Hardware configuration of network interconnection

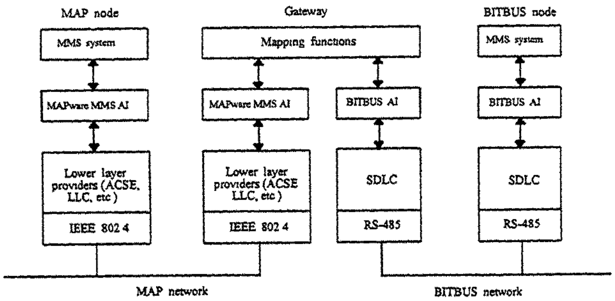


Fig.5: Information model of network interconnection

The implementation of MMS interconnection of MAP and BITBUS is characteristic of interoperability between two MMS users in networks. In other words, no matter where the requester and the responder are, they can communicate with each other using MMS services. Realizing the differences in network architectures and protocols of MAP and BITBUS, we constructed a gateway interconnecting the two kinds of networks in the application layer level. A minimal hardware configuration of network interconnection is shown in Fig.4. The Gateway, MAP node and BITBUS node are all based on personal computers running MS-DOS. The information model of network interconnection is shown in Fig.5. The mapping functions of the Gateway include : ① implement address mapping and identification between networks by establishing a link table of MAP and BITBUS ; ② implement message storage, forwarding and flow control ; ③ perform format conversion between MMS messages and BITBUS ones.

4. CONNECTION OF THE ROBOT TO THE NETWORK SYSTEM

In this project, we choose the JRB-1 robot for prototype implementation. The JRB-1 robot is a five-degree-of-freedom manipulator developed at the CIMS Center of Southeast University. The five joints of the robot are the base (waist), the shoulder, the elbow, the roll and the pitch. They are all driven by AC servomotors. The JRB-1 robot has a 10 kg lifting capacity and the repeatability error of its end-effector is less than one millimeter. As an universal one, the JRB-1 robot can find a place in material handling, welding, assembling and so forth. The JRB-1 robot system is made up of four parts : the robot controller, the servo computer, the servo unit and the main body. The robot controller is responsible for job management, coordinate conversion, and the produce of locus. The servo computer is responsible for closed-loop control of all joints of the manipulator. A parallel communication link is adapted in implementing data exchange between the robot controller and the servo computer .

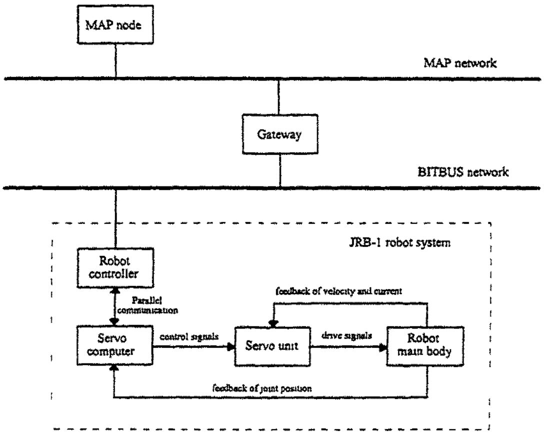


Fig.6: Connection of the JRB-1 robot system to the network system

In order to adapt the JRB-1 robot to be MMS compatible one and to control it via the network system , at least four conditions must be satisfied : ① the robot controller must be physically connected to the network system ; ②the VMD model of the robot must be established and a set of MMS services aiming at the robot must be implemented in the robot controller ; ③ the robot control program must be developed as VMD program invocations , which can be loaded , executed and monitored using MMS services ; ④ a MMS mapping interface must be designed to support remote communication and control of the robot .

The connection of the JRB-1 robot system to the network system is shown in Fig.6. By installing a iPCX 344 board, we can turn the robot controller into a BITBUS node and connect it to BITBUS network via transmission media (twisted pair wires). According to the ISO/IEC 9506 specification[1,2,8], we have developed a set of MMS services in the robot controller. In addition, to support local or remote operations on the robot, the means of mapping the VMD model of the robot to the real robot system is necessary. Hence, we designed a MMS mapping interface which can map the MMS variables to the commanded joint values of the robot, and reflect the data and status of the robot.

Fig.7 shows the flowchart of the JRB-1 robot control program. The program may be downloaded in the robot controller using MMS domain management services. In the MMS VMD model of the robot, the robot control program and other application programs represent separate VMD domains. Each domain may either be kept resident in the robot controller memory or be downloaded from disk files as needed via MMS services. Program invocations consisting of one or more domains can be created, started, stopped and monitored either locally from the robot controller, or remotely from another application process using MMS services over the network system.

5. CONCLUSIONS

1) PPI and VMI are different application interfaces provided by MAPware MMS. The Gateway performing protocol conversion needs a complete control of MMS messages. Therefore, PPI must be used in the process of MMS software programming in MAP network.

2) To conform to the MMS standard and make it easy for the Gateway to perform protocol conversion, data structures defined by MAPware MMS should be applied in the process of MMS software programming in BITBUS network. In addition, when a BITBUS node provides only a subset of MMS services, the service negotiation between two MMS users is required.

3) The Gateway, which plays the role of the network interface between MAP and BITBUS, is the basis of network interconnection. The construction of the Gateway in the application layer level and the implementation of mapping MMS services in MAP to those in BITBUS may simplify network interconnection.

4) Because MMS standard describes a general manufacturing device in software terms, the MMS services required by other machines are very similar to the services used by the robot. With the method we have described, it is not difficult to connect other computer-based devices to the network system.

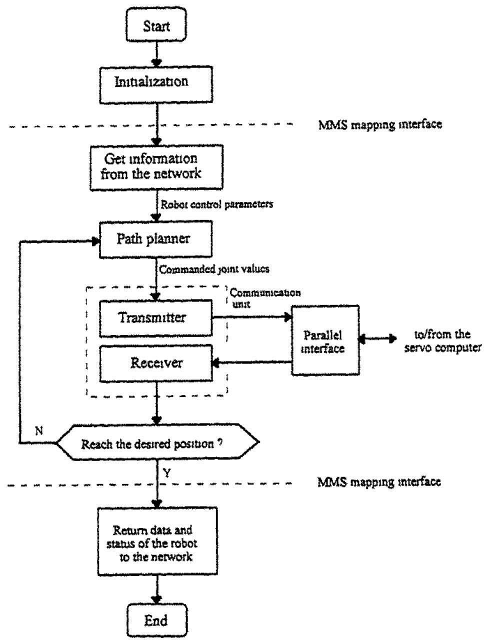


Fig.7: Flowchart of the JRB-1 robot control program

6. REFERENCES

1. ISO/IEC 9506, Industrial Automation Systems — Manufacturing Message Specification , *Part 1 : Service Definition* , 1990
2. ISO/IEC 9506, Industrial Automation Systems — Manufacturing Message Specification , *Part 2 : Protocol Specification* , 1990
3. N.Laurance , The Use of MMS for Remote CNC Control . *IEEE Trans. Ind. Electron.*, vol.34, no.4, pp.457-462 , 1987
4. X.Ding and J.Wu , Interconnection of MAP and BITBUS Based on MMS. *Journal of Southeast University*, vol.12, no.2, pp.18-24, 1996
5. M.Brill and U.Gramm, MMS : MAP Application Services for the Manufacturing Industry . *Computer Networks and ISDN Systems* , vol.21 , no.5 , pp.357-380 , 1991
6. S.G. Shanmugham *et al.*, Manufacturing Communication : A Review of the MMS Approach . *Computers Ind. Engng.*, vol.28 , no.1 , pp.1-21 , 1995
7. MMS Software Programmer's Reference Manual , pp. 40-42 , Concord Communication Inc. , 1990
8. ISO/IEC 9506 , Industrial Automation Systems — Manufacturing Message Specification , *Part 3 : Companion Standard for Robotics* , 1991

CAPP SYSTEMS IN CIM AND TRADITIONAL PRODUCTION ENVIRONMENT

N.K.Huang,
Dept. of Manufacturing Engineering, Northwestern Polytechnical University, Xi'an, P.R.China
N.Cui,
Machine Shop 16, Xi'an High Voltage Apparatus Works, Xi'an, P.R.China
Z.M.Zhang, J.X.Xu and Q.J.Fan
Dept. of Manufacturing Engineering, Northwestern Polytechnical University, Xi'an, P.R.China

ABSTRACT

As new advanced manufacturing technologies emerge and develop rapidly in industry, some newer and higher requirements on CAPP are prompted. In this paper, several key issues in the trends of integrated and intelligence system are discussed, and the scheme and techniques proposed or adopted by the authors in developing the NPURCAP and FA-CAPP system are introduced. In addition, some ideas on the building tools for CAPP systems are envisaged as well.

KEYWORDS

CAPP, Form Feature, Rule Element, Machining Element

1. INTRODUCTION

There has been significant progress in computer aided process planning (CAPP) since its origin more than 25 years ago. The principle of CAPP systems has evolved from its original retrieval approach, to the variant and then generative approaches, and to the current approach which incorporates the application of artificial intelligence techniques. But in spite of tremendous efforts have been made in developing CAPP systems, several questions remain unanswered, many issues must still be resolved, and the implementation of CAPP systems in industry lags behind the rate of development of new systems and the introduction of new ideas in the field[1,2]. These facts indicate that manufacturing industries impose higher and higher requirements on CAPP, motivating further applied research and development towards the directions of integration, intelligence and being building tools. This paper discusses these issues based on the authors research experience.

2. INTEGRATION OF CAPP WITH CAD AND CAM

The integration of CAPP with CAD and CAM is one of the key technologies to achieve CIMS. In this paper, emphasis will be placed on the interface between them. A brief discuss on both aspects of integration are as follows.

2.1 Interface using Features between CAD and CAPP

Naturally, CAPP has a close interface with CAD, since its input is the design model or the drawing of a component. Conventionally, transforming a drawing of a component by hand into the component data model is time consuming and obstructs the automation of process planning function. In the other hand, a product design model generated by a CAD system usually has not got sufficient information and convenient data structure for CAPP. The concept of features and their relevant techniques developed in recent years provide a better basis.

Although there are several kinds of features to form the information model for a machine component, form features are of most importance. One definition of form feature is " a geometric entity that represents a shape pattern that has some significance"[3]. From different viewpoints, there are different form features: such as functional, geometric and manufacturing etc. For form features of a machined components sometimes functional and geometrical meaning can be joined together with manufacturing features, and thus greatly facilitate the goal for features to succeed in being a true communication medium between CAD, CAPP and CAM. So far, the approaches to achieve this goal can be classified into three categories.

(1) Feature recognition. Feature recognition is an approach which provides a design interface for extracting (obtaining technological information) and recognizing (obtaining the form information) feature information (high level information) from a geometric representation of a solid model (low level information). The development of feature recognition techniques has emerged from the simple visual comparison between the raw material and the final product (cavity recognition to volume decomposition) to symbolic representation of the designed product (syntactic pattern recognition to feature grammar) and graph-based representation schemes (attributed adjacency graph)[4]. Despite of that extensive efforts has been made, some main limitations still remain to be overcome. For example, highly complex features may not be recognized[1], and the algorithm are too complex to be useful in practice[5]. In addition, technological information regarding the features has to be added interactively after the recognition process has been completed.

(2) Feature based modeling and design. Feature based modeling and design are the most recent and promising techniques regarding the interface between CAD and CAPP. They are of high level ones than the modeling techniques used in traditional CAD systems. Most feature based systems are built upon a geometric modeller, such as Boundary Representation (B-Rep) solid modeller, Constructive Solid Geometry (CSG) modeller, or a 2D representation modeller[6]. Recently, feature based parametric CAD systems have become commercially available, among which the most popular ones are Pro-Engineer, CADDS 5 and Unigraphics. However, none of them has actually implemented a sufficient and generally accepted feature based product representation for process planning application[7].

(3) Interactive feature definition. As neither of the approaches mentioned above are matured enough to be used in industrial applications, it would be preferable to use an interactive feature definition method and/or a predefined feature library for generating a suitable product model. The method might include interactively recognition of form feature geometrical attributes from a CAD model of a component, and adding technological information such as tolerance and surface finish etc. This approach is not an fully automatic one, but at present, is a practical method in many industrial environment.

Employing whatever of above mentioned approaches, an interface data file for a component model is of necessary to transfer it to CAPP system. Some systems have adopted component feature based model data in IGES or STEP neutral standard formatted files and have achieved some extent of success. As an earlier standard for product data conversion, IGES mainly describes the information about the product low level geometry which can not meet the needs for CAD/CAPP interface requirement. Based on the product information which cover the entire product life cycles, STEP international standard (ISO 10303) for product definition data exchange provides the fundamental and promising way of solving the integration problems. But up to now, some important parts of STEP standard, such as the application protocols, the general integration resources etc. still take time to complete and improve. Consequently, using some kind of specific interface files adapted to actual CAD system available and the hardware platform for CAD and CAPP implementation in many cases is still an effective way.

In recent years, we have developed several practical CAPP systems employed both in CIM and traditional environment. The feature-based component model[8] used by us in these systems consists of management feature, machining (form) feature and geometry features as in Fig.1.

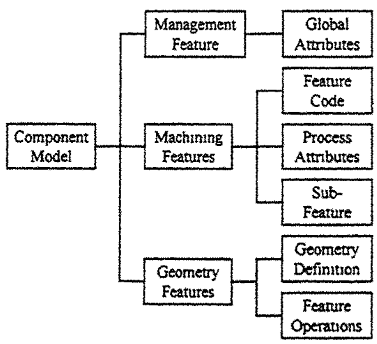


Fig.1 Featured-Based Component Model

The machining features which provide the form type code and technological attributes for decision making in CAPP are interactive defined. The geometry features are based on the available

CAD system in actual situation. It consists of low level geometry (i.e. edges, vertices and faces) to meet the need of CAPP to generate operation sketches and the need of CAM to do NC programming. Both machining and geometry features have the same identification code.

FA-CAPP, an integrated system for aircraft machined structural components, is a sub-project of FA-CAD/CAPP/CAM system in an industrial CIM environment. It achieves the information sharing, including all the management, machining and geometry features, with CAD through a specific interface data file and a geometrical data base called RMI (Root Modeling Interface) which is developed by Beijing Aeronautical Manufacturing Technology Research Institute.

2.2 Interfaces between CAPP and other CIMS Sub-systems

In most CIM application engineering enterprises, CAD system provide CAM systems with geometrical and dimensional information, while CAPP systems pass machining information to CAM. The FA-CAD/CAPP/CAM system uses this reasonable approach under the support of the RMI geometrical database. But for the NPURCAP, another similar system for rotational component, due to the significant shape and dimension changes of machining features on rotational components from original blanks to intermediate and final products, the geometrical information is also transferred from CAPP system through a specific CAPP/CAM interface data file.

In CIMS engineering project, the integration problems also include the interfaces between CAPP, MAS (Manufacturing Automation System), PDMS (Process Document Management System), MIS (Management Information System) and CAQ (Computer Aided Quality assurance system). In implementing NPURCAPP and FA-CAPP, we use a distributed supporting engineering data base systems built upon commercially available ORACLE system to share all the required information with other functional systems in CIMS environment. For example, the FA-CAPP system architecture is shown in Fig.2, and the data flow of aircraft machined structural components is shown in Fig.3. The decision making module of FA-CAPP queries remotely cutting tool base and machine tool base located in MAS machine shop, and also remotely commits the process plan into distributed ORACLE data base. The PDMS system manages the process documents and sends them through the same distributed data base to MIS, CAQ sub-systems as well.

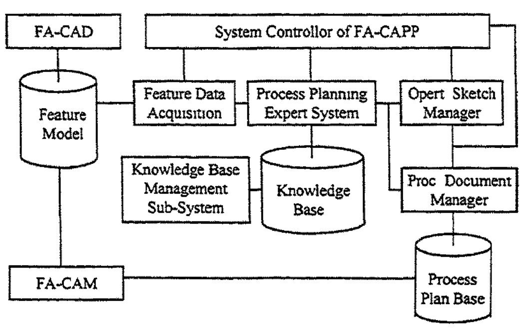


Fig.2 Architecture of FA-CAPP

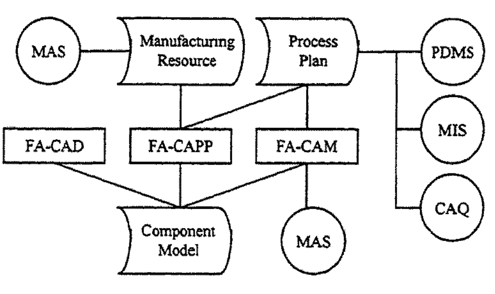


Fig.3 Data Flow in CIMS

3. INTELLIGENCE IN CAPP

Process planning is a mixture of complex and interrelated tasks, and hence involved in tremendous amount of heterogeneous knowledge. In CAPP research and development community, the intelligence of CAPP is considered to include the knowledge processing and decision making methodologies and is supposed to make use of expert system or other knowledge-based techniques in artificial intelligence field.

3.1 Knowledge Processing

The knowledge related to process planning activity can be divided into process engineering domain knowledge and meta knowledge, and the former can be further divided into factual and decision making knowledge. Therefore, there exist the following three levels of knowledge.

(1) Zeroth-level knowledge, i.e., factual knowledge, including components information, manufacturing resources information, all kinds of engineering standard data, etc.

(2) Primary knowledge, i.e., decision making knowledge or procedures. This includes the heuristic knowledge of part blank design, processing methods selection, process sequencing, operation planning, tool selection, etc. It also includes the knowledge of operation dimensioning, cutting path planning, operation sketch generation and drawing, machining time calculation and information processing.

(3) Meta knowledge also called secondary knowledge, including the knowledge of decomposition of process planning tasks and sub-tasks scheduling, process planning procedure control, knowledge sources management, administration, etc.

Obviously, it is very difficult to process the above large quantity of processing knowledge using a single knowledge expression method.

For the zeroth-level knowledge, generative systems need to create information models of components and process planning, along with models of the relations between them. The application of frame systems or object-oriented methods based on these models has a very promising future. For the meta knowledge, the production rules are the predominant method.

For the process decision making knowledge of the primary level, which consists of mostly heuristic knowledge, the production rule expressions are often used. Although production rules have many obvious advantages, they have bad structures which are hard to manage and maintain. Therefore there exist layering-structured production rule set methods and can be combined with the frame system approach. For the engineering calculations or operations of the primary knowledge, functions and procedures (subprograms) are often used.

In FA-CAPP system, the combined expression method of frames and production rules is used. The knowledge schemata are divided into internal one and external one. The internal schema is further divided into three levels: rule element, rule, and rule set frames (in Fig.4). This dramatically increases the efficiency of the storage and usage of the process decision making knowledge. The external knowledge schema is expressed using production rules and frames. This makes it easier to acquire knowledge from user experts. Currently, this approach is in the process of improvement and completion to become a better generic expression method.

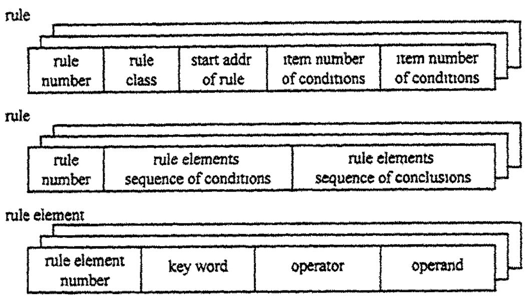


Fig.4 Internal Schema of Knowledge

3.2 The Decision Making for Process Planning

A number of process planning problems can be defined and solved analytically, and in those cases a solution is obtained by applying algorithms and technological constraints. In other cases, problems depend mainly upon the knowledge or expertise within a particular domain. There exist abroad some models and algorithms for local decision makings, such as the recursive approach and table of anteriorities[9]. Domestically, relying on the practical experiences of the process planning experts, people mostly use the method which is based on the selections of the processing series by experience for processing machining form features.

In FA-CAPP, an important concept of fundamental machining element is introduced and a procedural model of process planning is created. The machining element is an information entity about the process of a kind of machining form feature. It consists of several data items at least, e.g. form feature, process series, machine tool, cutting tool, allowance, etc.

On the basis of the machining form feature information of components, process planning starts with creating machining element automatically, and then combines machining elements into steps and steps into operations, queries machine tool and cutting tool data base, designs each operation in detail, arranges the process route of this component, and eventually forms practical process documents.

4. THE REQUIREMENTS FOR BUILDING TOOLS OF CAPP

Due to the vast diversity and multiformity of industrial products, manufacturing resources, planning expertise and conventional techniques in every particular case, it has been very difficult to develop applicable generic software for CAPP system, even though there has been a large number of prototype systems developed at home and abroad. From the industrial stand point both in CIM and traditional production environment, the need for building tools of CAPP system is becoming more and more pressing for all researchers in this field over the world.

There were already some similar ideas and attempts in the early stages of CAPP system development[10]. In CIMS engineering production environment, the requirements faced by CAPP are not only shared with the process planners and the planning procedure itself, but also imposed by CAD, CAM and production schedulers, including machinability verification and improvement in product design, process plan adjustment triggered by the feedbacks resulting from exceptional events in product manufacturing and scheduling, etc. These requirements cause much more variety in functions and activities of CAPP system. In traditional production environment, the process documents management functions, the various planning approaches and human being interacts are more prevailing requirements. Consequently, it may be seen that the future CAPP system should consist of modules with generic basic functions, modules with different kinds of enabling or auxiliary tools, modules of corresponding data and knowledge management functions. This is the basis for making CAPP systems as building tools in production environment, which should impose much more requirements: The software structure of expert system, i.e. with three layers of data, knowledge and control programs should be used. All the modules mentioned above should share common knowledge bases and have a common interface to and working collaboratively with each other. The knowledge base should contain multiple knowledge sources, multiple reasoning mechanisms and combined procedural algorithms for analytical engineering tasks. Based on feature-based component model, data models with certain generic usage should be created for products, manufacturing environments, process plan and planning procedures. Multiple approaches of planning, such as the retrieval, variant and generative principles should be incorporated in single system of hybrid architecture. And finally the system should also have open structures to allow the utilization of the work of different groups.

Once these requirements are met, different application software packages can be easily assembled particularly for CAD, CAM, production scheduling and control. In summary, the future architecture of CAPP systems may consist of three levels as illustrated in Fig. 5.

5. CONCLUDING REMARKS

The authors have long been dedicated in the research and development work for CAPP applications, and several systems built by us have been used for new products development in domestic industry. With the ever changing requirements from quite different production environments and the facts of that the industrial exploitation and practical engineering application lags behind the rate of development of new prototype systems, to be building tools of different CAPP functions becomes the most urgent need. Based on some key issues discussed and introduced in section 2 and 3, we have done some primary development in this direction, and it is believed that these ideas will soon be appreciated by other researchers, and become reality in the near future.

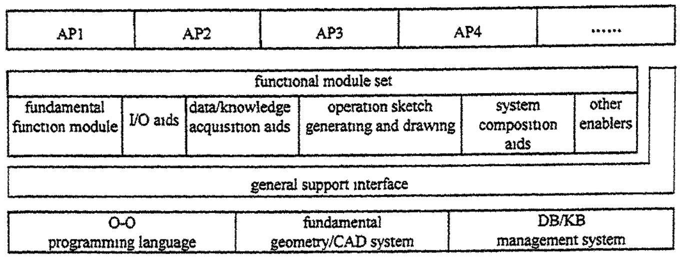


Fig.5 Architecture sketch of CAPP development tools

6. ACKNOWLEDGMENTS

We would like to thank the National Natural Science Foundation and the Aeronautics Science Foundation of China for financial supporting this work.

7. REFERENCES

1. Maropoulos P G, " Review of research in tooling technology, process modeling and process planning/Part II: Process planning " , Comtr. Integr. Syst. Vol.8, No.1: 13-20, 1995.
2. ElMaraphy H, " Evolution and future perspectives of CAPP " , Keynote papers, annals of the CIRP, Vol 42/2: 739-751, 1993
3. Gindy N N Z, " A hierarchical structure for form features " , Int.J.Prod. Res., Vol.27, No 12, 2089-2103, 1989
4. Senthil kumar A, et al., " Automatic recognition of design and Machining Features from prismatic parts " ,Int. J. Adv. Manuf. Technol., No.11; 136-145, 1996.
5. Chang T C, Expert process planning for manufacturing. California, Addison Wesley publ. Co.,: 67-69, 1990
6. Case K, "Using a design by features CAD system for process capability modeling", Comptr Integr Manuf. Sys., Vol 7, No 1: 39-49, 1994
7. Gao J X, Huang X X, " Product and manufacturing capability Modelling in an integrated CAD/process planning Environment " , Int J Adv Manuf Technol., No 11: 43-51, 1996
8. Xu J X, Zhang Z M, Huang N K, " FA-CAPP: A CIM Application Oriented Feature-Based CAPP System " , Proc. of INCOM'95, IFAC, October 11-13: 250-254, 1995.
9. Weill, R et al, " Survey of Computer-aided Proces Planning Systems " , 38th CIRP General Assembly, Tokyo, Japan, August 22, 1988
10. Link C H, " Computer-aided Process Planning (CAPP)" , SME Technical Report, MS 78-216, 1978

A PRELIMINARY STUDY OF A FUZZY CLUSTERING AND ASSIGNMENT PROBLEM-BASED CELL FORMATION ALGORITHM

Susanto, S., Kennedy, R.D., and Price, J.W.H.
Department of Mechanical Engineering
Monash University - Australia
(Email: sani.susanto@eng.monash.edu.au)

ABSTRACT

The ideal situation of CM is achieved when each part-cluster together with the machine-cluster for the manufacturing cell is independent from the rest of the manufacturing cell; that is exceptional elements are non existent. However, such a situation is rarely achieved. In this paper, a cell formation algorithm that identifies part-clusters and machine-clusters separately is introduced. The algorithm is based on fuzzy clustering and the assignment technique. Computational results show this new algorithm can give superior performance, especially for higher density machine-part incidence matrices.

KEYWORDS:

Cellular Manufacturing, Manufacturing Cell, Cell Formation, Part-cluster, Machine-cluster.

1. INTRODUCTION

Group Technology (GT) is a manufacturing-related philosophy that is gaining popularity[1]. This philosophy exploits the similarity among the attributes of given objects [2]. One of the popular applications of GT is called *cellular manufacturing* (CM). In CM, parts that require similar machines are grouped into the same *part-cluster*, while machines needed by a part-cluster are grouped to form a *machine-cluster*. A *manufacturing cell* is part of a manufacturing system where a machine-cluster is placed to process a part-cluster. The most distinctive feature of CM is that it contains dissimilar machines that are located in one area

The first problem which must be addressed when considering a cellular manufacturing system is that of *cell formation*. Data for cell formation is organized into a *machine-part incidence matrix*. This is a binary matrix with 0 or 1 entries. A '1'(one) entry in row i and column j of the matrix indicates that machine-type i needs to operate on part-type j, while a '0'(zero) indicates it does not. Mathematically, the cell formation problem involves converting the machine-part incidence matrix into a block diagonal form in which '1' entries are concentrated in blocks along the diagonal of the matrix. Each block represents a manufacturing cell [3].

Extensive work has been done by many researchers to provide new techniques for solving this problem. The cell formation problem is complicated by the existence of exceptional parts or exceptional machines[4]. An *exceptional part* is a part that requires processing in another machine-cluster. An *exceptional machine* is a machine that processes parts from a different part-cluster. Both exceptional parts and exceptional machines cause *intercellular movement* of parts. Ideally a part-cluster is processed in a single machine cell for its entire set of operations, however, it is a very rare[5]. Thus, the more '1' entries are concentrated in the block diagonal the more effective a cell formation technique is. Numerous research papers have appeared in the literature for cell formation. These methods are based on the following approaches: coding and classifications, machine-component analysis, similarity coefficients, mathematical programming and heuristic methods, knowledge-based and pattern recognition methods, neural networks and fuzzy clustering[2].

It is the last approach (ie. fuzzy clustering) that we focus on in this article. Very few papers have appeared in the area of fuzzy clustering techniques applied to cell formation. The application of this technique in cellular manufacturing was originally proposed by Chu and Hayya [6]. and was followed by Ponnambalam and Aravindan[7] with a minor modification in the initialization. Despite their contributions, there is still a room for improvement, especially in the effort of increasing the '1' entries to be concentrated in the block diagonal of the initial given machine-part incidence matrix. The main feature of this paper is to demonstrate a new algorithm to effect this improvement as outlined in Section 3. In Section 2, the basic techniques of the algorithm are discussed. Those techniques are *the fuzzy c-means* (fuzzy clustering) and *the assignment problem*. In Section 4 some computational results are presented. A summary and suggestions for further research are presented in Section 5.

2. A BRIEF REVIEW

2.1 Fuzzy Clustering Problem

Clustering is a process which is common and basic to human understanding. This process, which plays a vital role in grouping of related objects, can be found in such diverse fields as statistics, economics, physics, psychology, biology, pattern recognition, engineering, marketing [8,9]. The primary objective of *clustering* is to partition a given set of objects into so-called homogeneous clusters (groups) [10-12]. Globally, the clustering problem can be divided into two main categories ie. *hard clustering*, in which an object belongs only to one cluster, and *fuzzy clustering*, in which every object belongs to all clusters with different degrees of membership. Definitions, theorems and algorithms involved in the discussion of the fuzzy clustering topic are as follows [13].

2.1.1 Definition and Theorem

Definition 1. Fuzzy c-Partitions

Let $O=\{o_1,o_2,\dots,o_p\}$ be a set of p objects; $X=\{x_1,x_2,\dots,x_p\}$ where $x_r=(x_{1r},x_{2r},\dots,x_{mr})^T$ is the *attribute vector* of object $o_r \in O$ (r = 1, 2, ..., p); and V_{pc} is the set of real p x c matrices.

Fuzzy c-Partitions space for X is the set

$$M_{fc}=\left\{U \in V_{pc} \,\middle|\, u_{ri} \in [0,1]\,\forall r,i;\ \sum_{i=1}^{c} u_{ri}=1\ \forall r;\ 0<\sum_{r=1}^{p} u_{ri}<p\ \forall i\right\} \quad (2.1)$$

Definition 2. Fuzzy c-Means Functional

Fuzzy c-Means Functional is the functional $J_f : M_{fc} x R^{pc} \rightarrow R^+$ defined by the relationship

$$J_f(U,v)=\sum_{r=1}^{p}\sum_{i=1}^{c}(u_{ri})^f (d_{ri})^2 \quad (2.2)$$

where 1) $U \in M_{fc}$ is a fuzzy c-partitions of X; 2) $v=(v_1,v_2,\dots,v_c) \in R^{mc}$ with $v_i=(v_{1i},v_{2i},\dots,v_{mi})^T \in R^m$ is the cluster centre of $u_i=(u_{1i},u_{2i},\dots,u_{pi})^T$, $1 \le i \le c$; 3) $d_{ri}=\|x_r-v_i\|$ and $\|\ \|$ is any inner product on R^m and 4) f is the weighting exponent (also called *the degree of fuzziness*), $f \in (1,\infty)$.

Theorem 1. (Bezdek's Necessary Conditions for Global Minimum)

Let $\|\ \|$ be any inner product norm on R^h; $f \in (1,\infty)$; X has at least c (c < p) distinct points, and for each $r \in \{1,2,\dots,p\}$, define the sets

$$I_r=\left\{i \mid 1 \le i \le c;\ d_{ri}=\|x_r-v_i\|=0\right\} \quad (2.3)$$

$$I_r^{c}=\{1,2,\dots,c\}-I_r \quad (2.4)$$

then if $(U,v) \in M_{fc} x R^{pc}$ is the global minimum point for J_f then

$$1)\ I_r = \varnothing \Rightarrow u_{ri} = \frac{1}{\left[\sum_{k=1}^{c}\left(\frac{d_{ri}}{d_{rk}}\right)^{\frac{2}{f-1}}\right]}, \tag{2.5}$$

$$\text{or } I_r = \varnothing \Rightarrow u_{ri} = 0\ \forall i \in I_r^{\,c} \text{ and } \sum_{i \in I_r} u_{ri} = 1 \tag{2.6}$$

$$2)\ v_i = \frac{\sum_{r=1}^{p}(u_{ri})^f x_r}{\sum_{r=1}^{p}(u_{ri})^f} \quad \forall i \in \{1,2,\ldots,c\} \tag{2.7}$$

2.1.2 Chu and Hayya's Algorithm

Chu and Hayya[6] were the first authors to apply the fuzzy clustering technique to cell formation. They have treated the attribute vector of all part type as an input to their algorithm. Their algorithm will result in two different matrices The first matrix, notated as U, *the degree of membership matrix*, is used for grouping parts into part-clusters. The second matrix, notated as v, *the cluster centre matrix*, is used for grouping machines into machine-clusters. The detail of their algorithm is as follows.

Let $X = (x_{qr})_{mxp}$ be the given machine-part incidence matrix of *m* types of machine and *p* types of parts; $x_r = (x_{1r}, x_{2r}, \ldots, x_{mr})^T$, namely the *r*-th column of X (*r* = 1,2, ..., p) is the *attribute vector of* the part type *r* . The *p* part types and the *m* machine types are to be grouped into *c* clusters.

Algorithm 1. Chu and Hayya's Algorithm for part-cluster and machine-cluster formation

__Step 1__ : Fix c, $2 \le c < \min\{m,p\}$; choose any positive value ξ for stopping criterion; choose any inner product norm $\| \ \|$ on R^m and fix $f \in (1,\infty)$; initialize $U^{(0)} \in M_{fc}$ and set $l = 0$.

__Step 2__ : Calculate the c fuzzy cluster centres $\{v_{(i)}^{(l)}\}$, where

$$v_{qi}^{(l)} = \frac{\sum_{r=1}^{p}(u_{ri}^{(l)})^f x_{qr}}{\sum_{r=1}^{p}(u_{ri}^{(l)})^f}\ ; \ i = 1,2,\ldots,c;\ q = 1,2,\ldots,m \tag{2.8}$$

__Step 3__ : $l \leftarrow l + 1$; calculate

$$u_{ri}^{(l)} = \begin{cases} \dfrac{1}{\left[\sum_{k=1}^{c}\left(\frac{d_{ri}}{d_{rk}}\right)^{\frac{2}{f-1}}\right]}, & \text{if } I_r = \varnothing \\ 0, & \forall k \in I_r^c \text{ if } I_r \neq \varnothing \\ \dfrac{1}{|I_r|}, & \forall k \in I_r \text{ if } I_r \neq \varnothing \end{cases} \tag{2.9}$$

where I_r and $I_r^{\,c}$ are as defined in (2.3) and (2.4) respectively, $|I_r|$ is the number of element(s) in I_r

__Step 4__ : If $\left\|U^{(l+1)} - U^{(l)}\right\| < \xi$ then stop else goto Step 2.

The final matrices U and v are used to determine the part-clusters and machine-clusters respectively, under the following rule:

Rule 1. (Part-clusters formation Procedure)
If $u_{ri} = \max_k \{u_{rk}\}$; $r = 1, 2, ..., p$ (p is the number of part-types) and $k = 1,2, ..., c$ (c is the number of part clusters), then part r is assigned to part-cluster i.
Rule 2. (Machine-clusters formation Procedure)
If $v_{qi} = \max_k \{v_{qk}\}$; $q = 1, 2, ..., m$ (m is the number of machine-types) and $k = 1,2, ...,c$ (c is the number of machine-clusters) then machine q is assigned to machine-cluster i.

As a result of Chu and Hayya's stopping criterion, two problems may arise. The *first problem*, there is a distinct possibility that their algorithm results in some empty part-clusters and/or empty machine-clusters (ex. see result displayed in Table 3. in [7]). Furthermore, their algorithm may end up with different numbers of part-clusters compared to the resulting number of machine-clusters, which in turn, does not have any practical meaning. The *second problem*, In case of a tie in Rule-1, that is $\max_k \{u_{rk}\}$ is achieved by more than one part-cluster, then in Chu and Hayya's algorithm [6] (or its minor modification by Ponnambalam and Aravindan [7]) the part-type r is directly assigned to the first part-cluster achieving the maximum value of u_{rk}. This is too restrictive in the authors' opinion, since it eliminates the opportunity for the other part-cluster that achieved the same maximum value of u_{rk} to contain that particular part type. Improvements to overcome these problems is discussed in Section 3.

2.2 The Assignment Problem

Consider the situation of assigning c part-clusters to c machine-clusters and define $c_{ik} = \sum_{r \in PC-k} \sum_{q \in MC-i} x_{qr}$ be *the degree of conformance* of part-cluster(PC) k and machine-cluster (MC) i. That is, the total number of '1'(one) entries in the block diagonal formed by those part and machine-clusters. The objective is to assign the c part-clusters to the c machine-clusters (one part-cluster per machine-cluster) to maximize the total degree of conformance. This situation is known as *the assignment problem*. (see also [14,15]).

The assignment problem, if applied in cell formation situations, can be expressed in mathematical model as follows. Let

$$\Omega_{ij} = \begin{cases} 0, & \text{if the part-cluster(PC) } i \text{ is not assigned to the machine-cluster(MC) } j \\ 1, & \text{if the part-cluster(PC) } i \text{ is assigned to the machine-cluster(MC) } j \end{cases} \qquad (2.10)$$

The model is thus given by

$$\text{maximize } z = \sum_{i=1}^{c} \sum_{j=1}^{c} c_{ij} \Omega_{ij} \qquad (2.11)$$

subject to

$$\sum_{j=1}^{c} \Omega_{ij} = 1, \quad i = 1,2,...,c \qquad (2.12)$$

$$\sum_{i=1}^{c} \Omega_{ij} = 1, \quad j = 1,2,...,c \qquad (2.13)$$

$$\Omega_{ij} = 0 \text{ or } 1. \qquad (2.14)$$

Having obtained the part-clusters and the machine-clusters, the application of the assignment problem to the cell formation problem will maximize the total number of elements in the block diagonals of the final permuted machine-part incidence matrix. Thus, at the same time, this will minimize the incidence of exceptional parts and machines.

3. THE SKP-1 (Susanto-Kennedy-Price Version 1) ALGORITHM

The SKP-1 algorithm is aimed at tackling the two problems identified in Chu and Hayya's algorithm [6] and at maximizing the total number of elements in the block diagonal of the final permuted machine-part incidence matrix. As classified by Wemmerlöv [16] (in [17]), the cell formation literature can be divided into four categories, according to the formation logic used: 1) grouping part-clusters only, 2) forming part-clusters and then machine-clusters or vice versa; 3) forming part-clusters and machine-clusters simultaneously, and 4) grouping machine-clusters only.

The SKP-1 algorithm belongs to the third category in the forementioned classification, and can be outlined as follows. Let $X = (x_{qr})_{mxp}$ be the given machine-part incidence matrix of *m* types of machine and *p* types of parts; $x_r = (x_{1r}, x_{2r}, \ldots, x_{mr})^T$, namely the *r*-th column of X (r = 1,..., p) is the *attribute vector of* the part type *r*. Let $y_q = (x_{q1}, x_{q2}, \ldots, x_{qp})$, namely the *q*-th row of X (q = 1,..., m) is the *attribute vector of* the machine type *q*. The *p* part types and the *m* machine types are to be grouped into *c* clusters ($2 \le c < \min\{m,p\}$) separately using the fuzzy clustering approach.

3. 1 The First Step of SKP-1 (Machine-clusters Formation)

The *p* attribute vectors of each part type (ie. $x_1, x_2, \ldots, x_p$) serve as the only known variable in the functional to be minimized $J_f(U,v)$ (as defined in (2.2)). In this minimization problem, we define a *successful machine-cluster formation solution* as one that 1) converges to *c* non empty machine-clusters; 2) results in unimodal solution for $\max_k \{v_{qk}\}$ (q = 1, 2, ..., *m* and k = 1, 2, ..., *c*).

Instead of using the criterion in Step-4 of Algorithm-1, we will use a given number of maximum iterations. Thus, up to the maximum number of iterations, it is possible to get more than one *successful machine-cluster formation solution*. The details of the first step in SKP-1 are as follows.

<u>Step 1</u> : Fix c, $2 \le c < \min\{m,p\}$; choose any inner product norm $\| \ \|$ on R^m and fix $f \in (1,\infty)$; initialize $U^{(0)} \in M_{fc}$; let *success* = 0; l = 0 and the number of maximum iterations be *max_iterations*.

<u>Step 2</u> : Calculate the c fuzzy cluster centres $\{v_{(i)}^{(l)}\}$ using (2.8).

<u>Step 3</u> : If $\forall q \in \{1,2,\ldots,m\}$, the set $A_q = \left\{\alpha_l^{(q)} \middle| v_{q\alpha_l(k)} = \max_{i=1,2,..,c} \{v_{qi}^{(l)}\}\right\}$ is singleton

then if $\bigcup_{q=1}^{m} A_q = \{1,2,\ldots,c\}$

then

if $l <$ *max_iterations*

then 1) *success* ← *success* + *1* 2) at the *l*-th iteration and at the *success*-th successful machine assignment, assign machine *q* to machine-cluster $\alpha_l^{(q)}$

else 1) *success* ← *success* + *1* 2) at the *l*-th iteration and at the *success*-th successful machine assignment, assign machine *q* to machine-cluster $\alpha_l^{(q)}$ 3) goto Step 6.

else if l = *max_iterations* then goto Step 6

else if l = *max_iterations* then goto Step 6

<u>Step 4</u> : Determine I_r and I_r^c as defined in (2.3) and (2.4), where r = 1, 2, ..., p

<u>Step 5</u> : $l \leftarrow l + 1$; calculate $\{u_n^{(l)}\}$ as defined in (2.9); goto Step 2.

<u>Step 6</u> : If *success* ≥ 1 then $MC = \bigcup_{s=1}^{success} MC_s$,where $MC_s = \{MC_{is} | i = 1,2,\ldots,c;\ s = 1,2,\ldots,success\}$ and

MC_{is} is the machine-cluster i obtained from the s-th successful machine assignment. else "this algorithm fails to converge to a successful machine-cluster formation solution".

3.2 Second Step of SKP-1 (Part-clusters Formation)

Basically the second step of SKP-1 is congruent to that of the first step. While the first step of SKP-1 exploits the part-attribute vectors x_r (r = 1, 2, ..., p) (which are column-vectors) to form the machine-clusters, the second step exploits the machine-attribute vector y_q (q = 1,..., m) (which are row-vectors), to form the part-clusters. We define a *successful part-cluster formation solution* as one that **1)** converges to c non-empty part-clusters and **2)** results in a unimodal solution for $\max_k \{v_{qk}\}$ (k = 1,..., c).

Again, instead of using the stopping criterion in Step-4 of Algorithm-1 (Section 2.1.2) we will use a given number of maximum iterations. Thus, up to the maximum number of iterations, it is possible to get more than *one successful part-cluster formation solution*. If, after the maximum number of iterations is achieved, this second step converges and results in *SUCCESS successful part-cluster formation solutions* (*SUCCESS* $\geq$ 1), then let us define:

$PC = \bigcup_{S=1}^{SUCCESS} PC_S$,where $PC_S = \{PC_{kS} | k = 1,2,\ldots,c;\ S = 1,2,\ldots,SUCCESS\}$ and PC_{kS} is the part-cluster k obtained from the S-th successful part-cluster assignment.

3.3 The Third Step of SKP-1 (The Assignment Problem)

To apply this step, it is necessary that both the first and the second step of SKP-1 converge to *success* successful machine-cluster solutions and *SUCCESS* successful part-cluster solutions.

Let

$$c_{MC_{is}PC_{kS}} = \sum_{r \in PC_{kS}} \sum_{q \in MC_{is}} x_{qr} \tag{2.15}$$

be the *degree of conformance* of part cluster k obtained from the S-th successful part-cluster formation and machine cluster i obtained from the s-th successful machine-cluster formation. The final assignment problem is the solution of the following optimization problem:

$$\text{maximize } z_{sS} = \max \sum_{i=1}^{c} \sum_{k=1}^{c} c_{MC_{is}PC_{kS}} \Omega_{MC_{is}PC_{kS}} \tag{2.16}$$

subject to

$$\sum_{k=1}^{c} \Omega_{MC_{is}PC_{kS}} = 1, \quad i = 1,2,\ldots,c;\ s=1,2,\ldots,success;S=1,2,\ldots,SUCCESS \tag{2.17}$$

$$\sum_{i=1}^{c} \Omega_{MC_{is}PC_{kS}} = 1, \quad k = 1,2,\ldots,c;\ s=1,2,\ldots,success;S=1,2,\ldots,SUCCESS \tag{2.18}$$

$$\Omega_{MC_{is}PC_{kS}} = 0 \text{ or } 1 \tag{2.19}$$

Note: $\Omega_{MC_{is}PC_{kS}}$ = 1, if the k-th part-cluster (obtained from the S-th part-cluster formation solution) is assigned to the i-th machine-cluster (obtained from the s-th machine-cluster formation solution),
= 0, otherwise.

4. COMPUTATIONAL RESULTS

In this section the performances of SKP-1 and Chu and Hayya's algorithm are compared, based on the observations applied to 16 matrices with the dimension of 40x50, with density ranges from 20 to 80%. The results are presented in Table 1.

4.1 Example of Computational Result

The machine-part incidence matrix XVI is displayed in Appendix 1. The results of using this matrix to Chu and Hayya's [6] and the SKP-1 algorithm are outlined in the following Section.

4.1.1 Result from Chu and Hayya's algorithm [6]

This algorithm was applied to matrix XVI with $c = 6, f = 2$, $\xi = 0.001$ and Euclidean norm as the inner product. The machine-clusters and part-clusters are as follows.

The Machine-clusters :

MC-1 = {4,7,15,17,35,37}; MC-2 = {6,19,20,21,23,24,30,31,38,40}; MC-3 = {8,10,36}; MC-4 = {3,22}; MC-5 = {1,5,9,11,12,14,25,26,27,32,33,39} and MC - 6 = {2,13,16,18,28,29,34}.

The Part-clusters :

PC-1 = {1,7,18,19,20,25,33,37,43,49}; PC-2 ={14,32,38,50}; PC-3 = {5,9,21,31,39,40,45}; PC-4={4,10,22,28,44,46, 47}; PC - 5 = {2,11,15,17,23,24,26,27,29,34,35,41} and PC - 6 = {3,6,8,12,13,16,30,36,42,48}.

The resulted Manufacturing Cells are :

Manufacturing Cell-*i* will consist of MC-*i* that processes PC-*i* (i = 1,2, ...,6).

There will be 333 of '1'(one) entries on the block diagonal of the final rearrangement of the matrix XVI which contains 1609 (ie. $\approx$0.805x40x50) '1' entries. Thus, Chu and Hayya's algorithm [6] results in 20.7% (ie. 333/1609 x 100%) of non exceptional elements. The final rearrangement by this algorithm is displayed in the Appendix 2.

4.1.2 Result from SKP-1 algorithm

The SKP-1 algorithm was also applied to matrix XVI, $c = 6, f = 2$, *max_iterations* = 200, and Euclidean norm as the inner product. Four successful machine-cluster formation solutions and five successful part-cluster formation solutions resulted from the treatment are:

The Four Successful Machine-cluster Formation Solutions :

The first successful machine-cluster formation solution:

MC-1 = {1,2,3,4,7,15,17,35,37}; MC-2 = {6,19,23,25,30,31,38,40}; MC-3 = {8,10,18,21,24,36}; MC - 4 = {20,22,27}; MC - 5 = {5,9,11,14,26,32,33,39} and MC - 6 = {12,13,16,28,29,34}.

The second successful machine-cluster formation solution:

MC-1 = {1,3,4,7,15,16,17,35,36,37,40}; MC-2 = {6,19,23,31,38}; MC - 3 = {8,10,21,24,30}; MC - 4 = {20,22}; MC-5 = {5,9,11,14,25,26,27,32,33,39} and MC - 6 = {2,12,13,18,28,29,34}.

The third successful machine-cluster formation solution:

MC - 1 = {15,16,17,20,35,36,37,38,40}; MC - 2 = {19,31}; MC - 3 = {7,8,10,12,24,30}; MC - 4 = {3}; MC - 5 = {1,5,9,14,21,22,25,26,27,32,33,39} and MC - 6 = {2,4,6,11,13,18,23,28,29,34}.

The fourth successful machine-cluster formation solution:

MC - 1 = {15,16,18,20,31,35,36,37,38,40}; MC - 2 = {19}; MC - 3 = {7,8,10,12,17,24,30}; MC - 4 = {3}; MC - 5 = {1,5,9,14,21,22,25,26,27,32,33,39} and MC - 6 = {2,4,6,11,13,23,28,29,34}.

The Five Successful Part-cluster Formation Solutions :

The first successful part-cluster formation solution:

PC- 1 = {6,7,13,22,25,32}; PC- 2 = {34,35,46,47,49}; PC- 3 = {10,14,15,17,26,27,30,31,43,44,45}; PC- 4 = {1,8,12,16,18,19,20,21,23,28,33,36,37,38,48,50}; PC- 5 = {11} and PC- 6 = {2,3,4,5,9,24,29,39,40,41,42};

The second successful part-cluster formation solution:

PC- 1 = {6,13,18,25,32}; PC- 2 = {34,35,47,49}; PC- 3 = {10,14,15,17,26,27,31,43,45}; PC- 4 = {1,7,8,12,16,19,20,21,22,23,28,33,36,37,38,42,48,50}; PC- 5 = {11,30,46} and PC- 6 = {2,3,4,5,9,24,29,39,40,41,44};

The third successful part-cluster formation solution:

PC- 1 = {4,6,7,13,18,22,25,32,38}; PC- 2 = {34,49}; PC- 3 = {10,14,15,17,26,27,29,31,43,45,47}; PC- 4 = {1,8,12,16,19,20,21,23,28,33,35,36,37,42,48,50}; PC- 5 = {11,30,41,46} and PC- 6 = {2,3,5,9,24,39,40,44};

The fourth successful part-cluster formation solution:

PC- 1 = {4,7,13,18,22,32,38}; PC- 2 = {3,34,49}; PC- 3 = {6,10,14,15,17,25,26,27,29,31,43,47}; PC- 4 = {1,8,12,16,19,20,21,23,28,33,35,36,37,39,42,48,50}; PC- 5 = {2,11,24,30,41,45,46} and PC- 6 = {5,9,40,44}.

The fifth successful part-cluster formation solution:

PC- 1 = {4,7,18,22,32,38}; PC- 2 = {49}; PC- 3 = {6,9,10,14,15,17,25,26,27,29,31,34,43,47}; PC- 4 = {1,8,12,16,19,20,21,23,28,33,35,36,37,39,42,48,50}; PC- 5 = {2,3,11,13,24,30,41,45,46} and PC- 6 = {5,40,44}.

The desired machine-clusters and part-clusters are obtained by pairing each successful machine-cluster and part-cluster formation solution, and applying the Assignment Algorithm to each pair. In the case of machine-part incidence matrix XVI, the desired machine-clusters are those obtained from the *fourth* successful machine-cluster formation solution and the desired part-clusters are those obtained from the *fourth* successful part-cluster formation solution.

The resulted Manufacturing Cells (M-Cell) are :
M-Cell 1 = (MC-1, PC-4); M-Cell 2 = (MC-2, PC-6); M-Cell 3 = (MC-3, PC-1); M-Cell 4 = (MC-4, PC-2), M-Cell 5 = (MC-5, PC-3) and M-Cell 6 = (MC-6, PC-5).

There will be 412 of 1(one) entries on the block diagonal of the final rearrangement of the matrix XVI out of 1609 entries. Thus, this algorithm results in 25.6% of non exceptional elements. The final rearrangement by this algorithm is displayed in the Appendix 3.

Table-1. Comparison of percentage of non exceptional elements resulted from SKP-1 and Chu and Hayya's algorithm [6]

MATRIX	Density of Matrix *)	SKP -1	CHU & HAYYA [6]	Comment on SKP-1
	20 - 40%			
I	0 207	37.7	35.7	better
II	0.236	38.3	31.6	better
III	0.254	32.3	34.5	worse
IV	0.260	36.0	42.5	worse
V	0.295	25.8	37.4	worse
VI	0.352	29.6	30.7	worse
	40 - 60%			
VII	0.404	31.4	30.4	better
VIII	0 431	28.3	26.0	better
IX	0.445	28.2	29.4	worse
X	0.459	31.8	24.8	better
XI	0.485	35.5	25.6	better
	60 - 80%			
XII	0.611	26.1	25.6	better
XIII	0.659	33.4	29.5	better
XIV	0.669	29.5	23.0	better
XIII	0.704	28.8	24.8	better
XV	0.730	24.4	30.2	worse
XVI	0.805	25.6	20.7	better

*) Density of a matrix as defined in [19]

5. SUMMARY AND SUGGESTIONS

In general, computational experience summarized in Table 1 shows that for higher density machine-part incidence matrices, the proposed new algorithm, ie. Susanto-Kennedy-Price Version 1 or SKP-1, has more chance in giving cell formation solutions with fewer numbers of exceptional elements.

However, further research is required to identify other attributes of the input machine-part incidence matrix, such as *the total bond energy* matrix (as defined in [19]), in causing exceptional elements in the final solution.

ACKNOWLEDGMENT

We would like to thank the AusAID for their support of this research, and the Monash University Research Training and Support Branch for providing additional funding. Also, special acknowledgment is directed toward Dr. P.A. Grossman, Department of Mathematics, Monash University, for his time in the informal discussions and personal correspondence relating to Fuzzy Clustering and the Assignment Problem.

APPENDICES

Appendix 1. The Machine-Part Incidence Matrix XVI

Machine-type Number	Part-type Number 00000000011111111112222222222333333333344444444445 / 12345678901234567890123456789012345678901234567890
01	11111111111011111101110111110111010111110111111111
02	11110111011111111111111101110111111111100111011111
03	00011110110111101111010011111101111111111111111011
04	11001011101111011111101110110001101011111110010111
05	01011101011011110111000111011111111110101011111111
06	11111001111111101001111110100111110100111110111111
07	11110110111101111111111111010111101011011011101010
08	11101011110111111110110110111011111011111111100010
09	11111001111100111110000111111111000111111001100011
10	10111111110101110101111111001110001111101111111110
11	11111010101111011011101111011111110111111111111111
12	01111101100110111111110110110111001110111110001110
13	11111110011101111111011101111011111111111110110101
14	01111110111111110111111111111111111111001001011110
15	10111110111111101111111111110111111011011010111011 11
16	10101111100111011111111100110111011111110100101111
17	10111011111101101111111111111111111111110011001111
18	01111101110111101111011111111111111101111111101101
19	10011011111111111011111110111111110111111111110101
20	11111111111111111001111100111111111101010101011111
21	11111111111111111111101101110111110111111111111111
22	11001101111111111111011011111111111111101101111101
23	11101001111111101011111111100101011101111111111111
24	11111111111111111101111010111101110001111111111011
25	01001111101011111111110111111111110110110111111011
26	11101101011111111101011111101010111111111111011111
27	01111011111111111111011101101110011110111111111101
28	11111111001110101111111111111110111111101111110111
29	11110111011110110110100110111110011111101111111101
30	11111001111101110011101111111101101111111111111111
31	01110111110111100011111111011011111111111101111111
32	11101011111111111000111111111110011011011101111110
33	11100100011111111111111111101110110011111111111111
34	10110111111111100101111110101111111111011111111111
35	10111111111111111111101111101101011101111101010110
36	11111111111011111111101110101111111111111111111011
37	11010111110111111111111011101111111101011111111111
38	10110111011101100111001101101111111011111111011111
39	10110111111011101000101001111110111010111111111110
40	11110011001101111111111100101111011111111111110111

Appendix 2. The Final Rearrangement of Machine-Part Incidence Matrix XVI by Chu and Hayya's Algorithm

Machine-type Number	Part-type Number					
	0011223344 1789053739	1335 4280	0023344 5911905	0122444 4028467	011122222334 215734679451	0001113344 3682360628
4	1111111111	1111	1110110	0001010	110111010011	0011110011
7	1111111111	1110	0111011	1111101	111111100011	1101011000
15	1111100111	1111	1111011	1111101	010111111110	1101111011
17	1111111111	1111	1111110	1111101	011111111110	1011001101
35	1111110111	1110	1111011	1111001	011101111111	1111110111
37	1111111111	1111	0111011	1111111	101110111110	0111110111
6	1001101001	1111	1111111	1110111	110011101011	1011111011
19	1111101110	1111	1111111	1111110	011011111011	0011111111
20	1101101001	1111	1111010	1111111	111011011110	1111111111
21	1110111111	1111	1111111	1110111	111101111111	1111111011
23	1011111111	1111	1110111	0111111	110011000101	1011111111
24	1111101011	1111	1111111	1101111	111111111101	1111100010
30	1001110111	1111	1111111	1101111	111011111111	1011010111
31	0111111101	1111	0111111	1111111	100011101111	1111100111
38	1111101111	1111	0001110	1100111	011011111111	1111001011
40	1111100111	1111	0011110	1011111	111110101111	1011011111
8	1110101011	1110	1111111	0101100	101111111111	1011110110
10	1110110111	1110	1111111	1110111	001011110000	1111011111
36	1111111111	1111	1111111	1100111	101111011111	1111111110
3	0111111111	1111	1101111	1111111	001100111101	0101101110
22	1010111110	1111	1111011	0111111	111101111111	0111111101
1	1101110111	1111	1111101	1110111	111101111101	1110111111
5	0010001111	1101	1001101	1111111	111111111111	0110101011
9	1010010101	1011	1101111	1101100	111111111011	1010011100
11	1101111111	1101	1111111	1000111	110111111111	1001111111
12	0011100011	0010	1111110	1000001	101111111111	1111111111
14	0111111011	1110	1111000	1111111	111111111111	1101101001
25	0111110011	1111	1101111	0011111	111111111110	0110111110
26	1001011101	1011	1011111	0111111	111111100111	1111111111
27	0110100110	1001	1111111	1111111	111111110111	1011111111
32	1100010011	1010	1111101	0111011	111111111111	1011111111
33	1011110011	1111	0011111	0110111	111111111111	1101111011
39	1100000011	1110	0111011	1101111	011110111111	1110101111
2	1111111111	1111	0011001	1110011	111101111111	1111111111
13	1111110110	1111	1011111	1010010	110101111111	1111111111
16	1111101101	1111	1111111	0011001	000110110010	1111111111
18	0011111110	1111	1111111	1111101	101001111111	1111111011
28	1111111111	1111	1010101	1011110	110011111111	1111111111
29	1101111110	1011	0000101	1101111	110111011111	1111111111
34	1111101111	0111	0111101	1111111	010011101111	1111111111

Appendix 3. The Final Rearrangement of Machine-Part Incidence Matrix XVI by SKP-1 (Susanto-Kennedy-Price Version 1) Algorithm

Machine-type Number	Part-type Number					
	00111222233333445 18269013835679280	044 504	001233 478228	4 9	00111122223344 69045756791437	001123444 231340156
15	10111111101010111	111	111111	1	11110101111111	011111010
16	11111111111111111	110	011111	1	11010101101001	010101010
18	01111110111011111	111	101111	0	11111011111111	110111110
20	11111111111100111	111	110111	1	11111000111101	111111001
31	01101111111111111	011	111111	1	11110011011101	110110111
35	11111110101110110	110	111111	1	11111111111111	011110110
36	11111111011111101	111	111011	1	11111110111111	110111111
37	11111111111110111	011	111111	1	11111111111111	100100011
38	11101101011011111	011	111011	1	10111001111111	011011101
40	11111111101111111	011	111111	1	00011101011111	111001101
19	11111111111111111	111	111111	0	01111001111010	001111111
7	10111111111010000	011	111111	1	11111111001011	111011110
8	11110111111101100	111	011011	1	01111101111110	110110110
10	11110111000111110	111	111111	1	11111011101011	010011011
12	01111111001101110	110	101001	1	11001101111111	110111100
17	11101111111111011	111	111111	1	01111111111111	011011000
24	11101111110001101	111	111011	1	11111011111111	111110111
30	11111111101111111	111	100011	1	01111011111111	111010111
3	00101100110111101	111	111111	1	11111111111111	000101111
1	11011110000111111	101	110111	1	11111111111111	111111111
5	01000000111101111	101	101110	1	10111101111111	101111111
9	11010001101111001	111	101001	1	01111111111000	111011110
14	00101111111000010	101	111111	1	11111111111111	111111101
21	11110110011011111	111	111111	1	11111111111111	111111111
22	11110110111110011	111	001111	0	11111111111111	101111111
25	01011101101101101	111	011111	1	11011111111111	101111011
26	11111011111111111	111	000101	1	10111111001101	111111111
27	01110111101111111	111	111100	0	01111101101111	111111111
32	11110011101101110	100	010101	1	01111111111111	111111111
33	10111111001001111	011	001111	1	10111111111111	111111111
39	11000011101100110	011	110011	1	11111101111111	011101111
2	11111110011110111	000	111111	1	10111111111111	111111111
4	11111111111011111	110	011011	1	01010110100010	101110101
6	11111111011001111	111	100111	1	01110001011001	111111111
11	10111111011111111	111	110010	1	01010111111111	111111111
13	11111110001111111	110	111111	0	10010111111110	111111111
23	11111111110111111	111	001111	1	01110010000111	111111111
28	11111111111111111	101	111111	1	10010011110110	111111111
29	11111101111111111	001	110001	0	10110110110111	111111111
34	11111111111111111	001	111111	1	11100001011111	011111111

REFERENCES

1. Wemmerlöv, U., Production Planning and Control Procedures for Cellular Manufacturing Systems, The American Production and Inventory Control Society, Inc., Falls Church, USA, 1988.
2. Singh, N., "Design of cellular manufacturing systems: An invited review", European Journal Of Operational Research, Vol. 69, p. 284-291, 1993.
3. Seifoddini, H., and Djassemi, M., "The threshold value of a quality index for formation of cellular manufacturing systems", International Journal of Production Research, Vol. 34, No. 12, p. 3401-3416, 1996.

4. Boe, J.W., and Cheng, C.H., "A close neighbour algorithm for designing cellular manufacturing systems", International Journal of Production Research, Vol. 29, No. 10, p. 2097-2116, 1991.
5. King, J.R., "Machine-component grouping in production flow analysis: An approach using rank order clustering algorithm", International Journal of Production Research, Vol. 18, p. 213-232, 1980.
6. Chu, C.H., and Hayya, J.C., "A fuzzy clustering approach to manufacturing cell formation", International Journal of Production Research, Vol. 29, p. 1475-1497, 1991.
7. Ponnambalam, S.G., and Aravindan, P., "Design of cellular manufacturing systems using objective functional clustering algorithms", International Journal of Advanced Manufacturing Technology, Vol. 9, p. 390-397, 1994.
8. Brown, D.E., and Huntley, C.L., "A practical application of simulated annealing to clustering", Pattern Recognition, Vol. 25, No. 4, p. 401-412, 1992.
9. Zhang,Q, and Boyle, R.D., "A new clustering algorithm with multiple runs of iterative procedures", Pattern Recognition, Vol. 24, No. 9, p. 835-848, 1991.
10. Ruspini, E., "A new approach to clustering", Information Control, Vol. 15, No. 1, p. 22-32, 1969.
11. Gitman, I., and Levine, M., "An algorithm for detecting unimodal fuzzy sets and its application as a clustering technique", IEEE Trans. Comput., Vol. C-19, No. 7, p. 583-593, 1970.
12. Bezdek, J.C., and Pal, S.K., Fuzzy Models for Pattern Recognition, IEEE Press, Piscataway, NJ, USA, 1992.
13. Bezdek, J C., Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.
14 Taha, H., Operations Research: an introduction, 5th ed., p. 214, Macmillan Publishing Company, Don Mills, Ontario, 1992.
15. Bunday, B.D., Basic Linear Programming, Edward Arnold, London; Baltimore; Md, USA, 1984.
16. Wemmerlöv, U., and Hyer, N.L., "Research issues in cellular manufacturing", International Journal of Production Research, Vol. 25, 413-431, 1987.
17. Chu, C.H., "Recent advances in mathematical programming for cell formation", Planning, Design, and Analysis of Cellular Manufacturing Systems, ed. Kamrani, A., Parsaei, H.R., and Liles, D.H., p. 3-46, Elsevier Science B.B , Amsterdam, 1995.
18. Miltenburg, J., and Zhang, W., "A comparative evaluation of nine well-known algorithms for solving the cell formation problem in group technology", Journal of Operations Management, Vol. 10, No. 1, January, 1991.
19. McCormick,W.T, Schweitzer, R.J, and White, T.W., "Problem decomposition and data re-organization by clustering techniques", Operations Research, Vol. 20, p. 993-1009, 1972.

HOW TO IMPROVE THE ACCURACY OF PROTOTYPE BUILT BY THE LAMINATED OBJECT MANUFACTURING SYSTEM

Liu Bin, Xiao Yuejia, Han Ming and Huang Shuhuai
Department of Plastic Forming Engineering
School of Mechanical Science & Engineering
Huazhong University of Science & Technology
Wuhan , P.R.China 430074

ABSTRACT

The building principle and composition of our HRP-I system, which is based on the principle of Laminated Object Manufacturing technology, are introduced in this paper. Accuracy is the cornerstone of rapid protopyping & manufacturing, but there is a lack of literature on the problem how to improve the accuracy of prototype built by LOM system. As the widespread application of this system, it is one of the problems with which users are often concerned. In this paper we will analyze the main factors that influence the accuracy of part's prototype from software and hardware two aspects based on our self-developed HRP-I system. Some corresponding methods to improve the accuracy will be also proposed.

KEYWORDS

Rapid Prototyping & Manufacturing Technology, Laminated Object Manufacturing, Accuracy, Prototype, CAD/CAM

1. INTRODUCTION

Rapid Prototyping & Manufacturing technology is a new high technology which has been developed since the middle of 1980s. It makes an important progress in the field of manufacturing. It adopts the layer manufacturing principle to make the CAD geometry solid model and the process of manufacturing share an integrated environment, so it is able to build the part's prototype directly from the three dimensional CAD solid data model. This kind of technology has more advantages than conventional one, so it has been paid high attention in the field of engineering, and developed rapidly in the world.

Because of the difference in the aspect of building material and process, it is well-known that the RP&M technology has been divided into various processes such as Stereolithography, Laminated Object Manufacturing, Selective Laser Sintering, Fused Deposition Modeling, Solid Ground Curing and so on. We have been engaged in the research of RP&M in HUST since 1991. Taking into account the building material and the laser technique, we place emphasis on the studies of Laminated Object Manufacturing technology, and first successfully researched and produced a new, practical, low-cost system — HRP-I in China, which is based on the principle of LOM technology and has the independent copyright. The system has built some complex shape prototypes, which have certain dimensional accuracy.

Nevertheless, accuracy is the cornerstone of Rapid Prototyping & Manufacturing technology. The whole point of part building with the new advanced methods is to generate accurate parts. The benefits of the process may be virtually moot if the prototypes are not accurate. The accuracy of SLA system is studied in many literature, but there is a lack of literature on the problem how to improve the accuracy of prototype built by LOM system. As the widespread application of LOM system, it is one of the problems with which users are often concerned. So, in this paper we will propose to fill this gap in the literature, and analyze the main factors that influence the accuracy of part's prototype from software and hardware two aspects based on our self-developed HRP-I system. Some corresponding methods to improve the accuracy will be also proposed in this paper.

2. THE BUILDING PRINCIPLE AND COMPOSITION OF HRP-I SYSTEM

2.1 The Building Principle of HRP-I System

The system makes use of the laser beam and thin material(such as paper, plastic, etc.) to building any complex object based on the principle of laminated (or layered) manufacturing. A very thin layer of heat-sensitive material is spread on one side of the building material. When the temperature of the material surface is up to some degree, the upper thin material will be rapidly cemented on the previous one. Using the laser beam to cut this kind of material and pile them up layer by layer. The detailed process as follows: CAD data (such as "STL" format) is inputted into the HRP-I system. A cross-sectional slice is generated by the slicing algorithm and the laser cuts the material according to outline of the cross-section. New layer of the thin material is bonded to the top of the previous cut layer. Next cross-section is prepared and cut. This automatic process continues until all layer are laminated and cut. The' finished object is then separated from the excess material. The principle of system is shown as figure 1.

2.2 The Composition of HRP-I System

The HRP-I system is mainly composed of computer, laser cutting subsystem, heated roller, the subsystem of delivering sheet material, the subsystem of measuring the height of Z-direction, servo-control subsystem, platform and software. A simplified schematic drawing of HRP-I system is illustrated in figure 2.

Through analyzing the composition of the system, the factors that influence the accuracy of part's prototype are: (a) CAD model accuracy; (b) the slicing accuracy in slicing procedure, i.e.

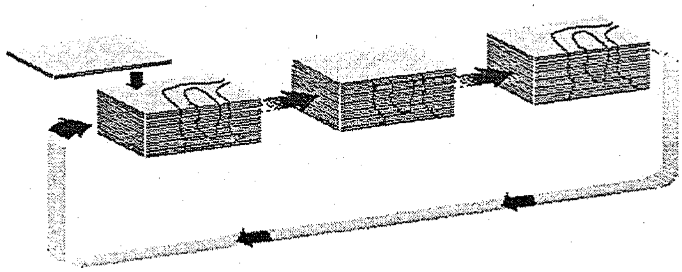


Fig.1: The principle of HRP-I system

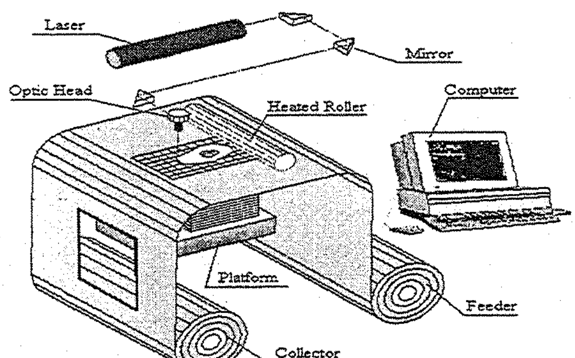


Fig.2: A simplified schematic drawing of HRP-I system

the contour accuracy of slicing cross-section layer; (c) the selection of slicing layer thickness; (d) the properties of the real-time control software; (e) the stability of laser power; (f) moving responding speed of laser cutting head; (g) the responding speed of laser beam's on/off; (h) the dimension of laser facula; (i) laser facula's scanning plainness accuracy; (j) movement control accuracy of servo subsystem; (k) the matching relationship between cutting speed and laser power; (l) the control accuracy of temperature and pressure of hot roller; (m) the thickness uniformity of sheet material; (n) the verticality between the surface of platform and the Z-direction; and so on.

3. THE MAIN INFLUENCING, FACTORS

We will put emphasis on analyzing some main influencing factors, such as:(i) CAD model accuracy; (ii) the contour accuracy of sliced cross-section layer; (iii) the selection of slicing layer thickness; (iv) the dimension of laser facula; (v) movement control accuracy of servo subsystem and (vi) the matching relationship between cutting velocity and laser power in this paper.

3.1 CAD Model Accuracy

In the field of RP, many systems widely adopt the STL file format to act as their input of solid data model. Meanwhile many CAD systems have also implemented 3D systems' stereolithography (STL) file format[1], which has become the *de facto* standard. The principle of STL tessellation is to approximate actual 3D object with a lot of planar triangular patches. Using this approach, an " accuracy" or offset parameter is input by the designer. This is the acceptable chordal error between the plane of a triangle and the actual surface it is approximating. This value has no relevance for the polyhedron, which is made of planar faces and the concept of tessellating such as object for file transfer to RP is quite efficient. However, Non-polyhedron is not this case, which a high accuracy requirement will result in a very large data file. In reality, of course, some methods can used such as splitting a model into separate pieces in order to keep file sizes down and vary chordal accuracy. Without question, the more the number of triangular facet is, the model which tessellates in STL will approximate the actual CAD solid data model better, and the model approximating accuracy is more high. But the grade of translating in some solid modeling system ,such as AutoCAD AME, is limited. When the grade is maximum, the numbers of triangular facet transformed by the translator is constant. If the geometry dimension of solid model increases, it will lead to increase the model's approximating error and decrease the CAD model accuracy, and influence the sequential prototype's accuracy. For example, in the 3D solid modeling system of AutoCAD AME 2.0, the maximum grade of translator is twelve. When the grade is twelve, the numbers of triangular facet transformed by the translator is also certain maximum constant. So ,in order to get high accuracy of prototype, first of all, it needs a high accuracy of solid data model. The method which improve CAD model accuracy should be to increase the grade of STL translator and to increase the numbers of planar triangular facet or to seek a new solid data model file format. For example, the file format may be the result of direct slicing of solid models.

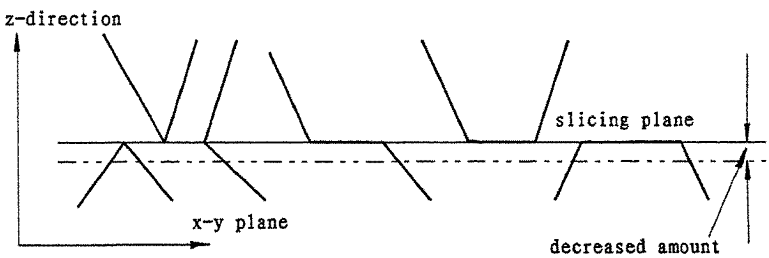


Fig.3: Some cases of the vertex of facet on the slicing plane

3.2 The Contour Accuracy of Sliced Cross-section Layer

When the CAD solid data model is represented with a lot of planar triangular facets, slicing procedure of this kind of model will result in the contour error of cross-section layer. The main reason is that the vertex of planar triangular facets is on the slicing plane. Some cases of the vertex of facet on the slicing plane are shown as figure 3. By the requirement of real-time of slicing algorithm, we have adopted a method, i.e. the Z coordinate of the slicing plane is decreased by a small amount, to avoid the complication of the slicing algorithm. So, the part sliced, and the resulting slice is translated upwards by the same amount. The distance is currently set to 0.001mm but can altered in the dialogue frame of the program. In case the vertex of planar triangular facet is still on the slicing plane, the procedure of decreasing the slicing height would be repeated until no vertices is on the slicing plane. This method can simplify the slicing algorithm, and enhance the speed of slicing, but at the meantime, it will result in contour error of cross-section layer. Since the contour of cross-section at the actual measured height and the one obtained by the slicing algorithm at the actual calculating height are often different, the contour error will arise. But only if the amount of decreasing the slicing height be not great, and the thickness of slicing layer be also not great, the contour error could be neglected for the sequential process. Certainly, if the thickness of slicing layer is more great, and the change of cross-section is also more great, the contour error of cross-section will seriously influence the sequential building accuracy. Thus, the thickness of slicing layer should be decreased in order to improve the accuracy of prototype. Certainly, this method will lead to increasing the amount of slicing layer and decreasing the building efficiency. In addition, on account of the error of the CAD model, slicing procedure will obtain some close polylines of the contour of the slice, not some real close curves, this also lead to the principle error. In order to obtain the exact contour of cross-section and reduce the error of the contour to minimum, it would be best to directly slice the CAD solid data models.

3.3 Movement Controlling Accuracy of Servo System

In the view of control, the processing of rapid prototyping and manufacturing, in fact, is the co-movement of the coordinate axis, which include x, y and z axis. The co-movement of x and y coordinate axis control the contour of the slicing cross-section, the z coordinate axis control the up and down movement of the platform. The step of z-direction must be equal to the corresponding thickness of slicing layer, it may ensure that the contour of cross-section is much closely approximate the actual contour of cross-section at the measured height. If errors exist in the displacement of z direction, it will directly influence the z-direction dimensional accuracy and shape accuracy of prototype.

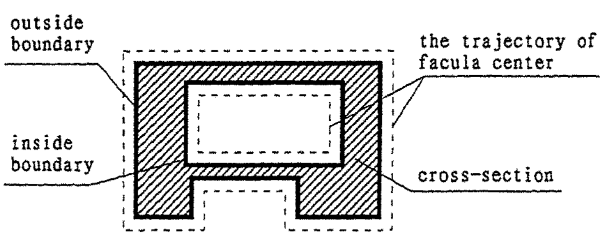


Fig. 4: The radius compensation of the laser facula

3.4 The Dimension of Laser Facula

As we know, the contour of cross-section layer is cut by the laser beam in our HRP-I system. In the actual process of building, the contour of cross-section layer is the theoretic path of laser beam, but the laser facula have a certain dimension. When the size of laser facula radius is too large which can not be neglected, the radius compensation of the laser facula in the process of cutting is needed, otherwise, the laser facula radius will also directly influence the contour error of cross-section layer and the accuracy of prototype. So in the process of cutting, on one hand, the radius of laser facula should be reduced, on the other hand, it should be measured in order to realize the real-time radius compensation of the laser facula and to improve the accuracy of cross-section layer. Thus, firstly, the inside and outside boundary among the solid cross-section contours should be automatically recognized[2], then it starts with the feature of the solid cross-section contours, and determines the radius compensation vector of laser facula in accordance with the concave and convex of the vertex, the direction of boundary and its inside and outside feature. We have realized the corresponding compensation for concrete boundary of cross-section[3] in the HRP-I system. The method of the radius compensation of the laser facula is shown as figure 4.

3.5 The Selection of Slicing Layer Thickness

The thickness of slicing layer also directly influence the surface rough, the dimensional and shape accuracy of z-direction of prototype and the building time and costs. It is one of the most important parameters in the process of RP. If the accuracy of prototype is chiefly considered, the thickness of slicing layer should be selected smaller value at the height which cross-section changes greatly. Otherwise it cannot ensure the accuracy of prototype and the staircase effect and the serious fuzzy phenomenon at the height will arise[4]. The staircase effect is illustrated in figure 5. In addition, an adaptive slicing algorithm, which select various slicing thickness at the various height in the light of the Z-direction contour of solid, should be adopted to eliminate the staircase effect and enhance the speed of building the prototypes. The staircase effect can be addressed in several ways. Firstly, good software tools can help minimize the problem. Secondly, post-treatment can be applied, and in this case, the part is usually polished. Finally, the processes can be improved to virtually eliminate the problem. For instance, the technology developed by laser 3D can use a layer thickness as low as 0.015mm resulting in parts with no noticeable staircase effect to the naked eye. Soligen claims that their process can also eliminate the staircase using different principles[5].

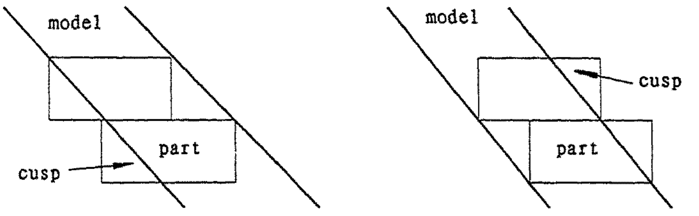


Fig. 5: Staircase effect and cusp height *c*

3.6 The Matching Relationship between Cutting Velocity and Laser Power

The matching relationship between cutting velocity and laser power is also a crucial problem which should be well resolved in the control subsystem. As the complication of the contour of cross-section is different, the appropriate cutting velocity(it is often variable) of laser beam must be fitted for the laser power in order to improve the accuracy of prototype. When the thickness of sheet material is a constant value, more high the laser power is, more fast the velocity of cutting of laser beam becomes, vice versa. The control principle frame is illustrated in figure 6.

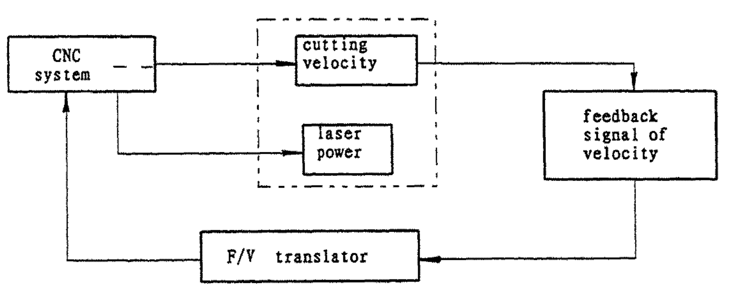


Fig. 6: The control principle frame of P/V

4. CONCLUSION

There are many factors that influence the accuracy of part's prototype in HRP-I system and their influencing principle is also complicated. Every one of these is not only possible, but also has been achieved by us, however, without accuracy, they are often meaningless. This paper has put emphasis on analyzing some main factors, so we are able to improve the accuracy of part's prototype. These analyzed results above will play the theoretical and directive role in improving the accuracy of part's prototype for the other system of RP&M technology in the same way.

Nonetheless, it is important to remember that the cornerstone of RP&M system is still accuracy. It is the responsibility of the RP&M system manufacturer to improve the accuracy of the system in a reliable, repeatable, and conclusive manner.

REFERENCES

1. 3D Systems Inc. Stereolithography Interface Specification. 1988.7
2. Liu Bin, Xiao Yuejia, Han Ming and Huang Shuhuai, "Research on Algorithm of Automatic Recognition of the Inside and Outside Boundary among the Solid Cross-section Contours", Journal of Huazhong University of Science and Technology, 1996(10), p23~25
3. Liu Bin, Xiao Yuejia, Han Ming and Huang Shuhuai, "Research on Algorithm for the Automatic Radius Compensation of the Laser Facula in the Laminated Object Manufacturing Technology", Journal of Huazhong University of Science and Technology, 1996(10), p26~29
4. A. Dolenc and I Makela, "Slicing procedures for layered manufacturing techniques", Computer Aided Design, 1994(2), p119~126
5. A. Dolenc, "An Overview of Rapid Prototyping Technologies in Manufacturing", further details of this paper are available from Dr. A. Dolenc. Institute of Industrial Automation, Helsinki University of Technology, Finland

STUDY ON MULTIFUNCTIONAL RAPID PROTOTYPING MANUFACTURING(M–RPMS) SYSTEM

YongnianYan, Guang Zeng, Wei Zhang, Haibin Guo, Liangwei Wu, Renji Zhang
Center of Laser Rapid Forming, Dept. of Mech. Eng.
Tsinghua University, Beijing 100084, P.R.C.

ABSTRACT

This paper introduces a new-style RP system: applicable in research, development, teaching and training. M–RPMS (Multi-functional Rapid Prototyping Manufacturing System) ensures an open environment to customers through both overall structure of software and function design. It is accomplished through decoupling of mechanical electronics system (MES) and coupling of mechanical electronics elements (MEE) is an effective way to realize function integration. Both of them are based on methods of mechatronics.

KEYWORDS

Rapid Prototyping (RP), Advanced Manufacturing Technology, Mechatronics

1. INTRODUCTION

RP is a rapid developing technology. Its application covers all of industrial and academic fields. However, these industrial and academic requirements are different: the industrial customers need high efficiency, accuracy and sufficient working space while the later emphasizes software openness, technical diversity and system structural recombinability with acceptable efficiency, accuracy and working space.

1.1 Openness of RP software

Openness of RP software means that the customers may select some of provided function modules and integrate their self-developed modules freely. A set of new software which has the customer required functions and features, is formed in the provided overall framework in our M–RPMS. It is necessary to open the following to customers in order to achieve the above requirements:

* main data structures
* detail analysis of function requirement
* major information transmission
* data flow and overall software structure.

Apparently, all commercialized RP systems can't reach these demands up to now. But these features are achieved in the M–RPMS.

1.2 Processes Diversity

There are at least tens of RP technologies, at the mature or the developing stages. The M–RPMS should provide several RP processes according to customers' interests, one of which should reach the present commercialized level. Each RP process has its disadvantages in application. Investigators of RP should break the narrow ancestry opinion of manufacturer of RP apparatus. To analyze problems with development of subject, it needs to be supported by manifold RP processes.

* Recombinability in structure. Analyzing from Modern Forming Science, RP technology belongs to discretization/stacking forming[1]. Discretization means to slice the 3-D CAD model in order to obtain the stacking paths and restrictions of materials. This process is accomplished by RP software while material stacking is accomplished by machinery structure and control system. Recombinability of structure is referred to provide a kind of flexible settings, which can be

recombined by customers on both machinery structure and control system in order to achieve their specific requirements of new RP processes development.

RP is the core content of advanced manufacturing technology in 21st century. If the customers focus on limited application areas mainly, it is apparent that mono-function RP apparatus are most competitive; if customers undertake the high-level mission on research and development of RP technology ,or requirements of application in a wide range of areas and professional training, then the M–RPMS is the most suitable system with the features of systematic openness, process diversity and structural recombinability.

This paper mainly studies on intention and achievement methods of open RP software , diversity of RP processes and recombined structure. The first M–RPMS apparatus has been developed according to above principles. The beta M–RPMS are being developed .

2. THE OPEN RP SOFTWARE SYSTEM

An outstanding software must have a friendly operating interface, reasonable module structure, extendibility, mentenability, fault freedom etc.. It is obvious that all these basic requirements must be embodied in the software system. However, what we more concern is how to make it convenient for the customers to use our provided functional modules and add their self–developed functional modules in order to reach a certain requirement. Customers don't only use but also participate the software system. The M–RPMS software's main feature is customer oriented.

2.1 Overall System Structure

Reasonable overall system structure is the basic assurance to realize the above openness. According to Dicretization/Stacking Forming principle, any RP process has two basic requirements for its software[2]:

(1) *To generate scanning routes and restrictions used to control material stacking through CAD model discretization.* This is a data treatment process, which has no direct relationship with particular RP process and NC system. Its input is STL file while its output is CLI file (see Fig. 1).

(2) *To generate NC Code according to CLI file, particular RP process and NC system.* Then NC system generates a series of control signals used to control every working unit of RP system.

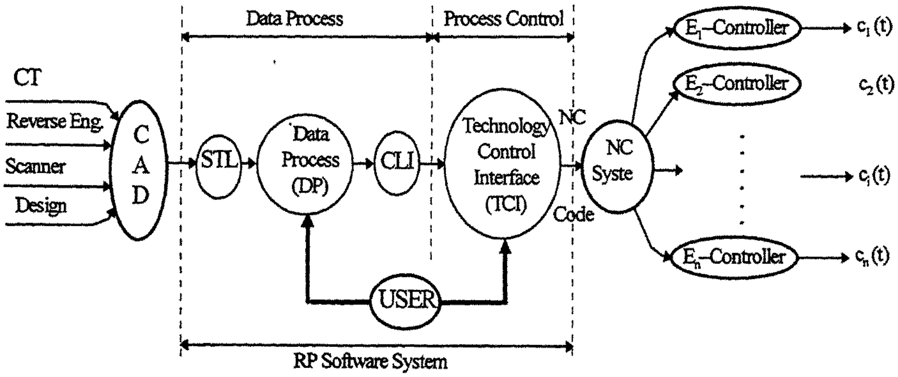


Fig. 1 Overall Structure of RP Software System

In Fig. 1, customers may get involved in DP (Data Process): selecting and recombining certain functional modules in order to obtain new data process function. For very particular process rules and demands, customers may get involved in through TCI (Technology Control Interface).

2.2 DP Function

Reasonable DP function division is not only the assurance to realize RP process, but also the basis for customers to get involved. Fig. 2 shows DP functional modules.

2.3 Data Structure

In order to make it convenient for customers to get involved in DP and TCI, the main data structure should be open, especially the STL topo-information and CLI data structure. The former is for customers to use DP modules better while the latter is for customers to modify some process parameters and rules through TCI.

3. RP PROCESS DIVERSITY

Several RP processes may be achieved in one RP machine through function integration. Decoupling of Mechanical Electronics System (MES) and coupling of Mechanical Electronics Elements (MEE) is an effective way to realize function integration.

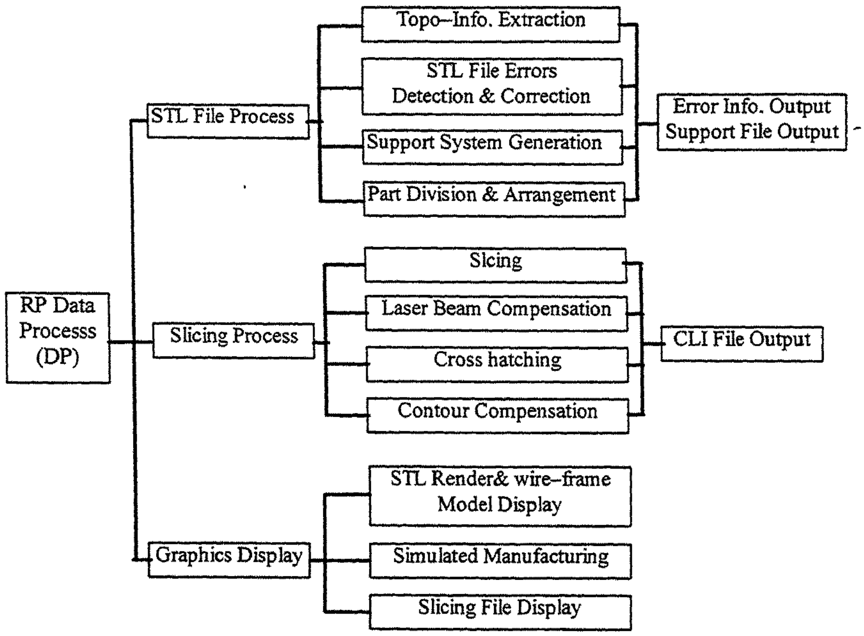


Fig. 2 Structure of DP Function Module

3.1 Three Step Structure of Mechanical Electronics System

RP machines are typical MES. It is composed of Mechanical Electronics Components (MEC) and MEE. Several MEE couple to a MEC (see Fig. 3)[3].

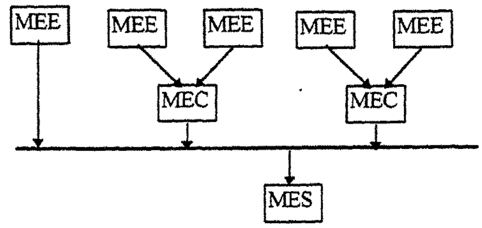


Fig. 3 MES structure

3.2 Mechanical Electronics Element (MEE)

MEE is the basic structural unit of MES. It possesses information process ability, achieves a certain physical function or output. MEE is composed of Controller—C, Diver—D, Executor—E, Sensor—S and Power—P (see Fig. 4).

The following are main MEE in RP machines:

* *D-MEE (Displacement MEE)* It automatically achieves a certain displacement(linear displace-ment or angle displacement) with desired velocity and acceleration according to instruction (Displacement , velocity or acceleration control).

* *T-MEE (Temperature MEE)* It automatically keeps the controlled object at a desired temperature according to instruction.

* *L-MEE (Laser MEE)* It automatically outputs a laser beam at a desired power according to instruction.

* *I/O-MEE (Input/Output MEE)* According to definition of MEE, input and output may also be regarded as a special kind of MEE.

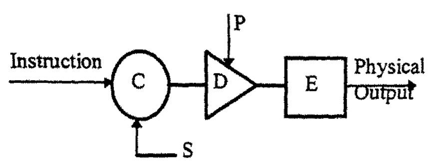


Fig. 4 MEE Structure

3.3 System Decoupling

System Decoupling (SD) is a process that converts complex, multi-factor MES into several simple, mono-factor units. SD is an inventive process. Its purpose is:

* *To solve engineering interference* MEE is a mono-factor unit and is fairly disciplinarian, and so easy to design and achieve.

* *To lower system cost* Taking relativity of every MEE into consideration, designing of new controller, driver or executor in order to obtain new MEE may lower system cost. Table 1 is the system decoupling result of SSM-500 developed by Tsinghua University.

Tab. 1 SSM-500 system decoupling result

Function Requirement	F1		F2	F3	F4	F5	F6	F7	F8	F9	F10	F11
Achieving MEE	D-MEE	D-MEE	D-MEE	D-MEE			T-MEE	P-MEE	I/O-MEE			AD-MEE
MEE No.	M^1	M^2	M^3	M^4			M^5	M^6	M^7			M^8

Explanation: F1: X-Y scanning F2: Z-stage moving F3: material(paper) feeding

F4: material collecting F5: material pressing F6: material heating
F7: laser F8: exhaust emission F9: panel operation
F10: working stage indication F11: Z height measurement

3.4 MEE coupling

Various MEEs constitute a more powerful MEC. This is also an inventive process. By software programming and hardware interface design, coupling of MEEs to become various MEC is the basic way to design RP machines. MEE Coupling (MC) is an packaging process. The purpose is:

* Obtaining the function that unattached MEE can not be provided.
* Using commercial controller products can save cost.
* Simplifying control software design.
* Solving the conflict between serial sending of control commands and parallel running/ corporation of various MEEs in RP machines.

There are four MEE coupling patterns.

* *Follow Coupling (FC)* One MEE's output always follows that of another MEE. The two MEE's outputs keep a fixed or relatively fixed ratio. The laser power, for example, should keep a proportion to the X–Y scanning speed in some RP processes.

* *Allocation Coupling (AC)* If the output ratio (k) of two coupling MEEs changes with time, then it will be allocation coupling. For example, X-Y scanning in RP apparatus is an allocation coupling of two D-MEEs. In RP technology, using STL data format could lead to that all slicing layers are composed by polygons. So each line has its own k value (slope of the line). When scanning of one line is completed, k valve is changed and another line will be scanned until every contour polygon of a layer has been scanned. NC interpolation card is a typical controller of allocation coupling.

* *Cooperation Coupling* Regarding the two coupled units of MEE, if one's input–output response is non-linear(lagging is an example), then the control signal of the two coupled units cannot change continuously, otherwise it will lead to dynamic following errors. In this case, controllers of the two units should be independent to each other. They control the lower drivers and executors separately. To ensure that the executors can work according to requirements of RP process, a kind of special cooperating signals must be used to set up the cooperation of the two units. This coupling pattern is so called cooperation coupling. Taking FDM process as an example, the angular displacement unit (filament feeding) and the X–Y scanning unit are typical cooperation coupling relationship. This coupling structure is used to solve the problem of material lagging when fused material is extruded from nozzle.

* *Sequence Coupling* If two units are ordered in action and have no other direct relationship, they are sequence coupling, which can be done by software of computer.

4. PROBLEM ON INTERGRATION OF FUNCTION

Function integration within RP can be completed by the methods of SD and MC, so several kinds of RP processes can be accomplished at one RP apparatus. Different RP processes only have its special demands for some special functions, but they have the same demands for MEE and MEC at the major aspects. We will discuss how to complete integration of function by the methods of SD and MC. The following discussion takes integration of FDM and LOM as an example.

Table 1 gives the result of SSM–500 decoupling. Same table of FDM can be obtained by the same method. Then function integration of LOM and FDM can be achieved by the methods of SD and MC. Figure 5 gives the coupling connection of M–RPMS which has the two functions: FDM and LOM.

5. RECOMBINABILITY OF STRUCTURE

Recombinability of structure is mainly reflected at the design of multi-functional forming chamber and multi-functional forming roller.

Multi-functional forming chamber is a recombined space which provides infrared laser beam with certain power and focus ability, precise X–Y scanning function, Z–axis movement control and detection, exhaust emission and temperature control system. Multi-functional forming roller has the functions of pressing, heating, powder reserve and leveling etc., which can be recombined also.

6. CONCLUSION

The M–RPMS is a new-style rapid prototyping apparatus developed for research, training and also industrial production. Its features include:

- Software opens to customers.
- Multi RP process can be done by function integration.
- Multi-function forming chamber and multi-function roller which can be recombined by customers to meet their demands.

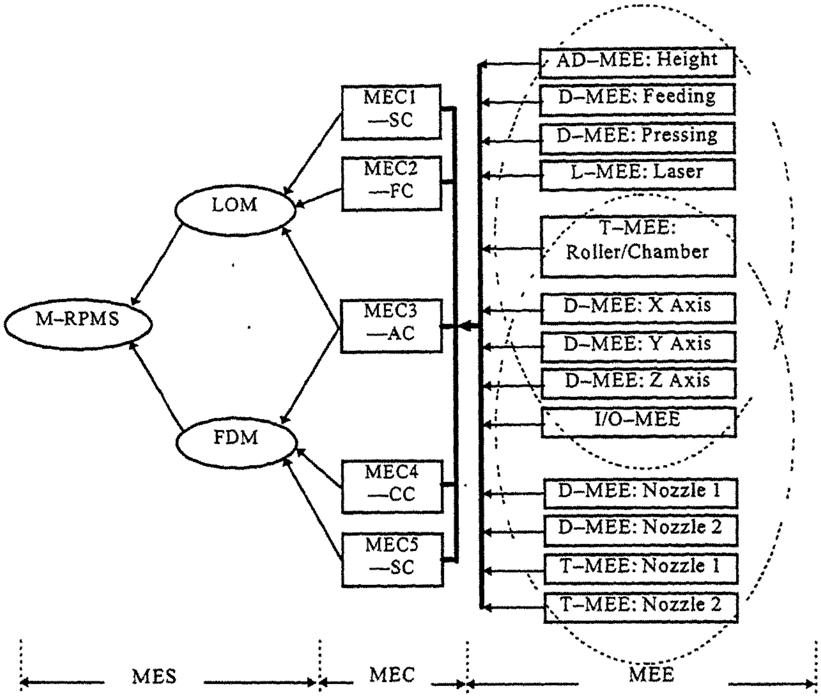


Fig. 5 M–RPMS Coupling Structure

REFERENCES

1. Yan Yongnian, Zhang Wei, et al. Principle and Development of RPM Based upon Discretization /Stacking Concept. China Mechanical Engineering, Vol. 5, No. 4, pp.64–66, 1994.

2. Yan Yongnian, Guo Haibin, et al., Study on Multifunctional Rapid Prototyping Manufacturing System(M–RPMS), Chinese Conference on Advanced Manufacturing Technology, Beijing, 1996. 9.

3. Zhang Wei, Principle and Application of System Design and Process Control of Rapid Prototyping, Ph.D. Dissertation, Tsinghua University, 1997.1.

DESIGN OF SLS (SELECTIVE LASER SINTERING) IN M-RPMS (MULTI-FUCTIONAL RAPID PROTOTYPE MANUFACTURING SYSTEM)

Renji Zhang, Guanghua Sui, Haibin Guo and Yongnian Yan
Dept. of Mechanical engineering, Tsinghua University, Beijing 100084, P. R. China

ABSTRACT

M-RPMS, the Multi-functional Rapid Prototype Manufacturing System, has the ability to practice SLS, LOM and FDM. The design of SLS in M-RPMS involves software design and hardware design. The software can deal with SLS data, monitor and control the machine to realize the SLS process efficiently. The hardware includes a powder hopper, a roller, which is also available in LOM to press the foil, a chamber and only one elevator. SLS and LOM use the same laser system. The SLS material in the M-RPMS are also discussed in this paper.

KEY WORDS

Selective Laser Sintering (SLS), Multi-functional Rapid Prototype Manufacturing System (M-RPMS)

1. INTRODUCTION

In the past decade, RP (Rapid Prototype or Rapid Part) technology has already been developed rapidly based on the developments in the fields of CAD, CNC, laser application and material science and engineering[1].

The SLS (Selective Laser Sintering) is a powder-based RP process. The SLS process builds three dimensional prototypes and/or parts in a layer-by-layer manner. The SLS facility has laser system, mechanical system, electrical system and appropriate software[2-4]. The other RP processes have machine system, electrical system and software too, even the laser system, such as LOM. Each process has its own features. Some processes have high precision. Some are better for direct metal prototype fabricating. Every RP machine can perform only one RP process. If a customer require another process, he usually buy another RP machine. It's really a big sum. For education or research of RPM, a machine with multitude functions, which means it can practice many RP processes, and an open platform may cut the ice. M-RPMS is the best choice in such a case. M-RPMS, the Multi-functional Rapid Prototype Manufacturing System, has the ability to perform SLS (Selective Laser Sintering), LOM (Laminated Objective Manufacturing) and FDM (Fused Deposition Modeling) processes, and also provides an open platform to be used to research other RP processes, such as SLA, 3D-P and so on in the future.

The design principle of M-RPMS is "the integration of the analogous components and the keeping of the different components". According to this principle, we analyzed SLS, SLA, FDM and LOM, found that the integration degree of SLS, FDM and LOM is higher than the other combination of three processes[5]. We designed the first M-RPMS which has the function of performing SLS, FDM and LOM.

2. DESIGN OF SLS IN M-RPMS

The design of SLS in M-RPMS involves software design and hardware design.

2.1 DESIGN OF SOFTWARE

Being a subset of M-RPMS software, the SLS software must deal SLS data and monitor and control SLS process safely and well. The SLS and M-RPMS software have such functions, as Fig. 1:

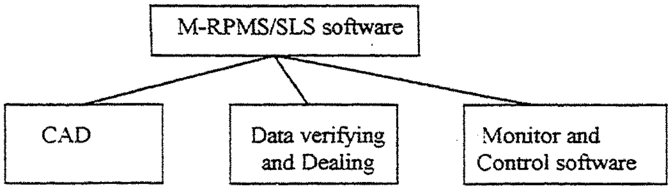


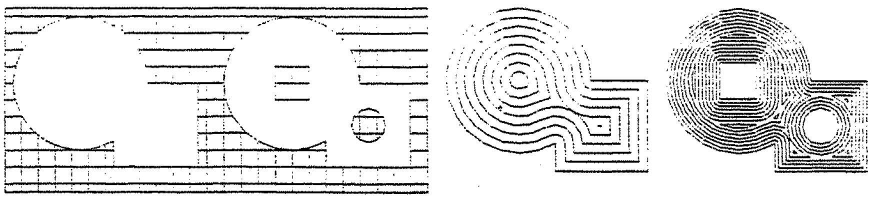
Fig. 1 : M-RPMS/SLS software

The functions of CAD involves designing geometric model and generating STL. data, which usually be done by a workstation.

Data verifying and dealing includes verifying and modifying STL. data, determining direction of model and shifting geometric and slicing object.

In SLS process, the laser spot affect the precision of the prototype, so the SLS software should compensate the effect. Compensating of LOM is on the contrary.

How to fill the cross section of a part is important. The filling pattern of SLS and LOM are showed as following (Fig. 2):



a. b.

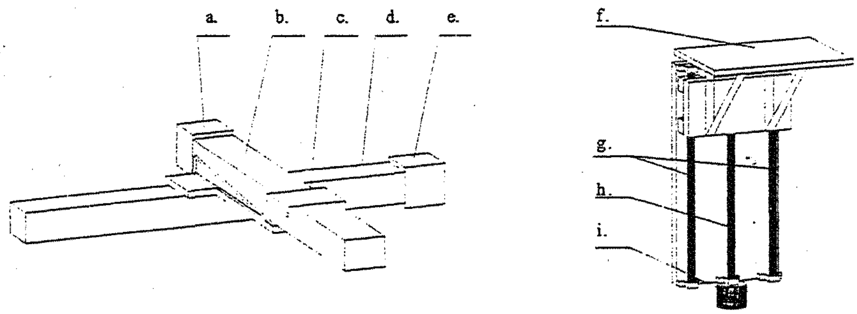
Fig. 2 Comparison of LOM filling pattern (a.) and SLS filling pattern (b.)

Such filling pattern can weaken the curling of the prototype and deconcentrate the stress.

The filling pattern of SLS is similar to that of FDM, but the SLS correlative control software is relative with the LOM system because the SLS and LOM use the same laser system.

2.2 DESIGN OF HARDWARE

The LOM, FDM and SLS have the same X-Y scanning system and the same elevator (Z axis) in the M-RPMS equipment. They are showed as following (Fig. 3):

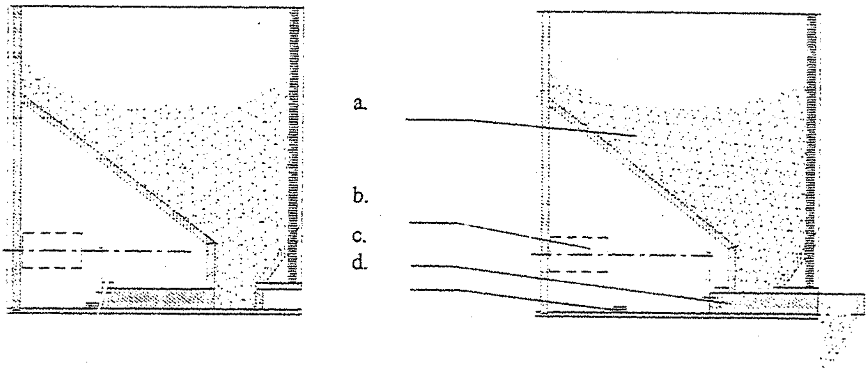


a. Motor b. Y guide c. Moving head d. X guide e. Motor
f. Platform g. Guide h. Leading screw i. Motor

Fig. 3 The X-Y scanning system and the elevator

Because of only one elevator and the limited working space in the M-RPMS equipment, the sending powder machinery in the M-RPMS equipment has small volume and easily-controlled characteristics and the platform should be compatible with SLS, FDM and LOM. We design a hopper used for powder spreading and a multi-platform for the other processes.

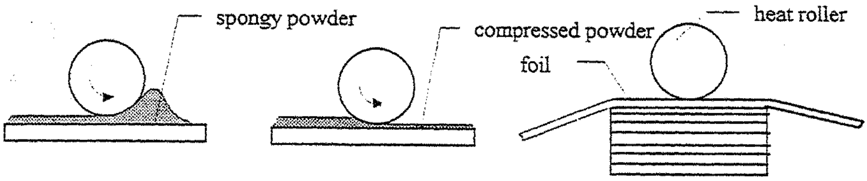
The hopper has an on-off element which can be controlled by the software to feed a fixed amount of powder or shut down the hopper in the right process. The hopper works as Fig. 4 :



a. Powder   b. Motor and guide screw   c. Drawer   d. Spacing element

Fig.4 The hopper

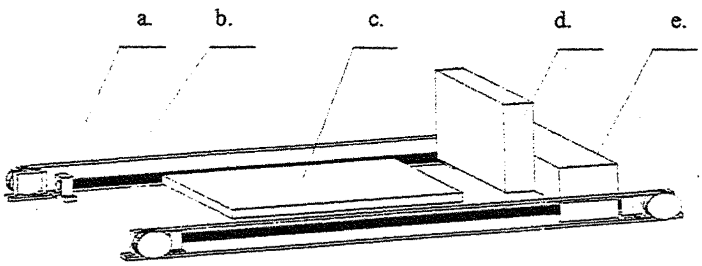
The heat-press roller, which is used in LOM process, is also available in SLS process to spread and press the powder. That is, the roller on the one hand can roll in SLS process, on the other hand can be heated in the LOM process.



a. spreading powder   b. pressing powder   c. pressing foil

Fig. 5: Functions of the roller

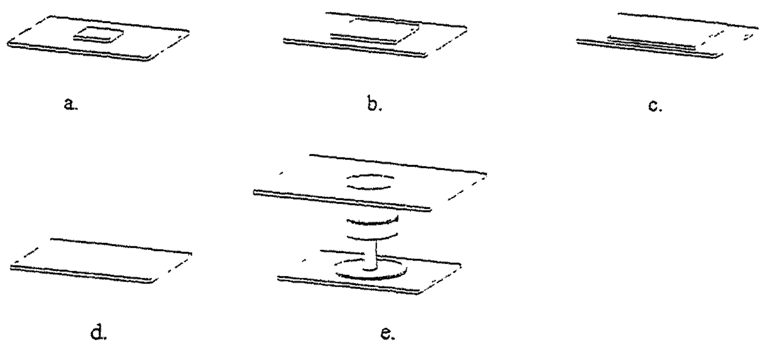
Hot pressing and powder sending system are showed as Fig. 6 :



a. Belt   b. Guide   c. Platform   d. Hopper   e. Rolle

Fig. 6 Hot pressing and powder sending system

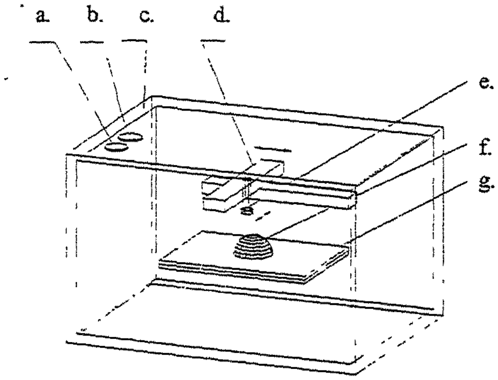
SLS, FDM and LOM has the same platform "port". Multi-platform economizes the limited resource in the M-RPMS and it is very convenient from a RP process to another. SLS, LOM and FDM fittings are showed as following (Fig. 7):



a. 100 × 100 mm LOM fitting b. 180 × 180 mm LOM fitting
c. 250 × 250 mm LOM fitting d. FDM fitting e. SLS fitting

Fig. 7 Fittings for multi-platform in M-RPMS

During the SLS and the FDM processes, to control the chamber temperature and atmosphere is very important, So we designed a thermostat in the building chamber to perform SLS or FDM processes (Fig. 8).



a. Gas-in b. Gas-out c. Heat preservation wall d. X-Y scanning system
e. Focus head (SLS or LOM) or Nozzle (FDM) f. Prototype g. Platform

Fig. 8 Multi-functional Chamber

However, since it may blows off irritant gas or needs inert gas during the sintering process, It's necessary to detect and control the atmosphere in order to avoid pollution or oxidation.

2.3. THE SLS MATERIALS

By now, since the limited chamber temperature and laser system, the available SLS materials are wax, polycarbonate and rosin etc.. The Fe-Cr-Ni alloy powders are also studied by using the M-

RPMS. However, the metal and ceramic powder can only be sintered with indirect process by M-RPMS. Parts created by using Fe-Cr-Ni powder have good corrosion and wear resistance and small degree of deformation. The SLS Fe-Cr-Ni powder is mainly composed of alloy powders and binders. The binders covered on the alloy powders are composed of special wax, polystyrene, polyethylene, engine oil, stearic acid etc.. The green parts were obtained by indirect SLS process by using M-RPMS.

3. CONCLUSIONS

The SLS software and hardware in M-RPMS are designed. The SLS software and hardware in M-RPMS have many characters different from those in mono-fuctional SLS equipment. The wax, polycarbonate, rosin and Fe-Cr-Ni alloy compound powder are available in indirect SLS process in M-RPMS.

4. REFERENCES

1. Yongnian Yan et al., Study of Multi-functional Rapid Prototype Manufacturing System (M-RPMS), Advanced Manufacturing Technology Conference, Beijing,1996. 9 . (颜永年、郭海滨，多功能快速成形制造系统的研究，先进制造技术学术会议，北京， 1996 年 9 月。)
2. Renji Zhang et al., Selective Laser Sintering Using Fe Based and Ni Based Alloy Powders, Advanced Manufacturing Technology Conference, Beijing,1996, 9.(张人佶等，铁基和镍基合金粉末的选择性激光烧结，先进制造技术学术会议，北京， 1996 年 9 月。)
3. Renji Zhang et al., Preparation of Fe-Ni-Cr Alloy Powders Used in SLS Process, 1996 March Meeting of the American Physical Society, St. Louis, USA. .18-22 March 1996.
4. Renji Zhang et al., Rapid Prototype with Selective Laser Sintering (SLS), Proc. VI Sino-Polish Conf. on CAD in Machinery, Warsaw, Poland, 17-18 July, 1996 .
5. Yongnian Yan et al., Study on Multi-functional Rapid Prototype Manufacturing System (M-RPMS) and Integration Degree. Proc. VI Sino-Polish Conf. on CAD in Machinery, Warsaw, Poland, 17-18 July, 1996.

Early Stage Rapid Feedback Design System Based on Rapid Prototype Technology

Qingping Lu, Li Bao, Jianguo Wang , Jun Zhu and Yongnian Yan
The Center of Laser Rapid Forming,
Department of Mechanical Engineering,
Tsinghua University.100084. Beijing. China

ABSTRACT

Early Stage Rapid Feedback Design (ESRFD) is a new design concept based on Concurrent Engineering. ESRFD emphasizes the understanding of the design iteration essence and the different design cycle within the design process, and makes the design process become controllable and predictable by making use of rapid physical and virtual prototyping technology to feedback the result of the design decision. The early stage rapid feedback design system configuration is based on rapid prototyping, virtual prototyping, reverse engineering and other design support technology. From the point of view of ESRFD, the design process dynamics and control features of the design process should be studied. The creative and innovation process in design makes the design dynamic up to higher level.

KEY WORDS

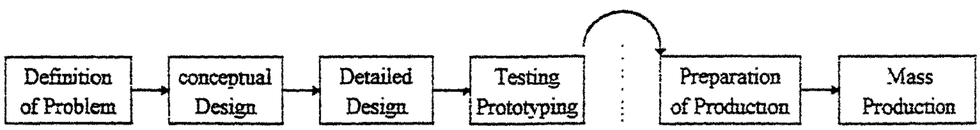
Early stage rapid feedback design, Concurrent engineering, Rapid prototyping, Design dynamics, Creative design

1. INTRODUCTION

When fabrication history is developing from mass production to post-industry production, design have changed greatly. The severe competition requires the enterprise to introduce new better and cheaper products in order to satisfy customer's requirements. A good design must consider the different contradicted needs. The conventional design method can not meet the requirements now. Many new design concepts have been proposed, from sequential design to concurrent engineering. Based on the newly developed technologies such as rapid prototyping and virtual prototyping, a new design concept of early stage rapid feedback design (ESRFD) has been discussed in this paper. Emphasizing on the dynamics of design process. ESRFD is more reasonable to describe the real design process[1].

2. Sequential Design Mode

Sequential design mode is a design mode formed in mass production stage. Fig. 1 shows its process:



Fag. 1 Sequential Design Scheme

Sequential design has the following features:

1) Generally up stage design sets a unique value as the input for down stage design. Design in every stage is oriented to its own stage. this kind of design mode cannot guarantee the overall quality of the design.

2) Design considers little about manufacturing. Completed design is sent to manufacturing department by " throw over the wall " method. This results are difficult in assembly and manufacturing.

3) The design process is uncontrollable: In large scale design activities, design cannot ascertain the result of his design decision. Many problems can only be found in prototype testing stage and manufacturing stage. If the design is not good, a large design iteration makes the design cycle become much longer.

Above problems can be analyzed by meaus of Design Structure Matrix(DSM). Design structure matrix, proposed by Steward[2],is a modeling tool to describe the relationships between different design tasks. If a design includes 5 design tasks represented by A, B, C, D, E, the correspondent DSM is in Fig. 2:

	A	B	C	D	E
A	×		×		
B	×	×	×		
C		×	×		×
D	×	×	×	×	
E		×		×	×

Fig. 2 example of DSM of a practical design process

	A	B	C	D	E
A	×				
B	×	×			
C	×	×	×		
D	×	×	×	×	
E	×	×	×	×	×

Fig. 3 DSM for sequential design process

In matrix, the marked elements within each row identify other tasks which must contribute information for proper completion of the design. The marked elements within each column identify the completion of the design and affect which other design tasks. For example, DSM in Fig. 2 show that task D has the results of task A, B, C as its input, and the output of task D is an input for task E.

Using DSM method for modeling sequential design process, the correspondent DSM must be a triangular matrix(Fig. 3).

However, the practical DSM of real design process should be in Fig. 2 instead of the ideal DSM in Fig. 3. This difference indicates that the sequential design method ignores some inherent relationships between some design tasks. So that some results by this ignorance can only be found in the very late stage of the design, that cause the unnecessary design iteration. This is basic reason of uncontrollability, unpredictability and large iteration in sequential design process.

3. CONCURRENT ENGINEERING

Concurrent engineering(CE) is a systematic design method for design production and has related processes integrally and simultaneously[3]. This method requires that new product designer considers all the elements in whole product life cycle comprehensively from beginning of the design, including quality,quantity cost, planning and so on. All the resource is available in the enterprise in order to satisfy the requirements of the customer ultimately. Its purpose is to gain the manufacturablility of the products, to reduce the time to market and enhance the competitiveness in market.

Concurrent engineering has following characteristics:

- Emphasizing team work and multi-disciplinary collaborative work:
- Considering whole product life cycle, emphasizing DFX such as Design For Assembly(DFA)and Design For Manufacturing(DFM).
- Increasing communication and cooperation by global network supported by product modeling and STEP protocol.

Fig. 4 shows the concepts of the concurrent engineering:

However,From the viewpoint of the DSMCE tries to parallel the design tasks, when DSM is a diagonal matrix all design tasks can be parallel completely (Fig. 5).This is impractical. How to parallel, to what degree the tasks can parallel, how to avoid the collide caused by over paralleling are still questions which needing further research. Presently CE emphasizes the utilizing framework and common blackboard technologies to share information among designer at right time.

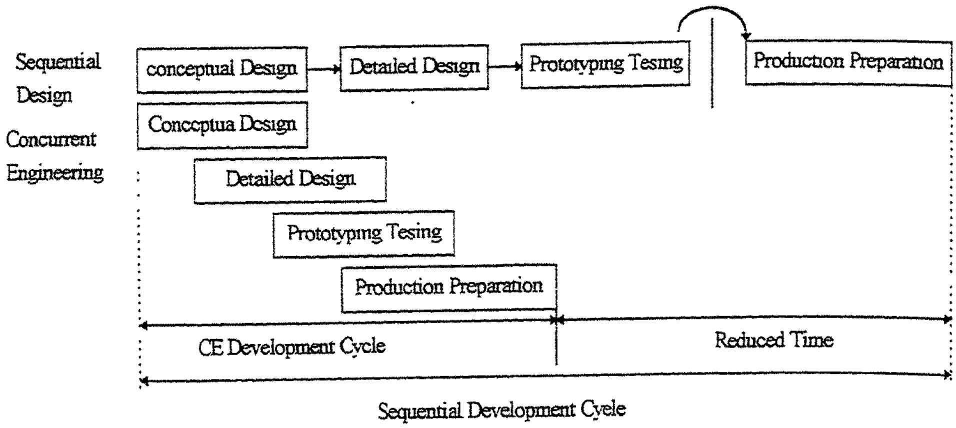


Fig. 4 Diagram Concept of Concurrent Engineering

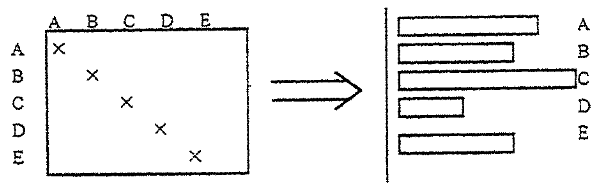


Fig. 5 ideal situation of concurrent engineering

4. RAPID PROTOTYPE TECHNOLOGY

Rapid prototyping (RP) technology uses the concept of isolating/piling to make parts[5]. Under the control of computer, it use the method of accurately piling material to make prototyping or parts according to the parts CAD model. It can automatically and rapidly realize the design through prototyping with certain structure and function. So that we can evaluate and modify the design of product. RP technology is the key technology of design of rapid feedback. The RP technology is one of important supporting technology of rapid feedback design.

5. Early Stage Rapid Feedback Design(ESRFD)

Design iteration is the main feature of the design process. Design iteration is the repeat of design tasks because of the appearance of new information. One reason is that error or unsatisfaction is found in down stage design tasks. Thus the up stage design tasks are repeated to solve this problem. Another reason is that the up stage design has some problems or the design is changed, then the down stage design is repeated. All these iterations form many cycle in design process.

In traditional design method. design iteration often appears in the late stage of the design. This causes many cycle in design.and affects the design cost and time greatly. The increasing cost and time in design , especially the increasing of development time, significantly impairs the competitiveness of the new product. According to a statistic material, if development time increases 6 month, the profits will lose 22 percent. In same case to keep development time unchanged but increase development cost by 60 percent, the profits only lose 10 percent. This shows that time and speed is the most important factor in new product development.

In this situation, the key is how to find problems in earlier design and to complete the design iteration as early as possible. The new concept is early stage rapid feedback design. That is stage the information get in the late stage of prototyping or manufacturing, can be found in the early stage of the design and feed them back to guide the design[4].

Design feedback is a process in which design concept is changed and design process is regulated under the limited use of time and resource. We need to simulate or test the design physically or virtually, in order to find the possible error or improper in design.

According to the early stage rapid feedback design concepts, a rapid feedback design system can be built with the support of various kinds of design technologies such as rapid prototyping, reverse engineering, virtual prototyping, DFX and QFD, as well as other design technologies. Its configuration is in following figures:

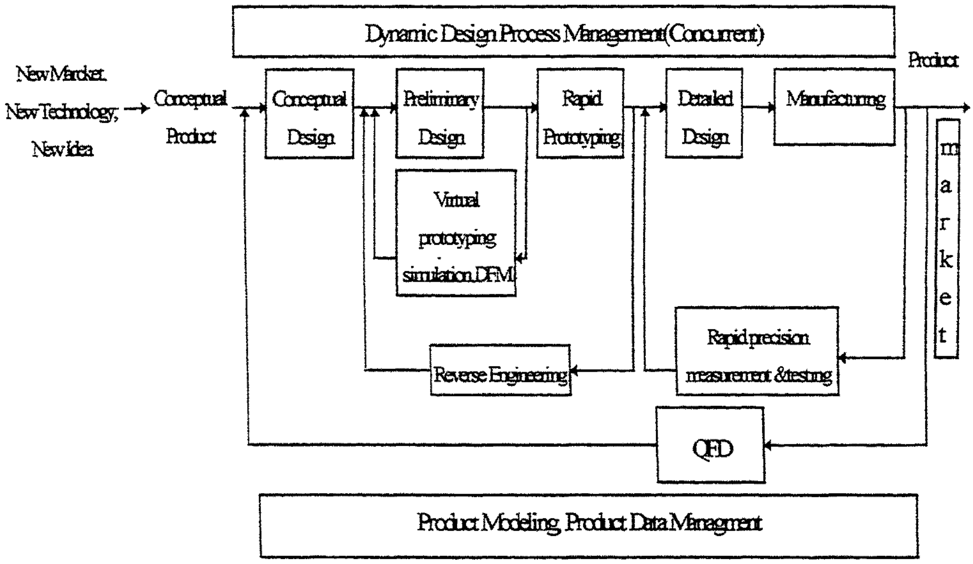


Fig. 6 Configuration of ESRFD system

Following is the explanation of the configuration:

- Although all design stage are depicted in a sequential format, the design method is not the traditional sequential design method. The different design stage can be performed concurrently.
- Fig. 6 gives some main feedback cycles. The design information and data generated by rapid feedback technologies can be feedback to the up stage design to verify or change the up stage design, which also can be fed forward to the down stage design to parallel the down stage design. However, these feedback cycles and fee-forwards cycles are not irregular. They should be performed according to the practical requirements and the law of the design.

The ESRFD is different from CE in following aspects:

1) ESRFD emphasizes the dynamics of design

Form the DSM in Fig. 2 you can see, practical design tasks often depend on each other. This kind of relationship indicates that all design tasks construct a dynamic design process. Both traditional sequential design method and concurrent engineering treat the real design process is a static process.ESRFD emphasize to understand this dynamic property of the design process, and wish to control whole design process according to this property.

2) ESRFD emphasizes the importance of prototyping in design

To make the whole design process be controllable, the result produced by a design decision must be known in advance. Concurrent Engineering emphasizes to consider whole product life cycle in product conceptual design stage. By making use of various kinds of DFX tools, CE design to "success one time".

This is a leap in design concept, but it is difficult to realize in real life. ESRFD emphasizes that prototyping, physical or virtual, can verify design concept and feedback design results rapidly, in order to coordinate and control the design process. It should be noticed that the design prototyping here is not a traditional prototyping, but the prototyping produced by prototyping technology and virtual prototyping technology.

3) ESRFD believes that the design iteration is unavoidable

Design iteration is the manufacture of the design process. In sequential design process large design iteration often causes significant increase in design cycle and cost. CE wishes to avoid the design iteration completely. This is in fact impossible. ESRFD thinks that the design iteration can be reduced but not be avoid completely. The design iteration feature should be studied deeply and try best to cut the product data generating cycle and design process management cycle.

6. A real ESRFD system

Fig. 7 shows a ESRFD system based on RP technology realized in our Center of Laser Rapid Forming. It is a part of the system is integrated by CAD, reverse engineering tools and rapid prototyping machines. This ESRFD system is developed and used in our laboratong.

- Reverse engineering tools:

One RE method 3D reconstruction from 2D photograph (self-developed TDS modeling system). It is suitable for shape design and is useful for car shape design and plastic moled making. We also apply this system in miniaturizing of the famous buildings that is meaningful for tourism souvenir manufacturing. Another method is non-contact grid coded illumination method. It is very useful for human face and body reconstruction. For a living person the laser scanning cannot betrayal because of the low speed and harm of laser to human. Grid coded illumination method is harmless for human and can capture the face data within several seconds with acceptable accuracy. Laser scanning method can be used for sculpture object and other hard object reconstruction.

- STL file processing:

The STL file generated from surface model often are invalid for RP. A method is to guarantee the validation of the STL file with triangulation. Another method is to verify and repair the STL file after its generation. Another problem is direct layer information generation without STL file. In European the CLI (Contour Layer Interface) file format has been developed for this purpose. It is high in precision and ready for RP fabrication. Nevertheless more work are needed to be done in areas such as volume generation, rendering and support generation. CLI is also very useful in medical appreciation of RP which is based on the CT or MRI layered information. A software system has been developed in our laboratory to cover whole process from 3D reconstruction from scanning data, STL file generation, STL file verification and repair support generation and slicing.

- Rapid prototyping machines:

The RP machines used in Fig. 7 are SSM-500 and MRPMS developed in our laboratory. This system has run in our laboratory. Some models such as car model and building model have been fabricated by this system, from photograph taking to surface model reconstruction, STL file generation, STL file processing, until model fabricanon.

Fig.7 shows the system structure of feedback design consisting CAD technology, RP technology and engineering. First we obtain geometry model of the design parts from CAD system. Then we change the geometry model into STL files. STL file is a popular-standard file used widely in RP technology field. CLI file is a developing layer format of files. We can get intersection information, which is changed into NC code controlling the machining of RP equipment of every layer, using STL file to slice the part to get information of each interface after parts are slicad. NC code is sent to RP equipment to form every kind of prototyping parts. We can use prototyping to evaluate design, revise the old design after feedback or reconstruct the 3D CAD drawing of the parts, after modifying the part directly then feedback to the geometry modeling. The whole process can also begin with the prototyping of course the prototyping mentioned here does not need to be machined by RP equipment. It may be product already existing. Then we use the reverse engineering method reconstruct the drawing and begin the new design cycle. (* TDS is a 3D reconstration method from 2D photograph)

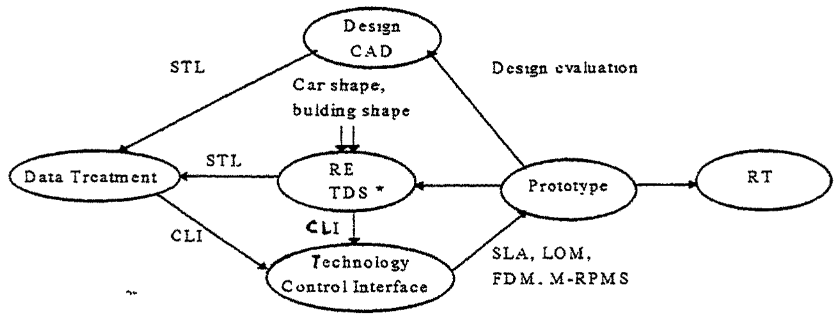


Fig. 7 ESRFD system based on RP technology

7. CONCLUSION

Early stage rapid feedback design concept and corespondent system configuration is discussed. From analysis you can see that design iteration is unavoidable. The rapid feedback in design makes the design iteration completed in the early stage of the design. Design process dynamics is the key to understanding the design creativeness. The utilization of the rapid prototyping and virtual prototyping technology are the guarantee of the rapid feedback.

8. REFERENCES

1. Bullinger Hans-Jorg, Wibler K. Worner K.RPD-a multi-disciplinary approach for design of innovative products, 3rd Conf. on Mfg. Tect. Dec.11-13, 1995, HK pp11-19
2. Steward D. V. The design structure system: a method for managing the design of complex system, IEEE Trans. on EM, Aug. 1981, pp71-74
3. Wang Jianguo, Yan Yongnian. New enabling technology for concurrent engineering: RP&M, Proc. ICIM ' 95, June 14-17, 1995 Wuhan pp350-353
4. Wang Jianguo. Yan Yongnian, Rapid feedback design system, 1st Chinese RP&M Conf. Tsinghua university, Beijing,Nov. 21-23, 1995
5. Kruth.J. P. " Material incress manufacturing by rapid praid prototyping techniques " CIRP Annals, Vol.40.No. 2.pp.603-14, 1991

AN ON-LINE CALIBRATION METHODOLOGY FOR ROBOT RELATIVE POSITIONING INACCURACY

Tien-Fu Lu and Grier C. I. Lin
Centre for Advanced Manufacturing Research
University of South Australia
tienfu@camr3.levels.unisa.edu.au
grier.lin@unisa.edu.au

ABSTRACT

For robots to successfully employ off-line simulation results, the physical settings of machines need to be calibrated accurately to match with the simulation conditions. Off-line calibration is normally time-consuming which needs extensive human efforts and measuring devices. To shorten the calibration time and reduce the need of human involvement, the development of an automatic on-line calibration methodology therefore became the goal of this research. In this research, a machine vision system and a 3D force/torque sensor were adopted and developed. A control strategy was designed to integrate and coordinate these components in order to enable a workcell robot to automatically calibrate for the relative positioning inaccuracy. The experimental results showed that the proposed methodology was highly successful and robust in calibrating the robot relative positioning inaccuracy without the need of off-line calibrations. Consequently, a significant amount of time in calibration can be saved.

KEYWORDS: Robot, On-line Calibration, Relative Positioning Inaccuracy

1. INTRODUCTION

Industrial robots are versatile, flexible and widely used in small batch manufacturing. In batch type production, workcell setting is always subjected to changes, such as the introduction of new machines[1]. To reduce the system setup time for changeover and increase the utilization of robots in dynamic changing environments, off-line programming and simulation tools are generally adopted to prepare robot programs and task descriptions at remote computers while robots are still in production on the proceeding job. For robots to successfully execute the programs generated off-line, the dimensions of real world robots and workcell components need to be modeled accurately. Otherwise, the differences will show up as relative position and orientation errors when the programs are executed on shopfloors.

To eliminate the relative position and orientation errors between robots and their environments caused by simulation inaccuracy[2], calibration is needed. For robot calibration, many calibration methods[3,4,5,6,7] have been developed, including the robot joint level calibration, kinematic calibration and dynamic calibration[8]. For relative position and orientation errors, two approaches[2] have been employed. The first is through the use of teach pendants to establish a number of reference or datum points. The second is through the use of workplace sensors to accurately measure the position and orientation of workplace elements and their dynamic characteristics. These measurements from sensors are then fed back to models for generating accurate programs, and/or off-line calibrating robots and their workplaces.

Although many robot calibration methods have been developed, most of them are dedicated to absolute positioning accuracy. These methods identify errors off-line through various calibration procedures. The common disadvantage of such calibration methods is the lack of flexibility in dynamic changing environments because of the need of complicated models, off-line and complex procedures, and accurate measurement systems.

In an integrated workcell, the relative positioning inaccuracy between robots and workcell facilities with which they operate is more critical for many applications than robot absolute positioning accuracy, such as assembly, loading and/or unloading tasks. However, very little consideration has been given to efficiently calibrating such integrated systems on-line. Therefore, the purpose of this research was to develop an automatic on-line calibration methodology for the relative positioning inaccuracy between a robot and the corresponding machines with which the robot works. To achieve the goal, a machine vision system and a 3D force/torque sensor were adopted and developed. A control strategy was designed to integrate and coordinate these components for calibrating the relative positioning inaccuracy

The experimental results showed that the proposed methodology was highly successful and robust in calibrating the robot relative positioning inaccuracy without the need of off-line calibrations. Consequently, a significant amount of time in calibration can be saved. Batch type production cells would then be more flexible and adaptable in accommodating to the workcell changes with less human involvement.

2. SYSTEM SETUP

The system configuration is shown in Figure 1 which consists of the following components:

- Six-axis industrial robot - A six-axis industrial robot, ABB IRB 2000 with Model M92 controller.
- ASM (Advanced Solid Modeller)[9] - ASM is a CAD package for 3D solid modeling with extended functions[10]. The ABB IRB 2000 robot workcell was modeled and the collision-free moving path for robot to safely move among workcell components were simulated, generated and stored in a common database.
- 3D Force/Torque sensor - The 3D Force/Torque sensor[11,12] was adopted to provide 3D Force/Torque information for robot control while the robot was interacting with the environment. This sensor had a passive compliance, (a rubber block), installed on the robot wrist.
- Sensor signal processing circuit - The signal from 3D Force/Torque sensor was amplified, converted into current for long distance transmission, transmitted and then converted back to voltage.
- Mutlifunction I/O board - The Lab-PC+ is a low-cost, multifunctional, analog, digital, and timing I/O board for IBM PC and compatibles. It was adopted for DC voltage measurement in this system.
- Vision System - ITEX 15040 series was adopted, including a monochrome camera with a monofocal lens, an image monitor, and the other modules[10]. The camera was attached to the robot final axis.
- 486 Personal Computer - This platform provided the environments for coding *C* programs, running Lab-PC+ board, fusing sensor information in order to make decision for adjusting robot movement.
- SUN Sparc 10 platform - This was used for running the vision system and the ASM package. It also stored the simulated robot moving path data.

3. ERROR ANALYSIS

To place the workpart precisely at the corner of a fixture located in a machine as illustrated in Figure 2, the major error sources may include:

- workpart feeding error
- robot absolute positioning error
- vision measurement error
- Force/Torque sensing error (3D Force/Torque sensor and rubber block)
- machine installation error (position and orientation mis-alignment)
- electronic or mechanical drift

- workpart manufacturing error
- simulation error

In this example, the vision system is developed to search, identify and guide the robot to pick up the desired workpart. The mathematical expressions of all the errors and their propagation effects is too complicated and difficult to be derived in dynamic changing environments. Besides, off-line calibration methods need accurate measurement systems and are very time-consuming for batch type production cells, which always need to be re-configured. Consequently, a different approach to calibration was developed for resolving the integrated relative position and orientation errors on-line in this research

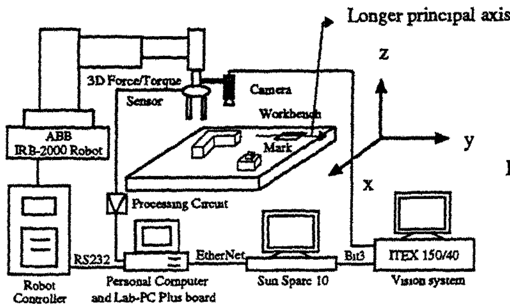


Figure 1 System Hardware Configuration

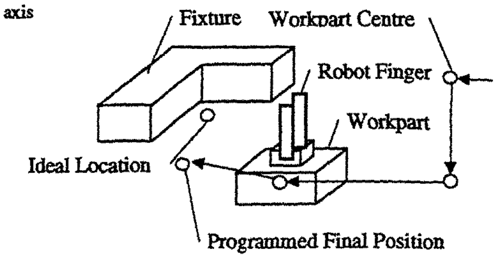


Figure 2: Place a Block at the Corner

4. POSITIONING ERROR CALIBRATION

To realise automatic on-line calibration, a robot control strategy was designed which coordinated the machine vision system, force/torque sensor, robot and the communication issues.

The vision system was developed for calibrating the relative orientation errors. It derived the machine orientation and computed the amount of relative orientation error for being compensated. The 3D force/torque sensor which includes an active sensing device and a rubber block was adopted for calibrating the relative position errors. The sensing device provided 3D force/torque information for measuring and monitoring the reaction forces between mechanical parts. Based on the information sensed, the robot detected the relative position errors while loading parts. The rubber block was used mainly for avoiding large impact forces and providing passive compliance in this study.

To calibrate the relative positioning inaccuracy, firstly the camera was moved to derive the machine orientation for calibrating the 'major orientation error'. Subsequently, the robot picked up the part and approached the destination by following the path simulated in the simulation environment to locate the part at the corner of the fixture. Due to the 'relative position and minor orientation error', the robot did not reach the corner accurately after the simulated path had been completely followed. Thus, the proposed strategies took control to automatically detect and compensate for the errors.

4.1 Major Orientation Error Calibration

The camera attached to the robot wrist was moved to the pre-defined position to find the designed rectangular marks on machines for deriving major orientation errors with respect to the global coordinate frame. The relevant data stored in the common database were then updated automatically according to the identified errors.

4.2 Relative Position and Minor Orientation Error Calibration

To place the part at the corner of the fixture, robot 3D motions are needed. However, it could be too complicated to model the 3D force-motion relations because of the coupling-effects and the integrated non-linearities from the robot, sensor device and rubber block Therefore, in the proposed guarded motion strategy, the 3D problem was decomposed into three 1D sub-problems.

<u>Motion Strategy</u>

In Figure 3 which is re-drawn from Figure 2 in C-space, P is an infinitesimally small part representing the centre of the part. The aim was to locate P at the corner G of the three faces X-Y, Y-Z, and X-Z. However, due to the location uncertainty, the initial location of the point P is within the ellipsoid S.

In the following descriptions, 'contact' stands for the plane-plane contact. The coordinate system adopted here is the absolute coordinate system shown in Figure 1. The calibration motion sequence was planned as:

- move part along the -Z direction until the part contacts the workbench (X-Y plane). The final Z coordinate is recorded and the part is shifted up 2 mm.
- move part along the -Y direction until the part contacts the fixture wall (X-Z plane). The Y coordinate is then recorded and the part is moved back 2 mm along the +Y direction.
- move part along the -X direction until the part contacts the fixture wall (Y-Z plane). The final X coordinate is recorded and the part is moved along the +X direction 2 mm.
- move to the recorded final location.

The part moving paths and the sequence are numbered in Figure 3. This motion strategy was coded in *C* programs. After this calibration, the robot moving path data will then be updated automatically

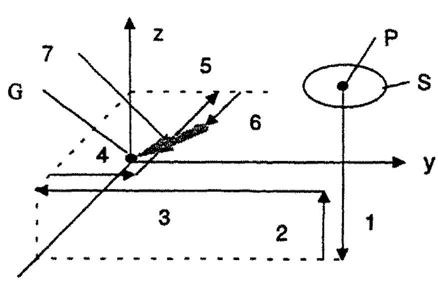


Figure 3: Guarded Fine Motion

<u>Termination Criteria</u>

If the following equation is met, the corresponding movement in the -Z, -Y or -X direction will be terminated.

$$|\Delta\xi_m(t)| = |\xi_m(t) - \xi_m(0)| \geq Threshold_m \quad (m=1,2,3,...,6) \qquad (1)$$

where the time-dependent sensor measurement of the *m* th channel was represented by $\xi_m(t)$. $\xi_m(0)$ was the zero-reference taken after the part was lifted by the robot. $Threshold_m$ was the threshold value for channel *m*. In this example, the 3D force/torque sensor readings collected through the process of contacting was employed to determine the threshold values which judge the occurrence of contact. The termination criteria were set as:

$$|\Delta\xi_{-z}(t)| \geq 1.5 \text{ (Volts)}$$
$$|\Delta\xi_{-x}(t)| \geq 1.2 \text{ (Volts)}$$
$$|\Delta\xi_{-y}(t)| \geq 1.2 \text{ (Volts)} \quad (2)$$

If the part does not contact with the environment, the voltage sensed equals to 0. Then, during every time increment, the robot will carry out a fixed amount of movement.

5. RESULTS AND DISCUSSION

This study attempted to develop an on-line calibration methodology for the relative position and orientation errors existing between robots and the facilities with which they operate. Some on-line running cases were conducted and the results are listed in Table 1 revealing that the values of *dz*, *dy* and *dx* are not fixed for various cases. This is mainly caused by the various initial positions of the workpart.

Apart from this finding, in every running case, there exists (dx, dy) = (-1, -1) which was designed to compensate for any possible under-movement below ±1 mm in X and/or Y directions. This ensures that the desired contact occurs and the workpart is located at the destination precisely. The small amount of over-movement, if it exists, is absorbed by the rubber block.

The limitation of implementing this methodology is that the allowable initial positioning inaccuracy along the X-Y plane is restricted by the dimensions of the corner block. If the contact between the workpart and the environment can occur in the presence of major location errors, the proposed strategies will function successfully. Similar constraints also apply to other environment conditions. The common limitations of this methodology are that the allowable orientation error can only be up to 89 degrees about the Z axis in the world coordinate frame, the capability of on-line calibrating the major orientation error is restricted by locating the specific marks on machines (without the mark, this vision based method will not be able to function), and various strategies need to be developed for different environmental conditions between the workpart and the destination.

Table 1: Guarded Motion for On-line Calibration

Case One			Case Two			Case Three			Case Four		
θ= 47.58°			θ= 92.34°			θ= -18.38°			θ= -61.74°		
dx	dy	dz	dx	dy	dz	dx	dy	dz	dx	dy	dz
		-6			-8			-6			-5
	-5			-4			-6			-6	
-3			-4			-7			-3		
-1	-1		-1	-1		-1	-1		-1	-1	

Note: The orientation, θ is the derived major orientation error in degrees
The dx, dy, and dz are in mm

6. CONCLUSION AND FUTURE TASKS

A different approach to resolve the propagated relative position and orientation errors in a robot-centered workcell is presented in this paper. Instead of modeling all possible errors, this research focused on developing an automatic method for on-line calibrating the relative position and orientation errors between robots and the equipment with which they operate. The developed methodology incorporates a vision system, a 3D Force/Torque sensor and control strategies with the robot. The automatic method is developed based on the external control loop closed around robot positioning control system. By implementing this methodology, complicated mathematical models and accurate off-line calibrations can be eliminated.

This methodology was experimentally proven to be robust and successful in calibrating the relative location errors of the workcells, subjected to dynamic changes, with fast speed and reliable accuracy. In conclusion, this methodology would enable the robot-centered batch type production cells to adapt to the dyanmic workcell changes in a more flexible and intelligent way with less human involvement than most of the current calibration methods.

6. ACKNOWLEDGEMENTS

The authors wish to thank the Australian Federal Department of Employment, Education and Training for supporting this project under its Targeted Institutional Links Program.

7. REFERENCES

[1] Groover M. P.: Automation, Production Systems, and Computer Integrated Manufacturing, Chapter 11, Prentice-Hall, 1987.

[2] Chan S.F., Weston R.H. and Case K.: "Robot Simulation and Off-line Programming". *Computer-Aided Engineering Journal*, Vol.5, Iss.4, pp.157-162, 1988.

[3] Veitschegger W. and Wu C. H.: "Robot Calibration and Compensation" *IEEE J. of Robotics and Automation*, Vol.4, No.6, 1988.

[4] Driels, M. R. And Swayze W. E.: "Automated Partial Pose Measurement System for Manipulator Calibration Experiments." *IEEE Transactions on Robotics and Automation*, Vol.10, No.4, pp.430-440, 1994.

[5] Takanashi, N.: 6 D.O.F. "Manipulators Absolute Positioning Accuracy Improvement Using a Neural-Network," IEEE, Int. Workshop on Intell. Robots and Systems, IROS, pp.635-640, 1990.

[6] Zhong X., Lewis J., and N-Nagy F. L.: "Inverse Robot Calibration Using Artificial Neural Networks." *Engng Applic. Artif. Intell.* Vol.9, No.1, pp.83-93, 1996.

[7] Watanabe, T.; Tokumaru, H, et al,: "The Calibration of Position and Orientation of Robot Manipulators Using a Neural Network." *JAPAN/USA Symposium on Flexible Automation*, Vol.1, ASME, pp.219-225, 1992.

[8] Qian G.Z. and Kazerounian K.: "Statistical Error Analysis and Calibration of Industrial Robots for Precision Manufacturing." *The Int. J. of Advanced Manufacturing Technology*, Vol.11, pp.300-308, 1996.

[9] Qikdraw systems Pty. Ltd.: User's Manual of ASM-Advanced Solid Modeling., 1993.

[10] Lu Tien-Fu and Lin Grier: CAD, "Vision and Sensor based Intelligent Robot Server." *J. of Computer Integrated Manufacturing Systems*, Vol.9, no.2, pp.91-100, 1996.

[11] Lin Grier C I. and Lu Tien-Fu: "Neural Network for Active Compliance Mechanism Calibration." *Proceeding of the Int. Conf. on Data and Knowledge Systems for Manufacturing and Engineering*, Hong Kong, 1994, pp.497-502, 1994.

[12] Lin Grier C. I., Lu Tien-Fu: "Neural Network for 3D Force/Torque Sensor Calibration and Robot Control." *Proceeding of the 3rd Int. Conf. on Computer Integrated Manuf. (ICCIM), Singapore*, Vol.2, pp.1471-1478, 1995.

Force Control Using Virtual Trajectory Generation

T. Fielding and J.K. Mills*
Laboratory for Non-Linear Systems Control
Department of Mechanical and Industrial Engineering
University of Toronto
Toronto, Ontario
Canada M5S 3B8

*email: mills@mie.utoronto.ca

Abstract

In this paper, we introduce a method to permit an existing PID controlled robot to be retrofitted with force control capability. A significant aspect of the proposed method is that the existing position servo loop can be left unaltered and the force controller can be implemented as a separate, external unit. Required in this proposed retrofit is the addition to the robot of a wrist mounted force sensor, and a data path for passing trajectory updates from the external unit to the robot. The algorithm presented here provides the user with a simple method of retrofitting a position controlled robot to enable it to control contact force and position when in contact with a compliant environment. A recursive least squares parameter estimation algorithm is used to estimate the environment parameters on-line in order to improve performance. This trajectory modification force control (TMFC) algorithm is implemented on the CRS Robotics Corporation A460 six DOF articulated industrial robot.

Keywords: robot force control, external force control loop

1. Introduction

Many different methods to control contact force have been proposed in the literature. The output of most of these control laws is a vector of desired torques that are sent to the robot joints. In order for a typical industrial robot to be controlled in this manner, significant modifications must be made to the existing controller structure in order for it to follow a commanded torque trajectory, rather than a position trajectory. Often, proprietary information is required in order to perform these alterations. If this information is unavailable, the entire controller would need to be replaced at considerable development cost. In addition, some method of torque feedback is often required in order to ensure that the commanded torque is actually followed. This can involve significant, expensive, or even impossible alterations to the existing robot hardware in order to generate a torque feedback signal.

Instead of attempting to implement a torque-based force control algorithm, adding the force control law as an external loop and retaining the inner position loop (the commercial controller) would be more a desirable approach. The overall control structure is shown in Figure 1.1. This configuration requires minimal alterations to the original robot and controller and therefore represents a simple method of adding force control capability to an industrial robot. This can result in an enormous cost savings when upgrading a robot from a purely position controlled device to one that can regulate both position and contact force.

This paper outlines a method of modifying the desired trajectory input to the inner position control loop. This inner loop represents the original position servo loop and is assumed to be a joint space PID loop. The outer force loopallows for the control of contact force. Physically, this outer loop would consist of a force/moment sensor (typically wrist mounted), and a controller which would feed the modifications into the inner loop. The trajectory modifications are generated by solving two sets of initial value problems given a desired force trajectory. Since the force controller is to reside external to the original controller, it will be limited in terms of update rate by the frequency of communication between the force and position controllers.

Uncertainties in the environment parameters are accounted for by a recursive least squares estimation algorithm. The resulting control law is then implemented on a CRS Robotics Corporation A460 articulated light industrial robot. The inner position loop is a joint level PID control law that runs at 1000 Hz. The outer force loop runs at 250 Hz.

2. Previous Work

Force control algorithms are generally broken into two basic categories. The first category is if the control law is executed in joint space or Cartesian space. The second category is based upon the nature of the signals used for servoing. The different types of signals that can be used are actuator torque, positions or velocities. It is concluded that the best method to use with respect to suppression of internal disturbances is position servoing. Internal disturbances are defined as those internal to the robot, such as gear cogging and friction.

Often control laws with and inner position and outer force loop use a hybrid control strategies to implement a trajectory modification scheme as in [2] and [3]. This allows the position and force controlled directions to be defined naturally in task space. One method of force control that seems to be growing in popularity is adding a trajectory modification algorithm to an impedance controller in order to achieve force tracking, as in [10] and [8]. In [1], a form of impedance control is used when implementing a trajectory modification scheme. An inner impedance control loop is usually employed to cancel the non-linear dynamics.

In [4], an adaptive control law was used in an outer loop to modify the position set points for force control. A big advantage of this controller is that no knowledge of the dynamic model of the robot is required. However, the inner position loop is a PD or PD plus gravity compensation in task space, which is not typical of industrial controllers. Ferritti et. al. [5] employ a inner loop, joint space PID control loop.

3. Control Law Formulation

In this section the TMFC algorithm is developed. The algorithm solves two dynamic equations in order to calculate the modification of the robot trajectory in order to generate the desired contact force. The robot dynamic equation for an n DOF rigid link and transmission manipulator, as derived from the Lagrangian equation, is given as:

$$\Phi(q)\ddot{q} + H(\dot{q},q) = u + J(q)^T F \qquad (3.1)$$

where $\Phi(q)$ is the nxn inertia matrix, $H(\dot{q},q) \in R^{nx1}$ contains the rate and configuration dependent terms, such as Coriolis and centripetal terms. This vector also contains the friction and gravity forces. $u \in R^{nx1}$ is the actuator applied force, and $F \in R^{nx1}$ is the contact force, expressed in task coordinates. $J \in R^{nxn}$ is the Jacobian of the transformation from joint space to task space. The environment can be modeled with a mass, spring, and damper so that the contact force between the robot and the environment is defined as

$$F = \begin{Bmatrix} F_c \\ 0 \end{Bmatrix} = \begin{Bmatrix} M_e\ddot{x}_c + B_e\dot{x}_c + K_e(x_c - x_o) \\ 0 \end{Bmatrix} \qquad (3.2)$$

where $M_e \in R^{mxm}$ is the mass, $B_e \in R^{mxm}$ is the damping and $K_e \in R^{mxm}$ is the stiffness of the contacted environment. m is the number of degrees of freedom that are in contact with the environment. $x \in R^{mx1}$ represents the position of the end effector in the contact DOFs, expressed in task co-ordinates. $x_o \in R^{mx1}$ is the location of the undeflected environment and assumed stationary $(\ddot{x}_o = \dot{x}_o = 0)$. $F_c \in R^{mx1}$, is the contact force between the robot manipulator and the environment.

The concept behind trajectory modification is that in order to increase the contact force, the end effector position is commanded such that the environment deformation increases, or the manipulator "pushes into" the environment. Conversely, the contact force is decreased by commanding the robot to "pull back".

The trajectory modification force controller (TMFC) presented here creates a new trajectory based upon the real time solution of two sets of differential equations. Equation (3.2) can be solved as an initial value problem to find the desired position of the actual contact point given a desired force. If the environment parameters and position of the environment are known, then the desired contact position could be generated off-line. However, this approach requires that all the parameters be known precisely, as well as the undeflected position of the environment. In order to address the problem of uncertain environment parameters, a recursive least squares (RLS) algorithm is used to estimate these parameters on-line based on real time feedback. In order to eliminate the problem of the uncertainty in x_o, a differential type of control law is used. This calculates a desired deviation from the current state, based upon the force error, in order to achieve the desired state. First, let the desired force be given by:

$$\begin{Bmatrix} F^d \\ 0 \end{Bmatrix} = \begin{Bmatrix} \hat{M}_e \ddot{x}_c + \hat{B}_e \dot{x}_c + \hat{K}_e (x_c - x_o) \\ 0 \end{Bmatrix} \quad (3.3)$$

where the ^ indicates an estimate of the actual value. The value of x_c is the desired position of the actual contact point. If the end effector physically follows x_c, then the correct contact force will result. If we assume that the estimates of the environment parameters are approximately equal to the actual values, the force error is then given by:

$$\begin{Bmatrix} F^d - F_c \\ 0 \end{Bmatrix} \cong \begin{Bmatrix} \hat{M}_e (\ddot{x}_c - x) + \hat{B}_e (\dot{x}_c - x) + \hat{K}_e (x_c - x) \\ 0 \end{Bmatrix} \quad (3.4)$$

Instead of solving (3.2) in order to find the total deflection of the environment from a reference point (x_o), equation (3.4) is solved in order to calculate a deviation from the current state. Once the desired position of the robot is determined in the *m* force controlled directions, this vector is augmented with the desired task space trajectories in the (*n-m*) position controlled directions. $x_{c*} = \begin{Bmatrix} x_c \\ x_u \end{Bmatrix}$ where $x_u \in R^{(n-m)x1}$ is the desired robot position in the non-contact directions. x_{c*} therefore completely represents the desired robot position during contact with the environment. This position is then transformed to joint space using the inverse kinematic relations: $q^c = L(x_{c*})$ where L( ) is the inverse kinematic function. The inner loop control law is assumed to be a PID controller. The control vector u then has the form:

$$u = G_p(q^d - q) + G_d(\dot{q}^d - \dot{q}) + G_i \int_0^t (q^d - q) d\tau \quad (3.5)$$

Substituting this into equation (3.1) yields:

$$\Phi(q)\ddot{q} + H(\dot{q}, q) = G_p(q^d - q) + G_d(\dot{q}^d - \dot{q}) + G_i \int_0^t (q^d - q) d\tau + J^T F \quad (3.6)$$

Equation (3.6) can be used to account for the robot dynamics and controller compliance by solving for q^d. Recalling that the required position of the robot end effector is given by q^c, then we replace q in (3.6) with q^c in order to generate the desired set points to be set to the inner position loop.

$$\Phi(q^c)\ddot{q}^c + H(\dot{q}^c, q^c) = G_p(q^d - q^c) + G_d(\dot{q}^d - \dot{q}^c) + G_i \int_0^t (q^d - q^c) d\tau + J^T F \quad (3.7)$$

q^d represents a *virtual trajectory*. This is not actually followed by the robot, but when this is input to the inner loop PID controller, the result will be that the robot will actually follow q^c, and thus generate the desired contact force (see Figure 3.1). Equation (3.7) can then be solved, as an initial value problem, in order to get the desired input trajectory.

4. Experiment

The TMFC algorithm was implemented on a CRS Robotics Corporation A460 6 DOF industrial robot.The first 3 joints are driven by harmonic drives with a 100:1 gear ratio. The CRS PID loop runs at a frequency of 1000 Hz. The outer force control loop modifies the inner loop set points at a rate of 250 Hz. A linear translational device under closed-loop control simulates a mechanical impedance (Variable Impedance Machine - VIM) with adjustable dynamic parameters to model equation (3.4).

The first series of experiments were conducted with the robot moving in free space. The complete path of the robot end effector is defined by a straight line interpolation algorithm, executed in cartesian space. Figure 4.2 shows the force tracking error during the motion of the robot in free space while solely under PID position control.

Simulations performed to test this algorithm were quite successful. However, when implemented on the CRS robot, it was found that if the inner position loop does not behave as modelled in equation (3.7) then errors will result in the force tracking. In free space, the errors of the PID position loop are quite large (see Figure 4.2). Solving equation (3.7) reduces some of these errors. In particular, the z space position errors are reduced from as much as 4.0mm (Figure 4.2) to less than 2.5mm (Figure 4.3). As the robot moves across the VIM with a constant commanded force, a steady state approximation can be used to estimate the error in the contact force due to a positioning error in the inner loop. This error is given by

$$\begin{Bmatrix} F^d - F \\ 0 \end{Bmatrix} = \begin{Bmatrix} K_e(x_c - x) \\ 0 \end{Bmatrix} \qquad (4.1)$$

In order to compensate for the inner loop positioning errors, a term based upon the position error was added in order to adjust the virtual trajectory. Not only did this improve the force tracking, but the position tracking as well. The second initial value problem now becomes

$$\Phi(q^c)\ddot{q}^c + H(\dot{q}^c, q^c) = G_p(q^d - q^c) + G_d(\dot{q}^d - q) + G_i\int_0^t (q^d - q^c)d\tau + J^T F + G_c\int_0^t (q^c - q)d\tau \quad (4.2)$$

Figure 4.4 shows the force tracking error with same inner PID loop, but with the TMFC algorithm modifying the commanded trajectory in an outer loop. The result of the addition of the trajectory modification loop is to reduce the position tracking errors from as much as 4 mm to less than 0.5 mm. For the contact experiments, the robot initially moves in free space, modifying its trajectory based upon equation (4.2) in order to improve position tracking. Once contact is detected (contact force exceeding a threshold value of 3N) the trajectory is modified based upon equation (3.4) as well. The robot then moves across the surface of the contact while following a desired force trajectory normal to the contact surface. The environment was set to a stiffness of 10000 N/m.

In order to judge the performance of the TMFC algorithm, a PI type of force control strategy was investigated. Modifications to the position trajectory were based upon the following error equations:
PI Force Control

$$\Delta x = G_p(F^d - F) + G_i\int_0^t (F^d - F)d\tau \qquad (4.3)$$

This type of control is similar to one strategy proposed in [3]. The PI control law was implemented in a hybrid type of control structure where corrections are made in task space and then converted to joint position set points via the inverse kinematic relations for the CRS A460 robot. The feedback gains were tuned through experimental trials. The results of this test are shown in Figure 4.5. The large spikes occur due to sharp changes in the desired position trajectory. Figure 4.6 shows the results of the TMFC experiment. The force response is much improved, and the sharp force spikes are noticeably absent. Figure 4.7 shows the algorithm's ability to track a varying force trajectory.

5. Conclusion

An algorithm has been presented which resides in an outer loop so that force control capability can be added to a position controlled robot as an external module. This algorithm does not make use of a common assumption that the inner position loop is perfect and therefore can be ignored. Instead, the position tracking in free space is improved by compensating for positioning errors based on the solution of the robot dynamic equations.

The reasons for the improvements in both the force and position tracking lie in the solution of the two sets of dynamic equations. These equations are solved on-line as initial value problems in order to generate the *virtual trajectory*. One set of dynamic equations is used to account for the contact dynamics.

From this set of equations, the desired position of the contact point is determined. The second set of dynamic equations account for the robot inertia, friction effects, gravity, centripetal and Coriolis effects, resulting in the creation of the *virtual trajectory*. By solving these equations on-line, not only is it possible to follow a desired force trajectory, but to improve the position tracking as well.

6.References

[1] Colbaugh, R. and H. Seraji, "Force Tracking in Impedance Control", Proc. 1993 IEEE Int. Conf. Robotics and Automation, 1993, pp. 499-506.

[2] De Shutter, J., "A Study of Active Compliant Motion Control Methods for Rigid Manipulators Based on a Generic Scheme", Proc. IEEE Int. Conf. on Robotics and Automation, 1987, pp. 1060-1065

[3] De Shutter, J., and H. Van Brussel, "Compliant Robot Motion II: A Control Approach Based on External Control Loops", International Journal of Robotics Research, Vol. 7, No. 4, Aug. 1988, pp. 18-33.

[4] Engelmann, A. and R. Colbaugh, "Adaptive Compliant Control of Manipulators: Theory and Experiments", Proc. Int. IEEE Conf. on Robotics and Automation, 1994, pp. 2719-2726.

[5] Ferretti, G., Maffezzoni, G. Magnani, and P. Rocco, "Decoupling Force and Motion Control in Industrial Robots", Control Engineering Practice, Vol. 1, No. 6, Dec 1993, pp. 1019-1027.

[6] Goldenberg, A., and G. Liu, "Robust Hybrid Impedance of Robotic Manipulators", Proc. IEEE Int. Conf. on Robotics and Automation, 1991, pp. 287-292.

[7] Khatib, O., "Dynamic Control of Manipulators in Operational Space", 6th World Congress on Theory of Machines and Mechanisms CISM-IFFTOM, New Delhi, India., 1983, pp. 1128- 1131.

[8] Lasky, T. A. and T. C. Hsia, "On Force-Tracking Impedance Control of Robotic Manipulators", Proc. 1991 IEEE Int. Conf. on Robotics and Automation, April, 1991, pp. 274-280.

[9] McClamroch, N. and D. Wang, "Feedback Stabilization and Tracking of Constrained Robots", IEEE Trans. on Automatic Control, vol 33, no 5, May 1988

[10] Mills, J. K., "Simultaneous Control of Robotic Manipulator Impedance and Generalized Force and Position", Mechanism and Machine Theory, Vol. 31, No. 8, pp. 1069-1080, 1996.

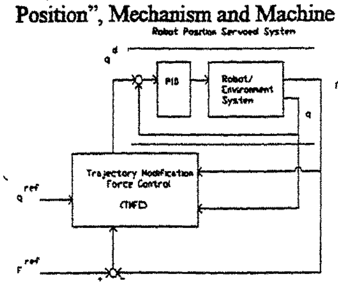


Figure 1.1: Configuration of Control Structure

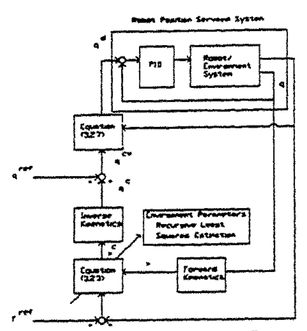


Figure 3.1: TMFC Control Structure

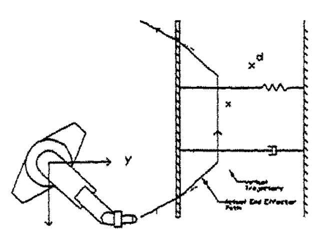


Figure 3.1: Virtual Trajectory Path of the Robot During Contact

Figure 4.1: Experimental Configuration

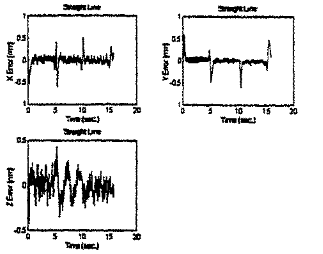


Figure 4.4: TMFC Algorithm Cartesian Space Position Errors.

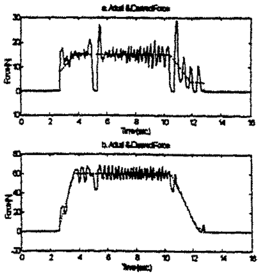


Figure 4.5: PI Force Control

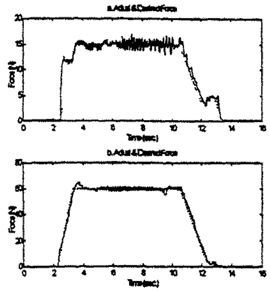


Figure 4.6: TMFC Force Control

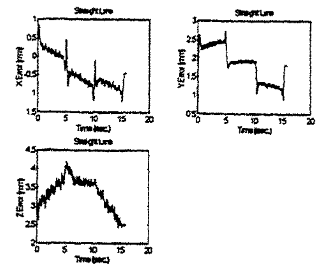


Figure 4.2: PID Position Servo Loop Cartesian Space Errors.

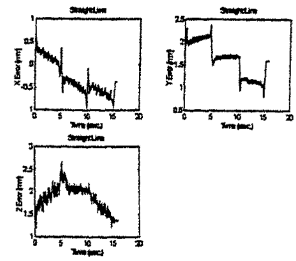


Figure 4.3: Positioning Errors with Model Based Compensation.

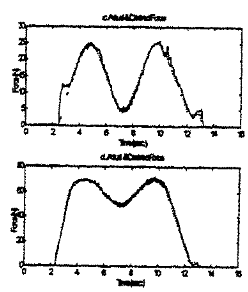


Figure 4.7: TMFC Force Control. Sinusoidal Desired Force

PAYLOAD ALIGNMENT CONTROL USING TASK MOMENT FEEDBACK FOR MULTI-ROBOT ASSEMBLY OPERATION

M. Laliberté and J.K. Mills
Department of Mechanical and Industrial Engineering
The University of Toronto
Toronto

ABSTRACT

In the problem of fixtureless assembly, two robotic manipulators are required to bring two flexible sheet metal parts in contact so that they can be bonded together. The first step of the assembly process requires each of the manipulators to pick up the parts, thus there is the possibility the manipulators will grasp the parts improperly. The present paper focuses on improper grasps that generates an angular misalignment between the contact surfaces of the two sheet metal part as they are brought in contact. The angular misalignment generates bending moments in the parts as well as poor contact between the contact surfaces of the parts. A control strategy is therefore proposed to eliminate the moments generated by the part's angular misalignment, and in doing so realign the parts for proper contact. The proposed control law is implemented on a multi-robot system where two robots bring sheet metal parts in contact for mating with one of the sheet metal parts improperly grasped. The controller successfully eliminates the moments and realign the parts.

KEYWORDS

Multi-robot control, hybrid control, moment feedback control, alignment, assembly.

1.INTRODUCTION

Typically, in manufacturing industries, hardware fixtures are widely used in assembly processes to position and orient parts, such that they are held in place in contact with one another so that the parts can be bonded. In situations where the number of parts to be assembled is large, the cost of the fixtures represents a significant portion of the total cost of the assembly operation. Due to the high annual cost of fixtures required in the assembly of finished products, it is desirable to find alternatives to assembly with hardware fixtures. Fixtureless assembly is proposed as such an alternative. It is defined as the assembly of parts without hardware fixtures. To implement fixtureless assembly on a production line, robotic manipulators are proposed to be used to perform the tasks previously fulfilled by fixtures. First, two robots will each grasp a single part. The robots will then bring the parts together such that they can be mated. Once in contact, the parts are forced together under prescribed forces such that permanent bonding can occur.

In a fixtureless assembly operation, an improper grasp of the parts results in alignment position and orientation error when the parts are brought into contact by the two robots. This implies that the parts will not be aligned properly for mating. For example, grasping the part with an orientation error generates a misalignment such that the surface of contact of the two sheet metal parts are not parallel when brought into contact (see Figure 1). Parts that are not parallel can result in poor contact between the parts and render the bonding operation unsuccessful. Unwanted bending moments can also be generated in the parts as they are brought into contact, as shown in Figure 1.

This paper studies the specific problem of controlling the orientation of parts each held by a robot so that the parts are correctly aligned for a subsequent bonding operation. The focus is on the mating of thin walled parts. Dynamic models of the robots that grasp these payloads are given where the unwanted forces and moments generated during part mating, due to an improper grasp of one payload, are modeled. With suitable dynamic model developed, a feedback control approach to eliminate part angular misalignment is proposed. With the proposed feedback control solution, the unwanted moment generated due to the angular misalignment is eliminated, and the part is brought into the correct orientation for mating. A test of the controller is conducted on commercial CRS Robotics Corporation six degree of freedom robots.

1.1 Literature Review

As stated, the work presented addresses the mating of thin walled parts, hence our review includes a short discussion of work relevant to flexible thin walled payloads. Modeling of these payloads is similar to the techniques used to model flexible link robots. There exists substantial literature available concerning the dynamic modeling and control of flexible manipulators [1,7], as well as on the control of multi-robot systems [8,10]. Also, some research has been performed regarding the assembly of parts by a multi-robot system [6]. The grasping and assembly of flexible payloads remains a fairly new research topic however, and the literature available in that area is not substantial. In [5], the dynamic modeling of a six degree of freedom robot manipulator holding a flexible sheet metal payload is addressed. The problem of robotic fixtureless assembly tasks is discussed in [4] where the control of two rigid manipulators holding two flexible thin walled payloads in contact is studied. In [6], a controller for two rigid manipulators, each holding a flexible payload, for the mating and control of generalized forces and positions of the robots is developed. The fixtureless assembly problem is also looked upon by [9]. They specifically examine the control of vibration when handling sheet metal parts as well as the control of contact alignment between the parts during assembly.

2. DYNAMIC MODEL

In the multi-robot system studied in this paper, the two sheet metal parts are not yet bonded to one another. Thus, the system does not form a closed kinematic chain and no rigid kinematic constraints exist between the two robots. Therefore, the two robot-payload systems are considered separately such that the contact forces and moments generated in the mating operation are treated as external forces and moments to each single robot-payload system. When developing the dynamic model, it is desired to separate the dynamics associated with the part orientation error, generated by an improper grasp, from the rest of the dynamics of the single robot-payload system. The procedure to obtain such a separation of the dynamics first requires to express the dynamics of the single robot-payload system in terms of a set of task coordinates which are based on the actual kinematics of the system, i.e. the kinematics that depict the robot holding the sheet metal part with an improper grasp. Then, by expressing these task coordinates as a function of both the kinematics of a system with perfect grasp and of the kinematics associated with the orientation error, the dynamic terms associated with the part orientation error can be expressed explicitly, and isolated form the rest of the model. The development of such a transformation is detailed in [2]. The resulting dynamic model for a n degree of freedom manipulator holding a flexible payload is given as,

$$\begin{aligned}\tilde{\mathbf{M}}_o(\mathbf{x}_o,\mathbf{q})\binom{\mathbf{x}_o}{\mathbf{q}} + \tilde{\mathbf{L}}_o(\mathbf{x}_o,\dot{\mathbf{x}}_o,\mathbf{q},\dot{\mathbf{q}}) + \binom{\mathbf{0}}{\mathbf{Bq}} + \binom{\mathbf{0}}{\mathbf{Kq}} \\ = (\mathbf{T}_o^T + \tilde{\mathbf{T}}_\Delta^T)\binom{\mathcal{T}}{\mathbf{0}} - \mathbf{T}_o^T\mathbf{J}_o^T(\mathbf{x}_o,\mathbf{q})\mathbf{F} + \mathbf{F}_\Delta,\end{aligned} \tag{1}$$

where $\tilde{\mathbf{M}}_o(\mathbf{x}_o,\mathbf{q}) \equiv \mathbf{T}_o^T\mathbf{M}(\mathbf{x}_o,\mathbf{q})\mathbf{T}_o \in \Re^{(m+N)\times(m+N)}$, $\tilde{\mathbf{L}}_o(\mathbf{x}_o,\dot{\mathbf{x}}_o,\mathbf{q},\dot{\mathbf{q}}) \equiv \mathbf{T}_o^T\mathbf{M}(\mathbf{x}_o,\mathbf{q})\dot{\mathbf{T}}_o^T\binom{\mathbf{x}_o}{\dot{\mathbf{q}}} + \mathbf{T}_o^T\mathbf{L}_o(\mathbf{x}_o,\dot{\mathbf{x}}_o,\mathbf{q},\dot{\mathbf{q}}) \in \Re^{m+N}$, with $\mathbf{M}(\mathbf{x}_o,\mathbf{q}) \in \Re^{(n+N)\times(n+N)}$ representing the inertia matrix of the robot-payload system, $\mathbf{L}_o(\mathbf{x}_o,\dot{\mathbf{x}}_o,\mathbf{q},\dot{\mathbf{q}}) \in \Re^{n+N}$ is the vector that results by combining the centrifugal terms, Coriolis terms, and the terms due to the interaction between $\dot{\mathbf{x}}_o$ and $\dot{\mathbf{q}}$, $\mathbf{x}_o \in \Re^m$ are the generalized task coordinates chosen with respect to the geometry of the task and which are based on the kinematics of a system with perfect grasp, $\mathbf{q} \in \Re^N$ is the vector representing the flexible states of the payload, $\mathbf{T}_o \in \Re^{(n+N)\times(m+N)}$ is the Jacobian as defined in [3,6] which relates $\mathbf{x}_o$ to the rigid joint coordinates $\boldsymbol{\theta}$ of the robot and the flexible states $\mathbf{q}$, $\tilde{\mathbf{T}}_\Delta \in \Re^{(n+N)\times(m+N)}$ is the Jacobian as defined in [2] that relates the dynamics associated with the orientation error to $\boldsymbol{\theta}$ and $\mathbf{q}$, $\mathbf{B} \in \Re^{N\times N}$ is the payload damping matrix, $\mathbf{K} \in \Re^{N\times N}$ is the payload stiffness matrix, $\tau \in \Re^n$ is the generalized input force of the rigid robot, $\mathbf{J}_o(\mathbf{x}_o, \mathbf{q}) \in \Re^{m\times(n+N)}$ is the Jacobian mapping the task space to the generalized coordinate space defined by $\boldsymbol{\theta}$ and $\mathbf{q}$, $\mathbf{F} \in \Re^m$ is the reordered vector of external forces and moments applied on the sheet metal part as it comes in contact with another sheet metal part, $\mathbf{F}_\Delta \in \Re^{m+N}$ represents the dynamics of the robot-payload system that are associated with the part orientation error, and N is the total number of flexible degrees of freedom of the payload.

3. CONTROL

Three specific control objectives are required for the controller: 1) regulate the contact force between the sheet metal parts, 2) regulate the position of the end-effectors of the robots as they force the sheet metal parts against one another, and 3) eliminate the moments that are caused by the angular misalignment that exists between the two sheet metal parts and thereby align the sheet metal parts properly. It is proposed to rewrite the dynamics of the system given by eq.(1) by making use of the geometry of the task such that the dynamics associated with the forces of contact are separated from the rest of the dynamics. This then allows to easily design a model based controller that uses force feedback and position feedback to fulfill the first two control objectives. It is also proposed to add a feedback correction to the desired orientation of the sheet metal part to eliminate the moments generated by the presence of an angular misalignment between the sheet metal parts. The following details how the separation of the dynamics is achieved and the feedback correction is designed. Note that the control law is designed based on the formulation of a rigid feedback, as in [4,6].

At this stage, a transformation is applied to the dynamic model that separates the dynamics associated with the contact forces from the dynamics associated with the position, as is done in [3]. Equation (1) is split into the following equations,

$$\begin{aligned}\mathbf{E}_1\tilde{\mathbf{M}}_o(\mathbf{x}_o,\mathbf{q})\mathbf{E}_1^T\ddot{\mathbf{x}}_{o_1} + \mathbf{E}_1\tilde{\mathbf{M}}_o(\mathbf{x}_o,\mathbf{q})\mathbf{E}_2^T\ddot{\mathbf{x}}_{o_2} + \mathbf{E}_1\tilde{\mathbf{M}}_o(\mathbf{x}_o,\mathbf{q})\mathbf{E}_3^T\ddot{\mathbf{q}} + \mathbf{E}_1\tilde{\mathbf{L}}_o(\mathbf{x}_o,\dot{\mathbf{x}}_o,\mathbf{q},\dot{\mathbf{q}})\\ = \mathbf{E}_1\binom{\tilde{\tau}}{0} + \mathbf{E}_1\tilde{\mathbf{T}}_\Delta^T(\mathbf{x}_o,\mathbf{x}_\Delta,0)\binom{\tau}{0} - \mathbf{F}_1 + \mathbf{F}_{\Delta_1},\end{aligned} \tag{2}$$

$$\begin{aligned}\mathbf{E}_2\tilde{\mathbf{M}}_o(\mathbf{x}_o,\mathbf{q})\mathbf{E}_1^T\ddot{\mathbf{x}}_{o_1} + \mathbf{E}_2\tilde{\mathbf{M}}_o(\mathbf{x}_o,\mathbf{q})\mathbf{E}_2^T\ddot{\mathbf{x}}_{o_2} + \mathbf{E}_2\tilde{\mathbf{M}}_o(\mathbf{x}_o,\mathbf{q})\mathbf{E}_3^T\ddot{\mathbf{q}} + \mathbf{E}_2\tilde{\mathbf{L}}_o(\mathbf{x}_o,\dot{\mathbf{x}}_o,\mathbf{q},\dot{\mathbf{q}})\\ = \mathbf{E}_2\binom{\tilde{\tau}}{0} + \mathbf{E}_2\tilde{\mathbf{T}}_\Delta^T(\mathbf{x}_o,\mathbf{x}_\Delta,0)\binom{\tau}{0} - \mathbf{F}_2 + \mathbf{F}_{\Delta_2},\end{aligned} \tag{3}$$

$$\begin{aligned}\mathbf{E}_3\tilde{\mathbf{M}}_o(\mathbf{x}_o,\mathbf{q})\mathbf{E}_1^T\ddot{\mathbf{x}}_{o_1} + \mathbf{E}_3\tilde{\mathbf{M}}_o(\mathbf{x}_o,\mathbf{q})\mathbf{E}_2^T\ddot{\mathbf{x}}_{o_2} + \mathbf{E}_3\tilde{\mathbf{M}}_o(\mathbf{x}_o,\mathbf{q})\mathbf{E}_3^T\ddot{\mathbf{q}} + \mathbf{E}_3\tilde{\mathbf{L}}_o(\mathbf{x}_o,\dot{\mathbf{x}}_o,\mathbf{q},\dot{\mathbf{q}})\\ +\mathbf{B}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{E}_3\tilde{\mathbf{T}}_\Delta^T(\mathbf{x}_o,\mathbf{x}_\Delta,0)\binom{\tau}{0} + \mathbf{F}_{\Delta_3} - \mathbf{J}_q^T(\mathbf{x}_o,\mathbf{q})\mathbf{F}.\end{aligned} \tag{4}$$

where $\tilde{\tau} = \mathbf{T}_o(\tau^T \mathbf{0}^T)^T$, $\mathbf{x}_{o_1} \in \Re^r$ are the task coordinates chosen to describe the geometry of the contact forces, $\mathbf{x}_{o_2} \in \Re^s$ are the task coordinates chosen to describe the task geometry for the position and the orientation of the sheet metal, $\mathbf{F}_1 \in \Re^r$ is the vector of forces in the directions of contact, and $\mathbf{F}_2 \in \Re^s$ is the vector corresponding to the moments generated by the presence of an angular misalignment between the sheet metals parts and to the forces in directions other than in the contact directions (which are all zero), with $m = r + s$.

It is expected that a control of $\mathbf{F}_2$ to make the moments vanish will realign the parts and a proper contact will result between the parts. A PID control approach is proposed to control $\mathbf{F}_2$ since it is a simple and robust control approach that ensures no steady state error. Therefore, a

PID error moment control is used to provide an angular correction to the orientation set-point of the sheet metal part. The correction applied to the desired orientation of the part is defined as,

$$\Delta \mathbf{x}_{o_2}^{\mathrm{d}} = \mathbf{G}_{\Delta_{p2}} \mathbf{e}_{\Delta_{F2}} + \mathbf{G}_{\Delta_{v2}} \dot{\mathbf{e}}_{\Delta_{F2}} + \mathbf{G}_{\Delta_{i2}} \int_0^t \mathbf{e}_{\Delta_{F2}}\, dt, \tag{5}$$

where $\Delta \mathbf{x}_{o_2}^{\mathrm{d}} \in \Re^m$ is the orientation correction, where $\mathbf{G}_{\Delta_{p2}}$, $\mathbf{G}_{\Delta_{v2}}$, and $\mathbf{G}_{\Delta_{i2}} \in \Re^{s \times s}$, are all diagonal gain matrices, where t represents the elapsed time, and where $\mathbf{e}_{\Delta_{F2}} = \mathbf{F}_2^{\mathrm{d}} - \mathbf{F}_2$ with $\mathbf{e}_{\Delta_{F2}} \in \Re^s$. $\mathbf{F}_2^{\mathrm{d}} \in \Re^s$ is the vector of desired forces, in directions other than the directions of contact, and of desired moments.

It is now possible to formulate a rigid control law based on the rigid dynamics of the system, and which makes use of the signal of the force/moment sensor to feedback the moments as well as the forces in the contact directions. The control law regulates the forces in the directions of contact by using r generalized coordinates of the task space, and regulates the position and orientation using the remaining s generalized coordinates. This allows to fulfill the first two control objectives set in this paper. The implementation of the orientation correction allows to fulfill the third one. The rigid hybrid control law is a computed torque type of control law and is given as,

$$\begin{aligned}
\mathbf{E}_1 \begin{pmatrix} \tilde{\tau} \\ \mathbf{0} \end{pmatrix} &= \mathbf{E}_1 \tilde{\mathbf{M}}_o(\mathbf{x}_o, \mathbf{0}) \mathbf{E}_1^T \left[\ddot{\mathbf{x}}_{o_1}^{\mathrm{d}} + \mathbf{G}_{v_1} (\dot{\mathbf{x}}_{o_1}^{\mathrm{d}} - \dot{\mathbf{x}}_{o_1}) + \mathbf{G}_{f_1} (\mathbf{F}_1^{\mathrm{d}} - \mathbf{F}_1) \right] \\
&+ \mathbf{E}_1 \tilde{\mathbf{M}}_o(\mathbf{x}_o, \mathbf{0}) \mathbf{E}_2^T \left[\ddot{\mathbf{x}}_{o_2}^{\mathrm{d}} + \mathbf{G}_{v_2} (\dot{\mathbf{x}}_{o_2}^{\mathrm{d}} - \dot{\mathbf{x}}_{o_2}) + \mathbf{G}_{p_2} (\mathbf{x}_{o_2}^{\mathrm{d}} + \Delta \mathbf{x}_{o_2}^{\mathrm{d}} - \mathbf{x}_{o_2}) \right] \\
&+ \mathbf{E}_1 \tilde{\mathbf{L}}_o(\mathbf{x}_o, \dot{\mathbf{x}}_o, \mathbf{0}, \mathbf{0}) + \mathbf{F}_1,
\end{aligned} \tag{6}$$

$$\begin{aligned}
\mathbf{E}_2 \begin{pmatrix} \tilde{\tau} \\ \mathbf{0} \end{pmatrix} &= \mathbf{E}_2 \tilde{\mathbf{M}}_o(\mathbf{x}_o, \mathbf{0}) \mathbf{E}_1^T \left[\ddot{\mathbf{x}}_{o_1}^{\mathrm{d}} + \mathbf{G}_{v_1} (\dot{\mathbf{x}}_{o_1}^{\mathrm{d}} - \dot{\mathbf{x}}_{o_1}) + \mathbf{G}_{f_1} (\mathbf{F}_1^{\mathrm{d}} - \mathbf{F}_1) \right] \\
&+ \mathbf{E}_2 \tilde{\mathbf{M}}_o(\mathbf{x}_o, \mathbf{0}) \mathbf{E}_2^T \left[\ddot{\mathbf{x}}_{o_2}^{\mathrm{d}} + \mathbf{G}_{v_2} (\dot{\mathbf{x}}_{o_2}^{\mathrm{d}} - \dot{\mathbf{x}}_{o_2}) + \mathbf{G}_{p_2} (\mathbf{x}_{o_2}^{\mathrm{d}} + \Delta \mathbf{x}_{o_2}^{\mathrm{d}} - \mathbf{x}_{o_2}) \right] \\
&+ \mathbf{E}_2 \tilde{\mathbf{L}}_o(\mathbf{x}_o, \dot{\mathbf{x}}_o, \mathbf{0}, \mathbf{0}) + \mathbf{F}_2,
\end{aligned} \tag{7}$$

where $\mathbf{x}_o^{\mathrm{d}}$, $\dot{\mathbf{x}}_o^{\mathrm{d}}$, and $\ddot{\mathbf{x}}_o^{\mathrm{d}} \in \Re^m$ are the desired position, velocity, and acceleration respectively, and where $\mathbf{F}^{\mathrm{d}} = ((\mathbf{F}_1^{\mathrm{d}})^T (\mathbf{F}_2^{\mathrm{d}})^T)^T \in \Re^m$ is the vector of desired forces and moments, with $\mathbf{F}_1^{\mathrm{d}} \in \Re^r$ the vector of desired forces in the directions of contact, and $\mathbf{F}_2^{\mathrm{d}} \in \Re^s$, as mentioned before, the vector of desired forces in the directions other than the directions of contact and of desired moments. Finally $\mathbf{G}_{p_2} \in \Re^{s \times s}$ is the position gain, $\mathbf{G}_{v_1} \in \Re^{r \times r}$ and $\mathbf{G}_{v_2} \in \Re^{s \times s}$ are velocity gains, and $\mathbf{G}_{f_1} \in Re^{r \times r}$ is the contact force gain, which are all diagonal matrices. Figure 2 gives the block diagram of the rigid hybrid control law with orientation correction.

4. EXPERIMENTAL SET-UP AND RESULTS

The Laboratory for Nonlinear Systems Control at the University of Toronto is equipped with two CRS Robotic Corporation A460 industrial robots, each with six degrees of freedom. An ATI six-axis force/moment sensor is mounted at the tip of each robot arm. The sensor can measure a linear force in the three Cartesian coordinate directions and a moment about each of these directions. The end-effector of the robot is affixed directly on the force/moment sensor. The sheet metal parts used in the experiments consist of curved sheet metal parts made of mild steel. At each straight edge of the payload resides a small rectangular lip. The lips present a flat surface area that facilitates the contact of the payloads as they come in contact. For pick up, the payload is positioned in a wooden fixture where, due to its curved shape, it sits on its two lips only which makes it very prone to generate improper grasps.

In order to test the control law, a multi-robot system composed of both industrial robots is used where two pieces of sheet metal are brought into contact to simulate an assembly operation, with one part improperly grasped. The two robots move the payloads towards each other while keeping a fix orientation of the parts until contact is made, at which point the controller starts to regulate the force and moments. This paper is only concerned with the control of the robots once contact is achieved. The payloads held by both robots are flexible, thus the multi-robot experiment can be seen as having the hybrid controller deal with a compliant environment. In Figures 3 to 5, the time history of the moments generated during the experiment is given. They are expressed with respect to the frame of the force/moment sensor. The figures show that that the orientation correction acts very rapidly and eliminates the moments. When the readings have reached their settling values, the figures also show that an oscillation is present in the results of the moments. The cause of this oscillation is believed to be the vibration of the robot at one of its natural frequency when it is positioned in the configuration required to push the payload against the surface.

5. CONCLUSION

In this paper, a hybrid controller was developed for the assembly of thin wall parts by using a multi-robot system as an alternative to fixtures. The proposed controller modifies the orientation of the walled parts so that they are mated properly. The controller resolved the specific problem of angular misalignment by using moment feedback to modify the orientation set-point of the payload. Experiments conducted on commercial CRS Robotics Corporation six degree of freedom robots demonstrated the ability of the controller to eliminate the moments generated by the misalignment and thus realign the parts.

REFERENCES

1. Book, W.J., "Recursive Lagrangian dynamics of flexible manipulator arms," Int. Journal of Robotics Research, Vol. 3, No 3, pp. 87-101, 1984.

2. Laliberté, M., "Dynamic Modeling and Control of a Multi-Robot System During Assembly of Flexible Payloads with Kinematic Uncertainty," M.A.Sc. Thesis, University of Toronto, 1996.

3. McClamroch, N.H. and Wang, D., "Feedback stabilization and tracking of constrained robots," IEEE Trans. on Automatic control, Vol. 33, No. 5, pp. 419-426, 1988.

4. Mills, J.K., "Fixtureless Assembly: Multi-robot manipulation of distributed parameter payloads," IMAC/SICE Int. Symp. on Robotics, Mechatronics and Manufacturing Systems '92, Kobe, Japan, Sept. 16-22, 1992.

5. Mills, J.K. and Ing, J., "Robotic fixtureless assembly of sheet metal parts using dynamic finite element models: modeling and simulation," Proc. IEEE Int. Conf. on Robotics and Automation, Nagoya, Japan, May 21-27, 1995.

6. Nguyen, W. and Mills, J.K., "Fixtureless assembly: multi-robot manipulation of flexible payloads," Proc. of the Sixth Int. Symp. on Robots and Manufacturing at Second World Automation Congress, Montpellier, France, May 27-30, 1996.

7. Siciliano, B. and Book, W.J., "A singular perturbation approach to control of lightweight flexible manipulators," Int. Journal of Robotics Research, Vol. 7, No. 4, pp. 79-90, 1988.

8. Uchiyama, M., Iwasawa, N. and Hakomori, K., "Hybrid position/force control for coordination of a two-arm robot," Proc. IEEE Int. Conf. on Robotics and Automation, pp. 1242-1247, 1987.

9. Yuen, K.-M. and Bone, G.M., "Robotic assembly of flexible sheet metal parts," Proc. IEEE Int. Conf. on Robotics and Automation, pp. 1511-1516, Minneapolis, 1996.

10. Zheng, Y. and Luh, J., "Optimal distribution for two industrial robots handling a single object," Proc. IEEE Int. Conf. on Robotics and Automation, pp. 344-349, 1988.

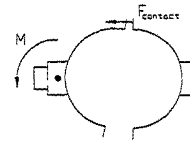


Figure 1: Angular misalignment

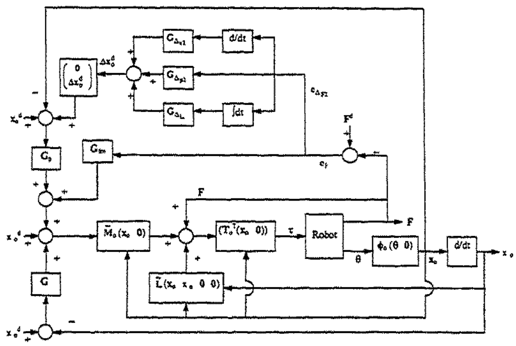


Figure 2: Block diagram of the hybrid control law.

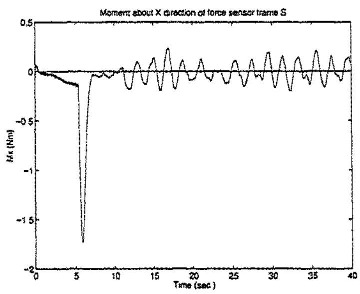


Figure 3: Moment about X-axis.

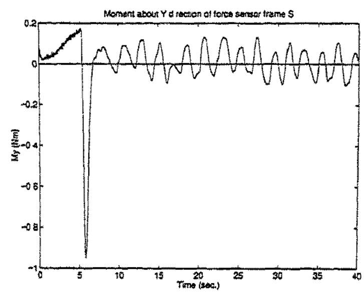


Figure 4: Moment about Y-axis.

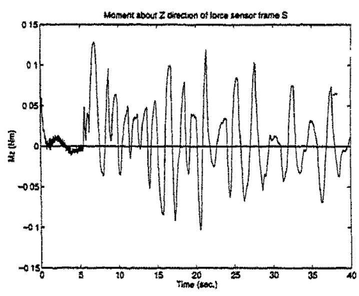


Figure 5: Moment about Z-axis

SHAPE GLASS CUTTING DIRECT DRIVE ROBOT

Chen Hui-tang, Jiang Ping, Wang Yue-juan
Department of Electrical Engineering
Tongji University, 1239 Siping Road
Shanghai, PRC

ABSTRACT

With the consideration of that D. D. robot can achieve high speed and high precision trajectory tracking and its simplicity in mechanical structure, this paper put forward the idea of developing a D. D. robot for shape glass cutting. This new idea has been proven to be successful by a model glass cutting machine. This glass cutting D. D. robot has two arms of 1 meter length respectively. It can cut the glass plate in size of 1.5m × 2.8m with maximum cutting velocity of 0.5m/s and the contour size error is less than 1 mm. The control algorithm is feedforward plus PID feedback with observer Due to the lag behind design of wheel cutter with a small free-rotation gap we can guarantee the compliance of the cutter. Also a cutting force compensation is added. As the result we simplified the constraint motion control problem and turn it into a quasi free space control problem. This machine equipped with vision guided curve tracking system for drawing curve input. It also include a CAD software package for on-screen designing cutting contour. Due to the simplicity of mechanical structure compared to X-Y cutting machine, cost of machine is quite low.

KEYWORDS

D.D. Robot, Trajectory Tracking, Vision Guided Tracking, Glass Cutting

1. INTRODUCTION

Up to now cutting windscreen of car in China is mainly rely on manual operation. Operator grasps the cutter and moves it along the inner contour of template. The drawbacks of this method are: (1) worker's labour intensity is very high; (2) manufacturing of template is time-consuming and highly cost; (3) production flexibility is low. With the consideration of that D. D. robot can achieve high speed and high precision trajectory tracking[1,2],due to direct linkage of motor and arm without gear reducer, and its simplicity in mechanical structure,authors put forward the idea of using D. D. robot for glass contour cutting. This idea has been proven to be successful by a model glass cutting machine. This glass cutting robot can cut the glass plate in size of 1.5 × 2.8m with maximum cutting velocity of 0.5m/s and contour size error less than 1mm. We have tried some advanced control algorithms such as feedforward, compute torque, adaptive. Experiments demonstrate [3] feedforward compensation brings about much higher accuracy than PD control. The accuracy can be improved by 3 ~ 4 times, but computation time only increases a little bit. Computed torque control has no obvious effect on performance improvement in our case. Adaptive control exhibits better performances only for various payload. Since there is no large payload variation in our case and also in order to use minimum computation time we selected feedforward for compensation of nonlinear coupling of robot dynamics. For the sake of simplicity tachogenerator is replaced by an observer. [4] Because wheel-cutter at the end of forearm is acting on the glass,force interaction with the environment must be considered, this leads to the lag behind design of wheel-cutter and cutting force compensation. As the result we simplified the robot motion constraint control problem and turned it into free space control problem. This machine equipped with vision guided curve tracking system for drawing curve input. Therefore it achieve the goal: "one machine with two usage,that is for cutting and for curve input." It also includes a CAD software package for on-screen design of cutting contour. Due to its simplicity in mechanical structure ,the cost is very low. Now it is manufacturing by Shanghai Glass Machinery Factory. This paper is organized as follows: Sec. 1 is an introduction ,Sec. 2 describes the robot structure ,Sec. 3 is about control system and control strategy,Sec. 4 discusses wheel-cutter design and compliance, Sec. 5 addresses vision guided curve tracking system, Sec. 6 brief in CAD software package and at the end in Sec. 7 conclusion is made.

2. STRUCTURE OF THE ROBOT

This D. D. robot imitates a human right hand has four degrees of freedom. Two of them are the rotation angles of upperarm and forearm (Fig. 1). Upperarm is direct coupled with driven motor, while the forearm is driven by motor through parallelogram mechanism. Maximum torques of the upperarm and forearm are 80N · M and 40N · M respectively. For high precision joint angle measurement, we use multipolar two speed resolver. The angle measurement resolution is 10 seconds. The remain two degrees of freedom are rotation and up-down displacement of wheel cutter. Except the up-down displacement of wheel cutter is driven pneumatically, for the other three degrees of freedom high precision servo control system are used as will be described below.

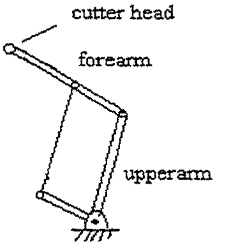


Fig.1: Schematic structur of glass cutting robot.

3. CONTROL SYSTEM & CONTROL STRATEGY

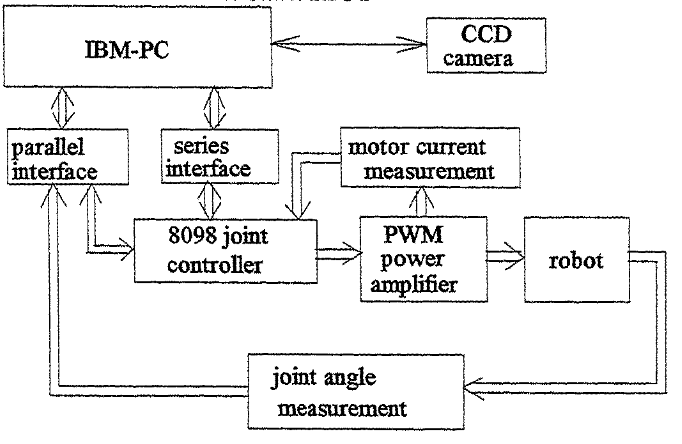


Fig. 2: Block diagram of control system.

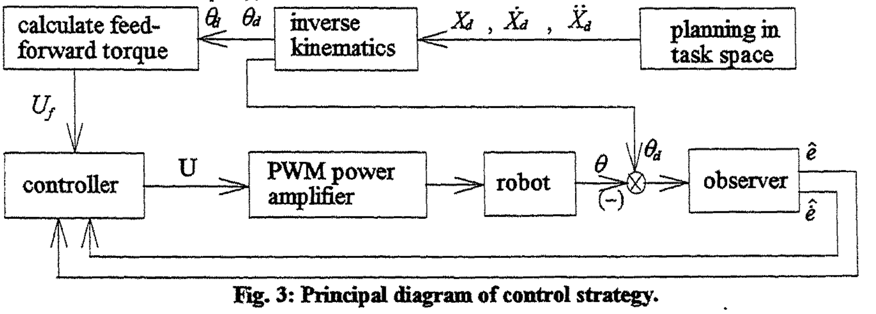


Fig. 3: Principal diagram of control strategy.

The block diagram of control system is shown in Fig. 2. It is a two level hierarchical structure. For upper level we use a personal computer IBM-PC, which realizes off-line estimation of kinematic and dynamic parameters, trajectory planning , feedforward calculation and data processing. For lower level three Intel 8098 single-chip computers are used. They accomplish joint-servo computation and PWM power amplifier current regulation. Data communication between PC and single-chip computer is realized through parallel interface 8255A, series interface realizes download control program into single-chip computer. Joint angle measurement is accomplished by multipolar two speed resolver. The control strategy is shown in Fig 3. PC executes task space planning and sends out $X_d, \dot{X}_d, \ddot{X}_d$. By inverse kinematic transform we get desired trajectory in joint space. Basing upon desired trajectory PC computes feedforward torque U_f. Using desired θ_d and measured θ, observer sends out $\hat{e}$ and $\hat{\dot{e}}$. Thus closed loop control of arm joint angle is accomplished For the purpose of feedforward compensation, first of all, we need off-line identification of the dynamics model of two arms, which is as follows: [3]

$$Y(\ddot{\theta},\dot{\theta},\theta)P = U_f$$

$$Y(\ddot{\theta},\dot{\theta},\theta) = \begin{bmatrix} \ddot{\theta}_1 & \ddot{\theta}_2\cos(\theta_2-\theta_1)-\dot{\theta}_2^{\,2}\sin(\theta_2-\theta_1) & 0 \\ 0 & \ddot{\theta}_1\cos(\theta_2-\theta_1)+\dot{\theta}_1^{\,2}\sin(\theta_2-\theta_1) & \ddot{\theta}_2 \end{bmatrix}$$

$$P^T = \begin{bmatrix} a & b & c \end{bmatrix}$$

Where U_f is 2×1 feedforward vector, Y is 2×3 sensitivity matrix. a,b,c is dynamic parameter. Then U_f can be determined as follows:

$$U_f = Y(\ddot{\theta}_d\,,\ \dot{\theta}_d\,,\ \theta_d)\hat{P}$$

Control algorithm for two arms is selected as follows:

$$U = K_P S + K_I I(s,t) + U_f \tag{1}$$

Where U is 2×1 control vector; S is 2×1 sliding error vector; K_P and K_I are 2×2 diagonal positive definite matrices; I(s,t) is 2×1 bounded integral vector with respect to S, it compensates slow varied uncertainties, the merit of bounded integrator is its improvement of dynamic performance compared to linear integrator, U_f is feedforward compensation torque.

Sliding error S is defined as

$$S = \hat{\dot{e}} + \lambda(e)$$

$$\lambda_i(e_i) = \frac{\lambda_m e_i}{1+K_e|e_i|} \qquad i = \{1,2\} \tag{2}$$

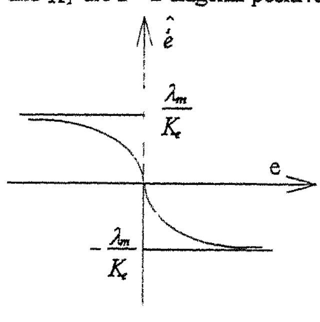


Fig. 4: Nonlinear sliding surface.

Where $e = \theta_d - \theta$ is joint angle error vector; θ_d is desired joint angle vector; $\hat{\dot{e}}$ is estimate of joint velocity error; λ_m and K_e are bounded positive constants. Sliding surface defined by eq(2) is an nonlinear sliding surface as shown in Fig 4 [1]. When e is small, λ (e) has a high slope this in turn guarantees a high tracking accuracy, while e is large, λ (e) is bounded and can avoid system saturation. Select a bigger λ_m is benefit to accuracy, and K_e is selected basing upon maximum linear range.

$\hat{\dot{e}}$ in eq.(2) is achieved using the following observer. [3]

$$\begin{bmatrix}\hat{e}(k+1)|k\\ \hat{\dot{e}}(k+1)|k\end{bmatrix}=\begin{bmatrix}I & TI_n\\ 0 & I_n\end{bmatrix}\begin{bmatrix}\hat{e}(k)\\ \hat{\dot{e}}(k)\end{bmatrix}$$

$$\begin{bmatrix}\hat{e}(k+1)\\ \hat{\dot{e}}(k+1)\end{bmatrix}=\begin{bmatrix}\hat{e}(k+1)|k\\ \hat{\dot{e}}(k+1)|k\end{bmatrix}+\begin{bmatrix}TK_{po}\\ TK_{vo}\end{bmatrix}\left[e(k+1)-\hat{e}(k+1)|k\right].$$

4. CONTROL OF WHEEL-CUTTER AND COMPLIANCE DESIGN

In order to keep the wheel-cutter always along the tangent of the curve, when it is cutting glass. In the process of planning, we must calculate the tangent of curve in the task space, and then convert it to rotation angle of wheel-cutter by homogeneous transform. This rotation angle θ_{d3} is then sent to the wheel-cutter servo system. Suppose some inaccuracy factors exits, which may cause the wheel-cutter in fact not really along the tangent of curve. Cutting force F acting on the wheel-cutter can be decomposed into two components, one is along the tangent direction of wheel-cutter, another is in the normal direction. Due to the constraint in the normal direction, normal direction force component may damage the glass or cutter itself. To avoid this, we design the center of wheel-cutter a bit lag-behind its moving center, and also allow a small free rotation gap.Fig 5. Then the normal direction force can cause the wheel-cutter to rotate around motor axis(or moving center) until align with the tangent of curve.The regulation process just like the rear wheel of bicycle align with the direction of front wheel. Theoretical analysis can reach the following conclusions.

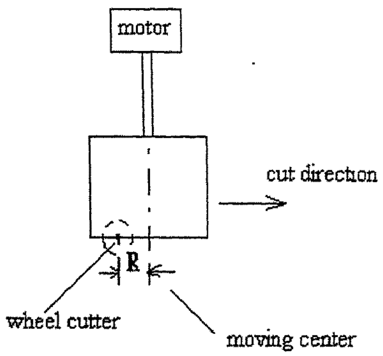


Fig. 5: Schematic diagram of lagbehind wheel cutter.

The motion of the wheel cutter in horizontal plane is shown in fig. 6, where the moving center is in o and the wheel cutter is in c with a tangent angle β. Suppose that the manipulator moves the center o to $o1$ with the angle θ and displacement ds. Constrained by the wheel cutter, the motion of it can be only along its tangent direction to $c1$ and changes the tangent angle of the wheel to $\beta 1$. From triangle $oo1c1$, the motion equation of the wheel cutter can be expressed as follow:

$$\frac{d\beta}{ds}=sin(\theta-\beta)/R \tag{3}$$

where R is the distance of lag behind, θ is the tangent angle of the moving center trajectory, and β is the tangent angle of the cutting trajectory.

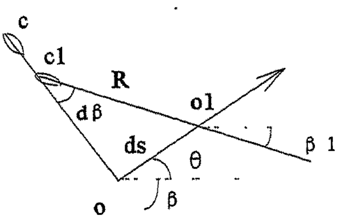


Fig.6: Motion analysis of cutter.

<u>conclusion 1</u>: If moving center moves along straight line, then cutting trajectory is globally asymptotically approach to this straight line.

proof: Suppose the tangent angle of desired line is θ and the angle of wheel cutter is β. Define a Lyapunov function as $V=(\theta-\beta)^2$, where θ is constant for the straight line.

Then

$$\frac{dV}{ds}=-2(\theta-\beta)\frac{d\beta}{ds}=-2(\theta-\beta)sin(\theta-\beta)/R<0;\quad \theta\neq\beta \text{ where } (\theta-\beta)\in[0,\pm\pi)$$

So $\lim_{s\to\infty}\beta=\theta$, the cutting trajectory is globally asymptotically approach to the desired straight line.

conclusion 2: If the angle of tangent of curve is θ, while angle of wheel-cutter is β, curvature of desired cutting trajectory is K, the amount of lagbehind between center of wheel-cutter and moving center is R. If the condition $K_{max} < \frac{1}{R}$ is satisfied, then cutting trajectory relates to moving center trajectory is local stable, where stable area $|\theta - \beta| < \pi / 2$, with accuracy $|\theta - \beta| < \arcsin(RK)$.

proof: Define the Lyapunov function as $V = (\theta - \beta)^2$, then

$$\frac{dV}{ds} = 2(\theta - \beta)\left(\frac{d\theta}{ds} - \frac{d\beta}{ds}\right) \le 2\left[|\theta - \beta|\left|\frac{d\theta}{ds}\right| - (\theta - \beta)\frac{d\beta}{ds}\right]$$

$$= 2\left[|\theta - \beta|K - |\theta - \beta||sin(\theta - \beta)| / R\right] \qquad \text{where } (\theta - \beta) \in [0, \pm\pi)$$

When $K < |sin(\theta - \beta)|/R$, dV/ds<0, system is stable.

From $K_{max} < \frac{1}{R}$, we can see a positive constant $e = arcsin(RK)$ exists such that $K < |sin(\theta - \beta)|/R$, if $\pi / 2 > |\theta - \beta| > e$. This means that cutting trajectory relates to moving center trajectory is local stable, where stable area $|\theta - \beta| < \pi / 2$, with accuracy $|\theta - \beta| < \arcsin(RK)$.

These conclusions derived above tell us that between cutting trajectory and desired trajectory there is a small deviation, the magnitude of which is depending on curvature of desired curve and the amount of lagbehind R. In order to raise accuracy, we can carry out compensation as follows:

$$\beta = \theta + sgn(\dot{\theta}) \arcsin(RK)$$

$$X_c = X - R\cos\beta$$

$$Y_c = Y - R\sin\beta$$

Where X,Y are desired trajectory, X_c, Y_c are moving center coordinate after compensation. In our case R=2mm.

We can see from the above analysis, adoption of lagbehind wheel-cutter can solve the problem of normal direction constraint. Experiments show tangent direction cutting force maintain constant with the same cutting pressure. Hence we can compensate if by adding it to feedforward.

$$U_f = M(\theta_d)\ddot{\theta}_d + C(\theta_d, \dot{\theta}_d)\dot{\theta}_d + J^T N_t$$

where $N_t = |N_t| \left[{dr}/{dx} \quad {dr}/{dy} \right]^T$. r is desired trajectory.

Up to now, we have solved the control problem of motion constraint that is eliminating normal direction constraint by using lagbehind wheel-cutter. Also we compensate cutting force by feedforward. As the result, we simplified the motion constraint control problem, and turn it into a quasi-free space control problem. Practice demonstrates the cutting trajectory is very smooth.

5. VISION GUIDED TRAJECTORY TRACKING FOR CURVE INPUT

A CCD camera is mounted of the end of forearm, it can guide the end of forearm to track the curve of a drawing. At the same time, the raw position datum of curve are recorded. These raw position datum are then pass through a data processor which filter off noises. Processed datum then store in computer memory. Should we cut glass, we simply download the stored datum and D. D. robot automatically cut down the curve as the drawing. This vision-guided process is implemented by the following tangent prediction algorithm. [5]. Basing upon image we calculate the tangent direction α_n and matching point r_{i-1} of two succeed image, then

$$r_i(i) = r_{i-1} + \frac{\partial r}{\partial s}\dot{s}(t_i - t_{i-1}) \quad , \quad \dot{r}(t) = \frac{\partial r}{\partial s}\dot{s}$$

where $\frac{\partial r}{\partial s} = [\cos\hat{\alpha}_t \quad \sin\hat{\alpha}_t \quad 0]^T$.

According to this equation we control the robot and hence realize vision guided tracking. After tracking the complete curve, we get the raw coordinate datum $r_d + N$, where N is noise term. Obviously simply connect these point, the first derivative & second derivative can change violently. As the result we can not use it to calculate feedforward compensation. For this sake, considering the curve always made of straight line and circular arc, we adopt the following steps to fit the curve.

(1). filter off the frequency component with maximum curvature higher than 1/R.

(2). map individual point (x_{di} , y_{di}) to (α_i , S_j) coordinate, where α_i is trajectory tangent angle, S_j is arc length, then we get

$$\alpha_i = constant \qquad \text{for straight line}$$

$$\alpha_i = K s_i + \alpha_0 \qquad \text{for circular arc}$$

hence if we find out straight line or circular arc points set, we can record it with straight line and circular arc on the drawing;

(3). except the straight line and circular arc,we fit remaining points with 3rd spline.

As the result, we realize curve input and cutting process on the same glass cutting robot, that is, one machine has two usage. This greatly reduce the cost of machine. This robot can also guarantee production flexibility. User no more need manufacturing and keeping a large number of templates. Moreover, if we need to change some parts of curve, we can do it on the computer screen, hence it is very convenient to the user.

6. CAD SOFTWARE PACKAGE

This glass cutting D. D. robot is equipped with a CAD software Package, which provides the performance of design any glass cutting contour on the computer. The contour can be made of straight lines, circular arcs and splines. For the purpose of optimizing, we adopt trajectory planning with torque constraint. Above mentioned every thing is done automatically, basing upon desired cutting contour and robot dynamic parameter we calculate optimal cutting trajectory and feedforward torque.

7. CONCLUSION

Model machine demonstrates that our new idea is feasible. Using this robot we have cut down circle and parallelogram and contour composite of straight and circular arcs, and also any curves inputted by CCD camera with speed 0.5m/sec and accuracy better than 1mm. The advantage of this shape glass cutting D. D. robot is its simplicity in mechanical structure, and its excellent performance is achieved by control and computer technique.

8.REFERENCE

1. Zhu, W.H., Chen, H.T., et al., A Variable Structure Robot Control Algorithm with an Observer, pp.486-492, IEEE Trans on Robot & Automation, Vol.8, No.4, 1992.
2. Cao, B.L. and Chen, H.T., PD-Based Trajectory Tracking Control for Robot Manipulators, Transactions of ASME, Vol.115, pp.566-569, 1993.
3. Chen, H.T., Jiang, P., et al., Comparative Study of D. D. Robot Control Algorithm for Trajectory Tracking, Proceeding of IFAC 12th World Congress, Vol.6, pp.157-160, Sydney,1993.
4. Jiang, P., Chen, H.T., et al., An Observer Based Robust Controller for Robotic Manipulator, Proceeding of IFAC 12th World Congress, Vol.6, pp.177-180, Sydney, 1993.
5. Jiang, P., Chen, H.T., et al., A Vision Guided Curve Tracking and Input System for Glass Cutting Robot, Proceedings of 2nd Asian Conference on Robotics and Its Application, pp.665-669, Beijing, 1994.

STEP BASED SEMANTIC PRODUCT MODEL

Luo Yan , Zhou Xinjian * and Mo Ming
Department of Plasticity Technology
* Department of Mechanical Engineering
Shanghai Jiaotong University
Shang hai

ABSTRACT

When the world enters the 90's, enterprise competition becomes more and more fierce, both product development cycles and product lifecycle will be greatly shortened, enterprises face how to improve product quality, shorten time, reduce production cost, etc.. So many scholars and constitutes present various manufacturing theories. This paper analyzes the requirements of Advanced Manufacturing Technology (AMT), and thinks that a common product model is necessary to support information share in enterprise and out of enterprise. Authors present a product model for injection mould. The paper describes in detail the development of a STEP-based Injection Mould Product Model(IMPM) which has been designed and implemented according to the integrated resources of STEP and supports a complete injection mould product information structure and standard data format. The IMPM consists of a generic information model, a main feature information model , a structural information model and a auxiliary information model.

KEYWORDS

Advanced Manufacturing Technology, Injection Mould Product Model, Lifecycle, Product Modeling

1. INTRODUCTION

Advanced Manufacturing Technology hasn't got an accepted concept. It is known to many scholars that it is based on traditional manufacturing technologies, continuously absorbs mechanics, electronics , information, material and modern management technologies, etc., and makes use of them in manufacturing processes to realize the highest quality, efficiency, low consuming, clean and agile production and get the ideal economical and technological result. So many theories have been presented, such as Concurrent Engineering(CE), Lean Production(LP), Agile Manufacturing(AI), Intelligent Manufacturing(IM) and Virtual Manufacturing System(VMS), LAF(Lean-Agile-Flexible), etc.. It insists on high information integration, high flexible and intelligence. The basic object of AMT is how to improve enterprise competence. A manufacturing enterprise must be agile enough to quickly respond to product demand changes. This requires that a suited data exchange mechanism should be established in enterprises and out of enterprises. Therefore a uniform product model is crucial to the successful application of CE and other advanced manufacturing technologies.

1.1 Modeling Requirements

The requirements of AMT are as following:

•Uniform product model

The entire enterprise share a common product model at any stage of the product lifecycle. It satisfies the need of AMT for integration and economics. User don't care the system interface when product model uses a united standard. So system's reconstruction and interchange will be improved.

•Supporting product lifecycle's activities

AMT asks integrated consideration of product lifecycle's information(including design, manufacturing, inspection and disposal etc.) and modeling for product lifecycle. That means that a data exchange mechanism should be built according to some standards, different system can exchange their information directly.

•Using a neutral mechanism

That means that a product model should be independent on any CAD/CAM system, Various systems transfer their data according to common rationale.

•Different implementation ways

Systems can exchange data by files, rules, intelligent base and so on.

Information integration is a key technology of CAD/CAM, then product information modeling is the foundation of information integration. Product modeling technology improved when AMT develops.

1.2 Reasons to Select Injection Model as research object in this paper

The paper selects mould as modeling object, because:

•The mold is suited to single production, it requires short date of delivery, high quality and low cost. This is identical with advanced manufacturing theories.

•The price of mold is high and there is a wide market.

•The design of mold must reflect the user's requirements. There is a high collaboration in the process of design and manufacturing, etc.

•There is a strict and great requirement of data transportation. Some complicated mold cannot be splendidly designed if CAD/CAM technology can not be used.

Injection mould is a large part of all moulds in quantity, so the research of injection mould is representative.

1.3 The Development of Product Modeling

Much of the research in product modeling assumes that the goal is to support geometric modeling, others use feature-based modeling, recently, some reports was found that product information exchange through a neutral file was the important implementation method to be standardized in STEP. But we found featured based modeling is lack of complete information and difficult to modify, because modification of one part of the model easily effects other parts of the model. Some STEP based model only describes simple part, such as rational compartment.

•Information is incomplete.

A product model is the mechanism to provide information at any stage of the product lifecycle.

•The structure of model don't satisfy the need of AMT.

•Not based on a uniform standard.

Though some try to build product model using international standard, such as STEP, and get much experience, these models can't suit a special application, for example, injection mould. We consume Part41~Part 49 of STEP and other related parts, and give a STEP based injection mould.

2. INTRODUCTION TO STEP

STEP(The ISO10303 standard) is an international standard that provides an unambiguous representation and an exchange mechanism for computer interpretable product information through the lifecycle of a product. In addition, it provides a consistent data exchange format and application interfaces between different application systems, The present Implementation methods for information transfer include file exchange, an application programming interface and database sharing. STEP is a worldwide federation of national standards bodies. It consists of Application Protocol, Integrated resources , Description methods, Implementation methods and so on.

2.1 Integrated Resources

It includes integrated generic resources and integrated application resources. The integrated generic resources contain EXPRESS defined entities which are independent of any application. The integrated application resources consists of entities related to special applications. In the application resources entities are constructed or referenced from the entities in the generic resources.

2.2 Data Descriptive Language

EXPRESS is both human and machine readable as it enhances human understanding and the generation of complete interpretable applications It is the name of a formal information requirements specification language.

EXPRESS-G is a formal graphical notation for the display of data specification defined in the EXPRESS language. The notation only supports a subset of EXPRESS language. It is represented by graphic symbols forming a diagram. In the paper, we describe the structure of the mould by EXPRESS-G, the complete EXPRESS descriptions are omitted because of length.

3. PRODUCT MODEL

The IMPM consists of a Generic Information Model(GIM), a Main feature Information Model (MIM) , a Structural Information Model(SIM) and a Auxiliary Information Model(AIM).

3.1 Architecture of IMPM

The IMPM provides a detailed description of structure of injection mould product information. It makes use of the integrated generic resources of STEP, different entities from STEP have been referenced by the product model. Fig. 1 gives the overview.

3.2 GIM

The GIM comprises all those properties which are suitable to characterize the injection mould and its functional capabilities in total. This information may be relevant to any application context or lifecycle. It pertains to the whole mould as a product rather than to any of its subsystems and components. It may still include an overview of the functional capabilities available in a certain mould.

3.3 MIM

The MIM expresses the information of geometry , topology , precision and material etc., Precision

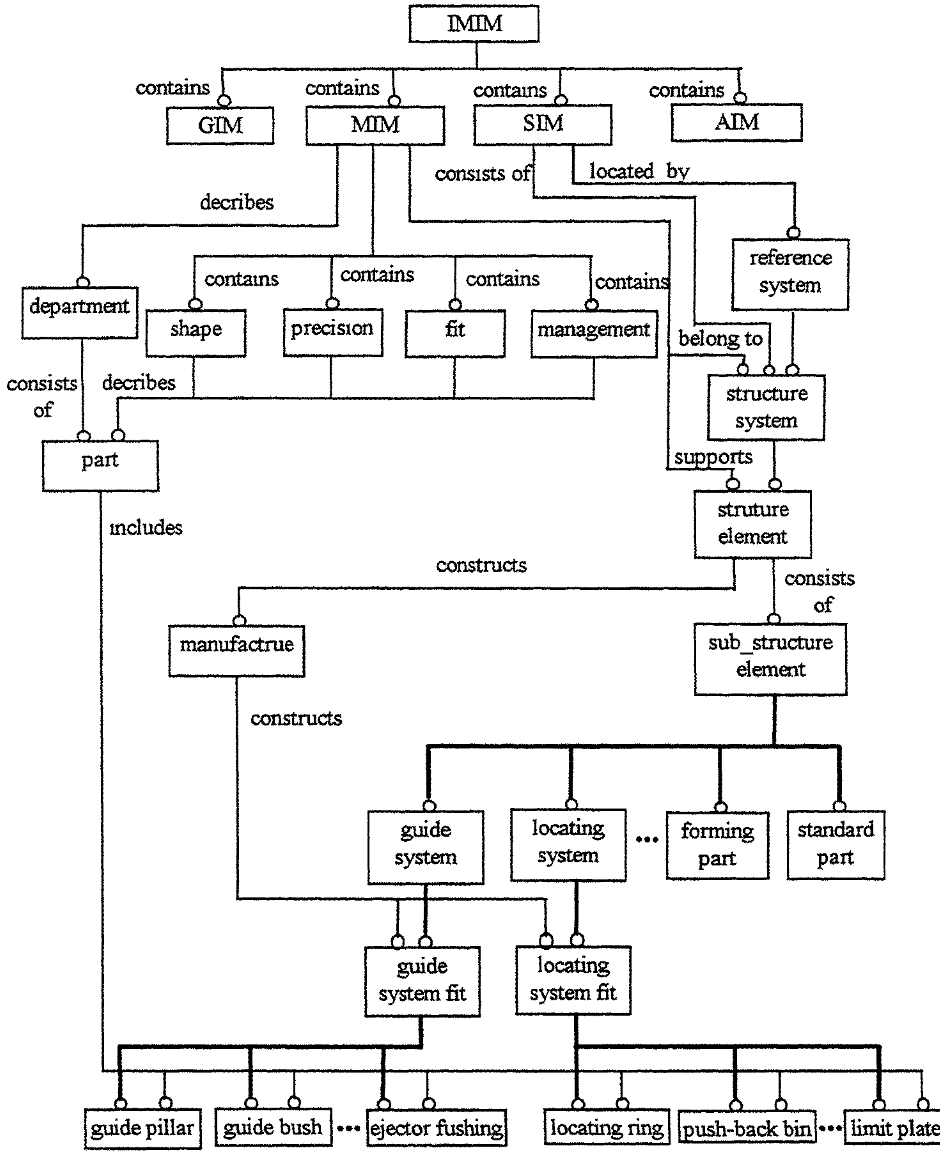


Fig. 1: Simple Injection Mould Structural Organization

includes dimension tolerance, shape tolerance(geometric tolerance and position tolerance) and fit tolerance. The model includes the concepts which support other models, such as, geometry and topology . It consists of a resources model and four application models.

3.4 SIM

This model defines the structure of the product as it is required throughout its lifecycle from design, process planning , manufacturing , fit , checking and maintain ,etc..

The SIM includes a resources model, a reference model, and seven applications.

•Resource model makes reference to STEP generic resources as geometry, topology, material, tolerance, fundamental of production description data and support and product structure.

•Reference model concentrate on product definition data and covers different lifecycle stages.

•Product information model describes the product physical structure and its components.

•Activities information model describes different activities during product lifecycle, such as, design and manufacturing.

•Design information model describes design activity and design information.

•Process planning information model describes definition, product and activities in process planning stage.

•Manufacturing information model describes definition, product and activities in manufacturing stage.

•Using and maintenance information model describes definition, product and activities in using and maintain stage.

•Management information model includes the definition of management activity.

3.5 AIM

The purpose of AIM is the definition of the injection mould's auxiliary department. The auxiliary department of injection mould includes cooling and heating facility . From physical structure these departments are characterized by their shape, location, material, etc.. From function they has one or several functions. The model consists of a auxiliary department reference model, a functional model, a geometric and topological model and a management information model.

•Reference model is a high level model which finishes the auxiliary's definition in product lifecycle stage.

•Functional model describes auxiliary and its function, this model defined here, but it can be used or referenced by other model.

•Geometric/topological model describes the geometry and topology of auxiliary department, its entities come from SIM and Part 42 of STEP.

•Management information model manages the related information of auxiliary department.

3.6 Integration of Models

The IMPM is divided into four majority parts, every is independent and constructs a schema. In order to make it clean, some models use the common resources, and now, we combine some submoulds and make it suit the requirements of STEP.

Injection mould reference model is built, it includes two model(Injection Mould Application Reference Definition Model and Injection Mould Application Reference Representation Model). The first mode consists of SIM's reference model and AIM's reference model, it defines the activities from lifecycle view and the specialization of shape, material, principal character and function, etc.. The second model inherit the product definition attributes from the first model and represent various properties of product. All application can share these attribute sets of the reference model. Fig. 2 gives the structure of integration of models which is three levels recommended by STEP.

•Injection mould application reference schema

•Injection mould application schema

•Injection mould integrated resources schema

Injection mould integrated resources schema consists of Injection mould integrated application resources schema and STEP integrated generic resources schema.

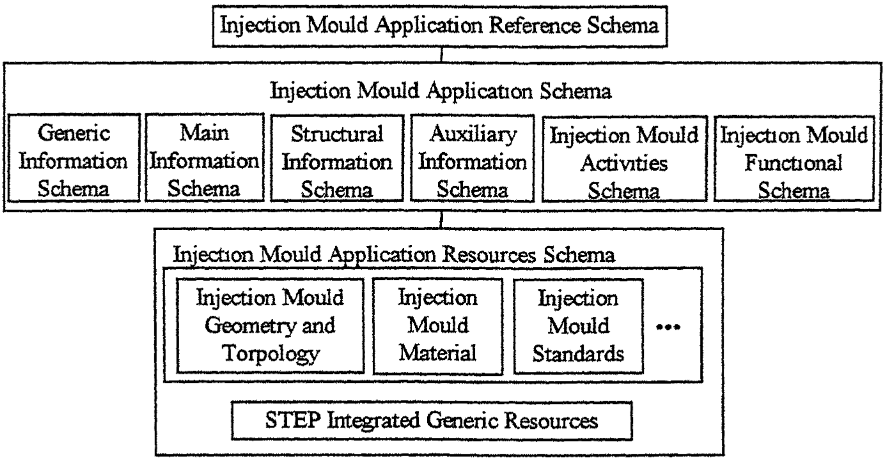


Fig. 2: Integration Structure of IMPM

4. CONCLUSION

This paper has described in detail the development of a STEP based injection mould product model, the model contains amount of information, and is fairly complex. STEP has the capability to provide a complete product lifecycle representation. IMPM can be used as the basic of integrated mould CAD/CAM system. We are studying currently IMPM based file exchange and database sharing.

5. REFERENCE

1. S.Finger, S.Konda and E.Subrahmanian, Concurrent design happens at the interface, pp.69-71, AI, Vol.9No.2, 1995.

2. Young.RIM and Bell.P, Machine planning in a product model environment, pp.2487-2519, Int.J. Product.Res.Vo.l30No.11, 1992.

3. Salomons.oow, Van Houten.FJA and Kals.HIJ, Review of research in feature-based, pp113-132, J. Manuf.Syst.Vol.12No.2, 1993.

4. Shah.JJ and Mathew.A, Experimental investigation of STEP form-feature information model, pp.282-296, Compute-Aided Des.Vol.23No.4, 1991.

5. Industrial automation system-exchange of product model data, Part 11:Descriptive method: EXPRESS language, National Institute of Standards and Technology, USA, 1991.

6. Industrial automation system-exchange of product model data, Integrated generic resources, National Institute of Standards and Technology, USA, 1991.

ASSEMBLY MODELING OF PRODUCT INFORMATION BASED ON SELF-ORGANIZATION

Tan Jianrong, Wei Xiuting and Huang Chao
State Key Laboratory of CAD & CG at Zhejiang University
Hangzhou 310027, P.R.China

ABSTRACT

Based on self-organization, this paper put forward a new method for modeling assembly information. It includes four conceptual models of products information: principle diagram model, extensive mechanism model, brief assembly diagram model, and assembly drawing model, and self-organizational mapping from brief assembly diagram model to assembly drawing. In different design phases, the above four models are alternatively used, and modified and iterated until a satisfactory scheme is gained. Based on the research done, It has been proved the four product models and transmission method for design information among them provide workable path to develop the CAD system supporting conceptual design.

KEYWORDS

CAD, Assembly Modeling, Conceptual Design, Self-organization

1. INTRODUCTION

Assembly modeling of product information is an important issue in computer aided design. The quality of the model has directly fluency on the structure, function and cost of products. Since the 1990s, many scholars have done lot of researches on assembly modeling. M.Mantyla[1] put forward the part-of graph which is consist of nodes and arcs, with the node representing design feature and the arc describing the relations among nodes. K. Lee[2,3] proposed an assembly model based on virtual-links in which two parts or two assemblies are linked by a virtual-link, and positioning of parts is determined by assembly relations. In this paper, we provide a new approach to model product assembly information based on self-organization of assembly information, and it has following aspects:

1) The conceptual modeling of product information for assembly. Reasonable representation of design process and data structure are the key to build a CAD system supporting the top-down design ideology. Now that the assembly drawings contain part information and matching information, parts must be simplified into symbols and the matching viewed as the relations among symbols in order to support conceptual design. And the product model is consist of those symbols and relations in conceptual design. At the same time, part information and matching relation can be defined and input in a comparatively free manner. To describe the process of conceptual design, brief assembly diagram works as representative tool.

2) assembly positioning of parts. In order to simplify assembly freedom, two match types: virtual-fit and virtual-against are introduced, which in fact are the extension of fit and against. Thus, parts can be positioned by solving a simultaneous non-linear equations. The number of equations and variables can be reduced by utilizing the sequence positioning in assembly process, with the result that the complex in solving assembly position is greatly decreased. In most cases, positioning of parts can be settled in brief assembly diagram.

3) Assembly modeling and redesign of parts. In engineering, product design is completed in three stages which respectively correspond to three model: brief assembly diagram, assembly drawing and detail drawing, and the design process is the iteration of these models, with the key technique being the information transmission through different design stages. Brief assembly diagram is the basis of assembly drawing, and assembly drawing is the foundation to design detail drawings. And brief assembly diagram might need modifying according to detail drawings. In three stages, important design information must be carried on.

2. THE CONCEPTUAL MODELING OF PRODUCT INFORMATION FOR ASSEMBLY

Generally speaking, designing product is an iterative process. To express designers' idea effectively, each of three design processes needs its own adaptive conceptual model. The assembly model , obtained from conceptual and assembling design, is a global model of product information and the basis for detail design. In the earlier design phases, designers always make decision as a whole and pay attention to functions, input, output and structure of a product. With the progress of design, the assembly model is given more detail information. The paper presents four product information models for assembly in order to describe the conceptual design process, as shown in Fig.1.

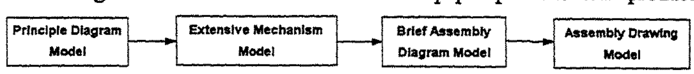


Fig.1

2.1 principle diagram model

In the beginning of conceptual design, designers determine the global scheme of a product according to its functions. Hence, design information is global and integral, i.e., functioning parts and their connection, input and output, with local details ignored. The model of product information is called principle diagram model in this design phase, as shown in Fig.2 where each block has a dynamic pointer linking a mechanism such as linkage and gear. Designers can use this model to make global design schemes and select a better one.

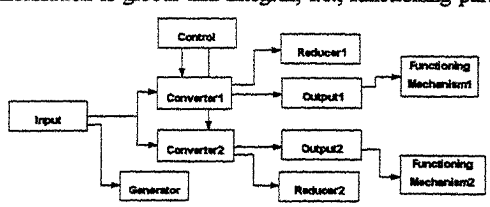


Fig.2

2.2 extensive mechanism model

Mechanism is a useful tool in concept design and described as a mechanism kinematic diagram. Extensive mechanism is derived from a mechanism by adding key structure information to it. For example, the linkage is a line through two link centers in mechanism while stood for by a structure graph through those centers in extensive mechanism, as shown in Fig.3. Designers can use extensive mechanism model to do kinematic and dynamic analysis, and attach some important information for structure design. Therefore, it is indispensable in connecting principal diagram model with brief assembly diagram model.

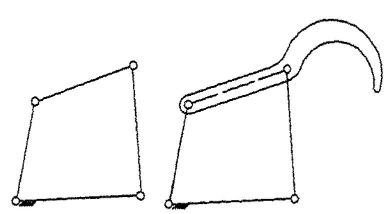

Mechanism Extensive Mechanism

Fig.3

2.3 brief assembly diagram model

Brief assembly diagram model is an important graphic tool to express designers' idea in the earlier design period, and contains parts and their positioning and matching information, instead of details, shown in Fig.6a. In this model, parts are symbols describing their structure and features. Generally, in conceptual design, parts with symmetry structure can be represented by their center line(called feature line), while others by their outline. In addition, standard parts and typical structures have their own symbols in some specification. As a whole, brief assembly diagram model expresses overall structure of a product and global design idea.

2.4 assembly drawing model

Brief assembly diagram and assembly drawing have mapping relations, and assembly drawing is detailization of brief assembly diagram. In assembly drawing, the structure, and the bases and positioning dimension of each part are completely determined, and so are the matching relations, shown in Fig.6c. In this model, feature line, outline and symbols in brief assembly diagram model have already been mapped into practical part shape.

The above-mentioned four information models cover the process from overall to detail in conceptual design and imitate the iteration progress from scheme design to structure design. The former three models can be further discussed as follows. Principle diagram model is a global design scheme and represents the information such as functions of a product and using what part or mechanism to realize them. Having been tested in kinematics and dynamics, extensive mechanism model ensure that the design scheme in principal model is correctly implemented. Brief assembly diagram contains the positioning and matching relations of product parts and the modification of these relations can be done more easily. Easy of modification is awfully wanted by iteration of design information in conceptual design. The whole design process is shown in Fig.4.

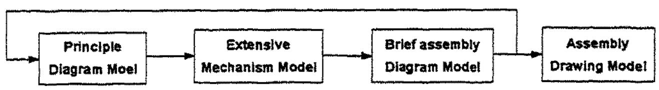


Fig.4

In different conceptual design phases, the above four models are used alternatively , and modified and iterated until a satisfactory scheme is gained. Because all the three models (principle diagram, extensive mechanism and brief assembly diagram) in cycle are conceptual model, modification and iteration can be conveniently achieved.

Above all, principle diagram including mechanism information and extensive mechanism containing important structure information make it possible to pass design information from one model to another. While design information is carried on in conceptual design, it also be developed from sketch to detail. The design scheme having been completed, assembling relation is determined, and the next task is to map from brief assembly diagram model to assembly drawing model.

3. SELF-ORGANIZATION MAPPING FROM BRIEF ASSEMBLY DIAGRAM MODEL TO ASSEMBLY DRAWING MODEL

3.1 Sequence solution to position of parts based on assembling sequence

After matching and positioning relations among parts have been determined and mapped to the assembly model, it is the key step to solve position of each part in the assembly drawing model. Essentially, positioning parts in assembly drawing is to adjust the part from one position to another, in order to meet the need of the assembly relation. The change of part position can be described as a 4 $\times$ 4 transformation matrix, and the method solving position of parts is usually called overall method[4].

Generally, the number of equations and variables in the overall method is very large. For example, determining a container assembly drawing made up of three parts involves 12 variables and 27 equations. With the increasing of assembled complexity, the number of equations and variables will be very large, and the solution to the equation group will be inconvergent. Besides, the superabundance equation will bring more difficulty to solution.

Having studied the assembling process of common engineering products, we learn that positions of most parts are determined automatically due to the assembling sequence. Therefore, we proceed a sequence method for positioning parts based on assembling sequence as follows:

If assembling positions of N matching parts(A1,A2,···,An) rely on its predecessor only, the position of each part can be determined according their assembling order by positioning one part a time. In other words, given A1's position, A2's assembling position can be solved according to matching relation between A1 and A2. Similarly, positions of A3 through An can be found out.

In assembly models, N parts, which can be positioned with sequence method, construct an open chain. Because of the complexity of products, open chains and close loops exist simultaneously in assembly models. If the assembling relations among N matching parts(A1,A2,···,An) construct a close loop, their positions are determined by all the N parts, and the overall method must be used to solve positioning problem.

Generally speaking, positions of most parts can be determined by sequence method, and the rest by the overall method. Combining the sequence method and overall method can greatly cut down complexity of the assembly positioning, with efficiency and reliability enhanced. For an assembly machine composed of N parts, after one part being fixed, positions of the other N-1 parts can be determined by using the algorithm in [6].

3.2 Construction of assembling drawing information and Self-organization of constraints

Based on the mapping between brief assembly diagram and assembly drawing, part drawing can be achieved by basic entities(line, arc, etc.) and graph in the graphic database. Currently, there are two methods for generating assembly drawing from part drawing: (1) According to assembly relations in assembly model, completing part drawings firstly and then using them to construct assembly drawing; (2) Using outline of parts to do boolean operation and to construct assembly drawing, then the hidden lines removed. Both methods require pre-treating part drawing manually before assembling. Moreover, whenever assembly relations or part dimension change, manual treatment is needed. To implement self-organization of the assembly model, outline information of parts must be automatically recognized, and the geometric and matching information of parts must be passed between models. When two parts suit the following requirement, there are geometric constraints between them:

1)two matching parts;

2)one part overlapped by the other partly or entirely;

3)two parts whose graphic elements are constrained by one dimension

Let Pi, Pj be two parts having graphic constraints and the algorithm for self-organizing graphic constraints can be described as follows:

Step1: Give feature code to each line of part Pi and Pj, if the coordinate x,y satisfy:

△ Y=0, it is a horizontal line, direction_code=0

△ X=0, it is a vertical line, direction_code=1

else, it is a general position line, direction_code=2 。

Step2: For each line segment of part Pi, identifying the constraints between it and line segments having the same direction-code in part Pj, if one of the following conditions is satisfied, there are constraints between the two lines:

Condition1: (direction_code=0)&&(y coordinates of two lines are equal)&&(x coordinate of one line is between two x coordinates of the other)

Condition2: (direction_code=1)&&(x coordinates of two lines are equal)&&(y coordinate of one line is between two y coordinates of the other)

Condition3: (direction_code=2)&&(one of end points of a line is on the other line)

Step3: Circulating each arc segment of part Pi and searching those having constraints with it in part Pj, if there are (center coordinates are equal)&&(radius are equal), the two arcs constrain each other.

3.3 Determining assembly model and redesigning parts

During modeling assembly information of products, determining assembly relations and designing part structure interweave each other and make a refining course. The design process can be expressed using Fig.5.

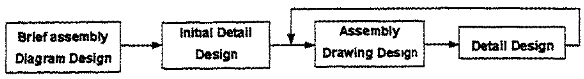


Fig.5

An excellent product assembly model should suit for the iterative course and make it possible and easy to transfer the information between assembly drawing and part drawing. On the one hand, the assembly information can be gained from part drawing, and then positions of parts are determined with the result that parts are assembled. On the other hand, part drawing could be taken apart from assembly drawing, with assembly information remained on parts.

The information transformation from part drawing to assembly model means that the assembly drawing will be automatically obtained from the given part drawing and their matching relations. The part positions can be solved with the method in 3.1. After matching relations and depth information among parts are given in interactive manner, assembly drawing can be automatically achieved by extracting outline information of parts, executing clip operation and removing hidden lines. The algorithms of recognizing part outline and clipping operation are discussed in[6].

The design for assembly drawing having been finished, structure and matching dimensions are determined. At this period, parts can be disassembled one by one, and their details can be proceeded based on assembly drawing, the machining and technology structure being thoroughly considered. Now that each part carries assembly information, the re-assembling for matching check is possible.

The inheritance of positioning and matching information can be realized by the extended entity data. The extended entity data is additional description(including number, symbol and mathematic expression) following the entity's geometry property(e.g., a line has end point coordinates, color and linetype, etc.). In the whole course of modeling product information, the transformation and inheritance of design information are passed through extended entity data.

4. CONCLUSION

The four conceptual models of products information(principle diagram , extensive mechanism , brief assembly diagram and assembly drawing), plus self-organizational mapping from brief assembly diagram model to assembly drawing model, make up an effective approach to model product assembly information and provide a new path to develop the CAD system supporting conceptual design. And ease of modification and implementation of assembly models shows that the methodology is workable in product design. Now, a set of algorithms modeling assembly information based on self-organization have been developed, and the example in Fig.6 is the design process of a reducer.

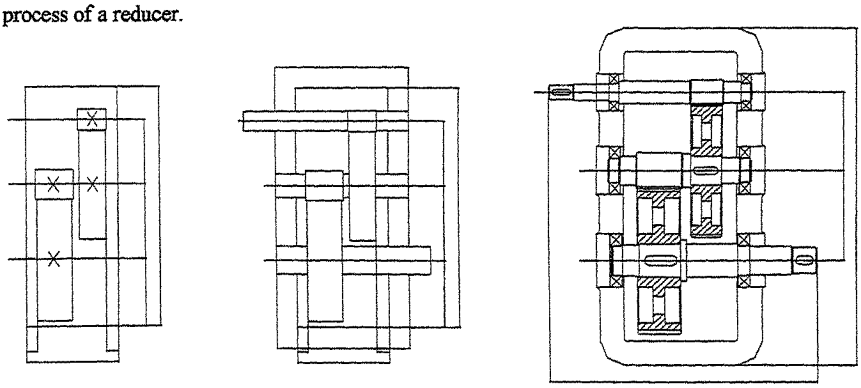

a) b) c)

Fig.6

5. REFERENCES

1. Mantyla,M. A Modeling System for Top-down Design of Assembled Products, pp.636-659, IBM J. RES. DEVELOP., Vol.34, No.5, Sep., 1990.
2. Lee, K. and Gossad, D.C. A Hierarchical Data Structure for Representing Assemblies:Part1, Computer-Aided Design, Vol.17, No.1, Jan./Feb., 1985.
3. Lee, K. and Gossad, D.C. A Hierarchical Data Structure for Representing Assemblies:Part2, Computer-Aided Design, Vol.17, No.1, Jan./Feb., 1985.
4. Rocheleau, D.N. and Lee, K. System for Interactive Assembly Modeling, Computer-Aided Design, Vol.19, No.2, Jan./Feb., 1987.
5. Turner, J.U. and Gupta, S. Constraint Representation and Reduction in Assembly Modeling and Analysis, pp.741-750, IEEE Transations on Robotics and Automation, Vol.8, No.6, Dec., 1990.
6. Chao, H., A Study on Conversion of Assembly Model to Detailed Model of Product Information, pp.37-46, pp. MD Thesis, Zhejiang Unervisity, 1996.

A UNIVERSAL GEOMETRIC INTERFACE IMPLEMENTED WITH STRUCTURED LANGUAGE

XU Shixin, CAI Hegao and WANG Yongjie
Robot Research Institute
Harbin Institute of Technology, P.R. China

ABSTRACT

It is an important issue for design and manufacture to transfer geometric data models among various CAD/CAM systems, and from CAD/CAM systems to NC systems. For this purpose this paper elucidates the design and implementation methods of a universal neutral geometric interface. Applying the already developed utilities of such interface, transferring tasks can be simplified and reduced greatly.

KEYWORDS

CAD/CAM, interface, transfer, compiler

1. INTRODUCTION

The geometric data models applied in various CAD/CAM systems are usually different. So the design results obtained in one system are difficult to be used in another. It is important to make communications among CAD/CAM systems, and also to standardize the design results obtained by different geometric data models. To tackle this problem, we put forward the idea of geometric interface and implemented it in our featured-based CAD/CAM system (fbCADM). We designed a universal fbCADM geometric interface (CADMgi), a complete geometric modeling language intended to function as a neutral geometric interface between the fbCADM system and other CAD/CAM systems. With the CADMgi, the fbCADM system is easily able to transfer geometric data models to and from other CAD/CAM systems. Thus the CADMgi facilitates the translation of geometric data from other systems, and the rendering of fbCADM data accessible to users of other systems.

2. DESIGN OF CADMgi

2.1 General Descriptions

A CADMgi file is identified by a name acceptable to the fbCADM system and which therefore meets two requirements. One is that the name is an alphanumeric string. The acceptable characters are numerals (6), lowercase letters (a), the period (.) and the underscore (_). The first character must be a lowercase letter. The other is that the last three characters of the name must be .gi .

A CADMgi file constitutes strings of ASCII characters together with various logic characters. These characters are reserved and may not be used for purposes differing from those described below. A slash (/) separates the name of an instruction from its arguments. Commas (,) are used as delimiters between the elements of a command. A semicolon (;) terminates a command. Newlines (NL) follow commas or semicolons to divide the file into strings having no more than 80 characters.

Three different token types occur in CADMgi files. They are integer numbers, real numbers and identifiers. Integer numbers are signed whole numbers; for positive values the + may be omitted. They are expressed as +/- xxxxxxxxxx , in which x is a digit between 0 and 9. Real numbers are signed decimal fractions either with or without an exponent; for positive values the + may be omitted. The decimal fraction must be expressed; thus 1.0 is a valid real value; 1 is not valid. Real numbers without an exponent may be expressed as +/- xxxxx.yyyyy , and real numbers with an exponent similarly may be expressed as +/- xxxxx.yyyyy E +/- zz . In these expressions x, y and z are digits between 0 and 9. An identifier is an alphanumeric string not exceeding 32 characters in length. Permissible characters are numerals, lowercase letters, and the underscore; the first character must be a lowercase letter.

2.2 Design of CADMgi Commands

For clarity, each command in a CADMgi file should occur on a separate line.The syntax of it is

COMMAND_NAME / *argument* [, *argument*];

The components of command syntax are: (i) The command name must be expressed in UPPERCASE type as shown. (ii) The slash / is required; it separates the name of the command from its arguments. (iii) Italic type indicates command arguments for which the user must supply a value. (iv) Commas , are required; they separate individual arguments. (v) Square brackets [] indicate optional arguments. (vi) A semicolon ; terminates the command.

2.2.1 Structure Commands

CADMgi files are structured as blocks. Start and finish commands delimit the entire file and also the blocks within each file.

(1) file delimiters

Every CADMgi file begins with a header command and concludes with an end command.

HEADER

[description] Constitutes the first line of every CADMgi file. It identifies the file with a date and the name of its author.

[syntax] HEADER / *mm, dd, yy, writer_name* ;

[arguments] *mm, dd, yy* --A date in the format month, day, year, where m, d, and y are digits between 0 and 9, and the commas are required. Entries should be restricted to legal dates.

writer_name --A name, usually intended to represent the author of the file, of type identifier.

END

[description] Constitutes the last line of every CADMgi file. It terminates the preprocessor.

[syntax] END ;

(2) component block

Every CADMgi file constitutes a single block of data used to define a component. Except for the two file start and termination commands, all CADMgi commands occur within component blocks. A component statement opens every component block, and an end statement terminates the block; these statements identify the component to which they pertain.

COMPONENT

[description] Introduces a block of data defining a component, and assigns a name to the component. It constitutes the first line of the block.

[syntax] COMPONENT / *name* ;

[argument] *name* --The name of the component, of type identifier, It may not have more than ten characters. The CADMgi compiler appends a .co suffix to this name.

END

[description] Terminates a component block and its last line.

[syntax] END / *name* ;

[argument] *name* --The name of the component as assigned by the component command, the first line of the block.

2.2.2 Solid Entities Description Commands

CADMgi depicts solid figures in a facetted boundary representation. This representation is facilitated by a logical structure of nesting blocks and commands which enable defining solid figures having many cavities as desired. A CADMgi solid figures is constructed as follows: (i) Points constituting the vertices of the faces are defined: the points array command. (ii) Boundaries are constituted from previously defined points: the polyloop command. (iii) Faces are constituted from one or more boundaries: the face block. The first boundary is the edge of the face; subsequent boundaries are voids in the face. (iv) Three-dimensional polyhedral figures are constructed from faces: the shell block. (v) The solid figure is constructed from one or more polyhedral figures: the polyhedron block. The first polyhedral figure constitutes the peripheral boundary of the polyhedron; subsequent figures are cavities within the polyhedron. (vi) The validity of each shell in the solid figure is verified to ensure that every edge appears twice, each time in the opposite direction. Files containing shells violating this rule are rejected.

(1) **polyhedron block**

The polyhedron block contains instructions and shell blocks which define a solid in a facetted boundary representation. A polyhedron may contain an unlimited number of non-overlapping internal shell blocks. The polyhedron block occurs only within a component block. It comprises an opening instruction, a single points array commands, one or more shell blocks, and a closing instruction.

POLYHEDRON

[description] Constitutes first line of the polyhedron block. It identifies the block with a name.

[syntax] **POLYHEDRON** / *identifier* ;

[argument] *identifier* --The name of the polyhedron, of type identifier.

END POLYHEDRON

[description] Constitutes the last line of a polyhedron block, and terminates the block.

[syntax] **END_POLYHEDRON** / *identifier* ;

[argument] *identifier* --The name of the polyhedron, of type identifier.

(2) **points array command**

CADMgi provides a facetted boundary representation of solids; the POINTS ARRAY command defines points intended to define the faces of these solids. Because a single POINTS ARRAY command defines all of the points used throughout the solid, it is used within the polyhedron block

POINTS ARRAY

[description] Defines a list of points by their Cartesian coordinates. These points are used by subsequently issued polyloop commands within face blocks, in order to define faces. This command may be used only once within a polyhedron block. It is placed ahead of the polyloop commands which utilize the points. The faces constructed from the points are used to construct a polyhedron.

[syntax] **POINTS_ARR** / *n, x0, y0, z0, ... , xn, yn, zn* ;

[arguments] *n* --The number of points constituting the vertices: an integer value. It must the same as the number of X, Y and Z coordinate values.

x0, y0, z0, ... , xn, yn, zn --The X, Y and Z Cartesian coordinates of each of the points intended to constitute the vertices of a planar face: real values separated by commas.

All points must lie on the same plane. All points must be used by subsequently issued POLYLOOP commands. The points are assigned index numbers incrementing from 0 in the order in which they appear in this command. The POLYLOOP commands refer to the points by their indices.

(3) **shell block**

The shell block contains instructions and face blocks which define all or a portion of a polyhedron. All pairs of adjacent faces must have a common edge; the direction of the edge of one face must be opposite to that of the corresponding edge of the other face. For each polyhedron, the first shell defines its peripheral boundary. Subsequent shells define internal boundaries within the first shell; these boundaries represent internal voids. The voids formed by the second and subsequent shells within the same polyhedron may not overlap each other, nor may they extend outside of the shape defined by the first shell. The shell block occurs only within a polyhedron block. It comprises an opening instruction, one or more FACE blocks, and a closing instruction.

SHELL

[description] Constitutes the first line of the shell block. It identifies the block with a name.

[syntax] **SHELL** / *identifier* ;

[argument] *identifier* --the name of the shell, of type identifier.

END SHELL

[description] Constitutes the last line of a shell block, and terminates the block.

[syntax] **END_SHELL** / *identifier* ;

[argument] *identifier* --The name of the shell, of type identifier.

(4) **face block**

The face block is a group of statements defining a planar area bounded by polygonal boundaries. These boundaries include one external boundary, either alone or together with one or more internal boundaries representing embedded voids. The face block currently can occur only within a shell block, where it is used together with other faces to constitutes a polyhedron.

FACE

[description] Constitutes the first line of the face block. It identifies the block with a name.

[syntax] **FACE** / *identifier* ;

[argument] *identifier* --The name of the face, of type identifier.

END FACE

[description] Constitutes the last line of a face block, and terminates the block.

[syntax] **END_FACE** / *identifier* ;

[argument] *identifier* --The name of the face, of type identifier.

(5) polyloop command

CADMgi provides a facetted boundary representation of solids; the polyloop command determines the boundaries for the faces of these solids. It is used within the face block.

POLYLOOP

[description] Connects designated points with straight lines to form a continuous loop constituting a planar, polygonal faces. Lines connect the points in the order in which they are listed, and a line also connects the last point in the list with the first point. Each loop thus has a sense of direction which proceeds from the first point through intermediate points to the last point, and then back to the first point. When a loop is later used as the face of a polyhedron, the sense of direction of the loop is counterclockwise as viewed from the outside looking into the solid. All of the loops must have the same direction.This command is used only within face blocks. If it is issued more than once in a single block, the first loop must enclose all of the subsequently generated loops. The first loop thus constitutes the face, and subsequent loops constitute voids within the face. If a face has no voids, a single face block contains only one polygon instruction. Loops may not cross other loops, nor may they cross themselves.

[syntax] **POLY_LOOP** / *i0, i1, ... , in* ;

[arguments] *i0, i1, ... , in* --Points previously defined by a POINTS_ARRAY command, listed by their respective index numbers separated by commas. This command constructs straight lines from *i0* to *i1*, from *i1* to *i2*, ... from *i(n-1)* to *in*, and from *in* back to *i0*.

The points may be listed in any order which satisfies all of the following criteria: (i) The loops so formed may not cross itself, nor other loops. (ii) The first loop defined within a face block must enclose all other loops defined in the same face block. (iii) Second and subsequent loops defined within a face block may not themselves enclose other loops defined in the same face block. (iv) When view from the same direction, the lines constituting all loops within the same face block must have the same sense of direction. (v) When the face is later used to construct a polyhedron, the sense of direction of the lines constituting the face is counterclockwise when viewed from outside looking into the polyhedron.

3. IMPLEMENTATION OF CADMgi

We designed two tools in the fbCADM system to implement the geometric interface, CADMgi : A compiler for files to be brought into fbCADM, and a generator which accepts fbCADM component files (the name of which have .co suffix) and translates them to CADMgi files. The compiler checks the syntax of the input files and generates calls to fbCADM libraries in order to produce fbCADM components. These utilities use and produce CADMgi files which may then be converted to the format of another geometric modeling system as desired.

The compiler's functions mainly include the lexical analysis, parsing, semantic analysis and corresponding error process, whose relations are shown in Figure 1. The main task of lexical analysis is to identify words from the character stream of the CADMgi file. The procedures of the compiler's lexical analyzer are shown in Figure 2. Based on the lexical analysis, parsing is to check the grammar of the CADMgi file. We adopted the LR method for the compiler, which is a kind of bottom-to-top method. Semantic analysis makes the computer understand the operations represented by statements. It involves the maintenance of symbol tables and code generation, etc.. The user's programs usually contain some errors. So error process is an important part for the compiler. The compiler devised an error recovery method. Once encountering an error, the compiler usually can report the error message and the position of the error, then it skips over the error and continues the analyses. On the basis of the above ideas we implemented the compiler with the yacc tool supplied by the UNIX system in the SGI workstation.

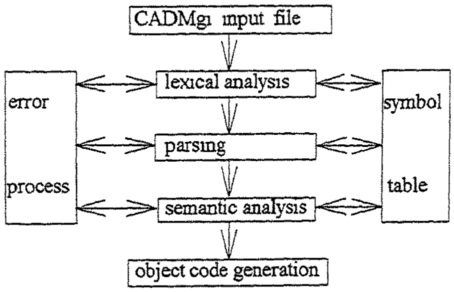


Fig. 1 The compiler model

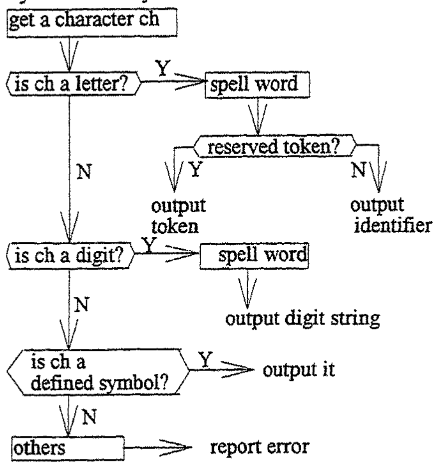


Fig. 2 Lexical analysis diagram

The generator is actually a converter for file format translating. It converts a component file (the name of which has .co suffix) into a CADMgi file (.gi file) according to their corresponding syntax. The method to implement the generator is omitted here since it is relatively simple.

4. EXAMPLE OF A CADMgi FILE

The following is a CADMgi file representing a prism pipe (shown in Figure 3) in the fbCADM system. We create the file by utilizing the generator tool that accepts the corresponding .co file.

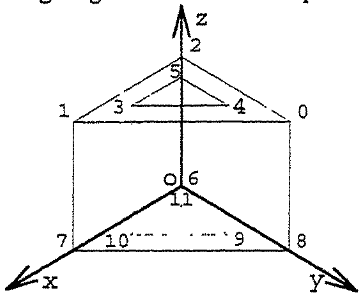


Fig. 3 A prism pipe

```
HEADER / 5, 20, 96, XU_Shixin ;
COMPONENT / prism_pipe ;
POLYHEDRON / bo1 ;
POINTS_ARR / 12, 0.0, 1000.0, 1000.0, 1000.0, 0.0, 1000.0, 0.0, 0.0, 1000.0, 646.5, 146.5, 1000.0,
146.5, 646.5, 1000.0, 146.5, 146.5, 1000.0, 0.0, 0.0, 0.0, 1000.0, 0.0, 0.0, 0.0, 1000.0, 0.0, 146.5,
646.5, 0.0, 646.5, 146.5, 0.0, 146.5, 146.5, 0.0 ;
SHELL / SH0;
  FACE / FA0;
    POLY_LOOP / 2, 1, 0 ;
    POLY_LOOP / 3, 4, 5 ;
  END_FACE / FA0;
```

```
  FACE / FA1;
    POLY_LOOP / 8, 7, 6;
    POLY_LOOP / 9, 10, 11;
  END_FACE / FA1;
  FACE / FA2;
    POLY_LOOP / 1, 7, 8, 0;
  END_FACE / FA2;
  FACE / FA3;
    POLY_LOOP / 2, 6, 7, 1;
  END_FACE / FA3;
  FACE / FA4;
    POLY_LOOP / 0, 8, 6, 2;
  END_FACE / FA4;
  FACE / FA5;
    POLY_LOOP / 9, 10, 3, 4;
  END_FACE / FA5;
  FACE / FA6;
    POLY_LOOP / 10, 11, 5, 3;
  END_FACE / FA6;
  FACE / FA7;
    POLY_LOOP / 11, 9, 4, 5;
  END_FACE / FA7;
END_SHELL / SH0;
END_POLYHEDRON / bo1;
END / prism_pipe;
END;
```

5. CONCLUSIONS

The geometric interface CADMgi offers important advantages: (i) Modifications to the fbCADM system do not affect the CADMgi user interface; user-written custom files intended for CADMgi remain compatible with subsequent versions of the fbCADM system. (ii) CADMgi is designed to be user friendly. Its English-like syntax facilitates its use by users of the fbCADM system. (iii) CADMgi enables a two-way data transfer between fbCADM and other CAD/CAM systems. CADMgi files can be generated automatically from fbCADM components, constituting them as neutral output files. (iv) CADMgi can be used either to write a simple application off-line, or to translate an external CAD database. (v) New entity types can be added to the CADMgi language with minor additions to the syntax and without changing existing instructions. (vi) CADMgi accommodates both CSG and boundary representations of solids.

6. REFERENCES

1. CHI Zhongxian, et al, Compiler Writing, Science Publishing House, P.R. China, 1992.
2. IRIS-4D Programmer's Guide (Volume I), Silicon Graphics Incorporation, 1990.

BUILD INTEGRATED CAD/CAM SYSTEM BASED ON SOFTWARE REVERSE ENGINEERING

Yiping Lu and Jianping Lin
Hebei University of Technology, Tianjin 300130, China
Ji Zhou
HuaZhong University of Science and Technology, Wuhan 430074 ,China

ABSTRACT

Because the quickly changing of computer technology, the realization of industry CAD/CAM is a progressive procession. Creating new system with advanced fuction upon old system is economic considering both time and financial expenses. Reverse engineering(RE) is the key point in this task In this paper, we generate a hypothetical constitution by analyzing the natural data relation and data flow of CAD/CAM system's object world. RE result is obtained by verifying and documenting items of the constitution.A mechanical integrated CAD/CAM system building upon a sole mechanical CAD system is given as an example.

KEYWORDS

Reverse Engineering, CAD, CAM , CAD/CAM

1. INTRODUCTION

In this paper , we will discuss building integrated/advanced CAD/CAM system based on current system's Reverse Engineering (RE). Our research is not a plain RE problem, it is a FE (Forward Engineering) based on RE.

Although RE can be defined a meaning clearly separated from FE[1], in application questions, the two conception often need together discussion. Chikefasky and the others[2] define RE in this way: the objectives of RE i.e. the production of documents that can increase the overall understanding of software system and that will be used in FE activities. While this definition draws a clear-cut between RE and FE, it also shows the connection among them. P. Benedusi and the others[1] point out: to guide RE, there should be a goal. In ours topic this goal is determined by FE.

The significance of FE in CAD/CAM lies on the practicing meaning of CAD/CAM study. As an Engineering technology, there is no problem that CAD/CAM should be researched for the goal of industry application. To cope with the quickly advancing computer technology, realization of industry CAD/CAM application have to be a progressive procession. There are two CAD/CAM software compositions: (1)general CAD/CAM software supporters, which are often some commercial CAD/CAM software packages (2)specified CAD/CAM software, which are often solely or nearly solely developed for the specific application. Renewing of the former is at the financial expense, the later at both financial and time expense. Sometimes the latter expense is too heavy to tolerate without a FE process. Practically, in most Chinese manufacture factories, confined by limited personnel technology lever, the development of CAD/CAM is gradual, building high lever CAD/CAM system upon old software resources is important. While we extol the efficiency of a new generation integrated CAD/CAM system, we cannot ignore its investment. Theoretically, it is relatively easier and somehow more reasonable to build an integrated/advanced system all from start than to form a counterpart by renovation and incorporate old software resources. But the former need much time/finance expense, and risk much system reliability since the new system may take time to be tested. Usually, current CAD/CAM software often contains very valuable resources, a new advanced system shell never blindly discard them.

RE serves a key point in RE based FE. We do our CAD/CAM software RE by focusing on the analysis of current CAD/CAM system's construction. After analysis, current CAD/CAM software's resource can be reformed and be reused.

CAD/CAM system has its own distinction created by CAD/CAM data's nature relations and their logic flow direction, this helps us to conceive a general data structure which we suppose the CAD/CAM system will reflect . We believe that this data structure is inherently within the objects of CAD/CAM, and will inevitably occur in the system, no matter what appearance they take. This conceived data structure is to be forward to form a hypothesized program construction of the current CAD/CAM software. Corresponding to a Goals/Models/Tools paradigm[1], this hypothesis program construction serves as one of our major tool.

Our example is a mechanical CAD system (without CAM function), which is to be reformed to a CAD/CAM integrated system.

2. CAD/CAM SOFTWARE REVERSE ENGINEERING

2.1 Consideration of the Reverse Engineering Goal

The first thing to do in CAD/CAM RE is defining the RE Goal. As a RE process serving a application FE, the CAD/CAM RE goal has practicing meanings. Our conception of FE activities is considered as software resources reuse and the reformation/addition of general software's function. To facilitate such FE activities, the goal of our RE should support the implementation lever, that means it should emphasis the recognition of the software resources, not only their existence and interrelations but also their form and their detailed formation description. The scope of RE can also be limited by specific FE aim.

The documentation should have detailed descriptions of software resources which is convenient for a implement lever usage.

2.2 CAD/CAM Software's Hypothetical Constitution

2.2.1 The General Hypothetical Constitution

From the viewpoint of software reusing, the general hypothetical constitution can be shown as following:

	Knowledge Bases	
Inputs,	Internal Mechanism Sets	
Human interference	Internal Function Sets	Outputs
	Static Data Bases	
	Interval Data sets	

Software reverse engineering can be attained by recognize systematically above constitution items. Recognizing difficulty increasing sequence and a better operation step can be:

Outputs, Inputs, Human Interference
|
\|/
Static Data Bases, Knowledge Bases
|
\|/
Internal Function Sets, Interval Data Sets
|
\|/
Fnction Internal Mechanism Sets

Among these items, the outputs, inputs, and human interference are the most transparent items, we just clearly document them. The items on the second lever of recognizing difficulty are static data bases and knowledge bases. Both of these tow items have the tendency to take relatively unattached implement forms, making them easier to be recognized. The interpretation of a knowledge base is little more difficult than that of a data base, since the organization of the former is often more complicate. The internal function sets and interval data sets are the items which have a greater recognizing difficulty. But the recognition this two types of items is often necessary, since a FE often need operations such as a data interception and redirection. Compares to static data bases, interval data sets may be entwined within internal functions. If this is the case, RE have to be extended to the task of breaking up some internal functions, i.e. the recognition of function internal mechanism sets. Function interval mechanism sets are the items which have the most recognition difficulty, often need source code lever RE.

2.2.2 The detailed Hypothetical Constitution

RE can be made as a hypothesis - verification/recognition - documentation process. The detailed software hypothesis constitution can be reached by first analyzing CAD/CAM object's entity-relations, then forward them to a logic data flow and finally a detailed logic system construction. This detailed logic system construction reveals the specified constitute items discussed in 2.2.1. Those items may take different appearance and may entwined among themselves in the real implementation of current CAD/CAM software, but they do logically exist in the implementation. A system with trim compositions will make great advantage for RE and FE process.

2.2.3 Constitution Verification and Recognition

The item recognizance work needs the cooperation of both the domain knowledge and programming knowledge. This work includes technical job such as hyperthyroid verification, data finding, data explanation, the difficulty depends the transparency of the resource.

2.3 REUSE OF THE RESOURCE

Current CAD/CAM software can be reused in these ways: (1)As reference for a totally new software; (2)Make use of abstracted useful databases and knowledge bases; (3)Reconstruction the resource by utilize the interval data or functions.

3. AN INTEGRATED CAD/CAM EXAMPLE

Here we give a machine tool shaft box integrated CAD/CAM system as an example. In this example, a shaft box CAD software with an output of engineering drawing is analyzed, it's static database is extracted and recreated, the interval data is captured. A new created CAM section is connected to form a new CAD/CAM software package.

3.1 FE Task

The CAD software is a combination machine tool shaft-box design software, with an output of the shaft-box's design drawing. The cube-shaped shaft-box is complicated by dozens of different shaft pivot holes. These holes will be machined by NC tools. Without a CAM module, the conversion from design drawing to NC code has to be down manually, induce not only time consummation but also inevitable error. To convert this CAD software to a new CAD/CAM software is our FE task.

3.2 Current software's Reverse Engineering

3.2.1 Current CAD Software General Function Analysis

The general function of current software can be determined by observation of the software's running procession, which is divided into four phases, showing in fig 1.

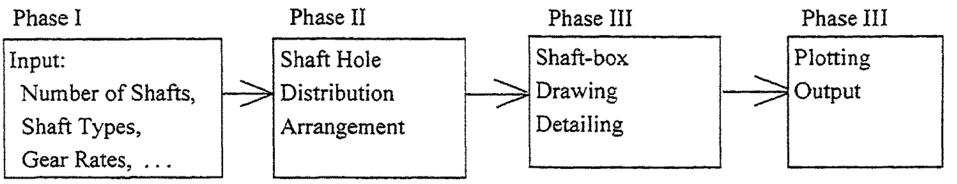


Fig. 1 : Current Software's Running and General Function

Among these four phases, the first and the fourth phase is the input and output phase. The second phase is apparent because it need some interactive human interference, and can be easily detected. The third phase is predicable because the requirement of phase IV. Our FE task need an interception of a data flow for CAM moudle's input. Obviously this data flow should between phase II and Phase IV. Phase II do not need our much attention, our RE goal is limited in phase III.

3.2.2 Current software's hypothetical constitution

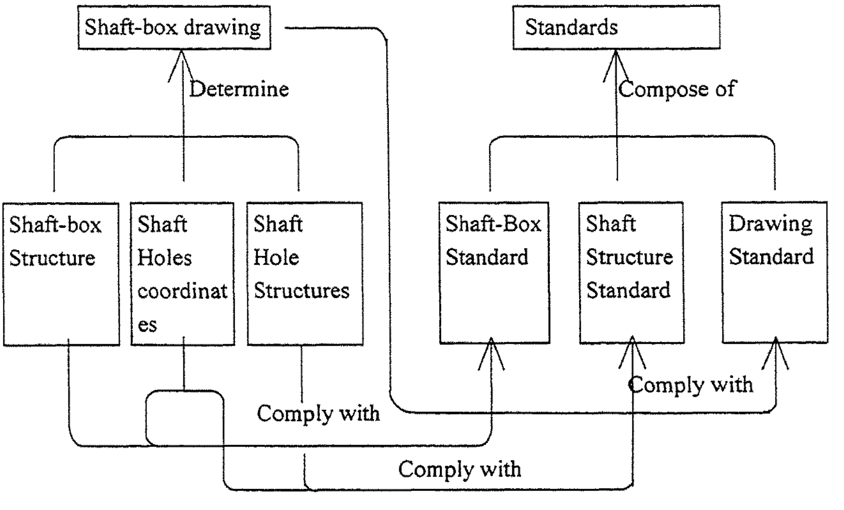


Fig. 2 CAD Objects' Entity-relations

The construction of the hypothetical constitution begins with the analysis of the CAD software's object entity-relations. Fig. 2 shows part of the entity relations. This figure reflects the design rule of the shaft-box. The shaft-box drawing is determined by the combination of shaft-box structure, shaft pivot hole coordinates and shaft pivot hole structures. Each of these three entities must comply with their relevant standards. After this analysis, we acquires enough confidence that there should exist data bases which correspond the three standard in the software's running process phase III. A hypothetical constitution showing in fig. 3 is constructed. The three ellipses: A, B and C, stand for three static data bases. Label a , b stand for interval data flows.

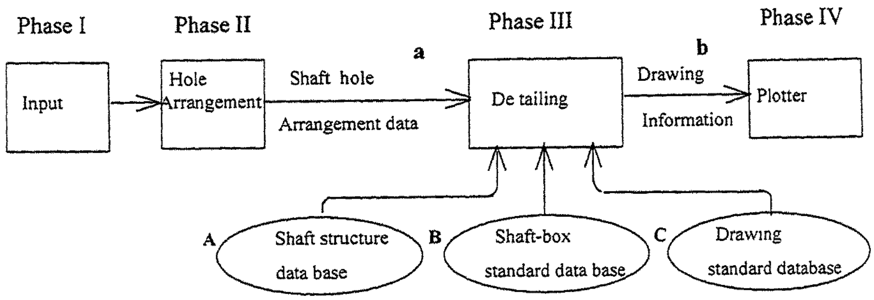


Fig. 4 Current Software Hypothetical Structure

3.2.3 The Recognition and Verification

The task of this step is to recognize and document the form of conceived items shown in fig. 4. Of all of the items, we are specially interested in static data base A, B and interval data a. Using base computer technology, we find static data base A and B take the form of one file, which are documented and interpreted as following lists:

[shaft diameter] [shaft type code] [shaft usage and structure descriptions]

```
20,T0721-41,7204,20*47*15-2,0,0,0,0,0,0,0,0,0,T0731-51,20*30.5,T0731-51,20*30 5,0,0,0,0,DB1027,M6-2,T0721-81,30,T0721-
87,30,T0721-82,30,T0721-83A,30,0,0,L=115,!
25,T0721-41,7205,25*52*16-2,0,0,0,0,0,0,0,0,0,T0731-51,25*29 5,T0731-51,25*29 5,0,0,0,0,GB78,M8*10-2,T0721-1,38,T0721-
87,38,T0721-82,38,T0721-83A,38,0,0,L=115,!
..      ...          ...      ..      .
```

The interval data a is also in the form of a file, which is documented and interpreted as follow:

```
u603-71END                                            Design Title
6,10,BZU603-7101W,630*630*4,0,0,!                     Shaft-box description
0,0,265.000,130.000,40,0,480,3250233,!                shaft arrangement data
1,T0722-41/0,185.000,112.000,20,42,375,2410352,!
2,T0722-41/0,185.000,227 000,20,42,375,2410352,!
.     ...     .     ..     ...
```

3.3 The Integrated CAD/CAM System

Using RE documents, a integrated CAD/CAM is created. The system's construction is showing in fig. 5.

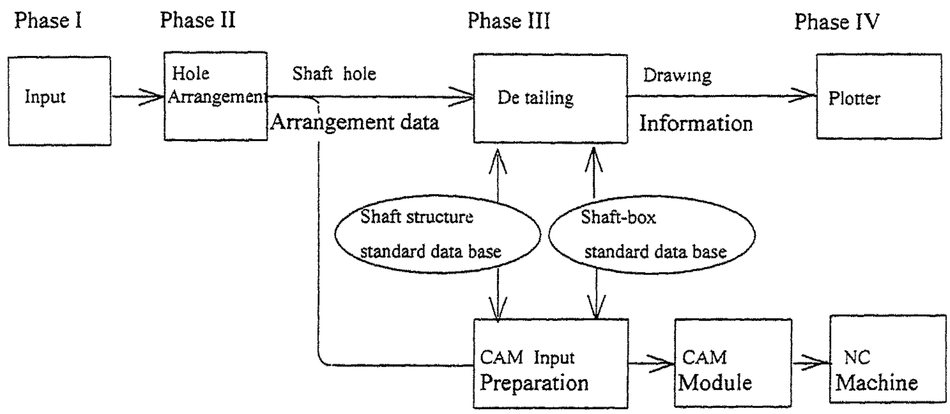


Fig. 5 The Integrated CAD/CAM Software Structure

4. CONCLUSION

In this paper we discussed software reverse engineering (SRE) in CAD/CAM system's building. SRE plays a key role in such FE activities. Not only the SRE be a way to quicken a system's creation but also it will be a way to enhance software quality and a way to spread qualified software techniques. SRE should be done systematically, the software hypothetical constitution method discussed in this paper is pragmatic. Some bottom data and program analyzing tools are also very helpful from the technique's viewpoint.

5. REFERENCE

1. P. Benedusi. A. Cmitile and U. De Carlini, 'Revere Engineering Processes, Design Document Production, and Structure Charts', J. Systems Software, 1992, Vol. 19, pp. 225-245

2. E. j. Chikofsky and J. H. Cross II, 'Reverse Engineering and Design Recovery, A Taxonomy', IEEE Software 1990, Vol. 7, pp. 13-17

3. Dalian Institute of Combined Machine Tool, Combined Machine Tool Design Manual, Vol. 1, pp637-664, Machinery Industry Press, Beijing, 1975

DESIGN AND DEVELOPMENT OF MATERIAL MANAGEMENT SYSTEM IN CIMS OF IRON & STEEL COMPLEX

Zhiwen Tang
Baoshan Iron & steel Group

Dingwei Wang
Northeast University
Automatic Control Department

ABSTRACT

This paper introduces the design and development of material management system in the CIMS of Iron & Steel Complex. The following content will be included: the main design idea, the system objectives, management transaction procedure and some problems which exist in the procedure, the structure and function of the system, the main technology which used in developing the system and the benefit which is achieved after the system's establishment.

KEYWORDS

Management information system; Network; Database technology; application integration technology; Business Process Reengineering

1.INTRODUCTION

CIMS is an abbreviation of Computer Integrated Manufacture System. In this system, we must integrate the automatic equipment in every different levels to let them work together smoothly. With CIMS, we can increase the labour productivity, improve product quality and get better management efficiency and economical benefit. In recent years, almost every large scale Iron & Steel enterprise in the world is engaged in developing the CIMS. The CIMS can be divided into three types: economical management, machine manufacture and material management. The leader of enterprise must (according to the production management requirement and the automation condition) decide to develop or not to develop the CIMS.

The construction project of Baoshan Iron & Steel Complex is divided into three periods. There are two phases in the development of the CIMS of Baosteel .The first CIMS phase is built for the first and second period of Baosteel's construction project. The second CIMS phase is built for the third period of the Baosteel's construction project. There are five subsystems in the first CIMS phase. They are production planning, quality analysis, equipment maintenance, raw material management and financial management. In this paper, we introduce the design and development of the material management system in the first CIMS phase of Baosteel.

2.THE BACKGROUND OF SYSTEM DEVELOPING

In the first and second period of Baosteel's construction project, we imported 69 process computers and 4 middle scale computers for local management. All these computers made contributions to Baosteel's production control. As a modernized Iron & Steel Complex, there are requirements not only for direct production control but also for business management. In our complex, the production is in large scale, the production procedure is automatic and continuous, so we must supply materials in time to ensure the steady and continuous production and satisfy the production requirements. Now, the material management in our complex is not efficient and there are too many management levels. In order to improve these conditions, we must integrate the modernized management idea and computer technology together and develop the computer-aided material management system.

3.THE MAIN IDEA OF SYSTEM DEVELOPING

At first, we investigated the management transaction flow. Then, we built the information flow according to the transaction flow and set up a database. At last, we realize the management function on the basis of information flow.

The main ideas are listed in following:

a) To establish a material computer management system which is suitful to Baosteel as a Modernized Iron & Steel Complex.
b) To combine the modernized material management idea closely with the information technology by using the principle of Business Process Reengineering.
c) In order to build this system, we utilized the modernized management idea, computer technology, network technology and Database technology.

4.THE OBJECTIVES OF THE SYSTEM

a) To produce only one account so we can share all the data in Baosteel's area .
b) To put terminals in every material management department(in the area of 12 square kilometres) to form a complete computer management network. So, the speed of information transmission is highly increased.
c) To separate the information flow, material flow and fund flow, to wipe out the false "in way" information, for example the materials has already come, but the information of the material can't be recorded on time because of the slow manual management, so the manager consider the arrived materials hasn't come. All of these make the inconsistency among the material flow, information flow and fund flow.
d) One data can be recorded only one time and the data must be inputted into computer where it is produced.

5. THE TRANSACTION PROCEDURE AND PROBLEMS IN MATERIAL MANAGEMENT

5.1 The Problems In The Original Management

The material management is necessary in the modernized Iron & Steel complex. In Baosteel, the types of the material are various and the amount is large. The types of hardware, chemical and metal materials are several ten thousands, the quantities are several hundred thousands. They are frequently used and occupy several hundred million RMB yuans floating capital. There are planning, managing, purchasing and warehouse department which are engaged in manage materials of Baosteel. Under the manual management, the information is transmitted by the hand-written sheets. Because these sheets are repeatedly made copy by hand, there are errors produced in these sheets. At the same time, the information can't be transmitted to the suitable Dept. on time, the fund circulates slowly and the stock data can't be shared. Because all of the above, the stock accounts are not accurate. According to these accounts, the manager can't make a perfect purchasing plan avoiding overstocking. These conditions don't tally with the requirements of a modernized Iron & Steel complex and Baosteel suffered heavy loses.

5.2 The Transaction Flow Of Computer Management

As the development of computer technology and communication technology, people apply the information technology almost in everything. Most enterprises in the world utilize the information technology for their own development. Hammer advanced a new idea about Business Process Reengineering in 1984[1].He believes that we must improve our transaction flow while using computer for management. Otherwise, we can't use the computer efficiently. As soon as this idea was put forward, it was prevalent in developed countries and it becomes the symbol of business management revolution in 1990's. The kernel of this idea is the management procedure. He considers that the basis of information system is not the structure of business departments but the management procedure. The important things in developing computer systems are improving the management procedure and producing a reasonable transaction flow. With BPR, we modified the material management transaction flow and formed a scientific integrated management transaction flow as following: accepting the purchasing plan, plan balance , forming the purchasing plan, purchasing management, material into and out of warehouse, stock management and settling financial accounts. As a result, information can be transmitted in time between plan, management, purchasing and stock department .

The transaction flow will be illustrated in figure1.

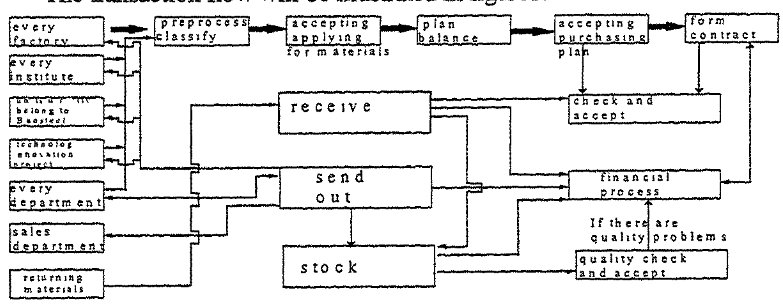


FIGURE 1 MANAGEMENT TRANSACTION FLOW

6.THE STRUCTURE AND FUNCTION OF THE SYSTEM

6.1 The Structure Of The System

Baosteel's material computer management system is developed and running on the UNISYS 2200/600.

UNISYS 2200/611 is a large scale computer. Its operating system is OS1100.This operating system has the following characteristics: real-time response, transaction process, designated time batch process, on-line interactive and batch process, etc.

We use MAPPER as Database management system for developing material computer management system[2]. MAPPER is a relational Database somewhat like a file management system. The structure of the database file is as following: the file is organized by mode, type and report. One mode has 8 types, one type has 2000 reports and one report can has unlimited data records[3,4].

There are 100 terminals in this system. They are distributed in various material management departments and these departments are scattered over Baosteel's area about 15 square kilometres. The system structure will be illustrated in figure 2.

6.2 The Main Functions Of The System

In order to meet the requirements of management and according to the principle of integrated procedure management, we developed 14 subsystems in the material management system. There are above 700 programs,34 file structures and1500 program screens. There are 8 computer automatically generated management number used in this system in order to make connections among modules and subsystems. The figure 3 will illustrate these functions further.

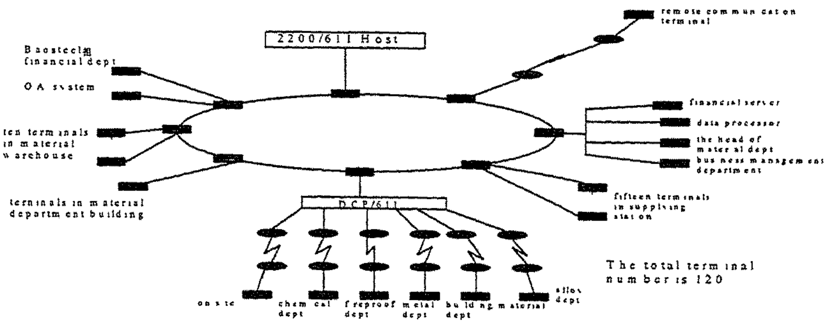


FIGURE 2 COMPUTER NETWORK STRUCTURE

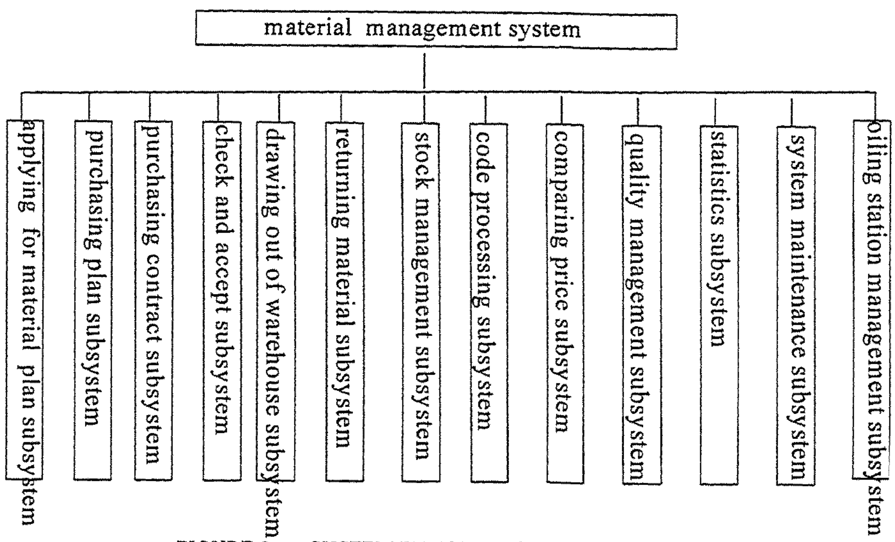


FIGURE 3 SYSTEM FUNCTION STRUCTURE

7.THE KEY TECHNOLOGY IN THIS SYSTEM

7.1 The Connection Technology Of Different Database Management System

The financial department use micro-computer to process their accounts. After all the other management department use UNISYS 2200/600,the manual accounts are wiped out. So, the financial department must get their original information directly from the MAPPER. The problem about the connection MAPPER and FOXPRO must be solved.

At first we must connect the financial server to CIMS main network to realized the physical connection between financial management system and UNISYS 2200/600. Then, by utilizing designer workbench, we can make some process programs by MAPPER and FOXPRO . Using these process programs, we can realize the data inter-transform between MAPPER files and FOXPRO files. The financial department can get the original information directly from the MAPPER, so they settle their accounts conveniently. At the same time, the other material management department and leader can get the financial information directly from the financial server in order to analysis and make decision. The inter-transform procedure is illustrated in figure 4.

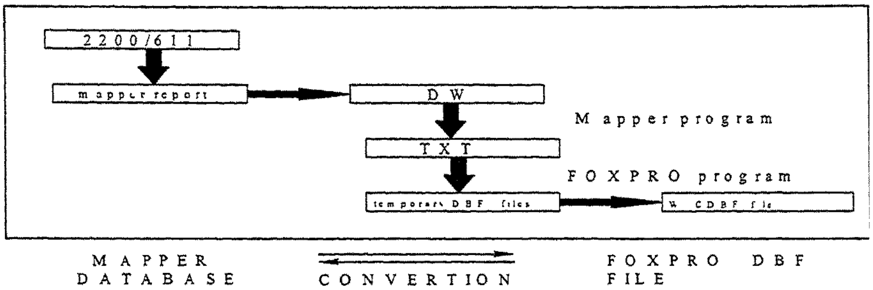


FIGURE 4 DATA CONVERTION

7.2 Uninterrupted Running Environment

We fully take the advantage of MAPPER while we design the material management system. We form two systems, one is the formal running system, the other is the debugging system. The two systems are independent. The developing and modifying programs are put in debugging system. Programs are tested in debugging system until they are proven correct. Correct programs are put in the formal running system. With these two systems, when we test or debug some program, there is no influence on the system running. In order to remove system breakdown, we develop a set of program which are used for record the breakdown information. With these breakdown information, we can find the breakdown reason quickly and fix them as soon as possible We process breakdowns in debugging system, then put the corrected program into the formal running system for users.

8. THE RESULT OF SYSTEM RUNNING

Since the formal running from 1994, this system has been proven that its functions are strong convenient to operate and it is highly reliable.

The advantages include:

a) Easy to operate

There are 14 subsystems in the material computer system, and there are above 700 functions. Users can use these functions on menu screen. After half a day's training, users can operate this system skillfully.

b) Information arriving in time

Every material's information including planning, purchasing, into and out of warehouse is stored in this system. All the information is integrated and can be shared. Users can get the newest information from the terminal. These information is the basis of reasonably organizing production.

c) Correctly processing

If users operate according to the demands of the function menu, the data will be processed at once. The material and fund accounts will express the correct fund occupied conditions at that time. Before this system is used, the accounts was manually processed by many persons and can't be correctly gathered.

9. The beneficial results of the system include:

The material management level has been raised since this system was put to run. The stock becomes suitable and a large sum of capital has been saved.

10. CONCLUSIONS

The material computer management system was successfully developed within one year. We use MAPPER as its Database. The MAPPER has following advantages : strong modularization, high transparency, conveniently shared data and good privacy.

We integrate the following technology in this system: the modernised management idea, advanced computer technology and up-to-date network technology.

In the management technology respect, we utilize the advanced management idea to form the management information flow. As a result, we get a unified and integrated management mode.

REFERENCE

[1] Hammer, Reengineering of design and manufacturing Process U.S.A ,1993
[2] UNISYS,MAPPER, System run design, Operation reference manual U.S.A,1990
[3] UNISYS,MAPPER, System operation guide U.S.A,1991
[4] UNISYS designer workbench operation guide U.S.A,1993
[5] Software engineering and software quality analysis in Chinese, Electrical industry publishing house
[6] Computer network introduction in Chinese, Nanking University
[7] Computer information system analysis and design in Chinese, National Defence Science & Technology University

MONITORING BEHAVIORAL EVOLUTION FOR MANUAL ASSEMBLY SYSTEM USING OBJECT PETRI NETS IN CIM ENVIRONMENT

F. Wu, W.H. Liu and T.Y. Xiao
National CIMS Engineering Research Center
Dept. of Automation, Tsinghua University
Beijing 100084, P.R.China

ABSTRACT

Monitoring design is one of basic problems during the development of applied CIMS. In this paper we study the monitoring problems for manual assembly system in CIM environment according to the CIMS architecture of Xian Aircraft Company(XAC), in Xian of China. We show that the problem of designing monitoring system for manual assembly system is reducible to modeling and synthesizing problem of monitoring system based on Object Petri Nets. Next the problem of implementing the monitoring system results in designing of structure and functions of the monitoring system. In particular, an efficient rapid prototype schema for implementation is presented. An application is documented about assembling aircraft vertical fin.

KEYWORDS

CIM, Object Petri Nets(OPN), Monitoring, Assembly Systems

1. INTRODUCTION

In recent years, Computer Integrated Manufacturing(CIM) has had a great influence on improving productivity by incorporating design, process planning, and shop floor control. As an effort to develop an efficient structure of CIM, a five level hierarchical control architecture has been proposed by CIM/OSA[1]. Construction of a real control architecture begins with diagnosis of complexity of an enterprise systems. A simple CIM architecture of the Xian Aircraft Company(XAC) manufacturing facility in Xian, China, as shown in Fig.1, may be viewed as a three level hierarchy: plant level, cell level and workstation level[2]. The plant level, the topmost level, controls one or more cells, each of which in turn controls several workstations each of which are comprised of a workstation controller and a supervisor. A supervisor behaves like an information interface between a workstation controller and manual assembly system which involve many operators, tools or equipment. So integration in XAC-CIMS control is achieved by interconnecting the different blocks involved through an efficient information exchange mechanism.

The plant level received so far has been responsible for long-term planning activities such as manufacturing resources planner(MRP), computer aided process planning(CAPP), inventory management, computer aided quality(CAQ), etc. MRP, which is implemented via computer software called IBM BPCS, obtains future requirements for finished products from a master production schedule and uses this and other information to generate the requirements for all the sub-assemblies, components and raw materials that go to make up the finished product. The second level in the hierarchy, the cell level, is the most interesting because it offers a wide range of problems which include scheduling, dispatching, monitoring, i.e., keeping track of the operations of the workstations, and reporting the cell status back to the plant[3]. The workstation level has the order that directly controls the assembly system, for example, the order to assembling front spar of vertical fin of aircraft. The order is issued by a supervisor back to manual assembly system.

The main concern of this paper is problems of the monitoring system in cell and workstation level of XAC-CIMS, i.e., monitoring structure, monitoring system modeling, monitoring functions, monitoring implementation, etc. It has been modeling using Object Petri Nets(OPN) developed by C.Sibertin-Blanc[4]. The development tool of monitoring system used is Rapid Application

Development & Integration Support system(RADISS 3.0) on a basis of rapid prototype methodology, which is recommended by CIMS Office of the National High-tech Program in China.

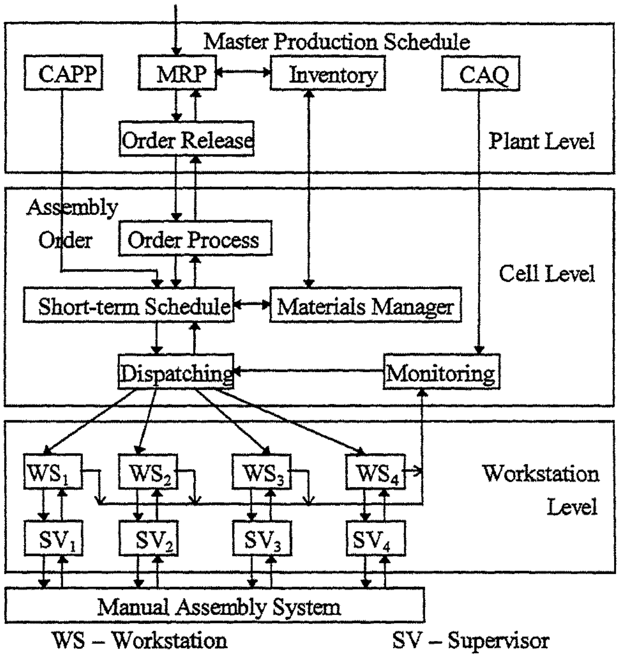


Fig.1: XAC-CIMS architecture

2. A BRIEF OPN RESEARCH AND APPLICATION

There has been much research in the theory of Petri nets and other application for modeling production systems[3-10]. Models based on Petri nets are becoming more and more complex for these systems. To avoid the proliferation of the number of states and getting too complex at interpretation level, and to obtain concise Petri nets models, it is essential to individually identify each token by assigning it a color, or several attributes and therefore a whole data structure. Several models, called "high-level Petri nets",have been defined:(1)Colored Petri nets,by K.Jensen[5],(2)Predicate/Transition Petri nets by J.Genrich[6],(3)Object Petri nets by C.Sibertin Blanc[4]. Colored Petri nets or Predicate/Transition Petri nets aim at both describing and simulating system dynamic behavior in a more concise way. So our main objective isn't to do them, but rather to take system data structures into account, for example, quantity and direction of materials handling at assembling node for assembly systems such as manual assembly systems. Therefore we decided to adopt the OPN as a basis for our work on specifying the monitoring system in cell level or workstation level. OPN, or more properly Petri nets with data-structure are Petri nets whose data structures are associated to the tokens. The tokens are called objects; their values can be tested and changed by methods defined for transition. This make OPN an even more powerful tool for monitoring behavioral evolution for assembly systems and permits the integration of scheduling aspects. We can refer to literature[7-10] about OPN applications. They discussed mainly modeling methods of FMS and performance analysis against this background using OPN so that rare researches are carried out how they can be applied in Chinese enterprises with the characteristics of dense manpower and less automated equipment. It is our distinguishing work that we will deal with these problems.

3. OPN FOR ASSEMBLY SYSTEMS

3.1 OPN Definition

Object Petri nets(OPN),devised by C.Sibertin-Blanc[4] are a high abstraction level model which can be considered as a specific case of predicate/transition nets. They make it possible to obtain a precise model of how operation sequences and data processing interact. OPN have three major components[8]: a data structure, a set of elementary operations and a control structure.

• In OPN, tokens are no longer undifferentiated, but are objects or tuples of objects. Objects are distinguished by the class to which they belong, that class defining their structure. A class of objects is defined as a list of attributes each with a type and domain of values. An object can therefore be taken to be a tuple of attributes.

• The elementary operations performed on objects correspond to indivisible transitions. A transition behaves like a function taking one object(at least) from each input and producing one object(at least) at each output. A transition consists of a precondition, an action and one or more emission.

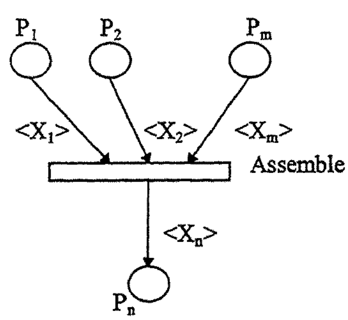
• The control structure defines the sequences of operations and objects involved. It is translated by the net structure using the standard components of Petri nets, i.e. places, transition and arcs.

This description is only an informal instruction to OPN. A formal definition is given[4].

3.2 Modeling Simple Assembly Systems

There are two problems considered as followed:

(1)Arc Inscription Most of the production cells manage parts which have to be assembled together. In such cases, an assembly operation consists in consuming a certain number of separate parts, and in creating a new part. Let us take the example of an operation of assembling m parts, leading to the nth part as shown in Fig.2.



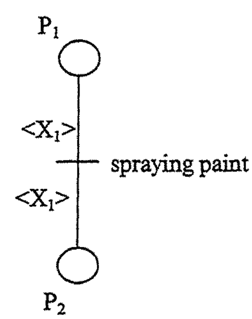


Fig.2: OPN of a simple assembly system   Fig.3: OPN of a simple un-assembly system

(2)Preconditions and Actions Attached to Transitions The interpretation attached to transitions enables us to make a selective choice of the objects to be removed from input places and modify the value of attributes when transitions are fired. In practice, this corresponds, for example, to perform un-assembled operations such as spraying paint, inspecting and the transition modifying the value of attributes of un-assembled classes. In this case, we can show key transition for the sake of conciseness, and represent it in Fig.3 if we only care the assembled operations.

4. MONITORING FOR MANUAL ASSEMBLY SYSTEM

4.1 Monitoring Structure

A workshop structure which monitors behavioral evolution in four sections of shop floor as shown in Fig.4, may be viewed as a two level hierarchy, cell level and workstation level according to XAC-CIMS architecture. The cell level, controls four workstations, is responsible for the analysis and putting data collected in order, and definition a certain number of mechanisms that unsolicited events such as lack spare parts, equipment breakdown, etc., must be able to alarmed, and for reporting to operators in some kinds of outputs. The workstation level, each of which includes one section, or supervisor, as an interface between a workstation computer and the section operators, not only receives loaded input orders from the cell level and then transfers to operators, but also feeds production status data from shop floor back to a workstation computer. All status data will be stored in system data base in client/server computation mode.

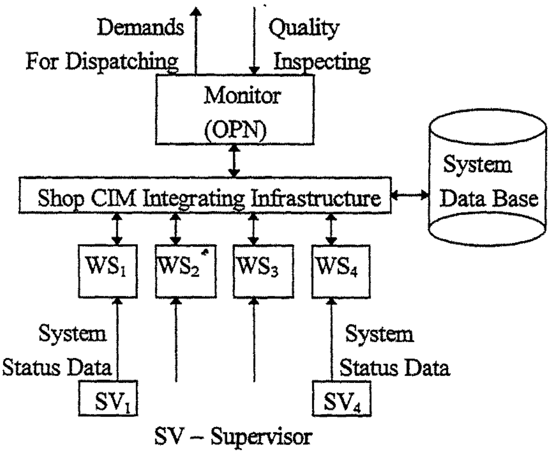


Fig.4: Monitoring structure of manual assembly system

4.2 Monitoring Functions

These different elements of monitoring functions will be discussed in more detail. There are also data capture, data analysis and decision support[11].

(1)The Data Capture Effective control of assembly process depends on accurate, timely recording of data from the shop floor. Ideally, the data should be stored in system data base, which will allow ease of access and manipulation of the data. The current method of data collection is manual mode in workstation level through a supervisor. There are a number kinds of data collected for assembling aircraft, i.e., assembling progress, inventory level, inspecting quality information, tools and equipment status, real man-hours, etc.

(2)The Analysis of Data The data analysis function of the monitor can be seen as a reporting mechanism in the cell level. The reporting mechanism can be in any or both of two modes, namely, real-time and historical. If it is real-time, data concerning materials, jobs, equipment or other should be collected real-time. All of this data can then been on screens provided which displays the information in OPN form. It is historical reporting, the historical reporting section of the monitor involves providing the operators with screen reports, graphical output and hard copy output for post production analysis of what has taken place. These enables the operators to see exactly where each job is, how long it has been there, etc.

(3)Decision Support The monitor also acts as a feedback mechanism in real-time, passing information to the other control functions such a scheduling and dispatching in the cell level. It has to keep track of what is happening so that if an unsolicited event, such as equipment breakdown, occurs dispatching model has immediately to change its strategy to allow for this. There are other decision supports such as suggestions for rescheduling, revision of CAPP, etc. These supports in the forms of some kinds of actions attached to transitions help the operators make decisions.

4.3 Monitoring Implementation

We have selected two basic tools according to the goals of XAC-CIMS, i.e., (1)RADISS[12], (2)Visual C compiler. RADISS, as a commercial production recommended by CIMS Office, the

National High-tech Program in China has shown special functions in rapid developing prototype oriented management information systems(MIS) of enterprises. It is a basis of implementing automatic generation of MIS software that have been standardized, automated and integrated, but show congenitally deficient in executing manufacturing process algorithms. It is Visual C that can make up the defect. In addition, a designer firstly specifies production problems into transition algorithm modules based on OPN, and then embeds them into applications with RADISS.

4.4 Monitoring Application

In vertical fin section of an assembly shop of XAC, we take an example of OPN monitor for manual assembling vertical fin of aircraft, shown in Fig.5. We omit net precondition, actions attached to transitions and variables to read easily. In each node, we define object classes and their attributes as shown in Tab.1.

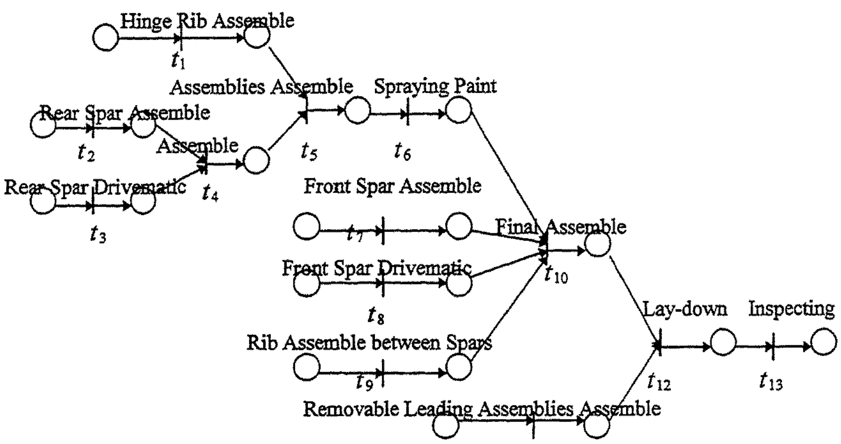


Fig.5: OPN for assembling vertical fin

Tab.1

class	attributes
"Assembly Status"class :	material.dwg.no, part.name, process.no, work.order.no, ship.no, begin.date, finish.date, current.progress.node
" Materials"class :	sequence, name, specification, quantity, quantity.per, status
"Tool & Equipment"class:	name, sequence, quantity.repaired, status.sode
"Operators and Man-hours" class:	work.order.no, operator.name, group, man-hours, real.date, quantity, fulltime.work/overtime.work
"Lack.Spare.Parts"class:	work.order.no, ship.no, superlier, part.no, kit.item.no, quantity.per, total, date

Those object classes are that we are concerning. We make some decision support depending on their value of attributes. The supervisor inputs all kinds of production data every day after manual assembly operations have been completed. All of this data can then be displayed by OPN monitor and be manipulated. If some of transitions of OPN have taken place and their value of transitions are modified, a real-time OPN will be refreshed which display assembly system status. By use of OPN for assembling process, we can make historical reports providing operators with graphical output and hard copy, for example, man-hour totals per month and waste product percentage. There are other indices such as productivity and man-hour utility percentage. By comparison of this data , real-time on shop floor can easily be manipulated by managers. It makes assembly production both in schedule and balance.

5. CONCLUSION

This paper has described how to design the monitoring system in XAC-CIMS environment. It began by discussing definition of OPN and its modeling applications in simple assembly systems, and especially the monitoring problems. The monitoring problems were shown to comprise of monitoring structure, monitoring functions, and monitoring implementation. In particular, an efficient rapid prototype schema for implementation is presented. An example is illustrated for monitoring application in vertical fin section of an assembly shop of XAC. These approaches and programs coming from enterprises are suitable for Chinese characteristics, and practical and available. By the way, the presented design methodology of the monitoring system can also be used in other wide fields.

6. REFERENCES

1. Kosanke, K., "CIMOSA – Overview and status", Computer in Industry, vol.27, no.2, pp.101-109, 1995.
2. XAC-CIMS/PMCS Design Group of CIMS-ERC of Tsinghua University, "A detailed report on Production Management and Control System of an assembly workshop of XAC", Tsinghua University, Beijing, 1996.
3. Kasturia, E., DiCesare, F., etc., "Real-time Control of Multilevel Manufacturing Systems Using Colored Petri Nets", IEEE Int. Conf. On Robotics and Automation, pp.1114-1119, San Francisco, California, 1988.
4. Sibertin-Blanc, C., "High Level Petri Nets with Data Structure", Ph.D thesis,pp.141-170,1985.
5. Jensen, K. "Colored Petri Nets and the Invariant Method", Lecture Notes in Computer Science, Brauer, W., Reisig, W., and Rozenberg, G., no.254, pp.248-299, Springer Verlag, Berlin, 1986.
6. Genrich, H.J., "Predicate/Transition Nets", Lecture Notes in Computer Science, Brauer, W., Reisig, W., and Rozenberg, G., no.254, pp.248-299, Springer Verlag, Berlin, 1986.
7. Ausfelder, C. Castelain, E. and Gentina, J.C., "A Method for Hierarchical Modeling of the Command of Flexible Manufacturing Systems", IEEE Trans. On Systems, Man and Cybernetics, vol.24, no.4, pp.564-573,1994.
8. Ahmed, S.B. Moalla, M. Courvoisier, M. and Valette, R., "Flexible Manufacturing Production System Modeling Using Object Petri Nets and Their Analysis", Robotics and Flexible Manufacturing Systems, Gentina and Tzafestas, S.G., pp.245-257, Elsevier Science Publishers B.V., North-Holland, 1992.
9. Cruette, D Bourey, J.P. and Gentina, J.C., "Description and Validation of Logical Operation Sequences in the Flexible Manufacturing System Context Using Object Petri Nets", Robotics and Flexible Manufacturing Systems, Gentina and Tzafestas, S.G., pp.269-279, Elsevier Science Publishers B.V., North-Holland, 1992.
10. Wu Yun, Yang Hanyu,etc., "A Study of Development Method based on Object Petri Nets", J. of Southeast University, vol.25, no.3A, pp.40-47, 1995.
11. Higgins,P. and Browne,J., "The Monitor in Production Activity Control Systems", Production Planning & control, vol.1, no.1, pp.17-26, 1990.
12. The National CIMS-ERC of Tsinghua University, "Rapid Application Development & Integration Support system(RADISS 3.0)", Tsinghua University, Beijing, 1996.

A TAXONOMY FOR INTERACTIVE AND BLACKBOARD BASED CAPP

G. Van Zeir, J.-P. Kruth, J. Detand
Katholieke Universiteit Leuven,
Department of Mechanical Engineering, Division P.M.A,
Celestijnenlaan 300B, B-3001 Leuven, Belgium.
Tel. (+32) 16 32 24 80, Fax. (+32) 16 32 29 87
e-mail : Geert.VanZeir@mech.kuleuven.ac.be

ABSTRACT

This paper describes the taxonomy of an interactive Computer Aided Process Planning kernel, based on a blackboard system architecture. The paper explains why an object-oriented blackboard system approach is well suited for solving the process planning problem and describes the different object-oriented data models for CAPP. The developed CAPP kernel provides the human process planner with assisting 'expert modules', each capable of performing a specific automated process planning task. The interactive nature of the CAPP kernel permits the operator to create process plans manually, ignoring the built in logic, or let them be generated or completed by means of the knowledge in the systems. The advantage of this approach is that it gives the process planner full control and great support in generating process plans and it allows an arbitrary order of doing the different planning tasks.

KEYWORDS

Computer Aided Process Planning, knowledge based CAPP, feature based CAPP, Blackboard architecture.

1. INTRODUCTION

Computer Aided Process Planning (or CAPP) has been a very popular research topic over the last two decades. Many researchers have been working towards software for carrying out process planning tasks or largely assisting human process planners. Despite the definite need for CAPP systems in almost any manufacturing sector today [1][2], their break-through in industry is still painfully slow [3]. If a CAPP system is used in a company, it is mostly done in isolation from the product design process, the production planning, scheduling and control activities. The static routing sheets, generated during process planning, do not consider actual resource utilisation. Therefore, practice shows that the large majority of all process plans is still created/adjusted manually, and detailed optimised plans are rarely produced.

Recent assessments of CAPP systems and technology by El Maraghy [3], and the limited degree of proliferation and impact they had on industry up to now, suggest the following desirable characteristics for an effective CAPP system. It should a/o :

- be extendible and adaptable to new applications and facilitate the inclusion of new data bases and knowledge, as well as being customisable (covering the complete part and process spectrum of the company),
- provide effective knowledge acquisition, representation and manipulation mechanisms as well as means to check the completeness and consistency of that knowledge,
- keep a human in the loop, to participate in some decision making, provide heuristics as needed and supplement the system's abilities,
- provide an excellent user interface to support effective interaction by facilitating inputs, producing outputs and reports in flexible formats and display the results graphically.

Starting from these requirements, the authors tried to come up with a suited architecture for CAPP applications. This architecture is presented in the following paragraph.

2. AN INTERACTIVE CAPP KERNEL

"Interactiveness" is one of the requirements identified in the previous section. This paragraph explains what the authors mean by interactiveness of the system and why they chose for this approach. Further it describes some important aspects that allow the implementation of such system.

2.1 The interactive approach

In traditional manual process planning, the planning tasks (a.o. the specification and interpretation of the part data, the process selection, the selection of machines, tools and fixtures,...) are carried out by the human process planner in a more or less arbitrary order. When a process planner inspects a certain part, (s)he could for instance immediately decide that it will have to be processed on a five-axis milling machine or that EDM will have to be used or that a special tool or fixture will be required. In order to comply with traditional planning, the new CAPP kernel should provide an interface or editor for each process planning activity, allowing the step by step, interactive construction of the process plan. The human process planner should be able to invoke these 'editors' at any time, to perform the different planning tasks manually though in a structured, formalised manner, through the assistance of the computer. Further, the kernel should provide reasoning mechanisms and knowledge sources. However, the process plan should not be generated by one monolithic black-box system. Instead, the kernel should have several knowledge sources (e.g. one for process selection, one for tool selection, one for set-up calculation, etc.; that could even consist of plugged-in commercial or legacy systems) that can be consulted in an arbitrary order to perform some specific planning task. This consultation or reasoning, will normally cause some process planning information to be generated by the CAPP system, rather than by a human process planner. However, any of this generated data should still be adaptable interactively by the human expert at any time by means of the editors.

The advantage of this approach is that it resembles manual process planning, because of the arbitrary order of doing the different planning tasks. Thus, the human process planner gets familiar with the system very quickly. Plans can be manually edited, or completed by means of the knowledge based systems. The computer never dictates a solution, but makes suggestions that can be accepted or overruled by the human expert. The next paragraph presents a software architecture that is neatly suited to implement the approach described above. It is called the *blackboard architecture.*

2.2 The implementation

Process planning deals with many diverse, specialised applications (call it process planning tasks) that have to be integrated in some way. Some planning tasks can be executed arbitrarily (e.g. selection of a specific tool before determining the machine or vice versa). However, the outcome of each task can (will) depend on the results returned by previously completed planning tasks (e.g. the selected tool can only be mounted on a limited set of machining centres). Furthermore, the knowledge representation in each planning application can be different (e.g. rule-base, Petri-net, neural net, fuzzy logic, table, algorithm,...). Such problem can typically be solved by a blackboard system [4][5].

The architecture of a blackboard system can be seen as a number of people sitting in front of a blackboard. These people are independent specialists, working together to solve a problem, using the blackboard as the workspace for developing the solution. Problem solving begins when the problem and initial data are written onto the blackboard. The specialists watch the blackboard, looking for an opportunity to apply their expertise to the developing solution. When a specialist finds sufficient information to make a contribution, he records his contribution on the blackboard, solving a part of the problem and making new information available for other experts. This process of adding contributions onto the blackboard continues until the problem has been solved. A manager, separate from the individual experts, attempts to keep problem solving on track and ensures that all crucial aspects of the problem are receiving attention.

Translating this metaphor into a computerised blackboard system, the distinct specialists should be considered as expert modules and the blackboard as a global database containing input data, partial solutions and other data in various problem solving states. For triggering and controlling these expert modules, a user driven approach is chosen [6]. It not only promotes the interactiveness or human involvement; it also makes the CAPP system transparent and facilitates the understanding of its structure, behaviour and outcome.

2.3 An object-oriented Blackboard model

Objects are used to represent the blackboard information. Each data-object has attributes (slots that contain information), a set of methods, a constructor, etc. for handling this object (figure 1). The information contained in such object can be supplied by an **expert module** that consult the appropriate knowledge source, or by the **human expert**, by means of an interface that is provided for each type of object.

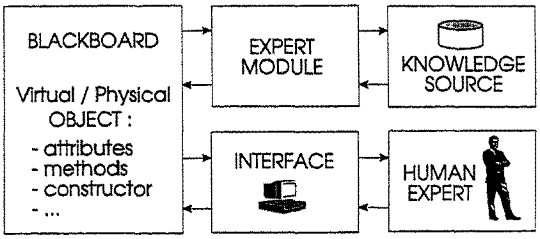


Figure 1 : Object-oriented data model for blackboard systems

Consulted expert modules will take into account the information that was added by the human expert (or by other modules). Moreover, the interface to the human expert allows him/her to verify, accept or alter the information generated by an expert module, at any time. Objects with partial information (empty attributes, attributes that describe intervals or constraints or multiple discrete values,...) are called *virtual objects*, while objects that are unambiguously determined by the information contained in their attributes are called *physical objects*.

2.4 An object oriented model for the developed Blackboard CAPP system

This paragraph illustrates how manufacturing data and knowledge have been incorporated in the CAPP system. The following data models are distinguished (figure 2) :

- The **blackboard** : the CAPP blackboard contains both part and process planning information in various states during the process plan generation. The part description can be considered as the initial input data (see metaphor). From this input, new blackboard objects are created by expert modules or by the human operator.
- A **part model** : the CAPP kernel is feature-based. Consequently, the part information, that serves as input to the CAPP system, should a/o contain a detailed description of the part's features. This model incorporates a/o company specific feature types.
- A **resource model** : the resource model embeds machine tools, fixtures, tools, and other auxiliary equipment, available in the factory and considered during process planning. It includes all data that are important to inquire about during the process planning task (e.g. power, accuracy, outer dimensions, axis data, etc.).
- A **process model** : this model contains the manufacturing processes that are used in the company (e.g. end-milling, face-turning, welding, laser-cutting, wire-EDM, etc.). Further, it embeds related

process parameters (cutting conditions, costs, accuracy, etc.), and associated geometric constraints and technical parameters (roughness,...).

- A **process plan model** : the blackboard CAPP system supports graph-based process plans that allow the modelling of alternative manufacturing sequences. Such process plans with alternatives are called non-linear process plans or NLPP's [7]. All process plan data that are required for further order processing, manufacturing and all administrative data are included in the model.

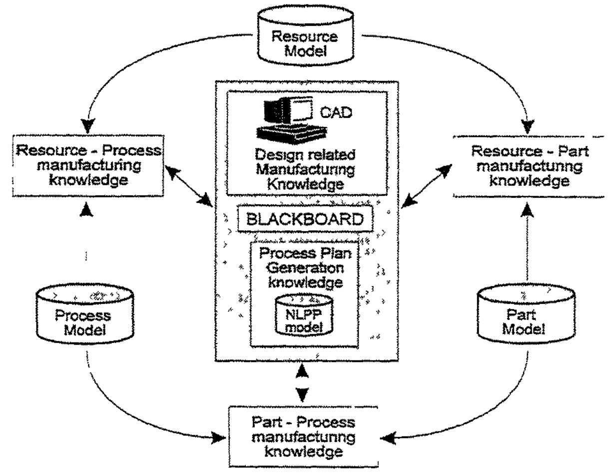


Figure 2 : Data models and manufacturing knowledge for CAPP system.

The outlined data classes are not just some isolated data structure but are interdependent and related to one another by some specific constructs. The manufacturing knowledge in a CAPP system holds the following relationships (figure 2) :

- The **design related manufacturing knowledge** is employed for instance by the 'CAD expert' module, which allows to extract/add process planning information from/to the part design.
- The **part-process manufacturing knowledge** associates the data content of the part model to the process model; it embodies the 'process selection expert' module which determines the different manufacturing steps to be undertaken on a certain part or feature, and the sequencing relationships between those manufacturing steps.
- The **resource-process manufacturing knowledge** associates the data of available resources to the process model; it embodies for instance the 'machine selection expert' which determines the candidate machine tools for the operations on a certain workpiece. However, this expert module will also have to consult the 'resource-part manufacturing knowledge' (e.g. because there are limitations on part dimensions for a certain machine tool)
- The **resource-part manufacturing knowledge** relates the part data with the resource data (e.g. the selection of a tool is influenced by the dimensions of the feature to be processed). This type of knowledge is consulted by for instance the 'tool selection expert' and the 'machine selection expert' modules
- The **process plan generation knowledge** encloses the knowledge that brings all other knowledge sources together. In this interactive CAPP kernel, the 'blackboard manager' that triggers the distinct process planning expert modules could be considered as part of this knowledge.

Generating a process plan requires the analysis of relevant part information, the selection of the right manufacturing processes and the appropriate resources thus building the objects on the CAPP blackboard.

3. OBJECT ORIENTED CAPP TAXONOMY

This section refines the data models shown in figure 2, and elaborates on how the information of each blackboard data-object is supplied by triggering the distinct expert modules. The part, process and resource model are illustrated first. They are completely user definable and thus adjustable to any company specific environment. The process plan model presents the type of plans that can be created by the developed CAPP kernel (namely non-linear process plans) and explains their advantage.

3.1 The part model

The part data, which serves as input for the developed CAPP blackboard system, contains the following information :

- the *overall part data* contain the part type (or family) and the parameters of this type of part (e.g. dimensions, lot-size, weight, surface treatments,...)
- the *feature data* contain all feature types, together with their specific parameters (e.g. geometrical parameters, roughness, position and orientation,...)
- the *feature relations* contain form and location tolerances; but also feature interference, adjacency, intersection, etc. (figure 3) can be entered in this model as user definable relations.

An important dilemma of feature based modelling is that many methods can be used for synthesising parts by features, and this makes the number of features virtually infinite. There are two options to overcome this infinity : (1) define a standard (fixed) feature catalogue or (2) use a customisable (extendible) feature catalogue. Many attempts have been made to define a standard feature catalogue, a/o the CAM-I process planning feature catalogue of John Deere [2] describing a hierarchy of process planning oriented features and the PDES/STEP catalogue (ISO CD 10303) containing design oriented features. Those catalogues are rather elaborate (up to over a hundred features), but still suffer from incompleteness, and are not always corresponding to the engineering practice of specific companies.

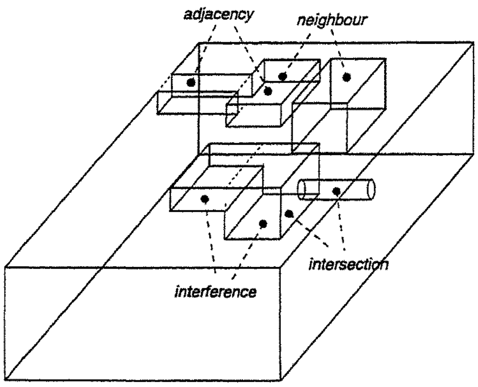


Figure 3 : Example of (user definable) feature relations

An extendible (versus a fixed) feature catalogue allows to define company specific design (and process planning) knowledge in a straight-forward and customised fashion (i.e. current practice can be retained) without too much information (i.e. about twenty features can cover 90% of all workpieces of one company) [8]. Moreover, an extendible feature catalogue allows to update the features according to the evolving machining technology (e.g. a new shape of cutter may have as consequence that other shapes can be machined, other NC programs will be used, ...). Since an extendible feature catalogue offers advantages, and since an extendible catalogue can always be made compatible to one of the above standard catalogues, this option was employed in the presented CAPP kernel.

Not only the feature types, but also the workpiece types and feature relation types (e.g. feature intersections in figure 3) can be edited in a *User Defined Specification File* : this is an ASCII file, in a readable syntax, that is maintained by the user. In this file, the user can create new types of workpieces, features and relations, or add new parameters to each of these types. The file is parsed when starting up the CAPP system, in order to initialise it with this company specific data [9].

3.2 The resource model

When setting up a CAPP system, it is vital to model all resources of the target manufacturing environment. The model should not only comprise the resource identification, and location (administrative data, logistic data), but should especially enclose their manufacturing capabilities (technical data). The resources can be categorised into the following subsets : *main manufacturing equipment, auxiliary manufacturing equipment* and *transport equipment*. The auxiliary manufacturing equipment includes fixtures, cutting tools, measuring probes, cooling fluid, NC-programs, etc., while the main manufacturing equipment includes machine tools, workbenches, painting room, etc. The transport equipment includes cranes, pallets, carriers, robots, AGVs, etc. and enables to transport the products and/or auxiliary manufacturing equipment from one place to the other, where each place corresponds to a waiting queue or a main manufacturing equipment device.

Each type of resource has its own list of parameters (technical, logistic and administrative information contained in the 'resource object'). Evidently, the resource taxonomy model is extendible and completely user definable, allowing to introduce new resource types and/or parameters. The resource data are stored in a relational database. Examples of the data content for machine tools is given in table 1.

Data type	*Examples*
Technical	Power, kinematic and dynamic data (spindles, axes, stiffness, accuracy,...), spindle speeds, feedrates, max. axis displacement, max. part dimensions and weight, clamping possibilities, controller type, tool magazine (capacity, clamping,...), data registration and sensor and monitoring capabilities, ...
Logistic	Shift model, availability, order allocation, maintenance plan, calibration periods, capacity,...
Administrative	Identification, location, classification, date of installation, cost model...

Table 1 : Machine tool data types and examples

The parameters stored with each type of tool (or fixture) describe some specific technical tool (or fixture) characteristics (e.g. insert shape, insert grade, tool holder type, modular fixture, etc.), their logistic data (e.g. availability, they may even be shared between different machine tools) and some administrative data [10].

3.3 The process model

The process model contains all kinds of basic manufacturing processes that can be carried out in the company's workshop. For the developed CAPP kernel, this model mainly includes metal cutting operations like milling, drilling, reaming, boring, grinding, honing, sawing, turning, etc. but also processes like hardening, deburring, inspection and quality control. The structure of the process model allows to enter new types of processes at any time, to update the model to the actual situation in the workshop (e.g. laser cutting, wire EDM, casting, forging). When the model is extended with new processes, it is evident that other data models (and knowledge sources) will have to be extended too (e.g. adding EDM machines to the resource model). Specific (user definable) process parameters that correspond to the company code of practice can be defined for each type of process : e.g., process type/classification, settings and control parameters, geometric shapes/constraints, technical limitations and reachable quality (roughness, IT grade,...), material properties/change (e.g. induced stress after grinding may require annealing), etc.

3.4 The process plan model

The information contents of a process plan consist mainly of two sections (Table 2) : organisational data (process plan header) and the activity description (technical) data. The process plan header can be further subdivided in order related and non-order related organisational data. The activity description is traditionally always non-order related. The order related organisational data are not created by the process planner, but are added by an overall PPC system in order to prepare the plan for scheduling purposes. A process plan can thus have identical data except for the order related organisational data. Table 2 gives an overview with some examples :

Data type	*Data sub-type*	*Examples*
Organisational / Administrative data (header)	Plan related, non-order related	Plan ID, Plan type (e.g. standard, alternative, test-run,...), Validity (e.g. date of issue, date of change, temporal,...), Plan origin, etc.
	Part related, non-order related	Part ID, Part name, Part classification, Drawing-ID, BOM (points to process plans of raw material, initialises orders for raw material)
	Order related	Order ID, Order type (stock, customer, spare part,...), Comments, cost centre, product unit, batch number, series number, specification for acceptance, product or subassembly number, actual batch size, manufacturing time per operation, set-up times, total amount of raw material, term data (ready time, due time, lead time, priority data).
Activity description data	Technical data (non-order related)	Operation sequence (or graph), Operation number, Operation description (incl. process conditions), Set-up-time per batch, machining time per part, required resources (machines, tools, fixtures, set-up tools, measuring tools, NC programs,...)
	Logistic data	Department data (ID, classification, location, hourly machine rate, related capacity group), personnel cost data, etc.

Table 2 : Process plan data

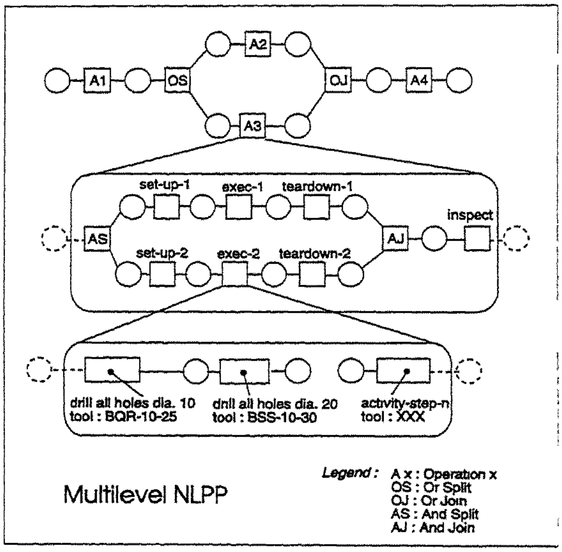


Figure 4 : Example NLPP.

The developed CAPP kernel generates non-linear process plans (or NLPPs). Such plans propose alternative manufacturing sequences by means of an AND/OR graph (figure 4). A NLPP is an aggregation of blackboard objects : elementary processes are grouped, forming operations on one or more features. Each operation is performed on a single machine tool; and several of these operations are structured in a graph : a non-linear process plan. NLPPs offer a significant benefit to schedulers. As they are often coping with unpredictable disturbances or rush orders, the NLPP can immediately offer a valid manufacturing alternative. Furthermore, when schedulers plan a job, they can choose the optimal path, taking into account the actual load of the machines on the shop floor. Elaborated research on the generation of NLPPs is illustrated in [7][11].

4. CONCLUSION

The developed CAPP system is based on a blackboard architecture : several expert modules can be triggered in an arbitrary order to perform a specific planning task, ensuring a very flexible way of performing the CAPP activities. Solutions generated by the assisting expert modules result in objects on the blackboard. These solutions can be adjusted or overruled by the operator, since each blackboard object has an interface to the human expert for manually adding or changing the information residing in the objects. In this way, the process plan is dynamically instantiated, stepwise built up by relating one or more types of objects together to create new objects. This complex generation mechanism increases the potential to represent manufacturing knowledge in the way humans reason and offers the user full control over all separate planning activities as the process plan is being finalised.

5. ACKNOWLEDGEMENT

Some of the work, reported in this paper, was supported by the projects ESPRIT 6805 "COMPLAN", GOA 95-2 on "Holonic Manufacturing Systems" and IUAP-50.

6. REFERENCES

1. Gu, P., and Norrie, D.H., Intelligent Manufacturing Planning, ISBN 0-412-46250-8., Chapt. 5, Chapman & Hall, London, 1995.
2. Butterfield, W.R., Green, M.K., Scott, D.C., Stoker, W.J., Part features for process planning, CAM-I feature catalogue, R-86-PPP-01, Deere & Company, Arlington Texas, 1986
3. ElMaraghy, H.A., "Evolution and Future Perspective of CAPP", Annals of the CIRP, Vol. 42(2), pp. 1-13, 1993.
4. Corkill, D.D., "Blackboard Systems", AI Expert Vol. 6(9), pp 40-47, 1991.
5. Carver, N., and Lesser, V., "The Evolution of Blackboard Control Architectures", CMPSCI Technical Report, no. 92-71, 1992.
6. Kruth, J.P., Van Zeir, G., Detand, J., "An Interactive CAPP Kernel based on a Blackboard System Architecture", Proc. of the ASME 1996 Design Engineering Technical Conf. and Computers in Engineering Conf., CD-ROM, Irvine (CA), 1996.
7. Kruth, J.P., and Detand, J., "A CAPP system for non-linear process plans" Annals of the CIRP, Vol. 41(1), pp. 489-492, 1992.
8. Carlier, J., Peters, J., A machining centre operation planning system, Annals of the CIRP, Vol. 34(1)., 1985.
9. Kruth, J.P., Van Zeir, G., Detand, J., "Extracting process planning information from various wire-frame and feature based CAD systems", Computers in Industry, Vol. 30(2), pp. 145-162, 1996.
10. Perremans, P., "Feature based Description of Modular Fixturing elements : The Key to an Expert System for the Automatic Design of the Physical fixture", Proc. of the 1994 ASME - Int. Comp. in Eng. Conf., Vol. 1, pp. 221 - 236, 1994.
11. Detand, J., A Computer Aided Process Planning System Generating Non-Linear Process Plans, PhD thesis, ISBN 90-73802-23-7, K.U.Leuven Division PMA, 1993.

SMART DRILLING - A FUZZY DECISION SYSTEM FOR PLANNING DRILLING OPERATIONS

Matthew Jia Ma and R. Du
Machine Automation Research Laboratory
Department of Industrial and Manufacturing Systems Engineering
University of Windsor
Windsor, Ontario N9B 3P4
Canada

Abstract

In this paper, a prototype Computer Aided Process Planning (CAPP) system - Smart Drilling is presented. The system is designed to generate process plans for multiple hole drilling operations. It contains four modules: (1) Tool selection module, (2) Machining condition design module, (3) NC code generation module, and (4) Simulation module. The tool selection module is developed based on fuzzy decision table. It groups the drills according to their sizes and shapes, and then select drills and drilling sequence that will minimize the number of drill changes. Based on fuzzy logic, the machining condition design module determines the drilling conditions that will minimize the production cost. The NC code generation module can automatically generate the NC code in APT format. Finally, the planned drilling process could be simulated using a commercial software VERICUT®. The simulation results show that Smart Drilling could save about 50% of cycling time comparing to that of manual planning.

1. Introduction

Of all the metal cutting operations, drilling is perhaps the most commonly used one. It accounts for one-third of all the metal cutting operations in industry. It is arguably the quickest and the most economical method for hole production. Its main advantages also include extended longer tool life. In addition, it is a preliminary operation to reaming, boring or grinding, where final finishing and sizing takes place.

This paper focuses on the process planning of multiple hole drilling operations. Many industry applications, such as mold and die making, require drilling multiple holes with different sizes and shapes. In particular, when drilling large holes, pre-drills are necessary in order to maintain accuracy and facilitate chip removal. The process plan for multiple hole drilling includes (a) the selection of drills and drilling sequence, (b) the path of drill motions, and (c) the design of cutting conditions. To determine an effective and efficient process plan, a large number of factors should be considered, such as the size and shape of the drills, the speed and the feed. Currently, multiple hole drilling process planning is done manually by manufacturing engineers. Its efficiency is often limited by the experience of the process planner. With the rapid advance of computer techniques, Computer Aided Process Planning (CAPP) has been widely used. For example, Adiga and Lin [1] developed an object-oriented knowledge-base CAPP system for scheduling. Lye and Yeo [2] developed a CAPP system for turning operations. Machining parts are usually divided as prismatic parts and rotational parts. The CAPP for machining prismatic parts are discussed in [3, 4]. The CAPP for rotational parts are discussed in [5, 6]. The methods of CAPP are systematically discussed in [7].

This paper introduce a prototype CAPP system for multiple hole drilling process planning, called Smart Drilling. The basic methods is presented in the next section followed by the simulation results in Section 3. Finally, Section 4 contains conclusions.

2. The Methods

2.1 Drill selection using fuzzy decision tables

After drilling pilot holes, the drilling operations should start from small holes to larger holes, with the same size holes drilled altogether. For small holes, the only choice is to use the drill that has the same diameter as the hole. However, as pointed out early, when drilling large holes, pre-drill (and multiple pre-drills) may be necessary. There could be several alternative choices of drills for pre-drill. To determine which drill would be used for predrill, a fuzzy decision table can be used.

The basic principles of fuzzy logic can be found in many monographs such as [8]. Briefly, an uncertain event A cam be evaluated by its membership function U(A). If U(A) = 1 then, it is certain that A will occur. On the other hand, if U(A) = 0, then it is certain that A will not occur. Fuzzy logic is different to the probability. The former deals with how likely an event will occur while the later deals with how it will occur. For example, the statement "the probability that a NBA player is seven feet tall is 0.9" implies that there is 90% chance the player being seven feet tall but there is 10% chance he may be just six feet. On the other hand, the statement "the fuzzy degree that a NBA player is seven feet tall is 0.9" implies that the player being most likely seven feet.

In the fuzzy decision table, each row represents a drill and each row represents a hole. The element U(i, j) represents the fuzzy degree that the ith drill is used to drill the jth hole, where i = 1, 2, ..., m (total number of drills) and j = 1, 2, ..., n (total number of holes). In particular, if U(i, j) = 0, then Drill i will not be used to drill Hole j. On the other hand, if U(i, j) = 1, then Drill i will definitely be used to drill Hole j. Following is the procedure of forming the fuzzy decision table.

Table 1. The procedure of forming the fuzzy decision table.

sort the drills and holes from small to large;
{the sorted drills are denoted $D_1, D_2, ..., D_m$}
{the sorted holes are denoted as $H_1, H_2, ..., H_n$}
set U(i, j) = 0, for all i and j;
for i = 1 to m do
    for j = 1 to n do
        if $D_i = H_j$, set U(i, j) = 1;
    end;
    if the $D_i > D_{th}$ (a diameter threshold) then
        for k = 1 to i do
            calculate the fuzzy membership function U(k, j) :

$$U(k,j) = \frac{U_k}{1 + B(D_i - D_k)}$$

            where: $U_k = 1$ if Drill k has been used before;
                $U_k = 0$ if Drill k has not been used;
                B is a weight factor depending on the depth and shape of holes;
                D_i is the diameter of current drill; and
                D_k is the diameter of Drill k.
        end;
    endif;
    find the number k* that corresponds the maximum fuzzy membership, that is:

$$k^* = \arg\max_k \{T(k,j)\}$$

    where "argmax" implies for argument maximum;
    chose Drill k* for predrill of Hole j.
end;

The detailed implementation also includes the case of drilling very large holes which require multiple predrills, and the case that all the holes require predrill. Table 2 shows an example of fuzzy decision table. As shown in the table, the last hole requires predrill and the best tool to use is the drill with diameter 0.8".

Table 2. An example of fuzzy decision

	Holes								
Drills	0.2"	0.2"	0.2"	0.4"	0.4"	0.6"	0.6"	0.8"	1.2"
0.2"	1	1	1						0.50
0.4"				1	1				0.56
0.6"						1	1		0.63
0.8"									0.72
1.0"									0
1.2"									1

It should be pointed out that using the fuzzy decision table ensures that the number of tool changes is minimized.

2.2 Drill travel path design using minimum distance method

When drilling several holes with a same diameter, one shall use a same tool. In this case, the drill travel path (the path that the drill follows to move from one hole location to another hole location) shall be designed as well. The optimal drill travel path problem is the same as the travel businessman problem. It is a NP-incomplete problem that requires exponential computation load [8]. As a result, it cannot be solved when the number of the holes are large.

However, a suboptimal solution can be found using the minimum distance method [8]. Its basic idea is to move from the current hole location to the closet hole location. Briefly, the procedure of minimum distance method is as follows:

Table 3. The procedure of designing drill travel path

```
set p = (0, 0) (the origin)
for i = 1 to n (the number of holes with a same diameter) do
    for j = 1 to n - i do
        find
            j* = arg max_j ||h_j - p||

        where, hj is the location vector of the jth hole;
        drill Hole j*;
        set p = hj;
        delete Hole j* from the hole list;
    end;
end;
```

Figure 1 shows a five (5) hole example. Using the minimum distance method, the drill travel path would be "Hole 3 → Hole 1 → Hole 4 → Hole 2 → Hole 5." This will reduce the drill travel time in comparison to travel following the hole number (i.e., "Hole 1 → Hole 2 → Hole 3 → Hole 4 → Hole 5.").

2.3. Drilling condition design using fuzzy logic

To determine the drilling conditions (speed and feed), various factors should be considered, such as drill size, tolerance, tool life, and the work material. In order to determine a drilling conditions, fuzzy logic was applied. Figure 2 illustrates an example of determining the cutting speed. First, the fuzzy membership functions of speed with respect to the drill size (d), the tolerance (t), the tool life (l) and the work material (m) are set as shown in Figure 2(a) - (d). The fuzzy membership functions can be found based on the information provided in the handbooks, such as [10] and other technical references, such as [11]. Next, the total fuzzy membership of speed is calculated using the max-min rule [8]. Let:

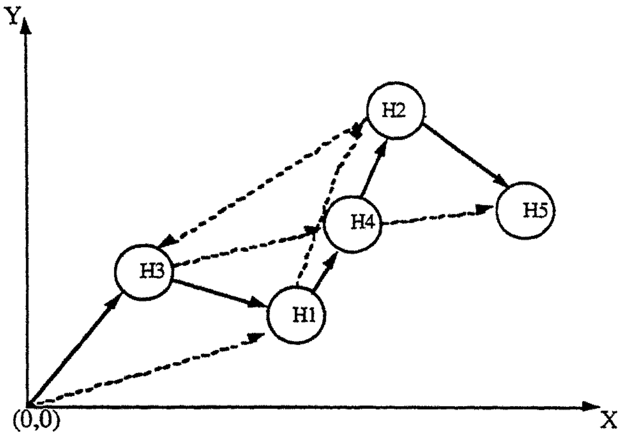
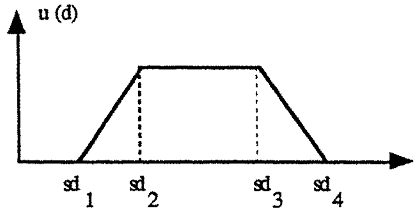
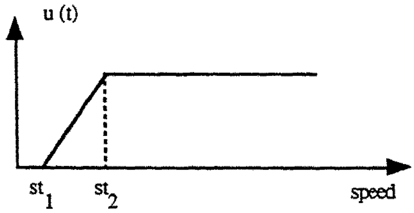


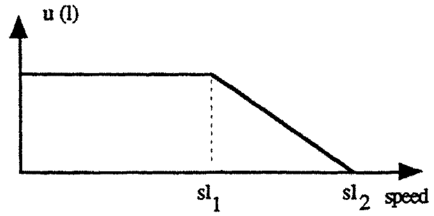
Fig. 1. Illustration of drill travels



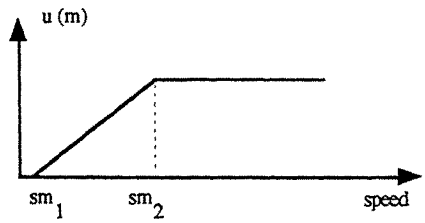
(a) the fuzzy membership function of speed given drill size



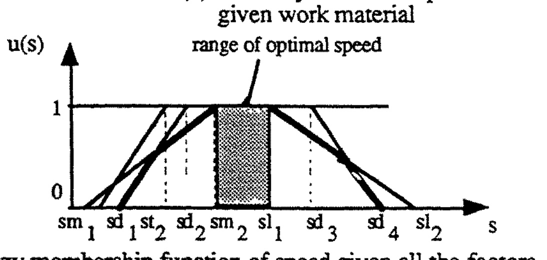
(b) the fuzzy membership function of speed given tolerance



(c) the fuzzy membership function of speed given tool life



(d) the fuzzy membership function of speed given work material



(e) the fuzzy membership function of speed given all the factors

Fig. 2. Illustration of fuzzy membership function of speed

$$u(s) = u(s,d) \cap u(s,t) \cap u(s,l) \cap U(s,m)$$

where, "$\cap$" represents taking minimum or intersection. The resulting fuzzy membership function of drilling speed is shown in Figure 2(e). The desirable cutting speed can then be found as it is corresponding to the maximum fuzzy degree. As an example, given d = 25.4 mm, t = 0.1 mm, l = 60 min. and m = steel, the cutting speed is calculated as between 0.4 m / sec.

3. The system and the simulation results

The Smart Drilling system is developed on an PC computer using C++ language. As shown in Figure 3, it consists of four module: (1) tool selection module, (2) drilling condition design module, (3) NC generation module, and (4) simulation module. The methods for tool selection and the drilling condition design are described in the previous section. The NC code generation module stitches the process plan and generates NC codes in APT format. The cutting simulation is conducted using a commercial software VERICUT®. To facilitate the system development, a knowledge base is used to store the fuzzy decision tables and fuzzy logic functions. Also, a data base is used to store the drills. Both the knowledge base and the data base can be easily modified allowing further upgrade.

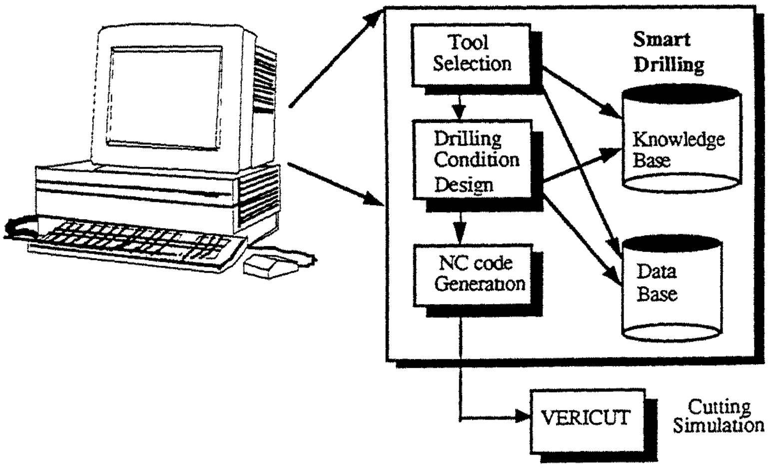


Fig. 3. The structure of the Smart Drilling system

To test the performance of the system, a number of simulations were conducted. During the simulation, holes are generated using a random number generator. For the comparison purpose, a manual drilling process plan is also simulated. In the manual process plan, the holes with a same diameter were drilled together. The drill travel follows the hole identification number and the drilling condition is designed based on the information provided in a handbook [11]. The simulation result is shown in Table 4. From the table, it is seen that Smart Drilling can save an average of 50% of cycle time. This is attributed to the reduced number of tool changes (the more the holes, the more the savings) and reduced drill travel time (independent on the number of holes). Also, Smart Drilling can significantly reduce the process planning time. While manual planning may take hours, Smart Drill design the process plan within minutes.

4. Conclusions

(1) A CAPP system, Smart Drilling, for planning multiple hole drilling operations has been developed. The system is capable of selecting drills and drilling sequence, determining drilling conditions and simulating the drilling operations.

(2) Smart Drilling generates process plan that can save about 50% of production time compared to manual sequential planning. It can also significantly reduce the process planning time.

Table 4. Summary of simulation results

Number of holes	Manual Drilling			Smart Drilling			Total Saving
	# of tool changes	travel distance	cycle time (min)	# of tool changes	travel distance	cycle time (min)	
10	21	447.11	14.70	11	290.15	7.97	45.8%
12	25	512.48	16.13	11	364.87	8.13	47.0%
15	28	788.75	17.98	11	440.72	9.23	50.1%
16	30	790.39	19.92	13	489.47	9.73	51.3%
18	32	940.54	21.37	13	510.56	10.09	52.7%
20	34	1033.80	23.49	15	640.86	11.13	53.6%
21	35	1054.92	24.99	15	650.43	11.45	54.0%
24	38	1090.51	26.58	15	644.57	12.29	54.8%
25	39	1129.59	29.19	15	550.65	12.58	56.1%
26	40	1298.89	32.33	15	752.45	13.01	60.0%

Acknowledgment

This work is partially supported by Manufacturing Research Corporation of Ontario. The authors would like thank for Mr. Dan Marinac and Mr. Uri Shakked of Cimatron Ltd., who challenge the authors to work on the multiple hole drilling planning problem.

References

1. **Adiga, S. and Lin, W. T., 1993**, "An Object-Oriented Architecture for Knowledge-Base Production Scheduling Systems", *Journal of Intelligent Manufacturing*, Vol. 4, pp. 139-150.
2. **Lye, S. and Yeo, S., 1992**, "Development of an Integrated CAD/CAPP/CAM System for Turning Operations," *Journal of Materials Processing Technology*, Vol. 29, pp. 103-117.
3. **Taiber, J. G., 1994**, "Development of an Optimization Method for Determination of Process Sequences Considering Prismatic Workpieces", *Computers in Engineering*, Vol. 1, pp. 271- 280.
4. **Zhang, Y. and Nee, Y., 1994**, "A hybrid approach to computer-aided process planning for prismatic parts," *Computers in Engineering*, Vol. 1, pp. 437- 442.
5. **Thakar, G. and Jain, V., 1993**, "An integrated process planning and NC part programming system for rotational components," *Computers in Industry*, Vol. 21, pp. 341-357.
6. **Vittal, V. N. and Jain, V. K., 1990**, "A computer-Aided Process Planning System for Rotational Parts (CAPP-RP) for an FMS Environment," *International Journal of Computer Applications in Technology*, Vol. 3, No. 2, pp. 61-69.
7. **Qiao, L.H., and Yang, Z. B., 1994**, "A Computer-Aided Process Planning Methodology," *Computers in Industry*, Vol. 25, pp. 83-94.
8. **Klir, J. G. and Folger, A. L., 1982**, *Fuzzy Sets, Uncertainty, and Information*, Prentice Hall.
9. **Garey, M. R. and Johnson, D. S., 1979**, *Computers and Intractability, A Guide to the Theory of NP-Completeness*, W. H. Freeman and Company.
10 **Cubberly, W. H. and Bakerjian, R. (edit), 1989**, *Tool and Manufacturing Engineers Handbook*, Society of Manufacturing Engineers
11. **Sandvik Inc., 1994**, *Metalworking Products Catalog*.

TOLERANCING, TOOL ADJUSTMENT AND MACHINING PARAMETER SELECTION PROBLEMS IN MULTI-PASS TURNING OPERATIONS

M. Liang[(1)], M. Zuo[(2)], and T. Yeap[(3)]
[(1)]Department of Mechanical Engineering, [(3)]Department of Electrical Engineering
University of Ottawa, Ontario, Canada K1N 6N5
[(2)]Department of Manufacturing Engineering, City University of Hong Kong, Hong Kong

ABSTRACT

This paper presents a two-stage approach to tolerancing and machining parameter selection decisions for turning operations. The first stage deals with design related tolerancing problem. The tolerance obtained in this stage is then used as an upper tolerance limit in the second stage. The manufacturing related issues such as machining parameters, dimension deviation for tool adjustment and number of passes for multi-pass turning operations are simultaneously determined in the second stage. The application of the proposed approach is demonstrated using example problems.

KEYWORDS

Tolerancing, Tool Adjustment, Machining Parameter Selection

1. INTRODUCTION

Tolerancing, tool adjustment, and machining parameter selection decisions have a great impact on overall production economics for turning operations. For example, making a part with a very tight tolerance is expensive due to high machining cost whereas a loosely specified tolerance is equally undesirable because of the possible scrap or rework cost. Similar dilemma exists in machining parameter selection. Though a high cutting speed is preferred for machining cost reduction, it may lead to increased tool and tool replacement costs. In addition, the number of cutting passes also affects the total production cost. Furthermore, tool adjustment decisions in the final pass should be made carefully to meet the tolerance requirements without jeopardizing the cost reduction objective.

Several methods have been proposed for design related tolerancing (e.g., Spotts 1973, Sutherland and Roth 1975, Dong and Soom 1990). In manufacturing stage, the machining parameter problem has also been investigated by researchers (e.g., Ermer *et al* 1981 and Wang 1992). However, it is noted that these problems are treated separately and therefore often lead to conflicting design and manufacturing decisions. In view of this, we propose a two-stage approach to integrate tolerancing, tool adjustment and machining parameter selection decisions.

2. PROBLEM ANALYSIS

The problems of tolerancing, tool adjustment and machining parameter selection are solved in two stages. The optimum part tolerance will be determined in the first stage and used as an upper limit of tolerance in the second stage. In the second stage, tool adjustment decision is incorporated in the machining parameter selection for a single-machine multi-pass turning operation to determine the optimum cutting speed, feed rate, depth of cut, dimension deviation and pass selection.

2.1 Notations

The notations are defined as follows:

j=cutting pass, $j,\ldots,J$;, δ=tolerance variable in design stage and maximum allowable tolerance in manufacturing stage (mm); δ_l=minimum feasible tolerance (mm); δ_u=maximum allowable tolerance (mm); v_j=cutting speed for pass j (m/min); f_j=feed rate for pass j (mm/rev); d_j=depth of cut for pass j (mm); d_{total}=total depth of material to be removed (mm); $d_j^{(u)}$= maximum depth of cut for pass j (mm); y =part dimension deviation in the manufacturing stage (mm); p_j= 1, if the cutting pass is selected, = 0, otherwise; D_0= initial stock diameter (mm); D_{j-1} =part diameter before the (j-1)th pass (mm); L =part length (mm); C_o=operating cost rate ($/min); C_t=tool change cost ($/edge); C_a=tool adjustment cost rate ($/min); t_r=tool replacement time (min); t_a =tool adjustment time (min); t_m=machining time (min); w=maximum allowed nose wear of a tool (mm);

T_L=tool life (min); η_m=efficiency of a machine tool; $F_c^{(u)}$=maximum cutting force (Kg); K_P =spindle power constant; $P_c^{(u)}$=maximum cutting power of the motor (KW); $S_F^{(u)}$=maximum allowable surface roughness (μm); $v_j^{(l)}, v_j^{(u)}$=minimum, maximum cutting speed for pass j (m/min); $f_j^{(l)}, f_j^{(u)}$=minimum, maximum feed rate for pass j (mm/rev); $d_J^{(l)}$, $d_J^{(u)}$=minimum, maximum depth of cut for the final pass (mm); α_T, β_T, γ_T, K_T=constants in tool life equation; β_F, γ_F, K_F=cutting force constants; α_S, β_S, γ_S, K_S=surface finish constants; g_1, g_2, g_3, g_4=constants in the production cost-tolerance function.

2.2 Design Stage

The aim of the first (design) stage is to determine part tolerance based on machining cost and quality loss criteria. A generalized model proposed by Dong and Soom (1990) is adopted to express the machining cost-tolerance relationship:

$$g(\delta) = g_1 e^{-g_2(\delta - g_3)} + g_4 \tag{1}$$

where $g(\delta)$ is the cost of producing a mechanical feature with a specified tolerance level δ. g_1, g_2, g_3 and g_4 are constants which are determined by curve-fitting based on experimental data.

A quality loss function suggested by Taguchi *et al* (1989) is used to describe the quality loss-tolerance relationship. The quality loss is a function of quality deviation from the target dimension:

$$L(x) = \frac{A}{e^2}(x-m_o)^2 \quad (2) \qquad \text{or} \qquad L_q(\delta) = \frac{A}{\delta_u^2}\delta^2 \quad (3)$$

where A is rework or scrap cost, e tolerance limit, x quality characteristic of the product, and m_o the quality target value. The sum of Eqs (1) and (3) will be used to formulate the tolerancing model.

2.3 Manufacturing Stage

The purpose of the second (manufacturing) stage is to simultaneously determine the optimum cutting speed, feed rate, depth of cut, dimension deviation for tool adjustment, and the number of passes in terms of production cost. The main production cost components are discussed below.

2.3.1 Machining Cost

Machining cost is the cost incurred during the actual cutting process which depends on machining time. Machining time and cost are given by:

$$t_m = \frac{\pi D L}{1000\, v f} \quad (4) \qquad\qquad C_m = \frac{\pi D L C_o}{1000\, v f} \quad (5)$$

2.3.2 Tool Cost

Tool cost is the cost per cutting edge which depends on tool life T_L (min) and machining parameters. The extended Taylor's tool-life equation may be written as (Wang 1992):

$$T_L = \frac{K_T}{v^{\alpha_T} f^{\beta_T} d^{\gamma_T}} \tag{6}$$

If C_t is denoted as tool cost per cutting edge, then tool cost for a single workpiece is:

$$C_{to} = \frac{t_m}{T_L} C_t = \frac{\pi D L}{1000 K_T} v^{\alpha_T - 1} f^{\beta_T - 1} d^{\gamma_T} C_t \tag{7}$$

2.3.3 Tool Replacement Cost

Denoting t_r as the time to replace a tool, the tool replacement time distributed to a part is

$$t_c = \frac{t_m}{T_L} t_r \tag{8}$$

Tool replacement cost is the product of tool replacement time and operating cost rate:

$$C_{tr} = \frac{t_m}{T_L} C_o t_r = \frac{\pi D L}{1000 K_T} v^{\alpha_T - 1} f^{\beta_T - 1} d^{\gamma_T} C_o t_r \tag{9}$$

2.3.4 Tool Adjustment Cost

Due to tool wear, the dimension of the part at the end of the process may be larger than that at the beginning of the process (Figure 1). To maintain the part quality within the specified tolerance limits, it may be desirable to adjust the cutter when or before the dimension deviation reaches the tolerance limit. Cutter nose wear is more relevant for tool position compensation (Du *et al* 1993). Tool adjustments are often done off-line which involve additional machine down time and labour cost. Therefore, a trade-off between the improved quality and increased cost has to be taken into account.

According to Narang and Fischer (1993), the amount of tool wear causing dimension deviation during a pass for a single workpiece is:

$$w_p = \frac{t_m}{T_L} w \tag{10}$$

where w is the maximum allowable nose wear of the tool. The number of tool adjustments during a single pass for a single workpiece is therefore:

$$n_{ad} = \frac{t_m w}{T_L y} \tag{11}$$

where y is the dimension deviation at which the cutter is adjusted. Accordingly, the tool adjustment cost distributed to each part will be:

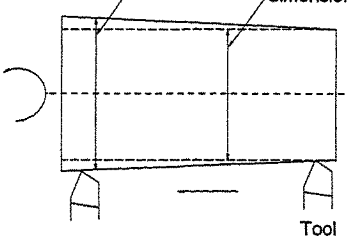


Figure 1 Effect of tool wear on dimension

$$C_{ad} = \frac{t_m w}{T_L y} C_a t_a = \frac{\pi D L}{1000 K_T} v^{\alpha_T - 1} f^{\beta_T - 1} d^{\gamma_T} \frac{w}{y} C_a t_a \tag{12}$$

2.3.5 Quality Loss

Referring to Equations (3) and (11), the quality loss per piece caused by dimension deviation is the product of the number of tool adjustments distributed to each workpiece and the cumulative quality loss between two consecutive adjustments, i.e.,

$$L_q = \frac{t_m w}{T_L y}\left(\frac{A}{\delta^2} y^2\right) = \frac{\pi D L}{1000 K_T} v^{\alpha_T - 1} f^{\beta_T - 1} d^{\gamma_T} \frac{w}{y}\left(\frac{A}{\delta^2} y^2\right) \tag{13}$$

Note that δ instead of δ_u is used simply because we are now dealing with a manufacturing stage problem and the value of δ is obtained in the design stage.

3. OPTIMIZATION MODELS

3.1 The First-stage Model

Referring to Equations (1) and (3), the model of this stage is formulated as follows.

Model 1

$$Min\ G(\delta) = \left[g_1 e^{-g_2(\delta - g_3)} + g_4\right] + \frac{A}{\delta_u^2}\delta^2 \tag{14}$$

subject to:

$$\delta_l \le \delta \le \delta_u \tag{15}$$

The objective of the model is to find optimal tolerance at the minimum total cost in the design stage. The optimal tolerance will be used as the upper limit in determining the dimension deviation for tool adjustments. Constraint (15) defines the region at which the tolerance is technically feasible. Fig. 2 below shows the effect of tolerance on machining cost, quality loss, and total cost.

3.2 The Second-stage Model

In machining operations, the depth of cut and the number of passes selected are often either too conservative leading to low productivity or too aggressive causing unnecessary tool breakages and part defects. In view of this, a second-stage model is developed in such a way that the pass selection decision is also made in parallel to tool adjustment and machining parameter selection decisions. An example of a four-pass turning operation is shown in Figure 3. The sum of machining cost, tool cost,

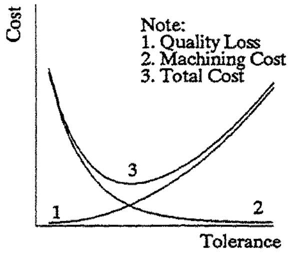


Figure 2 Machining cost, quality loss, and total cost

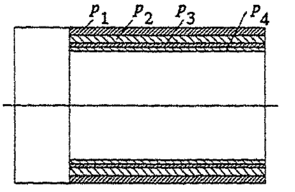


Figure 3 A four-pass turning operation

tool replacement cost, tool adjustment cost and quality loss of all passes will form the objective function. The tool adjustment cost and quality loss will be considered only in the last cutting pass where the final dimension of the part is obtained.

Model 2

$$Min\sum_{j=1}^{J}\frac{\pi D_{j-1}L}{1000K_T}\left[\frac{C_oK_T}{v_jf_j}+v_j^{\alpha_T-1}f_j^{\beta_T-1}d_j^{\gamma_T}(C_t+C_ot_r)\right]p_j+\frac{\pi D_{J-1}L}{1000K_T}v_J^{\alpha_T-1}f_J^{\beta_T-1}d_J^{\gamma_T}\frac{w}{y}\left(C_at_a+\frac{A}{\delta^2}y^2\right)p_J \quad (16)$$

where

$$D_{j-1}=D_0-2\sum_{q=1}^{j-1}d_q \quad (17)$$

subject to:

$$v_j^{(l)}\le v_j\le v_j^{(u)} \quad \forall j \quad (18)$$

$$f_j^{(l)}\le f_j\le f_j^{(u)} \quad \forall j \quad (19)$$

$$d_j\le d_j^{(u)} \quad j=1,2,\ldots,J\text{-}1 \quad (20)$$

$$d_J^{(l)}\le d_J\le d_J^{(u)} \quad (21)$$

$$0\le y\le \delta \quad (22)$$

$$p_J=1 \quad (23)$$

$$\sum_{j=1}^{J}d_j=d_{total} \quad (24)$$

$$\sum_{j=1}^{J}d_jp_j=d_{total} \quad (25)$$

$$\frac{d_j}{d_{total}}-p_j\le 0 \quad \forall j \quad (26)$$

$$0\le p_j\le 1 \quad \forall j \quad (27)$$

$$K_Ff_j^{\beta_F}d_j^{\gamma_F}\le F_c^{(u)} \quad \forall j \quad (28)$$

$$K_Pv_jf_j^{\beta_F}d_j^{\gamma_F}\le P_c^{(u)} \quad \forall j \quad (29)$$

$$K_Sv_J^{\alpha_S}f_J^{\beta_S}d_J^{\gamma_S}\le S_F^{(u)} \quad (30)$$

The objective of Model 2 is to minimize the total production cost. For a given tool-workpiece material combination, constraints (18) through (20) specify the lower and upper bounds for cutting speed, feed rate, and depth of cut respectively. Constraint (21) restricts the depth of cut for the last pass. Constraint (22) specifies that the dimension deviation at which the tool is adjusted should be within the tolerance specified in the design stage. Constraint (23) shows that at least one pass should be performed and it is the final pass. Constraint (24) states that the sum of depths of all passes should be equal to the total material to be removed. The purpose of constraint (26) is to use continuous variables P_j. Constraints (24) through (27) all together guarantee that p_j will take a binary value 0 or 1. As p_j is a continuous variable, the computational time can be significantly reduced.

The limiting cutting force and spindle power are given respectively in constraints (28) and (29) as suggested by Iwata *et al* (1977). The values of coefficients β_F, γ_F and K_F depend on tool-part combination. The empirical coefficient K_P for cutting power is the function of coefficient K_F and mechanical efficiency of the machine tool (Iwata *et al* 1972):

$$K_P = \frac{K_F}{6120\,\eta_m} \tag{31}$$

Constraint (30) reflects the relationship between surface finish and machining parameters (Bhattacharyya *et al* 1970). α_S, β_S, γ_S and K_S are constants and are tool-part dependent.

4. ILLUSTRATIVE EXAMPLES

Example for Model 1

Consider a low carbon steel shaft to be machined from a rotational stock. The length and diameter of the stock are respectively 250 mm and 90 mm. A portion of the shaft is 200 mm long and is to be machined to a diameter of 80 mm. The optimal tolerance of the final dimension is determined based on the data given by Dong and Soom (1990). For $40 \leq D < 500$ mm: $g_1 = 3.96$, $g_2 = 21.65$, $g_3 = 0.00$, $g_4 = 1.04$. The lower tolerance limit is $\delta_l = 0.002$ mm and the upper limit is $\delta_u = 0.08$ mm. The maximum quality loss is \$ 1.0. Using the LINGO software, the first-stage problem is solved in less than a second on a 486 PC. The optimal tolerance is 0.0659 mm.

Example for Model 2

As given above, $D_0 = 90$ mm and $L = 200$ mm. The total depth of cut is 5 mm. The tolerance limit is $\delta = 0.0659$ mm which is obtained from Model 1. The required surface roughness $S_F^{(u)} = 1.6$ μm. A maximum of four cutting passes is considered. Other data are given in Tables 1 and 2. The data related to cutting force and cutting power are taken from Wang (1992).

It took less than 3 sec. to solve this problem. The results are summarized in Table 3. As shown in the table, only two passes are selected. The optimum dimension deviation for tool adjustment is 0.051 mm. The total production cost for the two passes is \$ 2.2345 per piece.

5. CONCLUSION

In this paper, tolerancing, tool adjustment and machining parameter selection problems in turning operations have been jointly solved using a 2-stage approach. The optimum tolerance is determined by the first stage model and the optimum machining parameters, dimension deviation for tool adjustment, and the number of cutting passes are provided by solving the second stage model. The two models are coupled by using the optimum tolerance from the first stage model as the upper limit for the optimum dimension deviation for tool adjustment. The consistency in shop floor planning can be achieved using the proposed approach.

REFERENCES

A. Bhattacharyya, R. Faría-González and I. Ham, 1970, Regression analysis for predicting surface finish and its application in the determination of optimum machining conditions. *J. Eng. Ind.* **92**, 711-714.

Z. Dong and A. Soom, 1990, Automatic optimal tolerance design for related dimension chains. *Manuf. Review*, 3, 262-271.

R. Du, B. Zhang, W. Hungerford and T. Pryor, 1993, Tool condition monitoring and compensation in finish turning using optical sensors. *SAMI Doc.*, Windsor, Ontario, Canada.

D. Ermer and S. Kromodihardjo, 1981, Optimization of multipass turning with constraints. *J. Eng. Ind.* **103**, 462-468.

K. Iwata, Y. Murotsu, T. Iwatsubo and S. Fujii, 1972, A probabilistic approach to the determination of the optimum cutting conditions. *J. Eng. Ind.* **94**, 1099-1107.

K. Iwata, Y. Murotsu, and F. Oba, 1977, Optimization of cutting conditions for multi-pass operations considering probabilistic nature in machining process. *J. Eng. Ind.* **99**, 210-217.

R. Narang and G. Fischer, 1993, Development of a framework to automate process planning functions and to determine machining parameters. *Int. J. Prod. Res.* **31**, 1921-1942.

M. Spotts, 1973, Allocation of tolerances to minimize cost of assembly. *J. Eng. Ind.* **95**, 762-764.

G. Sutherland and B. Roth, 1975, Mechanism design: accounting for manufacturing tolerances and costs in function generating problems. *J. Eng. Ind.* **98**, 283-286.

G. Taguchi, E. Elsayed and T. Hsiang, 1989, *Quality Engineering in Production Systems*, McGraw-Hill Book Company, New York.

J. Wang, 1992, Multiple-objective optimization of machining operations based on neural networks. *Int. J. Adv. Manuf. Tech.*7, 1-9.

Table 1 Input parameters for Model 2

Parameters	Values
Operating cost rate C_o	\$ 3.0/min
Tool cost C_t	\$ 5.5/edge
Tool replacement time t_r	0.5 min
Maximum allowed Nose wear w	0.1 mm
Tool adjustment cost C_a	\$ 3.0/min
Tool adjustment time t_a	0.2 min
Rework cost A	\$ 1.0
Efficiency of machine tool η_m	0.8
Constants for tool life α_T, β_T,γ_T, K_T	1.70, 1.55, 1.22,1570000
Maximum cutting force $F_c^{(u)}$	20.0 Kg
Constants for cutting force β_F, γ_F, K_F	1.18, 1.26, 1.38
Maximum cutting power $P_c^{(u)}$	2.0 KW
Constants for surface roughness α_S, β_S, γ_S, K_S	-0.25, 0.72, 0.23, 1.17

Table 2 Minimum and maximum allowed machining parameters

	Cutting speed (m/min)		Feed rate (mm/rev)		Depth of cut (mm)		
Pass	$v_j^{(l)}$	$v_j^{(u)}$	$f_j^{(l)}$	$f_j^{(u)}$	$d_j^{(u)}$	$d_j^{(l)}$	$d_j^{(u)}$
1	90	120	0.8	2.0	5.0		
2	90	120	0.8	2.0	5.0		
3	90	120	0.8	2.0	5.0		
4	168	210	0.13	0.50		0.3	1.0

Table 3 Results for Model 2 example

Pass	Machining parameter	Value	Pass	Machining parameter	Value
	v_2	120.000		v_4	210.000
2	f_2	2.000	4	f_4	0.500
	d_2	4.361		d_4	0.638
				y	0.051

THREE-VALUED CALCULI IN THE PROBLEM OF CONVERSION FROM CSG TO BOUNDARY MODELS

A.G.GORELIK
Laboratory of Geometric Modeling
The Institute of Engineering Cybernetics of Belarusian Academy of Sciences
Minsk

ABSTRACT

This paper describes a computer method for conversion from Constructive solid geometry (CSG) models to Boundary representation (B-Rep). CSG-model of the object is represented in an algebraic form as a function of arbitrary length. The conversion from CSG-model to B-rep is performed in one step only. For solving this problem , the Indexing Three-valued Calculi (ITC) are used

KEYWORDS

Constructive Solid Geometry, Boundary Representation, Conversion, Three-valued Calculus.

1. INTRODUCTION

In the geometric modeling systems, CSG and B-Rep-models are mostly used. The many papers were devoted to problem of the conversion from one model to another [1-10]. As a rule, CSG-model of the object D is represented as a tree structure, which is derived step-by-step in the process of object construction. The nodes of this tree are solids, and the arcs of the tree are operations on the solids.

In most cases construction of B-Rep of new object D from CSG is fulfilled at every step with help of two operands connected with a set-theoretic operation. Sometimes, the CSG-model of object D is created by a method other than step-by-step construction of the object D. Then we have an arbitrary Boolean function of the form

$$D = F(D_1, D_2, \ldots D_n) \tag{1}$$

in which the set-theoretic operations of intersection, union, difference, and complement are applied to initial operands $D_1, D_2, \ldots, D_n$. In this case we have deal with the function of an arbitrary length. Of course we can convert all such functions to a tree structure, and then all subsequent operation problems may be performed step-by-step with help (n-1) conversions using this tree structure.

It is very interesting to consider the different approach when a new object is constructed on base of CSG-model represented in an algebraic form (1), and the conversion from CSG to B-Rep is fulfilled one time only just after the designer is contended with the results of his work. For solving above-mentioned problem, in [11, 12] a formally justified method has been proposed for constructing efficient algorithms for membership tests on CSG models represented in an algebraic form (1). The method is based on the geometric interpretation of three-valued calculi.

2. GEOMETRIC INTERPRETATION OF THE INDEXING THREE-VALUED CALCULUS

Following are brief description of the geometric interpretation of the indexing three-valued calculus. The interpretation has been detailed in [12].

Let us consider a three-valued calculus

$$\begin{aligned} x_1 \vee x_2 &= \max(x_1, x_2) \qquad & x_1 \wedge x_2 &= \min(x_1, x_2) \\ x_1 \setminus x_2 &= \min(x_1, \bar{x}_2) \qquad & \bar{x} &= -x \end{aligned} \tag{2}$$

defined on a set of elements $\{-1,\ 0_{i,j}\ ,\ 1\}$. Here $i, j,..$ are arbitrary digital indices. These indices may be negative. Operations on these quantities are defined by functions (2). During the performance of operations it is assumed that $0_{i,j,}$ has a zero value. Furthermore, operations (2) are defined in such a way that, if an operation yields a zero value, this result should have the same subscripts as the operands of the original formula:

$$
\begin{aligned}
&-1 \vee 0_{i,j,} \quad = 0_{i,j,} \\
&1 \wedge 0_{i,j} \quad = 0_{i,j,} \\
&0_{i,j,} \ \vee 0_{k,l} \quad = 0_{i\ j\ \ ,k,l} \\
&\bar{0}_i = 0_{-i} \\
&1 \setminus 0_i = 1 \wedge \bar{0}_i = 1 \wedge 0_{-i} = 0_{-i}
\end{aligned}
$$

However, the subscripts may vanish

$$-1 \wedge 0_{i\ j} \quad = -1 \qquad\qquad 1 \vee 0_{i\ j} \quad = 1.$$

Below we shall deal only with regular objects

$$D = \left[Int(D)\right] \tag{3}$$

(here $[D]$ = closure of D, Int = interior) and regularized set-theoretic operations [13, 14]:

$$
\begin{aligned}
&D_1 \cup^* D_2 = \left[Int(D_1) \cup Int(D_2)\right] = D_1 \cup D_2 \\
&D_1 \cap^* D_2 = \left[Int(D_1) \cap Int(D_2)\right] \\
&D_1 \setminus^* D_2 = \left[Int(D_1) \setminus Int(D_2)\right] \\
&\overline{D} = \text{ closure of the complement of } D.
\end{aligned} \tag{4}
$$

Using operations (4) on the solids (3), the function (1) is converted to a regularized logical function (rlf)

$$D = F^*(D_1, D_2, \ldots, D_n). \tag{5}$$

Let us define a characteristic function $I(D_i, M)$ for an arbitrary point M and regular set D_i:

$$
I(D_i, M) = \begin{cases} -1, \ M \notin D_i \\ 0_i \text{ or } 0_{-i}, \ M \in Fr(D_i) \\ 1, \ M \in Int(D_i) \end{cases}
$$

If D_i appears in (5) as $\overline{D}_i$ then, in the case when $M \in Fr(D_i)$, we have $I(\overline{D}_i, M) = 0_{-i}$.

In (5) we shall replace D_i by $I(D_i, M)$, $1 \le i \le n$, and replace the intersection, union, difference, and complement operators respectively by the operations from the indexing three-valued calculus (2). Then we obtain the expression:

$$I(D, M) = F(I(D_1, M),\ I(D_2, M), \ldots,\ I(D_n, M)) \tag{6}$$

Statement.

1. If the rules for indexing three-valued calculus, applied to (6), result in $I(D,M)=1$ then $M \in Int(D)$.
2. If $I(D,M)=-1$ then $M \notin D$.
3. If $I(D,M)=0_i$ then $M \in Fr(D)$ and, in the vicinity of point M, the boundary of the object D coincides with the boundary of the object D_i.
4. If $I(D,M)=0_{-i}$ then $M \in Fr(D)$ and, in the vicinity of point M, the boundary of the object D coincides with the boundary of the object $\overline{D}_i$.
5. If $I(D,M)=0_{i,j,}$ (i.e. there are more than one index for zero value) then the location of point M relative to D is undetermined.

Example. Let the following regions be specified on the plane XOY:

D_1: $-x \ge 0$

D_2: $-y \ge 0$

D_3: $-x^2-y^2+4 \ge 0$

D_4: $(x-0.5)^2+(y+0.5)^2-2.5 \ge 0$

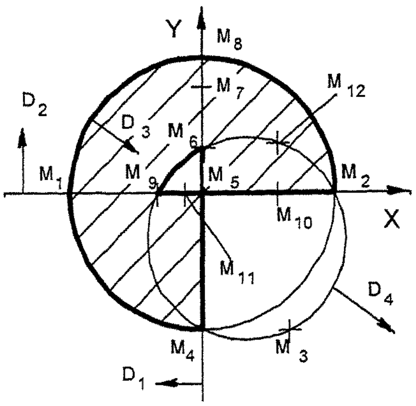


Fig. 1: Regions D_1, D_2, D_3, D_4

Then the hatched area D in Fig. 1 may be described by:

$$D=(D_1 \cup^* D_2) \cap^* D_3 \cap^* (\overline{D}_1 \cup^* \overline{D}_2 \cup^* D_4) \qquad (7)$$

For the function (7) we obtain at points M_1, M_7, M_{10}, M_{11}:

$$I(D,M_1)=(1 \vee 0_2) \wedge 0_3 \wedge (-1 \vee 0_{-2} \vee 1)=0_3$$
$$I(D,M_7)=(0_1 \vee 1) \wedge 1 \wedge (0_{-1} \vee -1 \vee 1)=1$$
$$I(D,M_{10})=(-1 \vee 0_2) \wedge 1 \wedge (-\overline{1} \vee \overline{0}_2 \vee -1)=0_2$$
$$I(D,M_{11})=(1 \vee 0_2) \wedge 1 \wedge (\overline{1} \vee \overline{0}_2 \vee -1)=0_{-2}$$

Thus, $M_1 \in Fr(D)$ and in the vicinity $\varepsilon(M_1)$ $Fr(D)$ coincides with $Fr(D_3)$; $M_{10} \in Fr(D)$ and in the vicinity $\varepsilon(M_{10})$ $Fr(D)$ coincides with $Fr(D_2)$; $M_{11} \in Fr(D)$ and in the vicinity $\varepsilon(M_{11})$ $Fr(D)$ coincides with $Fr(\overline{D}_2)$. Of course in set-theoretic sense $Fr(D_i)$ coincides with $Fr(\overline{D}_i)$. But for us it is interesting the orientation of these boundaries. In this sense $Fr(D_i)$ does not coincides with $Fr(\overline{D}_i)$ because they have opposite orientations.

The region shown in Fig.1 can also be described by another logical function:

$$D=((D_1 \setminus^* D_2) \cup^* (D_2 \setminus^* D_1) \cup^* (D_1 \cap^* D_2 \cap^* D_4)) \cap^* D_3 \qquad (8)$$

For this function we get at points M_1, M_7, M_{10}, M_{11}:

$$I(D,M_1)=((1 \setminus 0_2) \vee (0_2 \setminus 1) \vee (1 \wedge 0_2 \wedge 1)) \wedge 0_3=(0_{-2} \vee 0_2) \wedge 0_3=0_{-2,2,3}$$
$$I(D,M_7)=((0_1 \setminus 1) \vee (1 \setminus 0_1) \vee (0_1 \wedge 1 \wedge 1)) \wedge 1=0_{-1} \vee 0_1=0_{-1,1}$$

$$I(D, M_{10}) = ((-1 \setminus 0_2) \vee (0_2 \setminus -1) \vee (-1 \wedge 0_2 \wedge -1)) \wedge 1 = 0_2$$
$$I(D, M_{11}) = ((1 \setminus 0_2) \vee (0_2 \setminus 1) \vee (1 \wedge 0_2 \wedge -1)) \wedge 1 = 0_{-2}$$

Thus, the location of points M_1, M_7 relative to D is not found. But if we use Regular ITC in this case we have (see [12]):

$$I(D, M_1) = (0_{-2} \vee 0_2) \wedge 0_3 = 1 \wedge 0_3 = 0_3.$$
$$I(D, M_7) = 0_{-1} \vee 0_1 = 1.$$

3. CONSTRUCTION OF THE OBJECT BOUNDARY

The boundary of the object D will be considered has been oriented relative to this object. Then the object D is unambiguously determined by this boundary [14]. The boundary of a regular object D can be considered as an oriented "graph" $G=(U,V)$. The arcs of this graph are the boundary lines of the object D, and the nodes are common points of these lines (Fig.2).

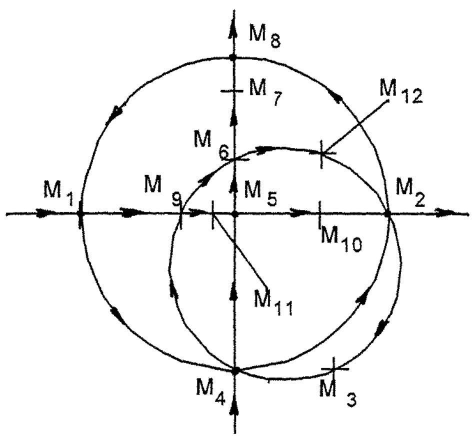


Fig. 2 : Oriented "graph" of the object boundary

For the rlf (5), we can construct a new oriented graph $G^+ = (U^+, V^+)$. The nodes of the graph G^+ are all intersection points of the boundaries of the $D_1, D_2, \ldots, D_n$, and the arcs of the graph G^+ are the parts of the lines connecting these nodes. From the above, it follows that the arcs of the graph G must be sought among the arcs of the graph G^+.

Thus for an arbitrary arc $u^+ \in U^+$ we must have a test to know

1) $u^+ \in U$ or $u^+ \notin U$,
2) $\overline{u}^+ \in U$ or $\overline{u}^+ \notin U$.

Here $\overline{u}$ coincides with a part of the line u but it has an opposite orientation.

The test may be constructed with help ITC or RITC. It contains the following steps:

1. Take an any inner point M on the arc u^+.
2. In this point, calculate the function value $I(D, M)$. Then we can have the following variations

2.1. $I(D,M)=1$ or $I(D,M)=-1$. It means that $u^+ \notin U$ and $\overline{u}^+ \notin U$. For example, in Fig.2 at points M_3, M_{12} $I(D,M_3)=-1$ and $I(D,M_{12})=1$. It means that $(M_2 - M_4) \notin U$ and $(M_6 - M_2) \notin U$. In the following, for a part of any line between points M_k and M_l, instead u^+, we will also use the designation $(M_k - M_l)$ and instead $\overline{u}^+$ we will also use the designation $\overline{(M_k - M_l)}$.

2.2. $I(D,M)=0_l$. It means that $u^+ \in U$. For example, in Fig. 2 at point M_{10} , $I(D,M_{10})=0_2$ It means that $(M_5 - M_2) \in U$.

2.3. $I(D,M)=0_{-l}$. It means that $\overline{u}^+ \in U$. For example, at point M_{11} we obtain $I(D,M_{11})=0_{-2}$. It means that $\overline{(M_9 - M_5)} \in U$.

2.4. $I(D,M)=0_{i,j,}$, i.e. there are more than one index for zero value. It means in U^+ there are a few coinciding arcs u_i^+, u_j^+,... and may be they have an opposite orientation. If all have the same orientation we can take an any arc from this arcs. Otherwise we can treat the vicinity of the point M relative to an any arc from u_i^+, u_j^+,..., for example relative to u_i^+. For that it needs to take 2 points: the point M' is to the left of u_i^+ and the point M'' is to the right of u_i^+. Now we calculate the function value $I(D,M')$ and $I(D,M'')$.

But in an explicit form we do not need to seek points M', M''. We can make it otherwise. Because we know in advance that the point M' is to the left of u_i^+ and the point M'' is to the right of u_i^+ then in these points

$$I(D_i,M')=1 \text{ and } I(D_i,M'')=-1 \tag{9}$$

Now if $u_j^+=u_i^+$, then for u_j^+ we have also

$$I(D_j,M')=1 \text{ and } I(D_j,M'')=-1. \tag{10}$$

If $u_j^+=\overline{u}_i^+$, then for u_j^+ we obtain

$$I(D_j,M')=-1 \text{ and } I(D_j,M'')=1 \tag{11}$$

Considering (9), (10), (11) we can calculate the function value $I(D,M')$ and $I(D,M'')$ for the arc u_i^+ using the point M throughout except $I(D_i,M)$, $I(D_j,M)$. Thus:

1) if $I(D,M')=1$ & $I(D,M'')=-1$ then $u_i^+\in U$,

2) if $I(D,M')=-1$ & $I(D,M'')=1$ then $\overline{u}_i^+\in U$,

3) otherwise the arcs u_i^+,u_j^+ do not form the boundary of the object D.

For example in Fig.2 for function (8) and for ITC (not for RITC) we have $I(D,M_7)=0_{-1,1}$. But we know in advance that in the points M_7', M_7'' $I(D_1,M_7')=1$ and $I(D_1,M_7'')=-1$. Then for function (8) we obtain

$$\begin{aligned}&I(D,M_7')=((I(D_1,M_7')\setminus I(D_2,M_7))\vee(I(D_2,M_7)\setminus\\&I(D_1,M_7'))\vee(I(D_1,M_7')\wedge I(D_2,M_7)\wedge I(D_4,M_7))\wedge I(D_3,M_7)=\\&=((1\setminus 1)\vee(1\setminus 1)\vee(1\wedge 1\wedge 1))\wedge 1=1\end{aligned}$$

Analogically for point M_7'' we obtain $I(D,M_7'')=1$. It means that $(M_6-M_8)\notin U$.

4. CONCLUSION

The method discussed above forms the basis to convert from CSG to B-Rep with help ITC. Of course there are the techniques to accelerate the solution of this problem. For example, you can calculate the values of the function (5) in the intersection points of the boundaries $D_1,D_2...,D_n$ and remove all points (together with their arcs) in which $I(D,M)=1$ or $I(D,M)=-1$.

You can also use RITC only [12]. Then algorithm discussed above will be more simpler. The item 2.4 of the algorithm do not need be carried out. But RITC is more involved for the calculations.

5. REFERENCES

1. Tilove, R.B., "Set membership classification: A Unified Approach to Geometric intersection Problems", IEEE Transactions on Computers, pp.874-883, 1980.
2. Yamaguchi, F. and Tokieda, T.A., "Unified Algorithm for Boolean Shape operations", IEEE Computer Graphics and Applications, vol.4, N6, pp.24-37, 1984.
3. Chiokura, H. And Kimura, F.A., "Method of Representing the Solid Design Process", IEEE Computer Graphics and Applications, vol.5, N6, pp.32-41, 1985
4. Weiler, K , "Edge-Based Structures for Solid Modelling in Curved-Surface Environment", IEEE Computer Graphics and Applications, vol.5, N1, pp.21-40, 1985.
5. Yamaguchi, F. and Tokieda, T.A., "Solid Modeler with a 4x4 Determinant Processor", IEEE Computer Graphics and Applications, vol.5, N4, pp.51-59, 1985.
6. Patnaik, L.M., Shenoy, R.S. and Krishnan, D., "Set-theoretic operations on polygons using scangrid approach", Computer-aided design, vol.18, N5, pp.275-279, 1986.
7. Toriga, H., Satoh, T., Ueda, K., Chiyokura, H., "UNDO and REDO Operations for Solid Modeling", IEEE Computer Graphics and Applications, vol.6, N4, pp.35-42, 1986.
8 Krishnan, D. and Patnaik, L.M., "Systolic Architecture for Boolean Operations on Polygons and Polyhedra", Computer Graphics Forum, vol.6, pp.203-210, 1987.
9. Segal, M. and Sequin, C.H., "Partitioning Polyhedral objects into Nonintersectioning Parts", IEEE Computer Graphics and Applications, vol.8, N1, pp.53-67, 1988.
10. Gorelik, A.G. and Cercovitch, Y.R., "Set-theoretic operations for 3D objects", preprint, N15, 30pp., Institute of Engineering Cybernetics, Minsk, 1990.
11. Gorelik, A.G., "Logical functions as the means of modelling geometrical objects", PROLAMAT'82, 16pp., Leningrad, 1982.
12. Gorelik, A.G., "Three-valued calculi in set-theoretic solid modeling", CSG-96, Winchester, 1996.
13. Requicha, A.A.G., "Representations of rigid solids: theory, methods and systems", Computer Serveys, vol.12, N4, pp.437-464, 1980.
14 Gorelik, A.G., "The elements of Boolean algebra of two-dimensional objects", "Computing technics in mechanical engineering", pp.3-32, Institute of Engineering Cybernetics, Minsk, 1969, december.

G¹ REFINEMENT OF DATA FOR RAPID PROTOTYPING

D.J. Walton
Department of Computer Science, University of Manitoba
Winnipeg, Manitoba R3T 2N2, CANADA
and
M. Yeung
Integrated Manufacturing Technologies Institute, National Research Council of Canada
Building M4, Montreal Road, Ottawa, Ontario K1A 0R6, CANADA

ABSTRACT

Rapid prototyping refers to processes for producing physical models of products or parts in a very short period of time. One of the technologies used to accomplish this is that of layered manufacturing using, for example, stereolithography apparatus (SLA). SLA may be used to build a plastic model of an object by curing liquid polymer into a solid with an ultra violet (UV) laser. SLA requires as input a triangulated geometric model of the object. Such a triangulated model may be obtained from a smooth geometric model, described as a free-form surface. Often, a crudely triangulated model is available from sampled data. The development of a smooth geometric model from a crudely triangulated model is discussed.

KEYWORDS

Surface Fitting, SLA, G^1 Refinement

1. INTRODUCTION

The Stereolithography Apparatus is suitable for modelling free-form surfaces. Unlike computer numerical controlled (CNC) machines, the shape of the surface to be produced is not limited by the dimensions of a cutting tool. By using SLA it is thus possible to produce models of objects with irregular surfaces. Such models are useful in health care for applications such as fitting of prosthetics, production of masks and moulds for medical treatment by X-rays or laser beams, preparation or training for surgery, and other purposes [1][2]. In reverse engineering, data are sampled from a physical object by laser scanning or using a coordinate measuring machine (CMM) [3]. Such objects may have irregular surfaces. To produce variations of the model, a surface suitable for manipulation in a computer-aided (CAD) system may be fitted to the data. After manipulation another physical model can be rapidly produced using SLA.

2. SURFACE FITTING FOR SLA MODELLING OF IRREGULAR OBJECTS

Traditional global surface fitting techniques such as bicubic B-splines are based on data sampled at the grid points of a network of parametric curves. Such methods are not convenient for fitting a surface to data sampled from an irregular object since it cannot be guaranteed that the sampling will be done along given parametric directions. Global surface fitting usually produce models that are quite smooth. This is a desirable feature in surface fitting, however, it may also cause essential features of an irregular object to be smoothed away. An alternative approach is to compose a surface using patches that are fitted locally. To obtain some visual smoothness, it is desirable that these patches join their neighbours in a manner that preserves positional as well as tangent plane continuity.

Much work has been done on surface construction [4] [5] [6] [7]. Resulting composite surfaces are often not of satisfactory quality [8]. The quality appears to be an artifact of the boundary curves [6]. A recent article [9] describes some results that use Bézier boundary curves based on a technique for point normal interpolation [2]. These boundary curves are such that their intrinsic curve normal vectors at sampled data points have the same directions as the corresponding unit normal vectors specified there. The positions of their interior control points are derived in a natural manner that is more intuitive than using a heuristic placement as done in most conventional methods. The triangular patch itself is based on the Gregory-like patch of Chiyokura [7]; the development of the boundary curves are given in Walton and Yeung [2].

3. POINT NORMAL INTERPOLATION

Given two points $\mathbf{P}_0$ and $\mathbf{P}_3$ with respective unit normal vectors $\mathbf{N}_0$ and $\mathbf{N}_1$, it is desired to find a cubic Bézier curve

$$\mathbf{Q}(t) = \mathbf{P}_0(1-t)^3 + 3\,\mathbf{P}_1(1-t)^2t + 3\,\mathbf{P}_2(1-t)t^2 + \mathbf{P}_3t^3, \qquad 0 \le t \le 1,$$

with endpoints $\mathbf{P}_0$ and $\mathbf{P}_3$, and with unit normal vectors parallel to $\mathbf{N}_0$ and $\mathbf{N}_1$ at $\mathbf{P}_0$ and $\mathbf{P}_3$ respectively.

Let $\mathbf{W}_0 = \mathbf{P}_1 - \mathbf{P}_0$, $\mathbf{W}_1 = \mathbf{P}_3 - \mathbf{P}_2$ and $\mathbf{D} = \mathbf{P}_3 - \mathbf{P}_0$. Then

$$\mathbf{Q}'(t) = 3\mathbf{W}_0(1-t)^2 + 6\,(\mathbf{D} - \mathbf{W}_0 - \mathbf{W}_1)(1-t)t + 3\mathbf{W}_1t^2$$

and

$$\mathbf{Q}''(t) = 6\,(\mathbf{D} - 2\mathbf{W}_0 - \mathbf{W}_1)(1-t) + 6\,(2\mathbf{W}_1 + \mathbf{W}_0 - \mathbf{D})t.$$

Hence at $t = 0$,

$$\mathbf{Q}'(0) = 3\mathbf{W}_0,$$
$$\mathbf{Q}''(0) = 6\,(\mathbf{D} - 2\mathbf{W}_0 - \mathbf{W}_1), \qquad (1)$$

and at $t = 1$,

$$\mathbf{Q}'(1) = 3\mathbf{W}_1,$$
$$\mathbf{Q}''(1) = 6\,(2\mathbf{W}_1 + \mathbf{W}_0 - \mathbf{D}). \qquad (2)$$

The principal normal direction of $\mathbf{Q}(t)$ is parallel to the vector $\{\mathbf{Q}'(t)\times\mathbf{Q}''(t)\}\times\mathbf{Q}'(t)$ [10] which, using vector calculus identities, may be written as $\{\mathbf{Q}'(t)\cdot\mathbf{Q}'(t)\}\mathbf{Q}''(t) - \{\mathbf{Q}'(t)\cdot\mathbf{Q}''(t)\}\mathbf{Q}'(t)$. Hence at $t = 0$,

$$\lambda_0\mathbf{N}_0 = \{\mathbf{Q}'(0)\cdot\mathbf{Q}'(0)\}\mathbf{Q}''(0) - \{\mathbf{Q}'(0)\cdot\mathbf{Q}''(0)\}\mathbf{Q}'(0)$$

and at $t = 1$,

$$\lambda_1\mathbf{N}_1 = \{\mathbf{Q}'(1)\cdot\mathbf{Q}'(1)\}\mathbf{Q}''(1) - \{\mathbf{Q}'(1)\cdot\mathbf{Q}''(1)\}\mathbf{Q}'(1)$$

or, from (1) and (2),

$$||\mathbf{W}_0||^2(\mathbf{D} - 2\mathbf{W}_0 - \mathbf{W}_1) - (\mathbf{W}_0\cdot\mathbf{D} - 2||\mathbf{W}_0||^2 - \mathbf{W}_0\cdot\mathbf{W}_1)\mathbf{W}_0 = \lambda_0\mathbf{N}_0/54 \qquad (3)$$

and

$$||\mathbf{W}_1||^2(2\mathbf{W}_1 + \mathbf{W}_0 - \mathbf{D}) - (2||\mathbf{W}_1||^2 + \mathbf{W}_0\cdot\mathbf{W}_1 - \mathbf{W}_1\cdot\mathbf{D})\mathbf{W}_1 = \lambda_1\mathbf{N}_1/54, \qquad (4)$$

which are 6 equations in the 8 unknowns, $\mathbf{W}_0$, $\mathbf{W}_1$, λ_0 and λ_1. The additional two degrees of freedom are used to set $\mathbf{Q}'(0)\cdot\mathbf{Q}''(0) = \mathbf{Q}'(1)\cdot\mathbf{Q}''(1) = 0$, or from Equations (1) and (2),

$$\mathbf{W}_0\cdot\mathbf{D} - 2||\mathbf{W}_0||^2 - \mathbf{W}_0\cdot\mathbf{W}_1 = 0 \qquad (5)$$

and

$$2||\mathbf{W}_1||^2 + \mathbf{W}_0\cdot\mathbf{W}_1 - \mathbf{W}_1\cdot\mathbf{D} = 0 \qquad (6)$$

which are the boundary conditions of a 'natural' space curve, i.e. an arc-length parameterised curve in which the above derivatives would be with respect to arc-length. This also simplifies subsequent algebraic manipulation. Equations (3) and (4) thus become

$$\mathbf{D} - 2\mathbf{W}_0 - \mathbf{W}_1 = \mu_0\mathbf{N}_0, \qquad \mu_0 = \lambda_0/\,(54||\mathbf{W}_0||^2)$$

and

$$2\mathbf{W}_1 + \mathbf{W}_0 - \mathbf{D} = \mu_1\mathbf{N}_1, \qquad \mu_1 = \lambda_1/\,(54||\mathbf{W}_1||^2),$$

from which it follows that $\mathbf{W}_0$ and $\mathbf{W}_1$ can be expressed in terms of μ_0 and μ_1 as

$$\mathbf{W}_0 = (\mathbf{D} - 2\mu_0\mathbf{N}_0 - \mu_1\mathbf{N}_1)\,/\,3 \qquad (7)$$

and

$$\mathbf{W}_1 = (\mathbf{D} + \mu_0\mathbf{N}_0 + 2\mu_1\mathbf{N}_1)\,/\,3. \qquad (8)$$

Substitution of $\mathbf{W}_0$ and $\mathbf{W}_1$ from (7) and (8) into (5) and (6) with subsequent simplification produces the following system of two equations in the two unknowns, μ_0 and μ_1.

$$2\mu_0{}^2 - \mu_0\mathbf{N}_0\cdot\mathbf{D} + \mu_0\mu_1\mathbf{N}_0\cdot\mathbf{N}_1 = 0$$
$$2\mu_1{}^2 + \mu_1\mathbf{N}_1\cdot\mathbf{D} + \mu_0\mu_1\mathbf{N}_0\cdot\mathbf{N}_1 = 0$$

For $\mu_0 \neq 0$ and $\mu_1 \neq 0$, this system of equations has the solution

$$\mu_0 = \{2\mathbf{N}_0\cdot\mathbf{D} + (\mathbf{N}_0\cdot\mathbf{N}_1)(\mathbf{N}_1\cdot\mathbf{D})\}/\{4 - (\mathbf{N}_0\cdot\mathbf{N}_1)^2\} \qquad (9)$$
$$\mu_1 = -\{2\mathbf{N}_1\cdot\mathbf{D} + (\mathbf{N}_0\cdot\mathbf{N}_1)(\mathbf{N}_0\cdot\mathbf{D})\}/\{4 - (\mathbf{N}_0\cdot\mathbf{N}_1)^2\}. \qquad (10)$$

This result is the same as that presented by Walton and Yeung [2], but using a different development. The resulting curve is independent of the orientations of $\mathbf{N}_0$ and $\mathbf{N}_1$. This is easily verified by, for example, replacing $\mathbf{N}_0$ by $-\mathbf{N}_0$ in (3), (9) and (10) which cause the signs of μ_0 and λ_0 to change, but (3) does not change. It can be observed that for $\mathbf{N}_0\bullet\mathbf{D} = \mathbf{N}_1\bullet\mathbf{D} = 0$, the solution is a straight line joining $\mathbf{P}_0$ to $\mathbf{P}_1$ regardless of the value of $\mathbf{N}_0\bullet\mathbf{N}_1$ because $\mathbf{W}_1 = \mathbf{W}_0 = \mathbf{D}/3$ since $\mu_0 = \mu_0$. A problem with the solution arises when $\mathbf{N}_0\bullet\mathbf{N}_1 = 0$, and either $\mathbf{N}_0\bullet\mathbf{D} = 0$ or $\mathbf{N}_1\bullet\mathbf{D} = 0$, in which case $\mathbf{W}_1 = \mathbf{0}$ or $\mathbf{W}_0 = \mathbf{0}$ respectively when $\mathbf{N}_0$, $\mathbf{N}_1$ and $\mathbf{D}$ are coplanar; for $\mathbf{N}_0$, $\mathbf{N}_1$ and $\mathbf{D}$ non-coplanar, the result is a planar (rather than space) curve whose principal normal vector is perpendicular to $\mathbf{N}_0$ or $\mathbf{N}_1$ respectively. To address this situation, regroup Equations (3) and (4) to obtain

$$\|\mathbf{W}_0\|^2(\mathbf{D} - \mathbf{W}_1) - \{\mathbf{W}_0\bullet(\mathbf{D} - \mathbf{W}_1)\}\mathbf{W}_0 = \lambda_0\mathbf{N}_0/54 \tag{11}$$

and

$$\|\mathbf{W}_1\|^2(\mathbf{D} - \mathbf{W}_0) - \{\mathbf{W}_1\bullet(\mathbf{D} - \mathbf{W}_0)\}\mathbf{W}_1 = -\lambda_1\mathbf{N}_1/54. \tag{12}$$

For $\mathbf{N}_0\bullet\mathbf{N}_1$ close to zero, $\mathbf{N}_0$ and $\mathbf{N}_1$ are not close to being parallel, hence $\mathbf{W}_0$ and $\mathbf{W}_1$ can be expressed as

$$\mathbf{W}_0 = a_0\mathbf{N}_0 + a_1\mathbf{N}_1 + a_2\mathbf{N}_0\times\mathbf{N}_1 \tag{13}$$

and

$$\mathbf{W}_1 = b_0\mathbf{N}_0 + b_1\mathbf{N}_1 + b_2\mathbf{N}_0\times\mathbf{N}_1. \tag{14}$$

The following four equations follow from (11) and (12):

$$\mathbf{W}_0\bullet\mathbf{N}_0 = 0, \tag{15}$$

$$\mathbf{W}_1\bullet\mathbf{N}_1 = 0, \tag{16}$$

$$(\mathbf{D} - \mathbf{W}_1)\bullet(\mathbf{N}_0\times\mathbf{W}_0) = 0, \tag{17}$$

$$(\mathbf{D} - \mathbf{W}_0)\bullet(\mathbf{N}_1\times\mathbf{W}_1) = 0. \tag{18}$$

It follows from (15) and (16) that $a_0 = 0$ and $b_1 = 0$ for $\mathbf{N}_0\bullet\mathbf{N}_1 = 0$. Hence, two equations, (17) and (18), in four unknowns remain. A discussion for $\mathbf{N}_0\bullet\mathbf{N}_1 = 0$ and $\mathbf{N}_0\bullet\mathbf{D} = 0$ in the planar case follows. An analogous discussion holds for $\mathbf{N}_0\bullet\mathbf{N}_1 = 0$ and $\mathbf{N}_1\bullet\mathbf{D} = 0$. Equations (7) and (8) yield $\mathbf{W}_0 = 0.5\,\mathbf{D}$ and $\mathbf{W}_1 = \mathbf{0}$. This planar case should not occur under a sensible triangulation since it means that a surface normal vector at the vertex of a triangle, is parallel to a side of the triangle that meets that vertex. However, for a robust implementation it is desirable to use the additional degrees of freedom to avoid $\mathbf{W}_1 = \mathbf{0}$. Since analogously, $\mathbf{N}_1\bullet\mathbf{D} = 0$ ($\mathbf{N}_0\bullet\mathbf{D} \neq 0$) gives $\mathbf{W}_1 = 0.5\,\mathbf{D}$, it is reasonable to choose $\mathbf{W}_1 = \pm 0.5(\mathbf{N}_1\bullet\mathbf{D})\mathbf{N}_0$. The sign, which determines the orientation of $\mathbf{W}_1$, would normally be chosen so that $\mathbf{W}_1\bullet\mathbf{D} > 0$, but in this case, $\mathbf{W}_1\bullet\mathbf{D} = 0$. Thus, to determine which sign to use, let $\mathbf{G}$ be the centroid of the immediate neighbours of $\mathbf{P}_1$. Choose the sign such that $\mathbf{W}_1\bullet(\mathbf{G} - \mathbf{P}_1) > 0$ as shown in Figure 1.

Note that since the curve joining $\mathbf{P}_1$ and $\mathbf{P}_2$ is independent of the orientations of the respective unit normal vectors $\mathbf{N}_1$ and $\mathbf{N}_2$, the opposite sign for $\mathbf{W}_1$ would cause a cusp at $\mathbf{P}_1$. For $\mathbf{N}_0\bullet\mathbf{N}_1 = 0$ and $\mathbf{N}_1\bullet\mathbf{D} = 0$ it follows analogously that $\mathbf{W}_1 = 0.5\,\mathbf{D}$ and $\mathbf{W}_0 = \pm 0.5(\mathbf{N}_0\bullet\mathbf{D})\mathbf{N}_1$. Thus $a_1 = \pm 0.5(\mathbf{N}_0\bullet\mathbf{D})$ and $b_0 = 0.5(\mathbf{N}_1\bullet\mathbf{D})$. For the planar case, $a_2 = b_2 = 0$. Consider now the non-planar case. For $b_0 \neq 0$, since $\mathbf{N}_0$, $\mathbf{N}_1$ and $\mathbf{N}_0\times\mathbf{N}_1$ form an orthonormal triad, it follows from (13) to (16) and (18) for $\mathbf{N}_0\bullet\mathbf{D} = 0$ and $\mathbf{N}_0\bullet\mathbf{N}_1 = 0$ that $a_2 = \mathbf{D}\bullet(\mathbf{N}_0\times\mathbf{N}_1)$ and $\mathbf{W}_1\bullet\mathbf{D} = b_2\mathbf{D}\bullet(\mathbf{N}_0\times\mathbf{N}_1)$. Since the orientation of the curve joining $\mathbf{P}_0$ to $\mathbf{P}_1$ is independent of the orientations of $\mathbf{N}_0$ and $\mathbf{N}_1$, choose the orientations of $\mathbf{N}_0$ and $\mathbf{N}_1$ for convenience such that $\mathbf{N}_1\bullet\mathbf{D} > 0$ and $(\mathbf{N}_0\times\mathbf{N}_1)\bullet\mathbf{D} > 0$ as shown in Figure 2. It is desirable that $0 < \mathbf{W}_0\bullet\mathbf{D}, \mathbf{W}_1\bullet\mathbf{D} < \|\mathbf{D}\|$. Substitution of a_2 into (17) yields, after some algebraic manipulation

$$a_1b_2 - \mathbf{D}\bullet(\mathbf{N}_0\times\mathbf{N}_1)a_1 + (\mathbf{N}_1\bullet\mathbf{D})\mathbf{D}\bullet(\mathbf{N}_0\times\mathbf{N}_1) = 0. \tag{19}$$

Equation (19) represents a hyperbola in a_1 and b_2, as shown in Figure 3, from which it can be seen that $b_2 \leq 0$ (hence $\mathbf{W}_1\bullet\mathbf{D} \leq 0$) for $0 < a_1 \leq \mathbf{N}_1\bullet\mathbf{D}$. Since $a_1 > \mathbf{N}_1\bullet\mathbf{D}$ means that $\mathbf{W}_0\bullet\mathbf{D} > \|\mathbf{D}\|$, a desirable solution is not available for $a_1 \geq 0$. For $a_0 = 0$ and $a_2 = \mathbf{D}\bullet(\mathbf{N}_0\times\mathbf{N}_1)$, it follows from (13) that $\mathbf{W}_0\bullet\mathbf{D} > 0$ for $a_1 < 0$ provided $|a_1| < \{\mathbf{D}\bullet(\mathbf{N}_0\times\mathbf{N}_1)\}^2/\mathbf{D}\bullet\mathbf{N}_1$. To see if an acceptable value for b_2 can be found, the value for a_1 that gives the smallest value for b_2, i.e. $a_1 = -\{\mathbf{D}\cdot(\mathbf{N}_0 \times \mathbf{N}_1)\}^2/\mathbf{D}\cdot\mathbf{N}_1$, is substituted into (19) to obtain

$$b_2 = \mathbf{D}\cdot(\mathbf{N}_0\times\mathbf{N}_1) + (\mathbf{D}\cdot\mathbf{N}_1)^2 / \mathbf{D}\cdot(\mathbf{N}_0\times\mathbf{N}_1). \quad (20)$$

Substitution of (20) into (14) yields

$$\mathbf{W}_1\cdot\mathbf{D} = \{\mathbf{D}\cdot(\mathbf{N}_0\times\mathbf{N}_1)\}^2 + (\mathbf{D}\cdot\mathbf{N}_1)^2 = \|\mathbf{D}\|^2$$

which is not desirable since then $\mathbf{W}_1\cdot\mathbf{D} > \|\mathbf{D}\|$ for $\|\mathbf{D}\| > 1$.

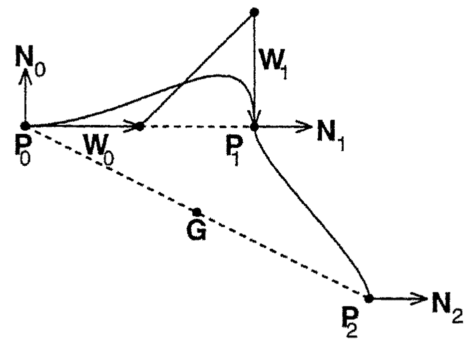


Figure 1. Special case for planar curve.

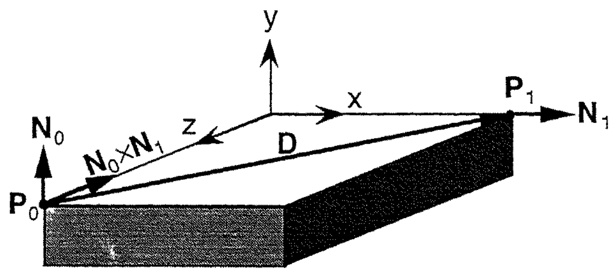


Figure 2. Special case for space curve.

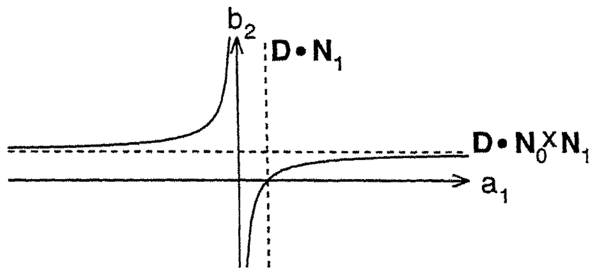


Figure 3. Plot of b_2 against a_1.

As an alternative, the requirement that the resulting curve have its principal normal parallel to $\mathbf{N}_0$ at $\mathbf{P}_0$ is relaxed; instead, it is only required that $\mathbf{W}_0$ be perpendicular to $\mathbf{N}_0$ at $\mathbf{P}_0$. Consistent with the planar case, the choice $\mathbf{W}_0 = 0.5\ \mathbf{D}$ is made. The choice $\mathbf{W}_1 = 0.5\ (\mathbf{D}\cdot\mathbf{N}_1)(\mathbf{N}_0 + \mathbf{N}_0\times\mathbf{N}_1)$ is made as a generalization of the corresponding value in the planar case; it degenerates to the same value as in the planar case when **D** does not have a component in the direction of $\mathbf{N}_0\times\mathbf{N}_1$.

4. EXAMPLES

4.1 Example 1

Data of a physical model of a donkey was obtained by laserscanning. The crudely triangulated model is shown in Figure 4. A refined geometric model was obtained using the technique described. The resulting more finely triangulated model is shown in Figure 5. SLA models of both triangulated models were made. Although 874 triangles are present in the crudely triangulated model, the special cases discussed did not occur; the refined model has 31,464 triangles.

4.2 Example 2

This is a single triangle used specifically to test one of the special cases. The vertices are (0,0,1), (1,0,0), (0,1,0) with respective unit normal vectors (0,1,0), (1,0,0), and (0,1,0). The system of axes used is as shown in Figure 2. The resulting surface is shown in Figure 6.

5. CONCLUSION

The technique presented seems suitable for developing geometric models for use in a layered manufacturing rapid prototyping environment, e.g. using SLA. The model appears reasonably smooth. However, cusps, such as that at the ankle of the donkey in Example 1, are not preserved. This situation may be remedied by marking such sampled points and treating the triangular patches that have such points as vertices, in a special manner. This is currently being investigated.

6. REFERENCES

1. Yeung, M., Allston, J., Orban, P. and Wells, W., "Stereolithography apparatus in health care," Proc. World Congress on Medical Physics and Bio-Medical Engineering, Kyoto, Japan, 1991.

2. Walton, D.J. and Yeung, M., "Geometric modelling from CT scans for stereolithography apparatus," New Advances in CAD & Computer Graphics (Proc. CAD/Graphics'93), Tang, Z. (Ed.), pp. 417-422, International Academic Publishers, Beijing, China, 1993.

3. Seiler, A., Balendran, V., Sivayoganathan, K. and Sackfield, A., "Reverse engineering from uni-directional CMM scan data,' Int. J. Adv. Manuf. Technol., Vol. 11, pp. 276-284, 1996.

4. Mann, S., Loop, C., Lounsbery, M., Meyers, D., Painter, J., DeRose T., and Sloan, K., "A survey of parametric scattered data fitting using triangular interpolants," Curve and Surface Design, Hagan, H. (Ed.), pp.145-172, SIAM, Philadelphia, 1992.

5. Loop, C., "A G^1 triangular surface of arbitrary topological type," Comput. Aided Geom. Design, Vol. 11, pp. 303-330, 1994.

6. Hansford, D., Barnhill, R.E. and Farin, G., "Curves with quadric boundary precision," Comput. Aided Geom. Design, Vol. 11, pp. 519-531, 1994.

7. Chiyokura, H., "Localized surface interpolation method for irregular meshes," Advanced Computer Graphics, Kunii, T.L. (Ed.), pp. 3-19, Springer-Verlag, Tokyo, 1986.

8. Peters, J., "Local smooth surface interpolation: a classification," Comput. Aided Geom. Des., Vol. 7, pp. 191-195, 1990.

9. Walton, D.J. and Meek, D.S., A triangular G^1 patch from boundary curves, Computer-Aided Design, Vol. 28, pp. 113-123, 1996.

10. Barsky, B.A., Computer Graphics and Geometric Modeling Using Beta Splines, Springer-Verlag, New York, 1988.

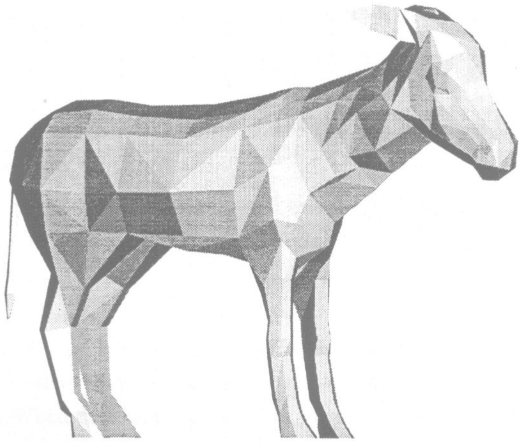


Figure 4. Crude triangulated model of donkey.

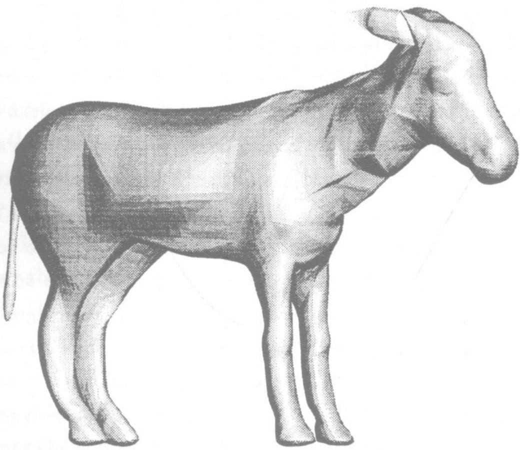


Figure 5. Fine triangulated model of donkey.

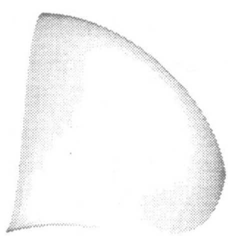


Figure 6. Special case of Example 2.

THE BOOLEAN OPERATION OF POLYGON WITH HOLES BASED ON EDGE RECOGNITION

Wu Yunxing Du Mike and Liu Shiping
Machinery Department, North China Institute of Water Conservancy & Hydroelectric Power
Zhengzhou

ABSTRACT

So far, the Boolean operation of polygon with no holes has been done successfully by using edge recognition algorithm. Based on this, this paper has applied it to the Boolean operation of polygon with holes.

KEYWORDS

Hole, Boolean Operation, Edge Recognition, Polygon

1. INTRODUCTION

The Boolean operation of polygon is a complicated problem in graphics. At present, we mainly talk about the Boolean operation of polygon with no holes, however, the Boolean operation of polygon with holes is rarely talked about. At the same time, even if the Boolean operation of polygon with holes is talked about, their algorithms are not so effective, because these algorithms are different from one another, and they have prescribed the line segments and their intersection points complicatedly. Why? This is because the result of Boolean operation of polygon is a collective of certain polygons, but the line segments of the polygons are not isolated, they are connected by the apex points naturally. At present the algorithms have not made full use of this so the line segments and intersection points have to be prescribed complicatedly. In this paper the algorithm of Boolean operation is based on the edge recognition algorithm. If we recognize along a certain line segment by using edge recognition algorithm, we are sure to recognize a single polygon . According to the above mentioned, the result of the Boolean operation is a collective of certain polygons. Therefore, if the primary parameter is chosen properly, it is likely to use the inside edge and outside edge recognition algorithm provided in paper 2 and paper 3 to do the Boolean operation. In fact, paper 1 has applied the edge recognition algorithm to the Boolean operation of polygons without holes successfully. This paper tries to apply it to the Boolean operation of polygon with holes.

2. BRIEF INTRODUCTION TO EDGE RECOGNITION

First, the application of the Boolean operation to edge recognition algorithm is introduced as follows:

The edge recognition algorithm includes inside edge recognition algorithm and outside edge recognition algorithm. Before solution, the line segments should be segregated to make all segments cross the other segments only at two apex points effectively. The purpose of edge recognition algorithm is to recognize certain closed polygons in any group of non—sequential segments (that is, the segments from DXF of Auto CAD). The formation segments of inside edge and outside edge will be arranged anti—clockwise. The storage structure of segments is as follows:

No	x1	y1	x2	y2	Fs	Foi	Fab	Fd

In which, No is the sequence mumber of segments; (x1, y1) and (x2, y2) are two ends of the segment; Fs=0 or 1 indicates that of the two ends of this oriented segment, which one is the starting end ? If Fs=0, then (x1, y1) is the starting end; if Fs=1, then (x2, y2) is ; F_{oi}=0 or 1 indicates that the segment belongs to inside edge and outside edge respectively; F_{ab}=0 or 1 indicates that this segment belongs to polygon A and polygon B respectively; F_d=0 or 1 indicates whether this segment has been deleted or not. Every flag bit can be given value when necessary.

Definition 1: if a segment crosses effectively with the directed line segment L at the end of this directed segment, the segment is the relation segment of the directed segment L.

The kernel of the edge recognition algorithm is to look for the segments on the most left or right of the directed segmentl.

Definition 2: When there are n relation segments on the left (right) side of the directed segment L (assuming its starting point is Ps, and the followings are the same), draw out line L′ on the left (or right) of segment L. Let L′//L and the distance is σ (σ is a small enough positive number, the followings are the same). Assume the effective crosspoint. of L' and these n relation segments on the left (or right) side of segment L is P'_i (i=1—n), if one point Pm' satisfies the following equation:

$$\overline{PsPm'}=\min\{\overline{PsPi'}\} \quad (i=1\sim n) \qquad (1)$$

Then the segment which includes point Pm' is the most left (or right) segment of L . When there are no relation segments on the left (or right) side of segment L, if there are n relation segments on the right (left) of side, then draw out line L' on the right (left) L, let L'//L, and the distance is σ. Assume the effective crosspoint of L and those n relation segments on the right (left) side of L is P_i' (i=1—n), if the following equation is correct:

$$\overline{PsPm'}=\max\{\overline{PsPi'}\} \quad (i=1\sim n) \qquad (2)$$

then, we can say that the segment which includes the point Pm′ is on the most left (right) side of the line segment L.

If we want to determine the different beginning segments, we can recognize the inside or outside edge by recognizing the most left segments . The algorithm is as follows.

Algorithm 1

Step 1: Determine the beginning segments, and take the beginning segments as the current dege segments.

Step 2: Look for the most left (right) segment of the current dege segments. This segment will belong to edge segments.

Step 3: Take the segment found in step 2 as the current edge segment.

Step 4: Compare the current edge segment with the beginning segments. If they are identical, go on to step 5; otherwise, go back to step 2.

Step 5: Finally the enclosed polygon composed of the segments found as above belongs to the inside (outside) edge.

3. THE BOOLEAN OPERATION OF POLYGON WITH HOLES

The formation segments of the inside and outside edge of the initial data of the Boolean operation of polygon are arranged anti-clockwise. Suppose there are two polygons, one is A and the other is B . The marks "+"、"*" and "−" are used to indicate the combination, cross, and difference of the polygon respectively. Before solution, it is required to mix A and B into a new data base and to be handled section by section (the handling of section by section is only to modify the coordinate of the end of the segments, not to change the value of the original marks) . According to the Boolean algebra, the cross and combination of a polygon can be changed as follows:

$$A+B=(A-B)+B \quad (3)$$

$$A*B=A-(A-B) \quad (4)$$

Therefore, provided A−B is found, then A * B and A+B can be done on the basis of A−B conveniently.

3.1 Solution of the Difference of Polygon

To look for the difference A−B of polygon is to look for the part of A which is located outside the outside edge of B and which is located inside the inside edge of B . If we want to recognize it by using edge recognition algorithm, the most important is to determine the initial segments. Although the segments in A or B are all likely to be taken as initial segments, using the segment in A as initial segment is feasible, because A−B is located outside of B, the selected segment in A should also be located outside of B, that is to say it should be located outside the outside edge of B and inside the inside edge of B. Thus we can use the most right side principle and most left side principle to look for the initial segments respectively. We can start from any segment of each polygon which is located outside edge and inside edge of B. But the found segment of A which is located outside of B should be deleted in time, shown in Fig 1.

Wherever it starts it is certain to search for a group of segments, but only one group at one time, it is impossible to search for another group. If a group of segments has been searched and deleted in time, we can go back to B for next search, thus, we can find the segments we want. These segments can be taken as the initial segments of recognizing A −B. The solution algorithm is as follows:

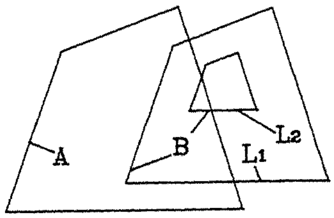


Fig.1 the initial segment

Algorithm 2

Step 1: Take any segment in the polygon of the outside edge of B as the initial segment and current segment.

Step 2: Let the current segment be the return segment, if the current segment belongs to the inside (outside) edge, then look for the most left (right) segment and let it be the current segment, see whether this segment and the initial segment are same, if yes, go to step 7:

Step 3: If the current segment belongs to A, go to step 4, if not, go to step 2.

Step 4: Delete the current segments, and keep the current segment additionally.

Step 5: If the current segment belongs to the inside (outside) edge, then look for the most left (right) segment and take it as the current segment.

Step 6: See whether the current segment belongs to B, if yes, take the return segment as the current segment, go to step 2. If not, delete the current segment and go to step 5.

Step 7: Is there any more polygon at the inside edge of B? If yes, take any segment from the next polygon at the inside edge of B, and let it be the initial segment and current segment, go to step 2.

Step 8: Keep all the segments, these segments are those we want to look for.

All the segments of A which are located outside of B have been found in algorithm 2 and kept specially. These specially kept segments can be taken as the initial segments to recognize the formation polygons of A—B. When solution, if the outside segments belong to the outside edge and inside edge of A, then we can use the most left (most right) principle respectively to recognize and also delete the used or recognized segments in time from the special file, otherwise, the recognized polygons will be recognized repeatedly. As in Fig 1, no matter where it starts L1 or L2, the recognized polygons are same. Thus the solution of difference of the polygons is as follows:

Algorithm 3.

Step 1: According to algorithm 2, look for all the segments in A which are located outside of B, and keep them in file 1.

Step 2: Open file 1, take the first segment and let it be the initial segment and current segment.

Step 3: If the initial segment belongs to the outside (inside) edge of A, then we apply the most left (right) side principle to recognize a polygon. This polygon belongs to A—B and we keep it in file 2.

Step 4: See whether in file 1 there are segments kept in file 2, if yes, delete them (modify Fd).

Step 5: See if file 1 comes to an end, if yes, go to step 7, if not, take the next segment in file 1.

Step 6: See if the taken segments have been deleted (judging from Fd), if yes, go to step 5, if not, take it as the initial segment and current segment, and go to step 3.

Step 7: File 2 is A—B.

3.2 Solution of Combination of the Polygon

According to equation 3, A+B is the combination of A—B and B. Because the difference A—B of polygon does not include B, A—B and B can not cross, therefore, A—B and B can be superimposed. However, between the polygons composed of A—B and B there might be coincide segments, one side of these segments belongs to A—B, the other side belongs to B, two sides all belong to A+B. However, as to the formation segment of graph edge, only one of its two sides belongs to area enclosed by the edges. Therefore, these coincide segments have lost the conditions of the formation segments of graph edges, they should be deleted. These segments belong to B, so $F_{ab}=1$. Because the segment of $F_{ab}=1$ in A—B must belong to B at the same time, but the segment of $F_{ab}=1$ in B does not necessarily belong to A—B, therefore, what should be deleted is the segment of $F_{ab}=1$ in A—B. Thus the solution of combination of the polygon is as follows:

Algorithm 3

Step 1: Look for A—B according to algorithm 2.

Step 2: Look for the segment which belongs to B ($F_{ab}=1$) in A—B.

Step 3: Delete from A—B and B the segments found in step 2.

Step 4: The polygon composed of the segments in A—B and B is A+B.

3.3 Solution of Cross of Polygons

According to algorithm (4), A * B is the difference of A and A—B.

Because the difference A—B of the polygon belongs completely to A, A—B can be deleted from A. As far as polygon is concerned, zoning delete is in fact the segment delete, that is to delete the shared segment of A—B and A. Because these segments belong to A, too, $F_{ab}=0$. The segment of $F_{ab}=0$ in A—B must belong to A at the same time, but the segment of $F_{ab}=0$ in A does not necessarily belong to A—B, what should be deleted is the segment of $F_{ab}=0$ in A—B. Thus the algorithm is as follows:

Algorithm 4

Step 1: Look for A—B according to algorithm 2.

Step 2: Look for the segment which belongs to A ($F_{ab}=0$) in A—B.

Step 3: Delete from A—B and B the segments found in step 2.

Step 4: The polygon composed of the segments in A—B and A is A * B.

4. CONCLUSIONS AND EXAMPLE

This algorithm is realized on computer by using C language. It has been tested by a great many examples. The result proves that, this algorithm is effective to any complicated graphes and the speed is fast. The ordinary graphes can be finished within several seconds.

5. REFERENCES

1. Wu Yunxing, The Boolean Operation of Polygon Based on Edge Recognition, Journal of Computer—Aided Design and Computer Graphics, Vol. 6, CN ISSN 1003—9767, pp. 260—265, 1994.
2. Wu Yunxing, Study on Automatic Recognition of Outside Edge of Two—dimension Graph, the Sixth Session of CAD & CG, A Collection of Acdemic Thesis. pp. 688—689, WuXi, 1990.
3. Wu Yunxing, Study on Automatic Recognition of Inside Edge of Two—dimension Graph, the 7th Session of CAD & CG, A Collection of Acdemic Thesis. pp. 467—472, ShangHai, 1992.

IMPLICATION OF DESIGN FOR INSPECTION

L. Wang and G. C. I. Lin
Centre for Advanced Manufacturing Research
School of Engineering
University of South Australia
The Levels, SA
AUSTRALIA, 5095

ABSTRACT

The determination of inspection datums is an important stage in the measurement process when using coordinate measuring machines(CMMs). The current inspection systems integrating design with CMMs require a high level of human intervention in measuring coordinate system alignment procedure. Even an experienced CMM operator may very often set up incorrect measuring datums. The paper describes the use of design datum information from a computer aided design(CAD) system to determine a correct datum for inspection. An AutoCAD development system (ADS) has been used to develop the design datum data-base. An automatic measuring coordinate system alignment (MCSA) method is defined. The initial testing indicated that the MCSA method which integrates a CAD system with a CMM inspection system can be used to achieve the coincidence of inspection datums with design datums.

KEYWORDS: CAD, Inspection, coordinate system alignment.

1. INTRODUCTION

The flow of information from design, manufacturing to inspection is made difficult due to a lack of communication. The problem in industry has been illustrated by a scene where designers and manufacturing engineers stand on either side of a tall brick wall [1]. A similar type of information gap exists between designer and inspector. To overcome the problem, the implementation of the coordinate measuring technique by using a CMM integrated with a CAD system plays a very important role in manufacturing.

Some developments on CMM integration with CAD systems with different foci are reported in the literature. One of them is the automated CAD-based inspection system called IPPEX developed by Brown[2], in which the feasibility of an automated inspection planning system was verified, the standard inspection techniques were established, and the product modeling system was integrated. Another development in this area is a personal computer-based system integrating a manual CMM with a CAD system to automate the inspection process developed by Gupta [3]. The greatest advantage of this system is that the method and programs were developed in the public domain, thus the manual CMMs can be enhanced without the addition of dedicated hardware and software. A software package that pre-processed CAD system output information into a neutral data-file format in support of off-line programming of CMMs was given by Pham [4]. Some other research on inspection planning based on the CAD models have been carried out [5][6][7].

However, in most of these developments the manual or self-teach method to do the measuring coordinate system alignment and to guide the CMM to find the real position and orientation of a workpiece on the CMM's worktable are still required. This alignment method has serious disadvantages for some precise components [8][9]. The purpose of this paper is to present how to use datum information from a CAD system to carry out an automatic measuring coordinate system alignment (MCSA).

2 DATUMS THEORY IN DESIGN AND INSPECTION

To achieve datum coincidence of design, manufacturing and inspection, it is necessary to investigate the datums in design.

2.1 Functional Theory of Datums in Design

Every feature of a part must be dimensionally related in three coordinate directions from an known point of origin. The position of a specific point in space can be properly dimensioned only from another point whose location is known. The known point is generally called the "datum point", and any point can be located from it in a Cartesian system. Datum features of a part are assumed to be exact for purposes of computation. The datum dimension is taken as a reference, and variations are specified in terms of the relation desired. A feature dimension is used to define the size of a feature. Position dimensions locate the various features of a component using feature surfaces or feature centerlines as datums. The most familiar use of positions is to locate points with coordinates[10]. A displacement screw method to find out the minimum geometrical deviation has been developed by Tanaka, et al. [11].

In the design procedure, datum features can be identified by the use of datum identifying symbols. Datum planes exist only with respect to datum features. A datum feature may be a flat surface on a part, or it may by a slot, a hole, or some other symmetrical feature. For example, when a datum feature is a flat surface, the corresponding datum plane is a flat plane that contacts the high points of the datum surface.

2.2. The Measuring Coordinate System Alignment Procedures

Using a coordinate measuring machine, the existing measuring procedures need to ensure that certain coordinates are met regarding the part. The coordinate measuring system alignment is one of the important procedures. The conventional alignment method performs alignment from measured features.

The measuring coordinate system is based on datum planes from which measurements are taken. Three mutually perpendicular datum planes, as illustrated in figure 1, make up a datum frame.

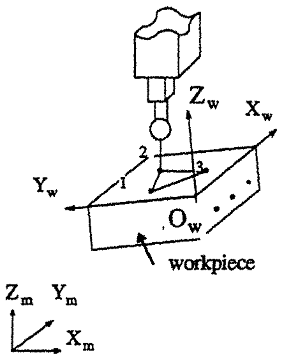


Figure 1. Coordinate systems in a CMM
(The X_wY_w plane as a work level)

where X_m, Y_m and Z_m are the axes of the datums planes. 0 is the intersection of the three planes.

The CMM's (machine's) coordinate system is the set of X_m, Y_m and Z_m axes parallel to its guideways. The origin of the CMM is at its probe holder, the end of the Z-rail, when the machine is homed. The machine knows its own $X_mY_mZ_m$ coordinate at any point in its envelope. Every workpiece also has its own coordinate system which inspectors use to measure features in accordance with the drawing. Thus, the workpiece's coordinate system must be related to the CMM's coordinate system before measuring the workpiece. This process defines the coordinate system transformation between the CMM (reference system) and the workpiece.

The procedure in setting the measuring coordinate system is to establish the origin. The origin which becomes a reference point for all the later measurements is the intersection of the set of workpiece axes X_w, Y_w and Z_w. The first step is to select a feature to level the part for definition of the working plane. A few points (at least three points) are usually determined to generate a datum plane. When acquired by measuring a surface on a part, the surface measured is known as the work plane.

In the $X_wY_wZ_w$ system, the measuring is performed straight down on the top of a cube. The top plane would be the workplane. The Z coordinate and its direction (opposite to the direction of probing) would be known. The 2D vector defines the major axis. The major axis direction comes from the order in which major axis points are measured. Another 2D vector defines the minor axis. It is defined as being perpendicular to the workplane and the major axis. The origin point is at the intersection of the workplane, major axis and minor axis.

The datum selection and identification requires a uniform basis for part location during manufacture and for measurement during subsequent inspections. This is especially necessary with parts with precise dimensions, that is, when dimensions have tolerances in the range of IT5 or finer[12].

Since datum features may not have perfect geometry, there can be some ambiguity in defining datum planes from these features by a touch trigger probe when using a CMM. The imperfect geometry of datum features is generally significant for precise work. When the imperfections of datums are unacceptable, it is necessary to define datums in a way that is independent of surface errors[13][14]. The information of the design datum can be used for this purpose.

3. THE ADS DEVELOPMENT

The fundamental requirements for using of the design datum data-base for inspection, an interface is needed. The AutoCAD Development System(ADS) is a programming environment in which all Autolisp functions can be used. Furthermore, the AutoCAD Development System (ADS) provides the C callable libraries which allow users to develop interface communication for external applications. Therefore, it supplies the link ability between the CAD system to other requirements. The ADS work flow chart for datum information transformation which is called measuring coordinate system alignment (MCSA) method is shown in Figure 2.

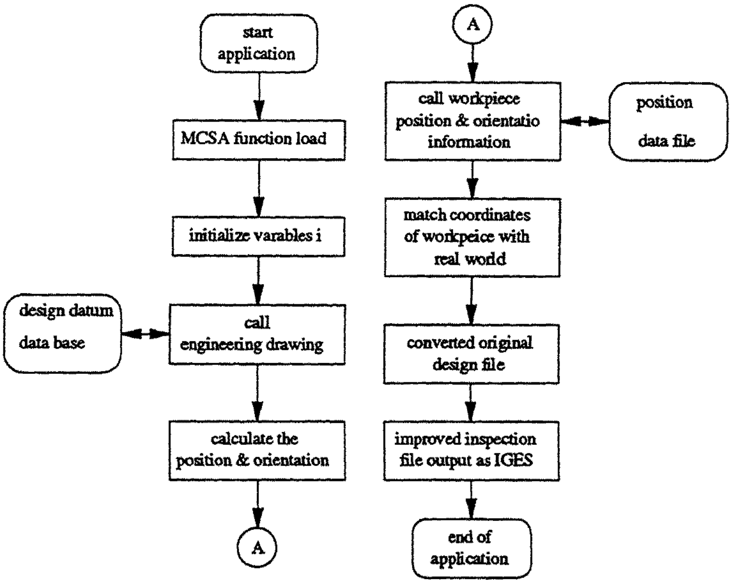


Figure 2. The flow diagram of the MCSA method

Where datum requirements are expressed as symbols arranged in particular ways on CAD-database.

The main function of the MCSA method is to match the coordinate system in the original design model with the position and orientation of the workpiece on the CMM's worktable. The ADS application using c program calling Autolisp file, such as datum symbols. The MCSA method executes in the AutoCAD environment. The variable *i* represents a current measured part ID which can be selected by the operator or be determined from the higher level database. It searches the engineering drawing in the drawing database according to the ID number. The method calculates the current engineering drawing datum position and orientation on the AutoCAD World Coordinate System. This execution will ensure the position and orientation of engineering drawing matching with those of the inspection system. The measuring coordinate system was set up when the probe was qualified. After the engineering drawing coordinates are aligned, the workpiece position and orientation information on the CMM's worktable can be received from the position data file. The information can be acquired by the machine vision system before the MCSA starts. Depending on the information of the workpiece in the real world the original design model is converted by matching the coordinates in it with the coordinates of the workpiece on the worktable.

4. THE USE OF DESIGN DATUM IN INSPECTION

Based on the fundamental requirements discussed above, the design and inspection link still needs the information of position and orientation of the workpiece. Therefore a machine vision system is involved.

A software package CADBRIDGES[15] which was developed by Automation Software was chosen for inspection execution It is a CAD-based geometric measurement programming package. It translates the high level commands required to measure parts into the detailed steps necessary to drive a CMM. It also provides several methods for transmitting both part programs and measurement data between various CAD systems and its measuring environment.

In the system as shown in Figure 2, the first step is to extract design information called the original design model from the a CAD environment. The coordinate system in the engineering drawing must be the World Coordinate System (WCS) that is used by CADBRIDGES. A machine vision system is used to determine the workpiece position and orientation on the worktable and a data file is created which stores the position coordinate information. Then the MCSA method will modify the original design model depending on information from the data file to create a converted inspection model. In which the position coordinates of the workpiece have been matched with the real coordinates of the workpiece on the worktable. The origin in the inspection model coincides with the origin of the measuring coordinate system. The coordinate system in this inspection model can be used without additional operation when CADBRIDGES executes inspection with the reference coordinate system. In the CADBRIDGES environment inspection will be executed depending on the converted inspection model and the information of the measured features. The complete processes of inspection can be implemented without any manual or self-teach touching. The flow diagram of the inspection system work environment and system set-up are shown in Figure 3 and 4.

5. CONCLUSIONS

The use of design datum information for inspection by using a CMM is described. The MCSA method has been developed for this purpose. The main task of the MCSA method is to integrate the AutoCAD system with the datum data-based in a inspection environment. The main advantages of this development are:

(1) The method supplies a uniform datum selection and identification during design and during subsequent inspections. This is especially necessary for measuring parts with precise dimensions.

(2) Using the information of design datum for inspection, the influence of the errors of datum feature's geometry on the measurement results can be eliminated since the alignment method defines datums in a way that is independent of the features. It is generally significant for precise work when the imperfections of datums are unacceptable.

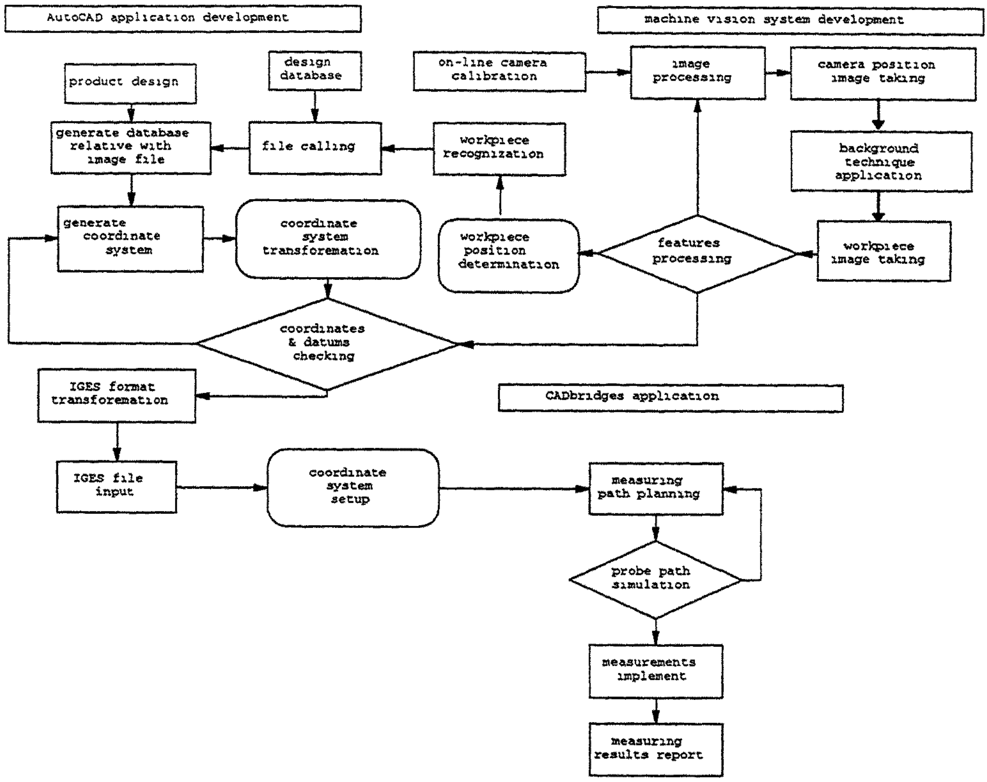


Figure 3 System flow diagram

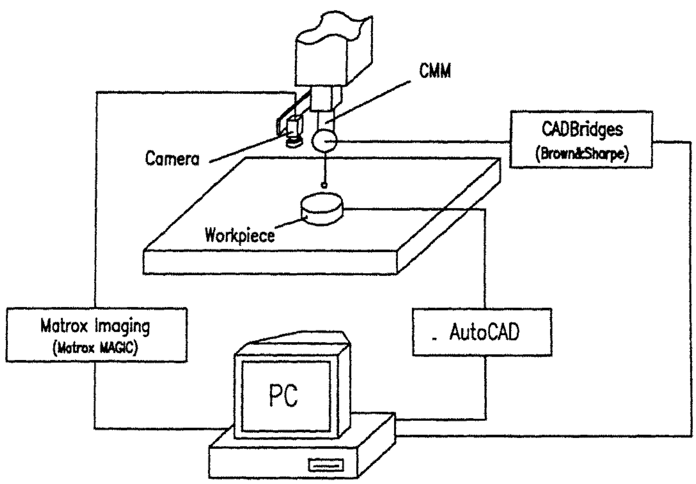


Figure 4 The system set-up

REFERENCES

1. Mark Richard Henderson "Extraction of Feature Information from Three Dimensional CAD Data" Thesis of Doctor of Philosophy. May 1984.

2. Brown, Curtis W. "IPPEX: An automated Planning System For Dimensional Inspection" CIRP, The International Institution for Production Engineering Research, 22nd CIRP International Seminar on Manufacturing Systems.

3. Gupta, V. K. and Sagar R., "A PC-Based System Integrating CMM and CAD for Automated Inspection and reverse Engineering", International Journal of Advanced manufacturing Technology 8: pp 305-310, 1993.

4. Pham, D. T., Martin K. F. and Khoo, L. P. "A Knowledge-Based Preprocessor generator for Coordinate-Measuring Machines", International Journal of Production Research Vol. 29, No. 4, pp 677-694, 1991.

5. Park H. D. and Mitchell, O. R. "CAD Based Planning and Execution of Inspection", IEEE CD-ROM, pp 858-863, 1988.

6. Marefat M. and Kashyap, R. L. "Planning for Inspection Based on CAD Models", IEEE CD-ROM, pp 608-611, 1992.

7. Merat F. L. and Radack, G. M., "Automatic Inspection Planning within a Feature-Based CAD System", Robotics & Computer-Integrated Manufacturing, Vol. 9, No. 1 pp 61-69, 1992.

8. Lin G and Wang L "Error Analysis of Effects of Workpiece Coordinate System Misalignment When Measuring on a Coordinate Measuring Machine" Proceedings of 1995 Annual Conference, Metrology Society of Australia, Sydney, Dec. 1995. pp19-22.

9. Hyman H. Katz "Handbook of Layout and Dimensioning for Production" New York The Macmillan Company.

10. Earlwood T. Fornrtini "Dimensioning for Interchangeable Manufacture" Industrial Press INC., 200 Madison Avenue, New York 10016.

11. Tanaka, F. Ikonomov, P et al. " Inspection method for geometrical tolerances using coordinate measuring machine". Proceedings of the 4th CIRP seminar on computer aided tolerancing. pp 325-326. 1995.

12. "ANSI/ASME Y14.5, Dimennsioning and Tolerancing", ASME, New York, NY, 1982.

13. Bosch, John A. "Coordinate Measuring Machines and Systems", Marcel Dekker, Inc. 1995.

14. Huong, C. "Study of Automatic Generation for the Trajectories of General Coordination Measuring Machines", Masters Thesis, National Taiwan University, 1994.

15. "CADBRIDGES user manual", Document Revision Level 2.0, Automation software.

A Neural Network Approach to Parametric Design

Y.H. Chen

Department of Mechanical Engineering
The University of Hong Kong
Hong Kong

ABSTRACT

Parametric design is an effective and productive tool for the definition and modification of geometric models. This paper presents an intelligent method for the creation of a parametric model. In the proposed method, an adaptive neural network model is built to map the set of dimensional parameters to a set of coordinates. Defining a parametric model is equivalent to teaching the neural network. The user needs only specify a set of dimensional parameters that defines the parametric model and teach the neural network how to react to the changes of the dimensional parameters. Once the neural net work is taught, any dimensional changes will result in corresponding coordinate changes. This novel method eliminate the need of programming or graphic interaction that are normally required by contemporary parametric design systems.

Key Words: CAD, Parametric Design, Variational Design, Artificial Neural Network

1. INTRODUCTION

Computer Aided Design systems define a geometric model in terms of coordinates and topologic data. These data barely has any engineering meaning. Higher level geometric model based on dimensions is therefore of key interest for CAD systems. Such systems provide a significant increase of efficiency in the design of generic part families where members of a family differ only in dimensions of the geometry. Moreover, in the early stages of design, dimensions are not always known initially and must subject to frequent changes. CAD systems with the capability to adjust the geometric model according to changes of one or more dimensions are often referred to as dimension-driven CAD systems[1,2]. Currently, parametric design[2,3] and variational geometry[6,7,8] are the two main techniques developed for dimension-driven CAD systems.

Parametric design is not a new technique. Most existing commercial CAD systems provide this capability. The underlying techniques of parametric design can be classified into two categories[1,2]:

* Define a geometric model through programming with dimensions as parameters in a high-level programming language[2,3]
* Interactive design of a master geometric model under a graphical environment. Generic variants of the master model can subsequently be generated[6,7,8].

The first approach requires the specific programming knowledge of a particular CAD system. It is therefore only suitable in situations where the extra training expense and time can be justified.

The second approach to parametric design has drawbacks in other areas: the handling of implicit constraints, such as tangential conditions, right angles, parallel lines etc., has not been satisfactorily solved. Manual assignment of a correct set of implicit constraints in a post processor is tedious and error-prune. It is also difficult to check weather a geometric model is over-determined or under-determined.

Even though variational geometry has been applied to parametric design, the two techniques are different. In variational geometry, the designer makes no assumptions about the way in which dimensions are combined. A parametric design system requires either that a predefined set of geometric constraints is used to represent a design solution or that the relation be connected in a causal ordering suitable for solution[11].

The approach to parametric design presented in this paper is totally different from the previously mentioned methods. In the definition of a parametric model, neither programming nor graphical interaction is required. Instead, a neural network model is applied to map a set of dimensions to its resulting point coordinates. When defining a parametric model, the designer need only to select a group of dimensions and teach the neural network how to react to the changes of dimensions.

2. NEURAL NETWORKS

Neural networks have been successfully used in pattern recognition, process control, and signal processing etc.[13,14]. The structure of the neural network implemented in this paper is shown in Figure 1. On the left is the input vector p connected to the neurons through a weight matrix W(S,R). All components of the bias vector b have initial value equal to 1. On the right is the output vector a. The Widrow-Hoff Learning rule is used to calculate changes to weights and bias which decrease the sum of squared errors of the neuron layer[12]. The network's error **e** is defined as the difference between the network's output **a** and target **t**.

$$\mathbf{e} = \mathbf{t} - \mathbf{a}$$

The Widrow-Hoff learning rule calculates small changes for a neuron's weights and biases in the direction that decreases the neuron's error. This direction is found by taking the derivative of sum-squared error with respect to these parameters. The derivative of the sum-squared error *sse* with respect to a weight **W(i,j)** for a single input vector **p** and target vector **t** is:

$$\frac{\partial sse}{\partial \mathrm{W}(i,j)} = \frac{\partial}{\partial \mathrm{W}(i,j)}\left(\mathrm{t}(i) - \sum_{j=1}^{R}[\mathrm{W}(i,j)\mathrm{p}(j) - \mathrm{b}(i)]\right)^2 = -2\mathrm{e}(i)\mathrm{p}(j)$$

The Widrow-Hoff rule is implemented by making changes to the weight in the opposite direction from the direction that error is increasing and absorbing the constant 2 into a learning rate *lr*.

$$\Delta \mathrm{W}(i,j) = lr.\mathrm{e}(j)\mathrm{p}(j)$$

This weight update expression can be restated in matrix form:

$$\Delta \mathrm{W} = lr.\mathrm{E}\,\mathrm{p}^{T}$$

This expression can be transformed into the bias update expression, by noting that biases are simply weights with constant 1 inputs.

$$\Delta \mathrm{b} = lr.\mathrm{E}$$

The above discussed neural network architecture and the learning rule is implemented for the parametric design described in the next section.

3. NEURAL NETWORK FOR PARAMETRIC DESIGN

Instead of writing a script for the definition of a parametric model, this section presents a novel method for parametric design. Figure 2 shows an parametric model defined by the dimension vector [L1 L2 L3 L4 L5]. This vector defines a family of generic parts. When any a dimension or a dimension combination is changed, the geometry must change correspondly. With the neural network method, the dimenion vector [L1 L2 L3 L4 L5] is chosen as the input pattern. In order to make the network respond to any dimensional change, the network is taught six times, each time with an altered dimension component. The six teaching patterns are shown in Figure 3. Corresponding to each of the input pattern, a target output pattern of the network must be defined. In the implementation, the network output is a layer with sixteen neurons representing the coordinates [X1 Y1 X2 Y2 X3 Y3 X4 Y4 X5 Y5 X6 Y6 X7 Y7 X8 Y8] of the 8 points p1 to p8 of the parametric model. Once the input and output vectors are defined, the structure of the network can be automatically constructed. Figure 4 is the network structure for the parametric model shown in Figure 2.

When teaching the network, the input and output patterns are presented as matrix P and T respectively. Point p1 is used as a reference point with X1=0 and Y1=0. After the teaching process is done, a weight matrix W and a bias vector B are generated. The network architecture combined with its trained weight matrix W and bias vector B defines a parametric model.

When the teaching process is over, the parametric model has been defined. Any dimension vector representing a generic part of the parametric model can be presented to the network. To illustrate this, a random dimension vector [120.9,80.5,60.3,20.1,80.0] has been presented to the network.The corresponding output is a vector [0, 0, 0, 80.5, 20.1, 80.5, 20.1, 20.2, 100.1, 20.2, 100.1, 80.5, 120.9, 80.5, 120.9, 0] indicating the point coordinates:

X1=0, Y1=0; X2=0, Y2=80.5;
X3=20.1, Y3=80.5; X4=20.1, Y4=20.2;
X5=100.1, Y5=20.2; X6=100.1, Y6=80.5;
X7=120.9, Y7=80.5; X8=120.9, Y8=0;

A scaled drawing of the part defined by the above dimension vector is shown in Figure 5. No error has been encountered during the test.

4. FUTURE RESEARCH

This paper has presented a successful implementation of neural network for parametric design. However, a number of issues have not been addressed in this paper which are still under research in the Department of Mechanical Engineering, The University of Hong kong. One of the important issue is how to cope with situations where dimension changes cause a change of topology of the parametric model. This problem may be solved by develop more teaching patterns which reflect the acceptable and unacceptable topology changes. The second issue is that the proposed neural network model is only applicable to parametric models where a linear dependence relation between the dimension vector and the coordinates can be formulated. This problem is now being solved by adopting a back propagation neural network model which can approximate non-linear relationships between a dimension vector and coordinate vector.

5. CONCLUSION

This paper has presented a novel method to parametric design through neural network. In the proposed method, parametric design is made easy. The designer needs not to do any programming or graphical interaction in defining a parametric model. Instead, the designer only need to teach the neural network with dimension vectors and coordinate vectors. The internal form of a defined parametric model is composed of a weight matrix and a bias vector. This method has the potential for the development of more user friendly parametric design systems.

REFERENCES

1. Roller, D ' An Approach to Computer-Aided Parametric Design' Computer Aided Design, Vol 23, No 5, 1991, pp 385-391

2. Roller, D, Schonek, F and Verrout, A 'Dimension-Driven Geometry in CAD: A Survey' in Strasser, W and Seidel, H-P (Eds) Theory and Practice of Geometric Modeling, Springer Verlag, 1989, pp 509-523
3. Cugini, U, Devoti, C and Galli, P ' System for Parametric Definition of Engineering Drawings' Proc. MICAD'85, 1985
4. Light, R and Gossard, D ' Modification of Geometric Models through Variational Geometry' Computer Aided Design, Vol 14, No 4, 1982, pp 209-214
5. Hillyard, R and Braid, I 'Analysis of Dimensions and Tolerances in Computer-Aided Mechanical Design', Computer Aided Design, Vol 10, No 3, 1978, pp 161-166
6. Gossard, D, Zuffante, R and Sakurai, H 'Representing Dimensions, Tolerances and Features in MCAE Systems' IEEE Computer Graphics & Applications, March 1988, pp51-59
7. Aldefeld, B 'Variation of Geometries Based on a Geometric-Reasoning Method' Computer Aided Design, Vol 20, No 3, 1988, pp 117-126
8. Chen, Y.H 'A Rule-Based 2D Geometry System for Engineering Design', Ph.D Thesis, The university of Liverpool, UK, 1990
9. Hinduja, S and Chen, S J' Geometry of 2D Components', Computer Aided Design, Vol 19, No 6, 1987, pp 323-328
10. Shah, J J'Assessment of Feature Technology', Computer Aided Design, Vol 23, No 5, pp 331-343
11. Fu, Z and Pennington, A D' Constaint-Based Design Using an Operational Approach', Research in Engineering Design (1993) 5, pp 202-217
12. Widrow, B and Sterns, S D 'Adaptive Signal Processing', New York, Prentice-Hall, 1985
13. Huang, S H and Zhang, H C 'Artificial Neural Networks in Manufacturing: Concepts, Applications, and Perspectives', IEEE Transaction on Components, Packaging, and Manufacturing Technology-Part A, Vol 17, No 2, June 1994, pp 212-228
14. Dayhoff, R E and Dayhoff, J E 'Neural Networks for Medical Image Processing', IEEE SCAMC Proceedings, pp 271-275

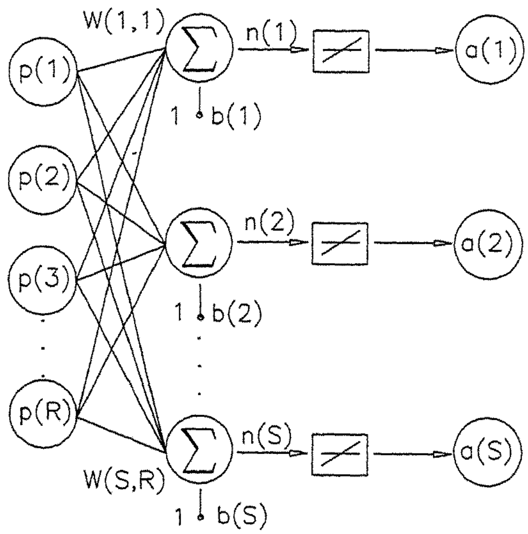


Figure 1 A typical neural network architecture

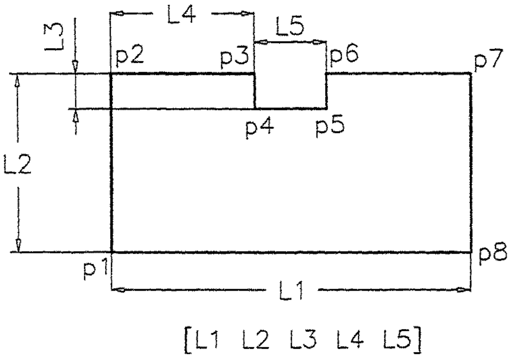


Figure 2 A parametric model defined by a dimension vector

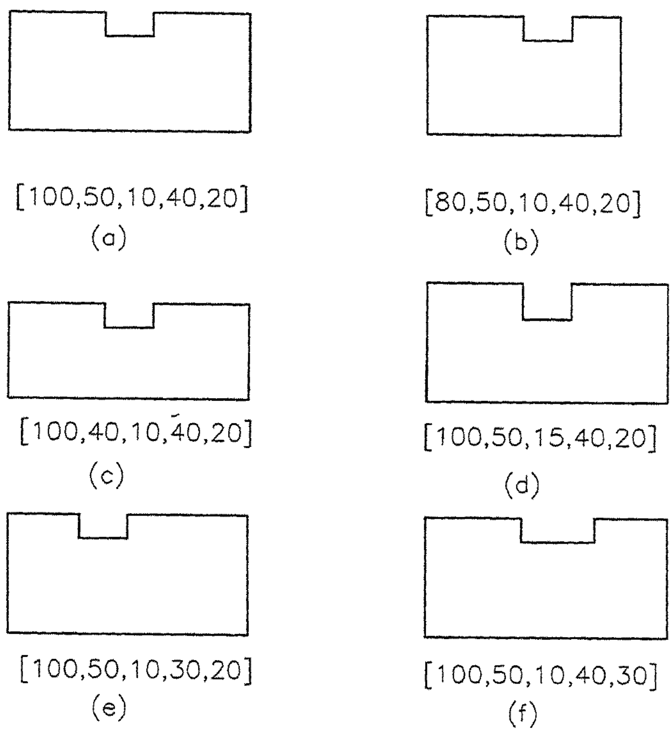


Figure 3 Six teaching patterns for a parametric model definition

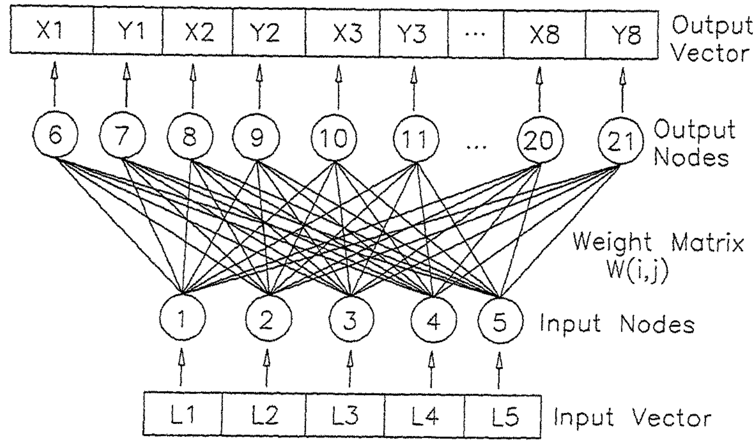


Figure 4 Neural network structure for a parametric model

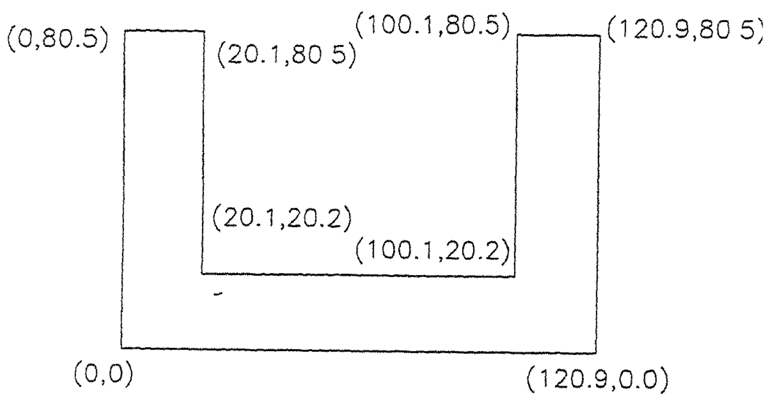


Figure 5 A generic part of a parametric model

ELEMENT METHOD FOR 2-D SHAPE TRANSFORMATION

K.C. Hui, Y.D. Li
Department of Mechanical & Automation Engineering
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong

ABSTRACT

This paper proposes an effective approach for interpolating different 2-D shapes in the industrial design process. This approach allows the blending result to retain the shape feature of the objects in the transformation. In this method, a 2-D shape is interpreted as a number of elements with different shape features or characteristics. Each of these elements consists of one or more natural segments of the object, such as line, arc or free-form curve. The correspondence between elements of the objects is established by comparing the similarity between their shape features. Several functions are provided for evaluating the similarity of the features. Experiments showed that this method works well for industrial design.

KEYWORDS

Element, Shape Feature, Transform, Correspondence

1. INTRODUCTION

Transforming one shape to another or blending between different shapes has been widely studied in the fields of computer animation, illustration and industrial design[1][2]. For 2-D shapes described by their boundary contours, the transformation is performed in two steps. Firstly, the corresponding relationship between vertices on the contours is decided. In the second step, the corresponding vertices are interpolated. The intermediate results can then be used for animation, or can be adopted as prototype in industrial design. The first step is more important in industrial design since different correspondence may give different output. Different methods have been proposed for manipulating the transformation process. This includes Least distance method[2], Physical based method[3], Start-skeleton method[4], Fuzzy approach[5] and the unions of circles approach[6]. All of them provide a smooth transformation process. Correspondence between different shapes are usually decided based on the vertex level without considering other geometry information so that shape feature may have been ignored. It may be difficult to generate intermediate results which retain the shape feature of the original objects.

In order to establish the correspondence between different 2-D shapes for generating suitable intermediate results to be used in industrial design, this paper presents an element method for establishing correspondence. Test results on an experimental system showed that this method is effective.

In the proposed element method, an arbitrary 2-D shape is composed of a number of elements. Each of these elements is a part of a natural segment, or one or more natural segments (e.g. line, arc and free-form curve) of the shape. Unlike other methods which usually establish the correspondence directly on the vertices level, the element method establishes the correspondence in two different levels. The first level correspondence is established by selecting the corresponding elements according to their similar shape features or characteristics. The second level correspondence is to find the corresponding points within corresponding element-pairs.

Figure 1 and Figure 2 give an example of a transformation. In Figure 1, the corresponding elements pairs of these two shapes includes *AG* and *ah*, *BC* and *bc*, *CD* and *cd*, *DE* and *de*, *EF* and ef, *FG* and *fgh*. Figure 2 is the interpolation result obtained by the element method.

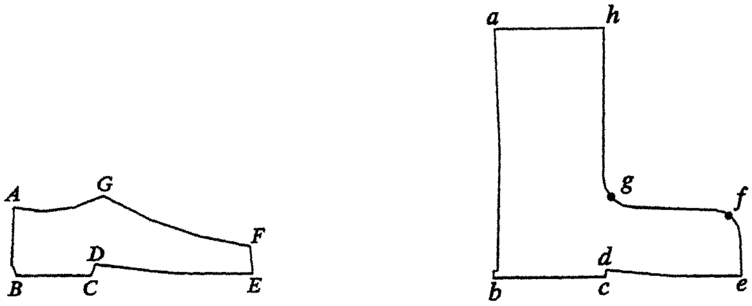


Figure 1   A shoe and a boot

Figure 2   Transforming a shoe to a boot

2. ELEMENT CORRESPONDENCE PROCESS

In the proposed method, elements are extracted from the objects, and are then selected to establish the element pairs as discussed in the following sections.

2.1 Locating the Elements

An element of a contour is part of a segment, or one or more continues natural segments. If the junction point between two adjacent elements is referred to as a node of the object, an element is then a part of a 2-D contour bounded by two nodes. Locating elements is thus the same as locating nodes of the objects. A node is defined as a location where there is a "sharp-turn" or an abrupt change in segment direction. A node may lie at the junction of two adjacent segments, or in the interior of a segment. To decide if a junction point between two segments is a node, the difference between the directions of the tangents at the junction is evaluated. If the absolute value of this angle is greater than a predefined value θ (normally, we set $\theta = \pi/4$), the junction point is considered as a node. A point lying in the interior of a segment is considered a node if the following criterion are satisfied.

1) the radius of curvature is smaller than a predefined value r_o

2) the change in tangent direction in the local area around the point is greater than the predefined value θ.

The point that satisfies criteria (1) is called a possible node, which becomes a real node if criteria (2) is satisfied as well. In Figure 3(a), the point *B* gradually becomes the sharp corner at point *A*, when the radius of curvature at *B* decreases to zero. Denoting the radius and arc length of the circle as *r* and *L* respectively, then $r/L = 1/2\pi$. If the corner at *B* is to be sharper than a circle, the ratio of the radius of curvature at B to the total arc length of a contour is less than $1/2\pi$. Hence, the radius of curvature ρ at a sharp corner must satisfy the condition $\rho < L/2\pi$. The radius of curvature r_o in criteria (1) is thus taken to be $L/2\pi$.

In Figure 3(b), the lines and arc are part of a closed 2-D contour. If the radius r of the arc satisfies $r < L/2\pi$, where L is the total arc length of the 2-D contour, there will be a possible node on this arc, which is defined to be the mid point of the arc.

Similarly, possible nodes on free-form curve segments can also be extracted. In this case, the points with the minimum radius of curvature ρ that also satisfy the criterion of $\rho < L/2\pi$ will be selected as the possible nodes (Figure 3 (c)).

A possible node will become a real node if the change in tangent direction in the vicinity of the point is greater than the predefined value θ. As shown in Figure 3(b), the change in direction is measured by the angle between the lines line-1 and line-2. For point C in Figure 3(c), the change in tangent direction is measured by the angle between the tangents at points B and D which are the locations of maximum radius of curvature closest to point C.

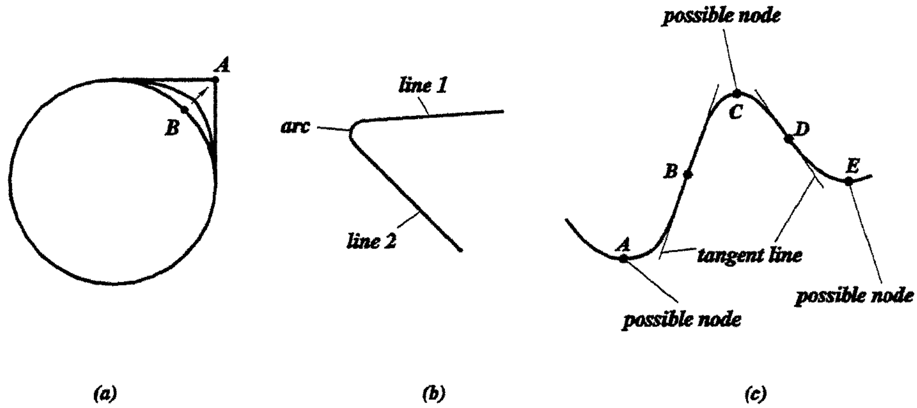


Figure3 (a) A circular arc gradually becomes a sharp turn
(b) Possible node on elements (c) Possible nodes on free-form curve

2.2 Selecting Corresponding Elements Pair

2.2.1 The process for establishing correspondence

The shape contours are first normalized to lie within a unit square. A pair of nodes on the two contours are then selected. This node pair is selected such that the degree of similarity between the nodes is the highest. Similarity of the nodes are measured in terms of the difference between their relative locations, tangent directions and the turn angles of the nodes. Corresponding elements are then selected based on five different functions that are constructed for measuring the difference between elements.

Consider an element *ele_1[i]* which is bounded by two nodes *node_1[i-1]* and *node_1[i]* respectively on an object. Assume that the corresponding element is to be selected out of q elements, namely, *ele_2[j]* (*j=1 to q*) of the other object with nodes *node_2[j]*.

Function 1 Comparing the change in tangent direction.

$$J_{1,j} = \frac{(t_{s,j} - t_{b,i}) - \mathrm{Min}\{t_{s,j} - t_{b,i} \mid j=1 \text{ to } q\}}{\mathrm{Max}\{t_{s,j} - t_{b,i} \mid j=1 \text{ to } q\} - \mathrm{Min}\{t_{s,j} - t_{b,i} \mid j=1 \text{ to } q\}} \qquad (1)$$

where $t_{b,i}$ denotes the change in tangent direction at the *node_1[i]*, and $t_{s,j}$ is the change in tangent direction at the *node_2[j]*.

Function 2 Comparing the location of nodes.

$$J_{2,j} = \frac{(l_{s,j} - l_{b,i}) - \text{Min}\{ l_{s,j} - l_{b,i} | j=1 \text{ to } q\}}{\text{Max}\{ l_{s,j} - l_{b,i} | j=1 \text{ to } q\} - \text{Min}\{ l_{s,j} - l_{b,i} | j=1 \text{ to } q \}} \quad (2)$$

where $l_{b,i}$ denotes the relative location of *node_1[i]*, and $l_{s,j}$ is the relative location of the *node_2[j]*.

Function 3 Comparing element direction.

$$J_{3,j} = \frac{(a_{s,j} - a_{b,i}) - \text{Min}\{ a_{s,j} - a_{b,i} | j=1 \text{ to } q\}}{\text{Max}\{ a_{s,j} - a_{b,i} | j=1 \text{ to } q\} - \text{Min}\{ a_{s,j} - a_{b,i} | j=1 \text{ to } q \}} \quad (3)$$

where $d_{b,i}$ denotes the direction of the element at *node_1[i]*, and $d_{s,j}$ denotes the direction of the other curve element at *node_2[j]*. Here, the direction of an element is the direction of the vector that starts from one of the element's bounding nodes to another(i.e. the vector connecting node[i-1] to node[i]).

Function 4 Comparing tangent direction of element at the first bounding node.

$$J_{4,j} = \frac{(st_{s,j} - st_{b,i}) - \text{Min}\{ st_{s,j} - st_{b,i} | j=1 \text{ to } q\}}{\text{Max}\{ st_{s,j} - st_{b,i} | j=1 \text{ to } q\} - \text{Min}\{ st_{s,j} - st_{b,i} | j=1 \text{ to } q \}} \quad (4)$$

where $st_{b,i}$ is the tangent direction at the element's first bounding node *node_1[i-1]*, and $st_{s,j}$ is the tangent direction at the first bounding node *node_2[j-1]* of the other curve element.

Function 5 Comparing tangent direction of element at the second bounding node.

$$J_{5,j} = \frac{(et_{s,j} - et_{b,i}) - \text{Min}\{ et_{s,j} - et_{b,i} | j=1 \text{ to } q\}}{\text{Max}\{ et_{s,j} - et_{b,i} | j=1 \text{ to } q\} - \text{Min}\{ et_{s,j} - et_{b,i} | j=1 \text{ to } q \}} \quad (5)$$

where $et_{b,i}$ is the tangent direction at the element's second bounding *node_1[i]*, and $et_{s,j}$ is the tangent direction at the second bounding node *node_2[j]* of the other curve element.

2.2.2 Initial matching

For each element and node of a curve, a set of possible corresponding nodes of the other object are selected such that they are less than a predefined distance from the current node. For each possible corresponding nodes, the function $f = f_{1,j} + f_{2,j} + f_{4,j} + f_{5,j}$ are evaluated. The elements with *f* values less than a threshold are then selected.

2.2.3 Exact matching

A list of unique one to one corresponding elements pair is then constructed from the result of the initial matching process. This includes eliminating cross-correspondence which will result in self-intersection in the interpolation process. Besides, one element of a shape may correspond to more than one elements of another shape. The "exact matching" process is thus to "purify" the corresponding elements list, namely to eliminate all those non-unique correspondence by selecting the best elements pair. This is based on the value obtained by the function $f = f_{1,j} + f_{2,j} + f_{3,j} + f_{4,j} + f_{5,j}$. The best elements pair is the elements pair with the minimum *f* values.

2.3 Generating Corresponding Points

The interpolation process is to be performed on corresponding points of each corresponding elements couple. The points are evenly distributed along the length of the elements. In this case, each of the two corresponding elements consists of an equal number of corresponding points, and hence the points generated for each of the two shapes are equal.

3. THE INTERPOLATION PROCESS

The matching process gives two sets of equal number of corresponding points from the two original shapes. Assume that these points can be expressed as $P_a=\{ p_1^1 , p_2^1 , \cdots p_m^1 \}$ for one of the shape, and $P_b=\{ p_1^2 , p_2^2 , \cdots p_n^2 \}$ for the other shape, the transformed shape is obtained by a linear interpolation between these two sets of corresponding points.

$$P = P_a \bullet (1 - t) + P_b \bullet t \qquad \text{where } 0 \leq t \leq 1 \qquad (6)$$

4. IMPLEMENTATION

An experimental system was implemented on a PC with a Pentium processor. A series of tests were performed. Figures 4 to 6 are example outputs of the system.

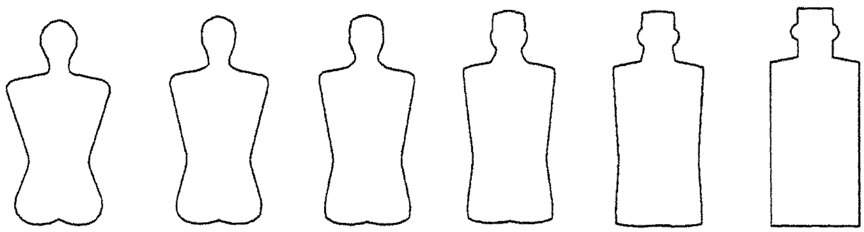


Figure 4 Transformation between bottles

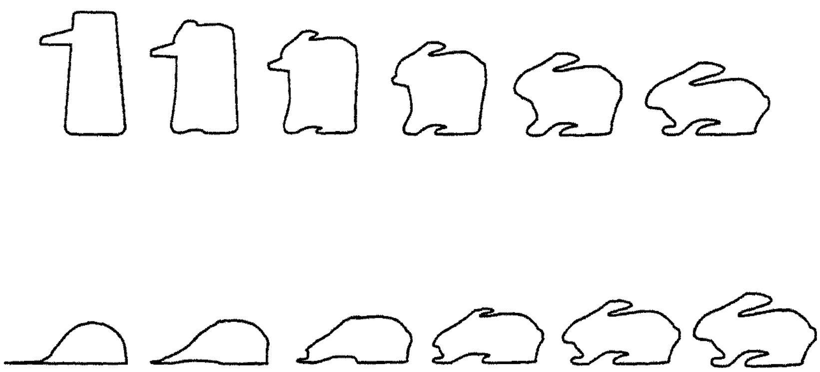


Figure 5 (upper) Transforming a kettle to a rabbit
(lower) From a cap to a rabbit

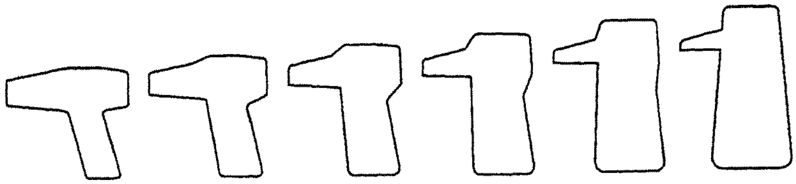


Figure 6 From a hair dryer to a kettle

5. CONCLUSION

To provide an effective approach for shape blending in industrial design, this paper proposed a feature based transformation method which allows the correspondence between two different shapes to be established automatically. The proposed approach establishes correspondence by selecting elements with the most similar shape feature. Five functions are constructed for evaluating feature similarity. The process runs automatically without user interaction. There is no restriction on the distribution of vertices on the shapes. The interpolation results embody or retain the shape characteristics of the original objects and thus can be adopted for industrial design.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to the University Grant Council for the financial support.

REFERENCES

[1] Hunghsiang Wang, "An Approach to Computer-aided styling", Design studies, vol.16, No.1, 1995, pp50-61.

[2] Shenchang Eric Chen and Richard Parent. "Shape Averaging and Its Applications to Industrial Design". IEEE Computer Graphics & Application, 9(1),1989, pp47-54.

[3] T.W.Sederberg and E.Greenwood, "A Physically Based Approach to 2D Shape Blending", Computer Graphics(proc. Siggraph), Vol.26, No.2, 1992, pp25-34.

[4] Michal Shapira and Ari Rappoport, "Shape Blending Using the Star-Skeleton Representation", IEEE Computer Graphics & Application, Vol.15, No.2, 1995, pp44-50.

[5] Zhang Y.F, "A fuzzy approach to digital image warping", IEEE computer Graphics & Applications, Vol.16, No.4, 1996, pp35-41

[6] V. Ranjan, A. Fournier, "Matching and Interpolation of Shapes using Unions of Circles", Computer Graphics forum, volume 15, No.3, 1996, pp C-130-C-142

INTERACTING WITH A VIRTUALLY ELASTIC OBJECT

K.C. Hui and M.C. Ma
Department of Mechanical and Automation Engineering
The Chinese University of Hong Kong
Shatin, Hong Kong

ABSTRACT

This paper discusses an attempt to interact with a virtually deformable object in a virtual world. In the proposed system, a CyberGlove™ is used for manipulating and deforming the virtual object. A virtual hand in the virtual environment, which is used for interacting with the virtual object, is constructed based on positional information obtained from the glove. Finite element analysis with tetrahedral element is adopted for evaluating the deformation of the virtual elastic object interacting with the virtual hand. The simulation system is described and experimental results on the interactive response between the virtually deformable object and the virtual hand is also addressed.

KEYWORDS

Deformations, Finite Element Method, Physical Modelling, Virtual Reality, Computer Aided Design

1. INTRODUCTION

Modelling of objects in the virtual world is one of the fundamental problems in virtual reality (VR) development. However, the object models used in contemporary systems are mostly rigid bodies. The systems often ignore the existence of deformable bodies in the virtual world. In other words, most VR applications in design and manufacturing are only capable of simulating rigid prototypes. Prototypes of deformable products cannot be created or tested under these VR systems. The development of deformable models tries to extend VR applications for prototyping deformable products.

A common practice for modelling deformable bodies is to anticipate deformation of the body subjected to the application of external forces. A new object is then constructed based on the original object to represent the deformed body. A major problem with this approach is the intensive computation required is essential for evaluating the deformation which have to be calculated using physical laws. However, this information for visualising and testing deformable model in for the design of soft objects.

Several models had been proposed previously to simulate deformable bodies. Terzopoulos et al [6] proposed a method based on an elastic model. Finite difference method was used for calculating the potential energy of deformation. Kang and Kak [4] proposed a hybrid model based on finite element method using cube element inside the object and plate element on the surface of the model. Gourret et al [5,7] proposed a method using eight nodes solid element. Unknown displacements was calculated based on given displacements on the body.

In this paper, an interactive system for manipulating and deforming a virtual object is proposed. The first section discusses the CyberGlove™ system which is used as the input device for the interaction. In the second section, the tetrahedral element model and the relation between local forces and displacements is described. The method for calculating unknown displacements is also presented. Finally, the implementation results are shown and discussed.

2. CYBERGLOVE™ SYSTEM

In the current implementation, a CyberGlove™ is used as an input interacting with the virtual object. A layout of the system is shown in **Figure 1**.

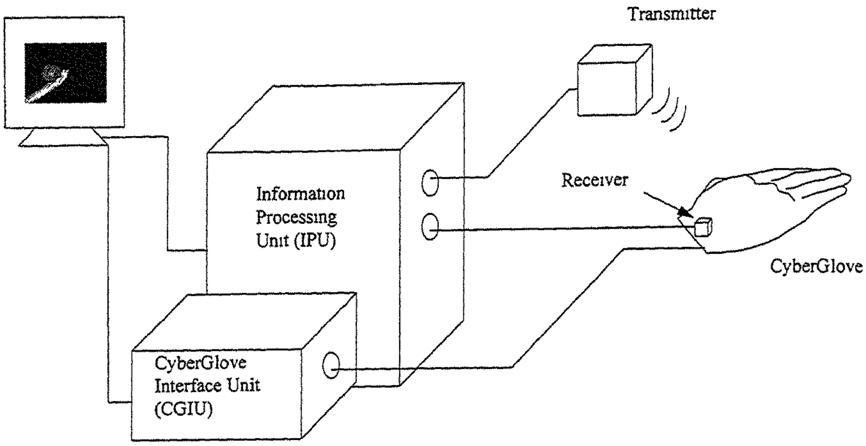


Figure 1 CyberGlove™ System

The position and orientation of the glove is detected by the three dimensional tracking receiver located at the wrist position of the glove. The receiver is a single assembly of three colocated, remote sensing antennas that detects magnetic field vectors sent by the transmitter [1]. The transmitter is a single assembly of three colocated, stationary antenna generating a near field, low frequency, magnetic field vectors. This signal is sent back to the information processing unit (IPU) . The IPU interprets the signal received from the receiver and calculate the three dimensional position and orientation of the glove relative to the transmitter. The position and orientation values are then sent to the computer for the display of the virtual hand.

The relative positions of the fingers are determined by the joint angles between each joints of the fingers [2] since the geometric model of the fingers are known. The sensors are responsible for determining the joint angles between the various sections of the fingers.

The joint angles are then sent to the computer where the virtual hand is constructed and displayed. The position and orientation of the virtual hand thus follow the actual movement of the glove, allowing the hand to interact with virtual environment.

3. SOLID DEFORMABLE MODEL

In this paper, the virtual hand is assumed to be rigid so that there is no deformation of the virtual hand while the object is deformed. It is also assumed that the virtual hand is always in contact with the object. The displacement of the nodes closest to the finger tips are taken to be the displacement of the corresponding finger tips. The objective is then to find the displacement of the other nodes based on these known displacements.

The object to be deformed is assumed to be elastic. The finite element method [3] with tetrahedral elements is adopted for estimating the deformation of the object. A typical tetrahedral element is shown in **Figure 2**.

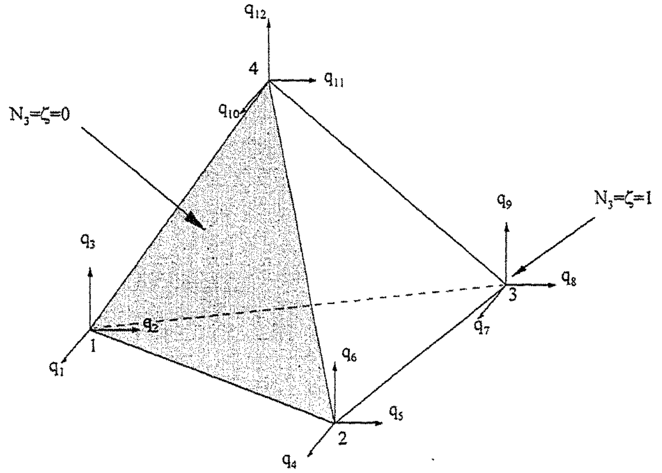


Figure 2 Tetrahedral Model

Since there are four vertices for each tetrahedral element, there are twelve degree of freedom (q_1 to q_{12}). Each degree of freedom corresponds to the displacement of a vertex in the x, y or z direction. The shape functions are defined such that N_j is 1 at vertex j while N_j is 0 on the opposite face of vertex j.

3.1 Stiffness Matrix

Based on the three dimensional elastic model, the stress-strain relationship for isotropic material is,

$$\sigma = \mathbf{D}\varepsilon \tag{1}$$

where **D** is a (6×6)symmetric matrix.

The strain-displacement relations is given by,

$$\varepsilon = \left[\frac{\partial u}{\partial x}, \frac{\partial v}{\partial y}, \frac{\partial w}{\partial z}, \frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}, \frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}, \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right]^T \tag{2}$$

where *u, v, w* are the displacements in the *x, y* and *z* directions respectively.

The relation between the derivatives in *x, y, z* and the derivatives in the normalised co-ordinates ξ, η, ζ are given by,

$$\begin{Bmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \\ \frac{\partial u}{\partial z} \end{Bmatrix} = \mathbf{A} \begin{Bmatrix} \frac{\partial u}{\partial \xi} \\ \frac{\partial u}{\partial \eta} \\ \frac{\partial u}{\partial \zeta} \end{Bmatrix} \quad (3)$$

The displacements *u, v, w* can be written in terms of the unknown nodal values as,

$$\mathbf{u} = \mathbf{Nq} \quad (4)$$

where $\mathbf{N} = \begin{bmatrix} N_1 & 0 & 0 & N_2 & 0 & 0 & N_3 & 0 & 0 & N_4 & 0 & 0 \\ 0 & N_1 & 0 & 0 & N_2 & 0 & 0 & N_3 & 0 & 0 & N_4 & 0 \\ 0 & 0 & N_1 & 0 & 0 & N_2 & 0 & 0 & N_3 & 0 & 0 & N_4 \end{bmatrix}$,

N_i are functions of ξ, η, ζ, and $\mathbf{q}$ is the local displacement vector of the nodes.

From equations **(2)**, **(3)** and **(4)**, the following relation can be obtained,

$$\varepsilon = \mathbf{Bq} \quad (5)$$

where $\mathbf{B}$ is a (6×12) matrix given by,

$$\mathbf{B} = \begin{bmatrix} A_{11} & 0 & 0 & A_{12} & 0 & 0 & A_{13} & 0 & 0 & -\tilde{A}_1 & 0 & 0 \\ 0 & A_{21} & 0 & 0 & A_{22} & 0 & 0 & A_{23} & 0 & 0 & -\tilde{A}_2 & 0 \\ 0 & 0 & A_{31} & 0 & 0 & A_{32} & 0 & 0 & A_{33} & 0 & 0 & -\tilde{A}_3 \\ 0 & A_{31} & A_{21} & 0 & A_{32} & A_{22} & 0 & A_{33} & A_{23} & 0 & -\tilde{A}_3 & -\tilde{A}_2 \\ A_{31} & 0 & A_{11} & A_{32} & 0 & A_{12} & A_{33} & 0 & A_{13} & -\tilde{A}_3 & 0 & -\tilde{A}_1 \\ A_{21} & A_{11} & 0 & A_{22} & A_{12} & 0 & A_{23} & A_{13} & 0 & -\tilde{A}_2 & -\tilde{A}_1 & 0 \end{bmatrix} \quad (6)$$

and $\tilde{A}_1 = A_{11} + A_{12} + A_{13}$, $\tilde{A}_2 = A_{21} + A_{22} + A_{23}$, $\tilde{A}_3 = A_{31} + A_{32} + A_{33}$.
A_{ij} are elements in $\mathbf{A}$ of equation **(3)**.

The element strain energy is then given by,

$$\begin{aligned} \mathbf{U}_e &= \frac{1}{2}\int_e \varepsilon^T \mathbf{D} \varepsilon \, dV \\ &= \frac{1}{2}\mathbf{q}^T\mathbf{B}^T\mathbf{DBq}\int_e dV \\ &= \frac{1}{2}\mathbf{q}^T V_e \mathbf{B}^T\mathbf{DBq} \\ &= \frac{1}{2}\mathbf{q}^T\mathbf{k}^e\mathbf{q} \end{aligned}$$

where the element stiffness matrix $\mathbf{k}^e$ is given by,

$$\mathbf{k}^e = V_e \mathbf{B}^T \mathbf{D} \mathbf{B} \tag{7}$$

and V_e is the volume of the element.

Individual element stiffness matrix are assembled to obtain the global stiffness matrix $(\mathbf{k}^g)$which is used in the relation,

$$\mathbf{F} = \mathbf{k}^g \mathbf{Q} \tag{8}$$

where **Q** is the global displacement vector of the nodes.

3.2 Solving the Linear System

The problem of deforming a virtual object with a virtual hand is thus to evaluate the unknown displacements of the nodes based on known displacements of those nodes corresponding to the points of contact between the virtual hand and object. Besides, the external forces are zero if the nodal displacements are unknown Rearranging the stiffness matrix with the displacement node orders and the force orders, relation **(8)** becomes,

$$\begin{bmatrix} 0 \\ F_u \end{bmatrix} = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} Q_u \\ Q_k \end{bmatrix} \tag{9}$$

where F_u is the unknown force vector, Q_u is the unknown displacement, Q_k is the known displacement and K_{11}, K_{12}, K_{21}, K_{22} are the submatrices of the rearranged global stiffness matrix $(\mathbf{k}^g)$.

Hence,

$$\begin{cases} Q_u = -K_{11}^{-1} K_{12} Q_k & \text{(10a)} \\ F_u = K_{21} Q_u + K_{22} Q_k & \text{(10b)} \end{cases}$$

Unknown displacements and forces can thus be found by equation **(10)**. By adding the nodal displacements to the nodes of the object, the deformed shape of the object can be found and displayed.

4. IMPLEMENTATION

In the current implementation, a virtual object with 267 nodes and 528 faces is used for the interaction test. The time required for rearranging the global stiffness matrix and inverting matrix K_{11} is 280 seconds.

In each frame of the display the virtual background is first drawn followed by the virtual hand. The known displacement is obtained by evaluating the displacement between the node of contact and the finger tip. By applying relation **(10a)**, the unknown displacements are calculated in real time. The deformed virtual object is then drawn. The experimental system implemented on an SGI Indigo2 workstation produces 16.8 frames per second with the above algorithm. **Figure 3, 4** and **5** are snap shots of the deformation process.

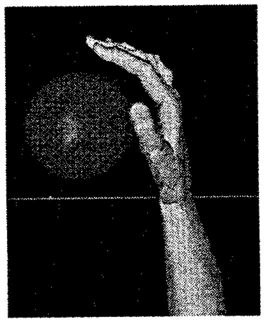


Figure 3 - An object attached to the virtual hand

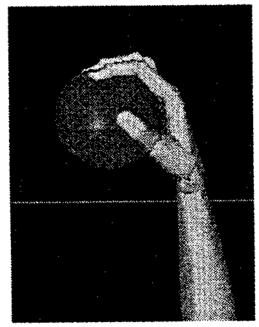


Figure 4 - Fingers collided with the object and the object undergo small deformation

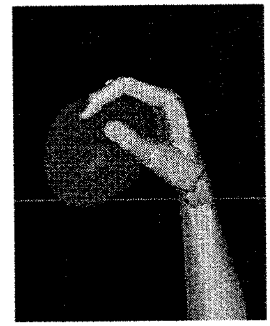


Figure 5 - Fingers moved further resulting in large deformation of the object.

5. CONCLUSION

A method to interact with a virtually deformable object using an instrumented glove is presented in this paper. Experimental results showed that the method gives acceptable interactive behaviour. Although the finite element method based on a linear elastic model is used for the deformation calculation, the results give a good visual perception of deforming a relatively soft object.

Since the calculation of matrix inverse is time consuming and cannot be performed in real time, the stiffness matrix is assumed to be constant. In other words, the points of contact are predefined.

If the point of contact are allowed to vary, the inversion of the stiffness matrix have to calculated in real time. In this case, parallel processing techniques may have to be adopted in order to speed up the process. This is being investigated by the authors currently.

6. REFERENCES

1. Polhemus 3Space® Users Manual, Chapter 1, Polhemus Incorporated Colchester, Vermont, 1994.
2. CyberGlove™ User's Manual, Chapter 4, Virtual Technologies, Palo Alto, 1994.
3. Tirupathi R. Chansrupatla, Ashoh D. Belegundu, Introduction to Finite Elements in Engineering, Chapter 1,2,3,9, Prentice-Hall, Inc., New Jersey, 1991.
4. HoSeok Kang, Avi Kak, "Deforming virtual objects interactively in accordance with an elastic model", Computer Aided Design, Volume 28, Number 4, pp. 251-282, 1996.
5. Nadia Magnenat Thalmann, Daniel Thalmann, "Finite elements in task level animation", Finite elements in Analysis and Design, Volume 19, pp. 227-242, 1995.
6. Demetri Terzopoulos, John Platt, Alan Barr, Kurt Fleischer, "Elastically Deformable Models", Volume 21, pp. 205-214, Proceedings of SIGGRAPH '87, Computer Graphics, ACM, Anaheim, 1987.
7. Jean-Paul Gourret, Nadia Magnenat Thalmann, Daniel Thalmann, "Simulation of Object and Human Skin Deformations in a Grasping Task", Volume 23, Number 3, Proceedings of SIGGRAPH '89, ACM, Boston, 1989.

2D BOUNDARY RECOVERY BASED ON MEDIAL AXIS TRANSFORM

Yujin Hu, M.M.F. Yuen and Ji Zhou*
Department of Mechanical Engineering
Hong Kong University of Science and Technology
*Huazhong University of Science and Technology

ABSTRACT

The medial axis transform(MAT) is an alternative representation of geometric models that has many useful properties for design and analysis modeling. In this paper, a method to recover the object boundary from its MAT is presented. First, the classification of the geometric entities associated with MAT based on the tangent continuity condition is outlined. Then the relationship between the MAT tangent and its corresponding boundary tangent is established. This is a key parameter in the process of recovering boundary from a given MAT. Moreover, the condition ensuring the geometric validity of boundary generated, that is no self-intersection, is discussed. Finally a few of examples demonstrating the boundary recovery of simple 2D object from its own MAT are given.

KEYWORDS

Medial axis transformation(MAT), boundary recovery, local validity

1. INTRODUCTION

The medial axis transform(MAT), also known as the skeleton, was first proposed by Blum[1] in 1967 for the purpose of describing and characterizing biological shapes. It is defined as the closure of the locus of centers of all maximal inscribed disks within the boundary of the object. The MAT possesses many interesting properties. One of its important properties is uniqueness and invertablity. That is, given an object boundary there is an exact MAT; and given a MAT, there exist a corresponding object boundary. For extracting MAT representation from a given object, a lot of algorithms have been developed[3-5]. In this paper, we will consider the inverse problem, that is how to recover the object boundary from a given MAT. In 1995, Gelston and Dutta[2] first investigated the boundary recovery for an object whose MAT is an individual geometrical entity with G^1 continuity. As the MAT of most engineering objects is not form from an individual geometrical entity, more complex cases in the process of boundary recovery need to be considered.

In this paper, the boundary recovery of simple 2D object are discussed, including the cases of boundary recovery of the MA points without G^1 continuity and boundary non-self-intersection. Section 2 provides definitions and problem description related to boundary recovery. Sections 3 and 4 discuss the boundary recovery from MAT made up of a mixture of individual curve segments with G^1 continuity and special points without G^1-continuity. Section 5 investigates the intersection behaviors in the process of boundary recovery, and proposes theorems to ensure valid boundary recovered from the MAT. Finally examples and conclusions are presented in section 6.

2. DEFINITION AND PROBLEM DESCRIPTION

Definition 1 A 2D simple object D is a bounded portion of a two-dimensional region R^2 satisfying the conditions:

a) D is bounded by a finite number of boundary curves and its interior is path-connected.

b) Each boundary curve segment is G^2 continuous except the two end points.

Definition 2 The Medial Axis (MA) of simple object D is the closure of the locus of centers of all maximal inscribed disks in D, which is completely contained inside D and not contained in any other inscribed disk. The radius function of the MA is a continuous real-valued function whose value at each point on the MA is equal to the radius of the associated maximal disk. The MAT of D is the MA together with its associated radius function.

By the two definitions above, the following lemma are proposed:

Lemma The MAT of a simple object D of R^2 is G^1 continuous except at finite points.

This lemma is easy to proof. As the boundary of D is G^2 continuous, so its MA is definitely continuous. By virtue of the definition of continuous function, the MA is G^1 continuous except at the finite points. Moreover, at these finite points, the one-sided tangent exists.

Definition 3 Normal point is one where the MAT is G^1 continuous and whose disk touches the boundary of D in two points.

Definition 4 Branch point is one where the MAT only has one-sided tangent (not G^1 continuous) and whose disk touches the boundary of D in three or more discrete points.

Definition 5 End point is one where the MAT only has one-sided tangent (not G^1 continuous) and whose disk has zero radius.

Definition 6 Finite contact point is one where the MAT only has one-sided tangent (not G^1 continuous) and whose disk touches the boundary along a finite segment of the boundary (or at continuous points).

In order to reconstruct the boundary from MAT, let $P = \{p_j | 0 \le j \le n\}$ be a set of special points consisting of branch points, end points and finite contact points of the MA; and let $S = \{s_i | 0 \le i \le l\}$ be a set of G^1 continuous curve segments which forms part of the MA. For each s_i in S, it has a corresponding radius function r_i. Assuming that the boundary curve corresponding to a s_i is b_i, then union of all b_i is the set $B = \bigcup_0^l b_i$. If the MA consists of S without special points, then the boundary B is formed from a set of continuous curve as shown in Fig. 1a. At the junction point of different s_i, then B may compose of disconnected boundary segments or self-intersected boundary segments as shown in Fig. 1b. Obviously, in this case B is an incorrect description of the resulting boundary corresponding to the MAT. In order to cope with this disconnected and self-intersecting problem, the boundary recovery at the end points of s_i must be specially considered. Generally, for the disconnected problem, circular arcs can be used to connect the disjoint boundary curve segments as in Fig. 1c. However, the self-intersecting problem is more complicated. The boundary intersections caused by points within neighborhood of any point on s_i are called local self-intersections, and the boundary intersections caused by points which are at a distant from each other are called global self-intersections. Fig. 1d shows an example of global self-intersection. In this paper, some of the above problems such as the boundary reconstruction of individual curve segment s_i, the boundary reconstruction of special point p_j and local self-intersection of the resulting boundary will be discussed.

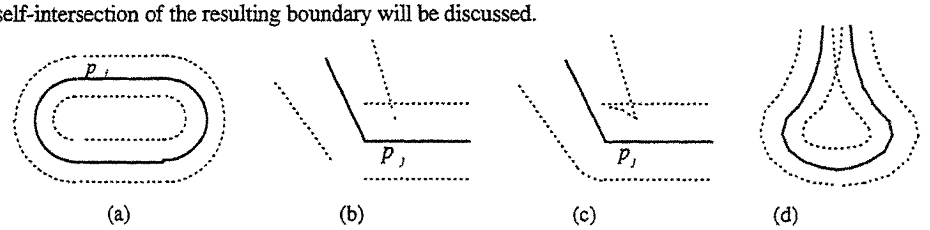


Fig. 1 Boundary curves b_i corresponding to MA curve segments s_i

3. BOUNDARY RECOVERY FROM INDIVIDUAL CURVE SEGMENT

The key in reconstructing boundary from an individual MA curve segment is to find the geometrical relationship between the MA and the boundary. Assuming that the curve segment s_i is described by a parameter curve $s_i(u)$; the maximal inscribed disk at any point of $s_i(u)$ exists, and the corresponding radius function is $r_i(u)$. Let p_1 and p_2 be two neighboring points on $s_i(u)$ with radius function r_1 and r_2 respectively as shown in Fig. 2. Further assuming that the curve segment $s_i(u)$ is G^1 continuous. The corresponding boundaries are b_i^i and b_i^o. q_i and q_o are two boundary points corresponding to the MA point p_1. For convenience, only one of the boundaries of $s_i(u), b_i^o$, is considered. Let $C(b_t^o)$ be the offset of the boundary curve b_t^o with distance r_1, τ_s and τ_b^o are the tangents of $s_t(u)$ and $C(b_t^o)$ respectively at the MA point p_1. p_2 is the neighbor of p_1 and the angle α between two tangents τ_s and τ_b^o can be written as

$$\sin\alpha = \lim \frac{r_2 - r_1}{\Delta s(u)} = \frac{r_t'(u)}{|s_t'(u)|} = \frac{dr}{ds} \qquad (1.1)$$

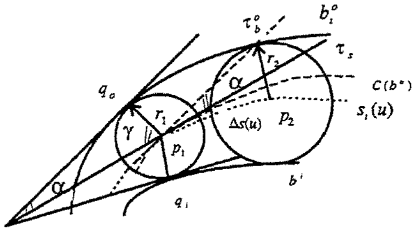


Fig. 2 Geometrical relationship between $s_t(u)$ and b_t

From the definition of the MA point, point p_1 is not only an MA point but the orthogonal projection of boundary point q_o. That is to say the line $\overline{p_1 q_o}$ is along the direction of the boundary normal of q_o. Calling the line $\overline{p_1 q_o}$ as the spoke of the maximal inscribed disk center at p_1 on $s_t(u)$ and defining the minimum angle between the spoke and the tangent of the MA at the point p_1 as the orientation angle of the spoke, γ ; the formula (1.1) can be rewritten in terms of the orientation angle:

$$\cos\gamma = \frac{r_t'(u)}{|s_t'(u)|} = \frac{dr}{ds} \qquad (1.2)$$

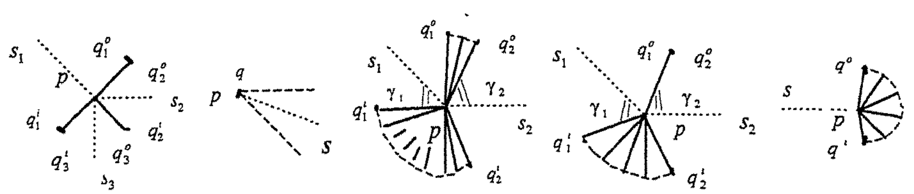
Obviously the MA curve segment $s_t(u)$ and corresponding radius function $r_t(u)$ have to satisfy the condition: $|\frac{dr}{ds}| < 1$ We generally call this condition as the maximal disk existence condition of the MAT. Once the MAT satisfies this condition and the angle γ is determined from equation (1.2), the boundary points q_t and q_o corresponding to a MA point on $s_t(u)$ can be defined.

4. BOUNDARY RECOVERY AT SPECIAL POINTS

The set of special points includes branch points, end points and finite contact points. At these points, although the MAT is not G^1 continuous, it has one-sided derivative. Let p be a branch point. According to the definition of the branch point, it belongs to at least three MA curve segments, such as s_1, s_2 and s_3 (in special cases such as a concave vertex boundary, it is the intersection of two MA curve segments). Thus there are at least six boundary points which can be recovered from the end point of the different curve segments respectively. Each of the three G^1 continuos MA curve segment is one-sided continuous at p. When the MA points on the different MA curve segments approach the branch point p from three different directions along s_1, s_2 and s_3, the two set of boundary points, q_1^o and q_2^o, q_2^i and q_3^o, q_1^i and q_3^i, will respectively become coincident as in Fig. 3a.

For the end point which is a point with radius equal to zero, the corresponding boundary points will be coincident with the end point as in Fig. 3b.

For the finite contact point with its maximal inscribed disk touching the boundary along finite curve segment or at continuous points, it is considered as the orthogonal projection of the finite segment of the boundary with the same radius function. The boundary corresponding to the finite contact point is certainly an arc segment. In addition, the finite contact point is also the endpoint of MA curve segments, such as an end point of one curve segment or the connecting point of multiple curve segments. Thus, the corresponding boundary point is at least one pair. Once the boundary points corresponding to all end points are determined, all boundary points with respect to the finite contact point can be described as an arc segment. Fig. 3c to 3e show several of the boundary configurations with finite contact point.



(a) (b) (c) (d) (e)

········· MA curve segment ——— Spoke of point p ------ Boundary curve

Fig. 3 Boundary points recovered from the special points

5. LOCAL VALID OF BOUNDARY RECOVERY

Self-intersection is closely related to the MA and its associated radius function. In this section, the complete definition of the spoke at any point and its local validity will first be given, then several theorems avoiding local self-intersection in the process of boundary recovery will be discussed.

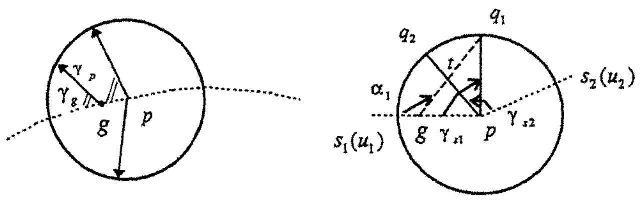
Definition 7 Let p be a MA point, q be its corresponding set of boundary points. Then the union of all points on the lines between p and q is called as the spoke of the MA point p, written as $Sp(p) = \cup\{w | w \in \overline{pq}\}$.

Obviously, the spoke of a branch point has more than two line segments from p to the corresponding boundary points. As for the finite contact point, its spoke has one or more sector region consisting points on all projection lines from points on the corresponding finite boundary segments as shown in Fig. 3 in solid lines.

For the case of local self-intersection of the boundary as shown in Fig. 1c, this is the result of a MA point boundary entering the boundary of another MA point because the MAT is not suitably defined. Based on the spoke concept, we can define the local validity of the boundary as follows:

Definition 8 Let p_1 and p_2 be neighboring points on the MA. If the MAT satisfies the maximal inscribed disk existence condition, and the intersection of the $Sp(p_1)$ and $Sp(p_2)$ is not located within the maximal inscribed disk of p_1, then the MAT at p_1 is local valid.

The neighborhood of p_1 is a finite small region around p_1. The first condition guarantees that there is a maximal inscribed disk at the point p_1. The second condition guarantees that there is no self-intersecting boundary at the neighbor of p_1. If the point p_2 in the definition above is not constrained in the neighbor of p_1, the self-intersection of the resulting boundary is called as the global self-intersection. For convenience, the discussion is limited to conditions avoiding the local self-intersection in the process of boundary recovery.



(a) (b)

Fig. 4 A point within the neighborhood of the

Let p be a normal point on an individual MA curve segment and g be another point on the same MA within the neighborhood of p as shown in Fig. 4a. If g is located to the left of p and the orientation angle of $Sp(g)$ is less than that of $Sp(p)$, then non-intersection of $Sp(g)$ and $Sp(p)$ occurs irrespective of the radius function at point g. Likewise when g is located to the right of p and the orientation angle of $Sp(g)$ is larger than that of $Sp(p)$, again non-intersection of $Sp(g)$ and $Sp(p)$ occurs.

Theorem 1 Let the MA segment $s(u)$ be a linear function and the radius function $r(u)$ be a linear or quadratic function. If they satisfy the maximal inscribed disk existence condition, that is, $\frac{dr}{ds} < 1$, then the MAT is the locally valid.

Proof: Without the loss of generality, we assume the MA and radius function are

$$s(u) = \begin{pmatrix} x(u) \\ y(u) \end{pmatrix} = \begin{pmatrix} au+b \\ cu+d \end{pmatrix} \qquad r(u) = Au^2 + Bu + C$$

then

$$\frac{dr}{ds} = \frac{r'(u)}{|s'(u)|} = \frac{2Au+B}{\sqrt{a^2+c^2}}$$

Obviously, $\frac{dr}{ds}$ is a linear function of u. If $|\frac{dr}{ds}| < 1$, the maximal inscribed disks at any point p on the MA exist, then the spoke at a point on MA $s(u)$ also exists. As $s(u)$ and $\frac{dr}{ds}$ are all the linear functions of u, any two spokes corresponding to two different MA points will not intersect with one another.

For an endpoint of a linear MA segment $s(u)$, the following theorems apply:

Theorem 2 Let p be the endpoint of one MA curve segment $s(u)$, irrespective of being an end point or a finite contact point, if the MAT satisfies the maximal inscribed disk existence condition, then the MAT is locally valid at p.

This result is obvious. If the boundary point at p exist, it is definitely locally valid.

Theorem 3 Let p be the connecting point of two or more MA curve segments $s_i(u)$ $i = 1,2,\ldots$. If $s_i(u)$ $i = 1,2,\ldots$ satisfy respectively the maximal inscribed disk existence condition at p, and the sum of the orientation angles corresponding to any two $s_i(u)$ at the connecting point p is not larger than the angle between their tangents at p, then the MAT is locally valid at p

Proof: Referring to Fig. 4b, for convenience, let p be a connecting point of two MA curve segments, $s_1(u_1)$, $s_2(u_2)$. $\overline{pq_1}$ and $\overline{pq_2}$ are respectively the partial spoke corresponding to $s_1(u_1)$ and $s_2(u_2)$ at point p. Their orientation angles are respectively γ_{s1}, γ_{s2}. Assuming that the MAT at p is non-locally valid, it means 1) the MAT doesn't satisfy maximal inscribed disk existence condition; 2) the $Sp(p)$ intersects with the spoke of any point within neighborhood of p. Result 1) is obviously contradicting with the assumption of theorem 3. However, for result 2), let g be the point on $s_1(u_1)$ within the neighborhood of p. From theorem 1, the $Sp(g)$ is impossibly to intersect with the spoke of any point on the $s_1(u_1)$. Let $Sp(g)$ intersects with the spoke of point p on $s_2(u_2)$, as shown in Fig. 4b, the intersecting point t together with g and p forms a triangle Δptg. Using simple relationship of the triangle, we can obtain the equation $\gamma_{s1} + \gamma_{s2} = \alpha + \angle q_1 p q_2$, where α be the angle between two MA curve segments $s_1(u_1)$ $s_2(u_2)$ at point p. As the spoke of g intersects with the spoke of point p on the $s_2(u_2)$, $\angle q_1 p q_2$ is always lager than zero. However, the relationship $\gamma_{s1} + \gamma_{s2} > \alpha$ is in contradiction with the assumption of theorem 3 and thus the theorem holds.

6. Examples and Conclusion

Based on above the theorems, an algorithm for the boundary recovery from piecewise linear MA is constructed and implemented. Two simple examples of boundary recovery are shown in Fig. 5. The MA curve segment in Fig. 5a is a single straight line segment from (0,0) to (10, 0). The corresponding radius function is a quadratic function $r(u) = 4(u-0.5)^2 + 1$. The MA curve segments and radius functions in Fig. 5b are given below,

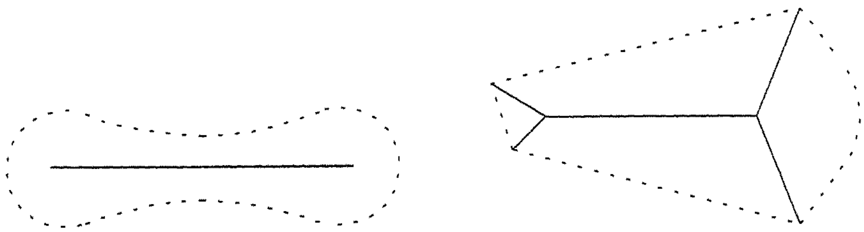
$$s_1(u) = \begin{pmatrix} x(u) \\ y(u) \end{pmatrix} = \begin{pmatrix} 2.4u - 2.4 \\ 1.5u - 1.5 \end{pmatrix} \qquad s_2(u) = \begin{pmatrix} x(u) \\ y(u) \end{pmatrix} = \begin{pmatrix} 1.6u - 1.6 \\ 1.6u + 1.6 \end{pmatrix}$$

$$r_1(u) = 2u \qquad r_2(u) = 2u$$

$$s_3(u) = \begin{pmatrix} x(u) \\ y(u) \end{pmatrix} = \begin{pmatrix} 10u \\ 0 \end{pmatrix} \qquad s_4(u) = \begin{pmatrix} x(u) \\ y(u) \end{pmatrix} = \begin{pmatrix} 2u + 10 \\ 5u \end{pmatrix}$$

$$r_3(u) = 2.5u + 2 \qquad r_4(u) = -4.5u + 4.5$$

$$s_5(u) = \begin{pmatrix} x(u) \\ y(u) \end{pmatrix} = \begin{pmatrix} 2u + 10 \\ -5u \end{pmatrix} \qquad r_5(u) = -4.5u + 4.5$$



(a) (b)

Fig. 7 Examples of boundary recovered from given MAT

A method to recovery boundary from the given MAT of a simple 2D object is presented. The ideas of utilizing the maximal inscribed disk existence condition of the MAT and the local validity conditions introduced in this paper can lead to an algorithm for determining the discrete boundary from the given MAT of a 2D simple object. Using this algorithm, some simple examples have been demonstrated. The local validity for non-linear MA cannot be checked directly with the theorems introduced in this paper, but linearization of the non-linear MA can be used as a solution. The global validity of boundary recovery from given MAT is more complicated, further investigation is necessary in the future work. Although the method presented in this paper is based on 2D simple object, it can form the basis to extend to the recovery of the boundary of simple 3D objects.

References

1. Blum,H., A transformation for extracting new descriptors of shape, in: W. Whaten-Dunn, ed., Models for the Perception of speech and Visual Form, MIT Press, Cambridge, MA. 1967.
2. Gelston, S. M. and Dutta, D., Boundary surface recovery from skeleton curves, Computer Aided Geometric Design, 12(1995) 27-51.
3. Dutta,D. and Hoffmann, C. M., On the skeleton of simple CSG objects, ASME Trans. J. Mech. Design,(1993) 87-94.
4. Sherbrooke, E. C., Patrikalakis, N. M., An algorithm for the medial axis transform of 3D polyhedral solids. IEEE Trans. on Visualization and Computer Graphics, 2(1996) 44-61.
5. Sheehy, D. J. and Armstrong, C. G., Shape description by medial surface construction, IEEE Trans. on Visualization and Computer Graphics, 2(1996) 62-72.

SYSTEM OF COST CONTROL AND ANALYSIS DECISION SUPPORT UNDER CIMS CIRCUMSTANCE

Baodong Xu,Bin Gao,Jiazhong Chang,Bo Xu
Northeastern University
P. R. CHINA

ABSTRACT

This paper states the basic concepts of cost control under CIMS circumstance in special steel enterprises. It describes the designing idea of integration, hierarchical structure, globe shared and local shared and closed-loop large-scale of cost decision support system, and it also puts forward analysis and design of model base, data base, algorithm base and knowledge base for cost control and analysis under CIMS circumstance, which aims at the cost control for special steel enterprises.

KEYWORDS

CIMS, Decision Support System, Cost Control

1. CONCEPTS OF COST CONTROL AND DECISION SUPPORT SYSTEM

1.1 Concept of cost control

Cost control [1] menus, during the forming procedure of manufacturing and operating cost, it directs, restricts and supervises all kinds of principles, policies and quota, and it also finds deviation in time and takes correcting measures to make detailed and total manufacturing expense controlled within stipulated range, at the same time , on the basis of adopting advanced measures and extending advanced experience constantly, it revises and establishes new cost goal and reduces cost so as to make cost reach superior level.

As to steel enterprises, enhancing cost control, finding efficient managerial methods, reducing cost, can not only increase managerial level, increase economic benefit but also be the key of improving interior manufacturing and reducing the short of capital. Using CIMS, can formulate reasonable cost plans, using feedback information can calculate and control cost, using computer can assist manufacturing management, shorten manufacturing period, realize computer inventory management and reduce capital occupation so as to reach the goal of reducing products cost.

1.2 Content of cost control decision support system

The object of cost control for steel enterprises is the whole process of enterprise capital motion, from investing capital , using capital, expending capital to taking back the capital. At the investing stage, the object is the investing cost control; at the product design stage, the cost control of production design; at the stage of stipulating technology scheme, the cost control of technology scheme; at the stage of products exploring, the control of materials stock; at the production stage, the cost control of products and quality; at the stage of product sale, the cost control of production sale and product life period.

Information is the basis of cost decision, there are a lot of indefinited random factors and need feedback information so as to make the best cost decision.

1.3 Mode of cost control

Cost control contains prior cost control, intermediate cost control, posterior cost control.

(1)Prior cost control refers to the control for production design, technology scheme design and material purchasing stock which can be controlled before putting into production.

(2)Intermediate cost control, refers to the control of the whole forming process of products from exploring, product investment of fixed assets, supplying to manufacturing and selling process.

(3)Posterior cost control, refers to analysis and checking comprehensively after formation of cost. Cost control runs through all the phases of manufacturing and operating of steel enterprises.

In steel enterprises, practical and efficient cost control is quite important. In view of modern control, cost control can be divided into feedforward control, feedback control, "Mode Identification" control model and preventive control [2].

"Mode Identification" control, means that there are many unknown factors which can affect cost, but people can identify system model to reach the goal of cost control according to the relation of related input and output, i.e. the relation of system input and system output. The model of "mode identification" control is shown in Fig 1.

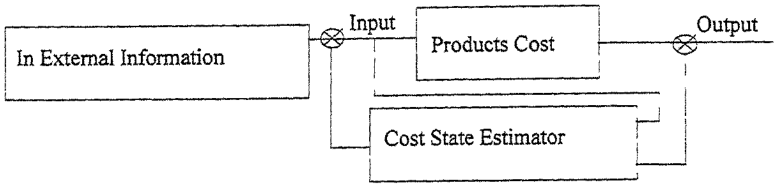


Fig1 cost "Mode dentification" System

2. ANALYSIS AND DESIGN OF COST CONTROL DECISION SUPPORT SYSTEM

2.1 Designing idea of cost control decision support system

The purpose of exploring cost managerial decision support system is mainly through the use of computers to provide efficient support for all kinds of decision motion. Owing to great variety of products and complex manufacturing technology which decide the complexity of cost management, it is difficult to be competent for this work only depending on person's experience and intuition. On the other hand, steel enterprises face many decisions including quantitative and qualitative factors, which should be integrated, that is what we called semistructural problem, i.e. fuzzily of decision.

(1)Designing idea of integration

As to the workshops (or subfactories), firstly, collecting according to the usage of expenses, and allocating all kinds of essential expenses; Secondly, collecting and allocating subsidiary manufacturing expenses; Thirdly, collecting and allocating manufacturing expenses and managerial expenses; Fourthly, allocating every variety of manufacturing expenses between complete products and uncompleted products in order to determine processing product cost. Generally, four procedures can be classified as initial allocating, collection, reallocating, forming reports.

CIMS system is concerned in information integration, physical integration, network integration and functional integration, while information integration is the core of CIMS. It concludes: ① integrating with manufacturing subsystem; ② integrating with subsidiary manufacturing subsystem; ③ integrating with functional department subsystem; ④ integrating with oversea and domestic information market subsystem.

(2)Designing idea for consideration of general use and special use system

In order to design cost control decision support system which has both general use and special use characters, designed system can use principle of data, model and method separation. As to model base and model base management system, use unified model skeleton and general function. Data base can be classified as globe shared, local shared, special data base. Between general use system and special use system, use friendly man-machine interface.

(3) Designing idea of hierarchical structure

CIMS is the integration of man, management, control, the productive technological equipment, designing technique of production and the circumstance of market, after making products by productive procedure, then entering into market, and forming a closed-loop large-scale system. It contains some related subsystems which are different in function and scale. These subsystems are integrated by computer network and constitute CIMS of the whole enterprise [3].

Cost control can be divided into three kinds of decision. The first kind is daily control decision, it mainly treats the arrangement of manufacture, control of expenditure of materials and labors, adjustment of task etc.; The second kind is tactical decision such as using resources, personnel transferring, capital circulation etc.; The third kind is strategical decision, involving operating target of enterprise, direction of operational developing, exploring new products and so on. The basis of hierarchical idea is: ① the managers of different levels have different requests for cost control. The leaders of enterprises are concerned about strategical decision, while the lower leaders are concerned about control decision; ② cost control of enterprise executes stage, from arranging products and ordering goods to sales of products and so on; ③ realizing hierarchical and multistage cost control under network circumstance. The whole system through the use of enable level (system enable and applied enable) supplies more advanced data management and the function of communication [4]. Describe the relation of every level of model through model codes, realizing the demands of managers of different levels through network structure.

(4) Design of closed loop cost control

If cost control in the process of production is to be realized, it is mainly through information feedback to reflect the deviation which appear in the forming process of cost, to adopt measures correcting deviation and ensuring the target of cost control to be realized.

2.2 Design of cost control decision support system

The quantity of cost calculating and management is large, and there are many links, the interconnection is complex, so the principles of designing cost control decision support system are: ① the function of data processing; ② emphasizing general and flexible feature; ③ the function of planning; ④ analyzing controlling and generating related target cost; ⑤ the function of decision.

2.3 Logical structure of cost control decision support system

In the process of designing cost control decision support system, system can be designed as four-base system according to the feature of problems and general designing idea, the logical structure which is shown [5] as Fig 2.

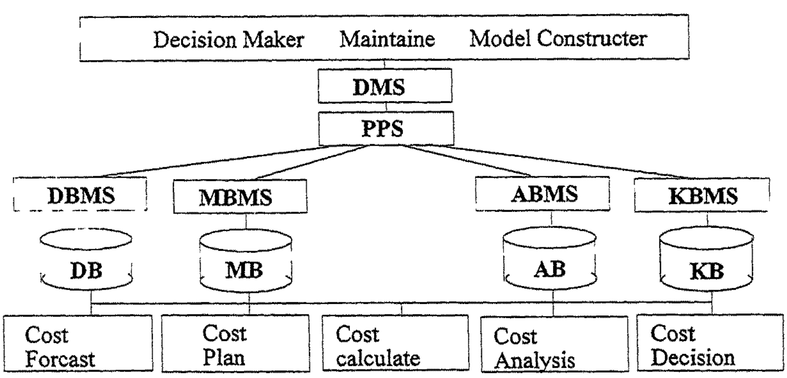


Fig 2 Cost Control DSS Logical Structure

Problem Process System (PPS) is in charge of screen menu management, coordinates and controls the other subsystems, and realizes user-dialogue management, it serves the function of general control of system.

Data Base System (DBS) is composed of DB and DBMS. DBMS can serve the function of managing data base, interface data and model, and dealing middle-data dynamically in decision process and so on.

Model Base System (MBS) is composed of MB and MBMS. It can afford the function of generating model automatically, producing data base structure automatically and maintenance of model.

Algorithm Base System (ABS) contains AB and ABMS, it can serve the function of adding algorithm, inquiring and performing synthetically and so on.

Knoeledge Base System (KBS) is composed of KB and KBMS. it has the function of adding, modification and depending.

3. DESIGN OF MODEL BASE AND MODEL BASE MANAGEMENT SYSTEM

The data base is mainly used in the stage of collecting data in the decision support process. As to the whole decision procedure, DSS should also provide a set of principle of using model and analyzing model. MBMS makes decision maker having enough ability of solving problems and analyzing problems through generating, analyzing, comparing different alternatives of problem. In cost control decision support system, model base management system is a very important part, such as establishing some fixed models, it is difficult to adapt the request of changeable, semi-structure or non-structure problems, so it is very significant to establish the models for efficient decision.

The model base should have the following features:

(1) Separating model from algorithm. Two parts shouldn't affect each other whether changing algorithm or changing model. So models of any structure can be established flexibly.

(2) Generating model automatically. System users can dialogue with computers, which can aid users to establish model automatically. If only the users input content about their own business, system will judge the content inputted and establish model structure and save it automatically.

(3) Establishing date base structure automatically. Different cost control decision problems have different data base structure, so data structure should be changed flexibly. Because model structure are related to data base structure closely, when model is being established, data base structure will be established automatically, thus solving the interface of model and data so that users can operate easily.

In model base system, how to save the model of cost control into computer so that decision maker can use model efficiently and get conclusion of problems quickly and reach the goal of decision is a key problem.

In general model can be represented in three forms, this paper only describes three levels framework. This method divides model into external scheme and internal scheme, external scheme is the basic external information of model, internal scheme is the detailed description of model.

3.1 Design of three levels framework of cost control decision model

Representation by framework was putted forward by M. Minsky in 1975, it is a kind of knowledge representation.

First level framework is index scheme of model, and it reflects external feature of model.Second level decries parameter scheme of model and internal scheme of model, cost control decision support system can adopt many kinds of scheme such as hierarchical analysis scheme, linear programming, goal programming and so on. Third level framework reflect entity data of model, which will be saved in the form of data base. Three level framework of model are realized by relation data base, first level framework structure is data dictionary scheme. It is suitable for any model, and belongs to general shared model structure, second level model describes parameter scheme. Its data structures vary with model, and belongs to local shared structure.

3.2 General structure of cost control and model generation

Because the model of cost control decision support system adopts representation of three level framework. Its general structure is shown as Fig 3 . It has following functions:

(1) Establishing new models. When the establishers face different models, system can fulfill the work of model generation conveniently;

(2) Editing model. In the process of man-machine interface, realizing the modification and deletion of model;

(3) Inquiring and indexing model, i.e. inquiring different models;

(4) Automatically solving models and output result;

(5) Establishing the third level framework storage base.

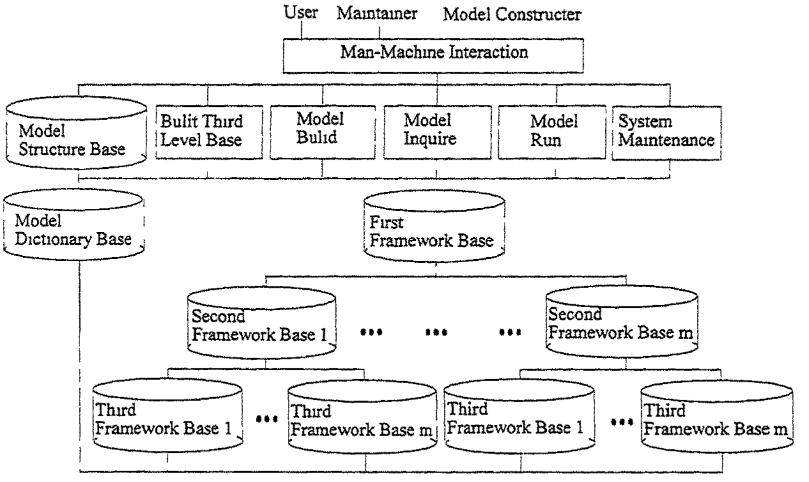


Fig 3 MBMS General structure of cost control Based on Three Level Framework

4. CONCLUSIONS

In the process of design and realization of steel enterprise cost control decision support system, usually concerning about semi-structure and non-structure problems. Semi-structure decision means that decision scheme can be described by suitable maths model, but can't get optimal answer; Non-structure decision means that decision scheme can't get through establishing suitable model, and it must be dealt with by solving method of analyzing, analogizing and judging etc. These need establishing knowledge base and knowledge base management system. Therefore the function integration, network integration, data base integration and human integration and so on [6], will serve the great function of realizing efficient cost control decision support system.

5. REFERENCES

1.Huailian Hou, Enterprise Cost Control Engineering,5-10,SHANDONG UNIVERSITY PRESS,1988

2. Junqiu Xie, "About the Cost Control Model and Method", accountist,No.3,pp.45-46,1991

3. Zhenyu Chen, "The development and prospects of CIMS of Iron-steel industry of our country", Metallurgical Automation, Vol.18, No.6,pp.4-5,1994

4. Ginsong Xue et al. , "Report on Design kf Computer Integrated Manufacturing System for Shenyang First Machine Tool Plant", pp.169-171,1994

5. Zihou Yang et al. , "Decision Support System for the Planning of steel product Mix", Met-alluvgical Industry Automation Vol. 14, No.6, pp.3-6, 1990

6. Zhichang Li, "Initial Research of CIMS System Integration Methods", Computer Applications, Vol. 14, No. 6, pp.1-3,1994

PRODUCT DATA MANAGEMENT SYSTEM

Hongyuan Xue, Ji Zhou and Yifang Zhong
CAD Center
Huazhong University of Science and Technology (HUST)
Wuhan, China

ABSTRACT

Supporting the capture, storage and presentation of design information is a vital research activity in today's increasingly electronic design environment. As the information infrastructure of the modern enterprise, Product Data Management system plays a more important role in utilizing large quantities of information from a wide range of sources. However, the current techniques of storage and presentation are relatively unstructured, badly indexed and poorly presented. These obstacles make it difficult for an individual both to find the appropriate data and to use it to support different design task. This paper mainly discusses three models to solve above problems and a prototype of Product Data Management system is presented to illustrate the usage of product reference model for enterprise, extended product structure model and task model.

KEYWORDS

Product Data Management, Product Structure, Design Process, Design Activity

1. INTRODUCTION

The traditional method of product design, which is usually sequential and rigid, has proved to be not only very time-consuming but also may lead to poor designs that are expensive to realize. Meanwhile, the support systems of product development used widely are characterized by island of automation, where most of application programs are stand-alone and operating on heterogeneous platforms. It would be difficult for product development to progress much further alone without seriously considering real work settings. Thus we see an emerging need to exchange information inter or intra computer aided information systems. Fortunately, the development of Information Technology (IT) has provided a basic framework for CAD, CAM applications in enterprise. In particular, Product Data Management (PDM) can be a virtual factor for successful transmitting and sharing of information.

PDM began to emerge as a distinct commercial industry in the early 1980s. Previous to that time, companies that wanted a PDM system generally had to custom-develop in-house solution. By the mid 1980s, several software developers began to offer relatively primitive PDM systems. The technology has evolved rapidly in the past few years. A few famous commercial PDM systems are available presently such as AutoCAD WorkCenter, Computervision EDM, PTC ProPDM and SDRC Metaphase, which target user's different requirements. PDM is one of the fastest growing technologies in the production industry. More and more companies are realizing the value of organizing, storing and managing design and development data using PDM systems. This important technology is emerging as one of the leading source of competitive advantage in an era of increasing change and shrinking product life cycles. Of course the success in today's global marketplace depends on your company's ability to develop and deliver quality products in the shortest time at competitive prices. PDM systems help you meet these goals by allowing your business functions to define, control, and access information using Concurrent Engineering (CE) practices[1].

2. COGNITION OF PRODUCT DATA MANAGEMENT

As a basic concept of PDM, product data concerns all descriptions of the product and the design process, such as requirements, specifications, drawing and physical models. PDM here encloses control, planning, organization and technology concerning product information as support of complete life cycle of a product to be developed. It is characterized by mixing various activities of design process with management. Since the activity of product development is related to different functional areas within an enterprise, a real PDM system provides a basic framework to organize more than one person and possible more than one discipline to work. To represent this, a simply model of a company is used, including four functional areas: Marketing, Research & Development, Production and Service & Sales. Generally, to accomplish development of complex product, the efforts of different departments are required and thereby the main project plan concerning the product to be developed must be defined. The work objectives for engineering the component are divided among departments of the enterprise. Consequently, a macro business process arises. That is, how can each department's task be managed so that it integrates well with the results of others. Furthermore, within a certain functional area, the works are usually routed to form different work flows which are referred to certain stage or aspect of complete product life cycle. All of above activities can be divided into three levels. The relationship is shown in Fig.1. A PDM system should support three levels of management objectives, and cope with data management and process management.

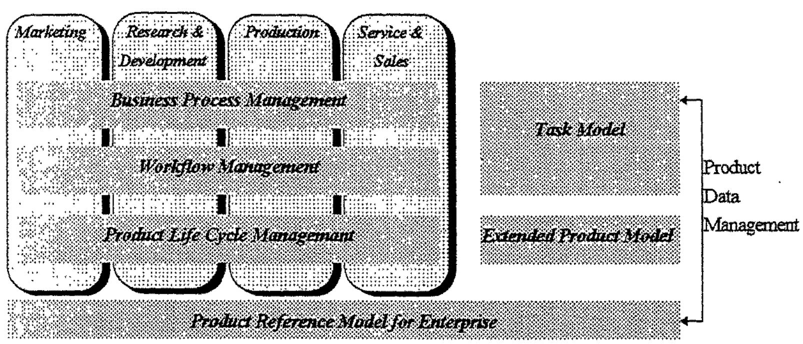


Fig. 1 : Levels and Models of Product Data Management

To understand PDM more fully, we have identified the following problems:

1. For more than one discipline is involved in the development of a product and its process, different product data views are required to meet the different member's requirements. This is worked out in section 3, where a product reference model is proposed that solves this problem.
2. During the product life cycle, product information can be stored in the different stages. This will result in product information as required, as designed, as planned, as produced, as used, as maintained and as demolished. On the other hand, the design process of a new product includes not only the final result in terms of CAD files, manufacturing instructions, *etc.*, but also information concerning intermediate design models, decision step and the overall reasoning process -- the "design history". In section 4, an extended product structure model is used to keep and access such records.

3. It is usually for multiple alternative solutions for the same design problems to be generated so that an optimal solution can be achieved. Furthermore, each step during the design process is related to personnel, organization, tools, project plan, *etc.* Section 5 addresses these issues, and it is shown how process is managed and controlled.

Evidently implementation of management objectives should structure product data and design activity. Hence, there are three models used here to fulfill a prototype of PDM system.

Implementation and conclusion are listed in section 6 and 7 separately.

3. PRODUCT REFERENCE MODEL

PDM is a vision in which networks of peoples will work collaboratively in the design. Their work will be digitally defined and distributed allowing them to work separately or together, anywhere and at any time. This network of production will includes suppliers, partners and customers in addition to the producing company's employees. For more than one discipline is involved in the development of a product and its process, different product data views are required to meet the different member's requirements. Therefore, a master product model is needed to support different applying environment including network, operation system and software, *etc.* The difficulty of data or model integration comes from the fact that product data or model has different aspects. In this paper, a product reference model is presented to support distributed team activities and multiviews during the whole product life cycle. A product reference model has four levels of functionality, i.e., Application level, Session level, Express level and Geometric level. See Fig.2(a).

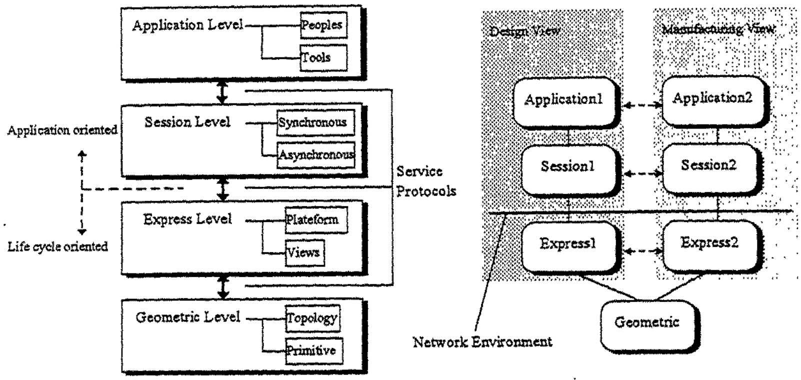


Fig. 2 : (a)Four levels of a product reference model (b) Support different views and users

Lower level provides a set of services for upper level. With geometric level, the former research on the shape form-feature is fundamental for the geometric level. Form-feature is usually defined as geometric classes with their own structure and constraints[2]. Within this level, main protocols are referred to geometric primitive and topological information. Such status may happen in other levels. Consideration of the heterogeneous platforms and domain specific features, express level is responsible to translate a view into another one. Session level maintains the communication among co-desingers and keeps work regular. Application level focuses on specific domain activity taken place in the distributed and networked environment. With the integrated usage of four levels, a product reference model can support different activities during developing a new product in a company. As shown in Fig.2(b), tow kinds of domain views are connected by reference model at geometric level and corresponding levels interact by means of services supported by lower levels. According to specific application environment and product data view, people can define corresponding application protocols. STEP standard is a good example.

4. EXTENDED PRODUCT STRUCTURE MODEL

During the product life cycle, product information can be stored in the different stages. This will result in product information as required, as designed, as planned, as produced, as used, as maintained and as demolished. On the other hand, the design process of a new product includes not only the final result in terms of CAD files, manufacturing instructions, *etc.*, but also information concerning intermediate design models, decision step and the overall reasoning process -- the "design history"[3]. A lot of product information is created, duplicated, modified and even edited during the design process. However, to keep such records is tedious and the stored information is difficult to access. It is a key issue to organize and manage the irregular information in a unified way. We have found that most of dynamic product documentation or records are related to a certain designing object such as part or assembly. It means that we can organize all of product information according to a kind of tree structure. A way used here is to employ an extended product structure model instead of product structure currently in Fig.3. The traditional product structure mainly depends on the assembly relationship of a product and usually is expressed in terms of bill of material (BOM). In order to support product development and its process, the related technology documentation is attached to corresponding product units, i.e., assemblies, subassemblies, and parts with respect to the progress of design process. The extended product structure model can provide the same function of the traditional product structure. Furthermore this model allows you to create, view and manipulate product structure and design information.

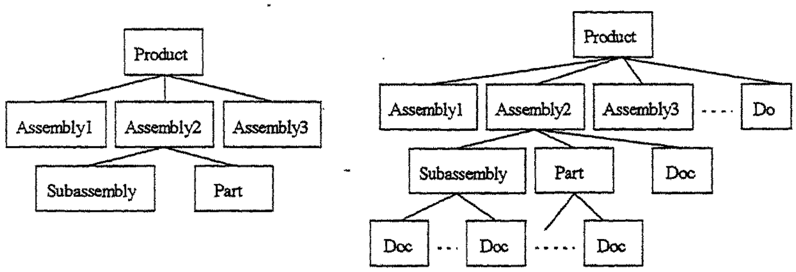


Fig. 3 : Extended Product Structure Tree

5. TASK MODEL

It is usually for multiple alternative solutions for the same design problems to be generated so that an optimal solution can be achieved. Furthermore, each step during the design process is related to personnel, organization, tools, project plan, *etc.* Generally, an "overall" design process model can be described as a process in which an abstract problem formulation in term of an "objective" is successively transformed into a manufacturable product description. The process can be divided into a number of major phase in which particular characteristics of the system are established. This includes a problem statement in term of requirements, concrete parameters, the search for alternative solutions, the selection of the "best" solution, and leads to decisions that influence later processes[4]. The design phase can further be broken into a series of more small activities. Each activity has its own tasks that reflect the decomposition of product to be developed with respect to function, objectives or constraints, *etc*. Just as the function/means tree reported by Svendsen, there exist causal relations between objectives and tasks of a design process. Hence, a design process can be represented in the objective/tasks tree. Here a task is defined as the work of a single or coherent piece of design effort generated in the distributed circumstance. Further on a task can be divided into multiple sub-task until easy to operate. The principle in the set up of the objective/tasks tree is a hierarchical arrangement of objective levels and task levels connected with lines that correspond to the casual relations. In a few words the design process is the process of task decomposition as well as design object evolution. Practically, only tasks can be considered and arranged according to the former experience of work or the progress of design. Each task has its own parameters expressed by input/output objects, feedback and exceptional jump status. The dynamic mechanism of decomposition, which allows to gather the best participants on the network according to the requirements of tasks, allocates design tasks within a team. When the design object handled by co-designers is acceptable, then a task will wake follow-up tasks, otherwise the task will activate one or more exceptional tasks. A networked task flow graph adopted here simulates the actual design process in this context. See Fig.4. This is to say that a design process can be represented with a dynamic task flow graph where the nodes denote task and subtask.

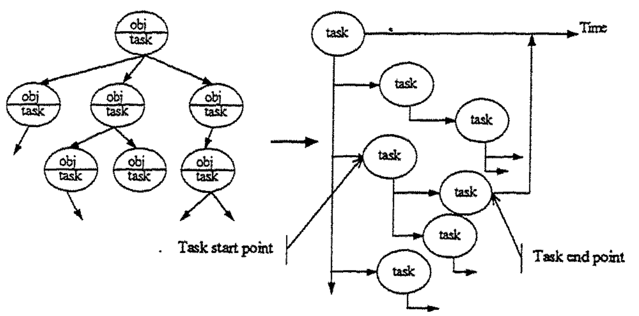


Fig. 4 : (a) Objective / Tasks Tree (b)A dynamic task & subtask

6. IMPLEMENTATION

The prototype of Product Data Management system was implemented using the object-oriented language. A snapshot of the developed system is illustrated in Fig.5. In the prototype of PDM system, mainly functions include:

1. finding and sharing information in a distributed environment
2. Integration of engineering and manufacturing systems
3. tracking design processes and design methods to reduce rework and reengineering
4. implementation electronic-based rather than paper-based engineering released and change processes
5. managing various resources used in design process

The representation of design processes in the prototype system is based on the task model. The activities are controlled and managed by a dynamic task flow graph. At the specified design stage, a user can work on a manageable design problem defined from the former pattern of problem solutions. According to the user's role and right, the system provides several entries from task browser, object browser and product structure browser. The module of product mapping rules or bases is the core of the prototype. The negotiation policy of task flow, work flow, product structure involved in design process is provided. Generally, users can define business process from Task Manager using graphic interface supported by system, and the exception results of process should be pointed out at the same time. According to the progress of the design process expressed by task flows, the finished corresponding product content is shown in the Product Structure Manager. Of course, if users define the product structure firstly, the available task flows can be provided automatically. For adopting object-oriented technology, the whole system platform is object management framework enclosed all of objects to be considered. Meanwhile, the system provides a common application programming interface to support user's developing work. ODBC module enhances the ability to connect different commercial databases.

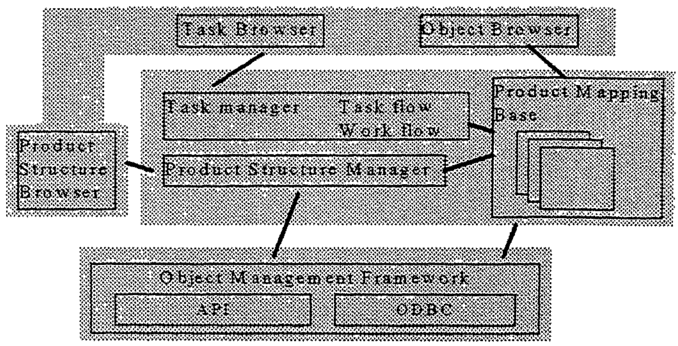


Fig. 5 : Illustration of architecture of the prototype system

7. CONCLUSION

As mentioned before, supporting the capture, storage and presentation of design information is a vital research activity in today's increasingly electronic design environment. Product Data Management (PDM) can be a virtual factor for successful solving above problems. In this paper, we point out that three levels of management objectives should be considered when constructing a Product Data Management system. Consequently, three models (i.e., task model, extended product model and product reference model for enterprise) are presented to structure the product data and design activity. Those models mainly focus on the problems of storing, accessing and sharing of product information during the complete product life cycle. Finally, we simply discuss the functions and architecture of the prototype system to illustrate how to integrate the product data and design activity.

8. REFERENCE

1. Georg Lohse, Methodical Product Development - Experience of A Practical Study, International Conference on Engineering Design'93, Proceedings, pp.200-207, Germany, Aug 17-19, 1993
2. Y.Garden and C.Minich, Feature-based models for CAD/CAM and thire limits, Computer in Industry, pp.3-13, Volume 3, 1993
3. G.Samaras, D.Spooner and M.Hardwick, Query Classification in Object-Oriented Engineering Design Systems, Computer-Aided Design, pp.127-136, Volume 26, Number 2, February, 1994
4. D.Karagiannis, F.J.Radermacher, B.Teufel and B.E.Wynne, Towards CSCW: Meta-Level Environments for Enhanced Group and Organization Effectiveness, Decision Making and Decision Support Systems, pp.132-155, Volume 4, Number 1, 1994

PROTOTYPING TOOL FOR FLEXIBLE MANUFACTURING SYSTEMS

Boudebous Dalila
Laboratoire d'Informatique du Havre
Place Robert Schuman 76610 le Havre France

ABSTRACT

Manufacturing systems are extremely difficult to design using traditional methodologies because the systems are inherently concurrent, asynchronous and nondeterministic. In this paper, we present a prototyping tool of manufacturing systems. Based on a conjoint simulation of both control system and controlled process, this tool rests on a Interpreted Petri Net specifying the control system behavior, together with an approach to the modeling of control system's environment which is easy to apply. In a design process, this tool is useful for control system's debugging, validation and dialogue to clarify end users requirements.

KEYWORDS

Rapid Prototyping, Interpreted Petri Net, Control System, Environment Modeling

1. INTRODUCTION

The design of manufacturing control systems continues to represent a crucial area of software engineering. The designer of manufacturing systems must therefore adopt appropriate methods which build on the inherent parallelism in a natural way, which help him to expose, identify and correct all design faults.

In this paper, we present a computer-aided design tool for rapid prototyping of manufacturing systems, based on off-site simulation of the control system described by Interpreted Petri Nets (IPN). In a design process, this tool can be used for : dialogue to clarify end users requirements, control system's debugging and validation, it can be used to verify that the behavior of control system is conform to the user's requirements, and that the control system have good properties such as liveness and safety.

This tool is based on two descriptions of the dynamic of manufacturing systems :

• The description of the control system's behavior is based on IPN. IPNs allow the specification of the service that control system provides to its environment by synchronizing Petri nets to an evolutive environment represented by a set of status variables . IPNs are very simple extensions of Petri nets [1].Petri nets are bipartite graphs, exceptionally well-suited for designing and modeling systems which contain concurrent processing considerations. They have been used successfully in modeling constraints of sequencing, of synchronization and of time which characterize the functioning of flexible manufacturing systems [10].

Moreover, the existence of numerous theorems and discussions make them a model which is particularly well adapted to formally guarantee the good properties such as the liveness and the safety of the system [4]. Lastly, this model lends itself very well to simulation [7] and to execution.

• The description of the environment's behavior which specifies, according to the characteristics of mechanical components and natural phenomena, the laws which govern the evolution of physical parameters. The aim of the modeling of the environment is to allow, in the course of a simulation of the control system, the automatic generation of values which modify the different entry variables and which account as faithfully as possible for the evolution of physical parameters present, without making necessary on the part of the user a specific programming. This last constraint is fundamental in context of rapid prototyping. With that aim, we rest on a generic types of event generators. Each of these types is a model which allows one to instantiate a set of event generators.

Section 2 describes the model used to specify the control system. Section 3 presents the principles of modeling of the control system's environment. In section 4 we will cover the principles of functioning of the prototyping tool.

2. THE SPECIFICATION OF THE CONTROL SYSTEM

The specification of the behavior of the system is based on Interpreted Petri Nets (IPN). A Petri net [2] [9] (see fig. 1 (a)) is a directed bipartite graph composed of two types of nodes, called places and transitions, connected by weighted directed arcs. Places are represented by circles and transitions by bars. Places can hold one or more tokens and serve as a mean of representing the status of the system being modeled. On the other hand, transitions indicate conditions for changing the system's status. According to Petri nets rules, a transition is firable only if all its input places hold a number of tokens at least equal to the weight associated with the arc connecting this place and transition. After a transition is fired, a new system status representation is obtained by removing tokens from all its input places and adding tokens to all its output places. The number of tokens removed from or added to a place is given by the weight associated with the arc that connects this place and the transition.

The behavior of a system is specified by synchronizing this Petri net transition rule to an evolutive environment represented by a set of status variables. This is done by associating conditions (characterizing environment state) and events (characterizing environment state change) with transitions, and by associating operations upon the environment with places (see fig.1 (b)). To have an unambiguous interpretation of the behavior specified using this model, the use of event occurrences in transition firing must be clearly defined : a transition is firable if its associated event e happens when its associated condition c is true and its input places are marked. This common definition is restrictive because only events happening when the associated condition c is true and the input places are marked, can be used for transition firing. This can lead to the overload of flow control for expressing new synchronization relations between transition firing and event detection in the environment. At the specification stage, it may be easier and more readable to associate with each event a communication type which characterizes this synchronization relation. So, to overcome the restrictiveness of the above definition, we consider transition as firable only if its input places are marked, its associated condition c is true, and if an occurrence of its associated event e can be used according to the communication type τ of this event (see fig. 1 (c)).

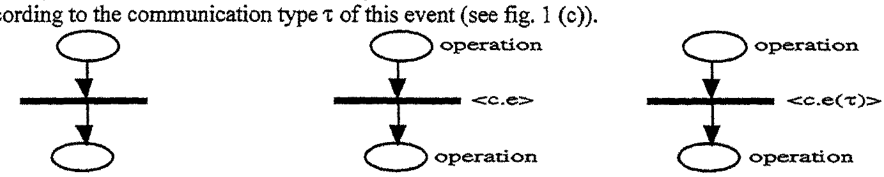


figure 1 (a) figure 1 (b) figure 1 (c)

2.1. Example

The example used to illustrate the use of IPNs in the modeling of the control system, concerns the control of a drainage pump [9].The level of water must be maintained below a certain threshold, and the supervision of the atmospheric environment, in an underground gallery of a coal mine must be ensured. For that we have at our disposal a drainage pump which can receive "on" working orders for its starting, and "stop" for its stopping, a level sensor which detects "high" and "low" thresholds of water in the gallery, and methane, carbon monoxide, and of aeration sensors which allow at any one time to know the levels of these gases. The control of the system is ensured on the surface by an operator who can put it into service (order "on") or off (order "stop"). Besides this, the operator must have the possibility of knowing the state of the pump and the gases by simple request, and must be informed by alarm signals of any exceeding of the threshold of these gases. The pump must work automatically once it has been started by the operator, according to the level of water in the gallery : starting up when the level of water reaches the "high" threshold, and stopping when the level of water goes back down to the "low" threshold. However, for safety reasons, the pump must not be started or continue running when the percentage of methane in the atmosphere exceeds a safety limit.

To simplify our presentation we will restrict ourselves to the net of figure 2 which specifies for the control part, how the starting up (CMD pump on) and the stopping (CMD pump stop) of the pump must alternate depending upon the evolution of two parameters ; the level of water (Lwater) and the level of methane (Lmethane) in the gallery.

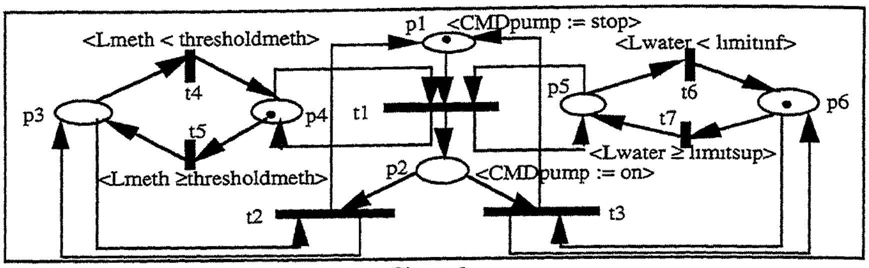


Figure 2

3. THE SPECIFICATION OF THE ENVIRONMENT

In the environment of a control system the events which characterize the evolution of the physical parameters can be regrouped into two categories : the endogenous events of which the occurrences depend on orders generated by the control system, and the exogenous events of which the occurrences can be considered to be independent of orders generated by the control system. The taking into account of these events by a control system is carried out by detecting the passage of a physical parameter by its nominal position, or by observing the current position of a physical parameter. In the first case, the acquisition takes place with the help of all or nothing sensors in the binary entry variables, and in the second case with the help of numerical sensors in the continuous entry variables. In the case of a binary entry variable, if the passage to nominal position leads to an unstable state which lasts for only a very short time, we will say that it is a fugitive binary variable. Inversely, if the passage to the nominal position leads to a stable state which lasts for a relatively long time, we will say that it is a non-fugitive binary variable.

3.1. Presentation of generic types of generators of events

3.1.1. The generators of exogenous fugitive binary events

A generator of exogenous fugitive binary events allows one to generate, according to a fixed or aleatory periodicity, events which force a fugitive binary variable to its significant value for the time of a simulations cycle. After the simulation cycle the variable in question regains its non significative value. A generator of exogenous fugitive binary events is instantiated by specifying the following parameters : the name of an entry variable of the control system, the significative value of this variable, and the law of distribution of the interval of the generation of events.

3.1.2. The generators of exogenous non-fugitive binary events

A generator of exogenous non-fugitive binary events allows one to generate, according to a fixed periodicity, events which force a non-fugitive binary variable to true according to a probability p and to false according to a probability 1-p. It is instantiated by specifying the following parameters: the name of an entry variable of the control system, the frequency of generation of events modifying this variable, and the probability with which the events force the variable to true or false.

3.1.3. The generators of exogenous continuous events

A generator of exogenous continuous events allows one to generate, according to a fixed periodicity, events which make the contents of a continuous variable vary by Δx which can be fixed or aleatory. It is instantiated by specifying the following parameters : the name of an entry variable of the control system, the frequency of generation of events modifying this variable, and the law of distribution of the Δx corresponding to variations of this variable.

3.1.4. The generators of endogenous fugitive binary events

Contrary of generators of exogenous events, the generators of endogenous events are subjected to conditions of generation which concern the condition of the control system. A generator of endogenous fugitive binary events allows one, each time that the condition of generation to which it is associated becomes true, to generate with a probability P an event of modification of fugitive binary variable. The variable v, to be modified, can be chosen with a probability p amongst n variables. The sum of the probabilities p, associated to n variables, must then be equal to 1. Each event which is generated forces a variable to its significative value at the end of a time Δt which can be fixed or aleatory. After a simulation cycle, the modified variable regains its non significative value. A generator of endogenous fugitive binary events is instantiated by specifying the condition in which the control system must be in order that an event may be generated (this is the condition of generation) the probability p with which an event is generated each time that the system reaches this condition, and for each entry variable of the control system being able to be modified : its name, its significative value, the probability of a generated event concerning it, and the law of distribution of time Δt at the end of which the generated event provokes the modification of the variables.

3.1.5. The generators of endogenous non-fugitive binary events

The generators of endogenous non-fugitive binary events differentiate themselves from the generators of endogenous fugitive binary events by the fact that the modifications on the entry variables which are generated by events persist beyond the simulation cycle. In other terms, the modified variables retain their new values after the simulation cycle. A generators of endogenous non-fugitive binary events is instantiated by specifying the condition in which the control system must find itself in order that an event might be generated (this is the condition of generation)the probability p with which an event is generated each time that the system reaches this condition, and for each entry variable of the control system being able to be modified : its name, the value which it takes, the probability of a generated event concerning it, and the law of distribution of time Δt at the end of which the generated event provokes the modification.

3.1.6. The generators of endogenous continuous events

The generators of endogenous continuous events are characterized on the one hand by a condition of activation and on the other hand by a condition of interruption of the generation, which have a bearing on the condition of the control system. Each time that the condition of activation to which it is associated becomes true, a generator of endogenous continuous events generates, according to a fixed periodicity, events which make the contents of a continuous variable of a fixed or aleatory Δx vary, until the condition of interruption becomes true. The contents of the entry variable can be forced at the beginning of each cycle of generation to a particular value. A generator of endogenous continuous events is instantiated by specifying : the name of an entry variable of the control system, the condition in which the control system must be found in order that the generation starts off (this is the condition of activation of the generation) the condition in which the control system must be found for the generation to stop (this is the condition of stopping of the generation), the initial value of the variable (optional) the frequency of the generation of events, and the law of distribution of the Δx corresponding to the variations of the variable.

3.1.7. The releasers of purges

A purge release allows, each time that a given event is generated, to provoke the purging of all the events which its occurrence invalidates. The date forecast for the occurrence of an event which is the subject of a purging must be superior to the date forecast for the occurrence of the event which provokes the purging. A releaser of purges is instantiated by specifying the following parameters : the name of the entry variable whose events of modification provoke the purging, and the list of entry variables whose events of modification must be purged.

3.2. Example

The example used to illustrate the use of generators of events in the modeling of the environment of a control system, concerns the control of a drainage pump (cf. 2.). The evolution of the "level of methane" parameter in the gallery is not directly linked to the orders generated by the control, although inversely the evolution of the level of water depends upon the functioning of the pump. Depending on whether the pump is in stop or working position, the level of the water in the underground gallery increases or decreases. The events of modification of the variable "methane

level" which report on the evolution of methane, can by this fact, be generated by a generator of exogenous continuous events, instantiated with the following parameters :
- name of the entry variable : "Lmeth" (methane level)
- frequency of generation of events of modification of the variable : every 5 units of time
- law of distribution of Δx corresponding to the variation of methane levels every 5 units of time : distributed evenly in -0.2 and +0.2.

As for the events of modification of the entry variable "Lwater" which serve to understand the evolution of the level of water, they can be generated by a generator of exogenous continuous events and by a generator of endogenous continuous events instantiated by the following method :

First generator :
- name of entry variable : Lwater (water level)
- frequency of generation of events of modification of the variable : every 5 units of time
- law of distribution of Δx corresponding to the increase in water level every 5 units of time : constant variation of +0.1

Second generator :
- name of the entry variable : Lwater
- condition of activation of the generation of events : M(p2) = 1
- condition of interruption of the generation of events : M(p2) = 0
- frequency of the generation of events of modification of the variable : every 5 units of time
- law of distribution of the Δx corresponding to the lowering of the water level every 5 units of time : constant variation of -0.45.

The first generator permanently generates events which report on the increase in water level, and the second generator generates those events which report on the lowering of the water level when the pump is in working state.

4. THE PROTOTYPING TOOL

The objectives of the rapid prototyping tool are presented first, then the implementation of this tool is described.

4.1. Objectives

It is difficult to develop a system well the first time, as users are unclear about their needs and often cannot communicate their wishes to the designer. For that raison the designer used this tool for the dialogue with the user. After a brief meeting with the user describes the control system with IPNs and the controlled system with a generic types of event generators, and presents working model to the user. After the user examines the prototype of his system and asks for improvements, the designer redevelops the system and presents another prototype to the user ...until the user is satisfied.

The second objective of this tool is the verification and the validation of the system. The use of Petri net for modeling the control system allows a detailed analysis to detect a large number of design errors before any implementation. Firstly, we verify that the Petri net representing the control system have some good properties (absence of deadlock, boundedness property, invariants of places and transitions). This first analysis is complemented by another one during the simulation of the system, if the first analysis (analysis of Petri net without interpretation) is not conclusive. An error detected at this level means that the system is not correct. For example a deadlock indicates the existence of some control sequences that can never be executed (if a deadlock occur during the network execution then the system is not live). Linear invariants of places and transitions are verified during the network execution. These invariants are useful to prove that some required specification constraints are implied by them. For example to prove the absence of contradictory commands or of dangerous control sequences.

4.2. Implementation

The prototyping tool is implemented in Pascal language [6]. It is composed of three parts : a graphic editor of IPNs, an interactive textual editor used for the description of the environment an player of IPNs. The graphic editor allows first the construction of the Petri net control structure and produce in memory tables representing the Petri net. Next the construction of the interpretation attached to the Petri net. The interpretation is set of structured declarations of variables (input

corresponding to sensors, output corresponding to actuators, variables of calculation and of temporization), conditions and events attached to the transitions and operations associated to the places. These declarations are analyzed and translated into tables representing the variables, and an internal form of the operations, conditions and events using pilosh notation.
The textual editor proposes to the designer a set of menus to display information about input variables, instantiate, save, restore, visualize, modify diagrams for the generation of events.
The player of IPNs relies on an events approach, it uses for this, a scheduler in which are stored the events coming from the external environment. Each time that a new event is taken into account the player causes the evolution of the net describing the behavior of the control of a stable state to another necessitating the taking into account of the following event. The events which depend on the order state (endogenous events) are generated throughout the research of this stable state. As far as the events independent of the order are concerned (exogenous events), an occurrence of event is generated according to a law of arrival following the taking into account of the occurrence of the preceding event. Operations associated to places marked are executed. During the simulation of the IPN modeling the system, the designer can observe the behavior of the control system (the values of input and output, the marking of the net),verify the properties of the system, modify values of variables or the marking of places of the net, insert or delete events from the scheduler to simulate dangerous situations, and re-start the simulation. It is particularly useful for the debugging and the dialogue between designers and users.

5. CONCLUSION

The design and implementation of manufacturing control systems must use tools for system specification, analysis and implementation. In this paper, we have presented a rapid prototyping tool founded on an adequate modeling of the control system and the environment. This tool is based on results from the theory of Petri nets. It is very easy to use, in a design process, this tool is useful for control system's debugging, validation and dialogue with the user so as to clarify his needs.

6. REFERENCES

[1] G. W. Brams
Reseau de Petri : Theorie at partique. Edition Masson 1983.
[2] M. Moalla
Réseaux de Petri interprétés et Grafcet TSI vol. 4, n° 1, 1985, pp 17-30.
[3] C.V. Ramamoorthy and Gary S. Ho
Performance evaluation of asynchronous concurrent system using Petri nets
IEEE Transactions on software engineering Vol. SE-6 N° 5 pp. 440-449 September 1980.
[4] REG G. Willson and B. H. Krogh
Petri net tools for the specification and analysis of discrete controllers
IEEE Transactions on software engineering Vol. 16 N° 1 pp. 39-50 January 1990.
[5] T. Bennani
Flexible production systems modelling, International conference on industrial engineering and production management Marrakech April 4-7, 1995.
[6] D. Boudebous and J.C. Derniame
"Method for a design of distributed control systems". COMPEURO 93 Computers In design, Manufacturing and Production. May 24-27 1993 Paris.
[7] Chiola G
A simulation framework for timed and stochastic Petri nets MASI, Octobre 1996
[8] E. Gressier
Astochastic Petri net for ethernet MASI, N°58, Juin 1995.
[9] Lister A. and J. Magee and M. Sloman and J. Kramer
Distributed process control systems : programming and configuration
RR n° 80/12 Imperial College, London 1980.
[10] J. Tankoano and J. C. Derniame
Structured design of distributed sytems using interpreted Petri nets COMPSAC 1989.
[11] R.Valette
Nets in production systems LNCS 225, pp.191-217 Springer-Verlag, 1986.

REFERENCE MODELS AS AN APPROACH FOR THE CONSTRUCTION OF FLEXIBLE PRODUCTION PLANNING AND CONTROL SYSTEMS

W. Dangelmaier, A. Kuhn, T. Langemann
Heinz Nixdorf Institute, University of Paderborn
Fuerstenallee 11, 33102 Paderborn
{whdlkuhnacltimo}@hni.uni-paderborn.de
Germany

ABSTRACT

In this paper we will present an approach for the construction of flexible Production Planning and Control (PPC-) Systems. Our approach is based on a method for modeling static and dynamic aspects of such systems. The created models can be executed by an object oriented framework as a PPC-System. In order to reduce time and costs for building systems which are best fitted to the processes within the organization, we are using so called reference models [1]. Cause of their high degree of reusability and flexibility these reference models are going much further than the usual parametrization and customization within standard software. This allows a relatively short development period and hence less expensive PPC systems which can additionally be adjusted to individual requirements.

KEYWORDS

Production Planing and Control Systems, Reference Models, Model-based System, MRPII

1. INTRODUCTION

Enterprises today see themselves opposed to new and manifold challenges. Especially shortening product-life-cycles and a continuously increasing dominance of purchasers are to be mentioned. Time-to-market is becoming a more and more critical factor of success [2]. In order to face these challenges there is not only a need for improved manufacturing technology but for new approaches to production planning and controls systems as well. As a solution for the changed challenges mentioned above rapid product development, flexible manufacturing systems, the fractal plant e.g. are often suggested. A sole optimization of construction and manufacturing processes or technology can't looked upon as sufficient. There is an additional need for an improvement of the organizational environment of the manufacturing system. This means the improvement of production planning and control processes.

Decentral information systems, that can be adapted to permanently changing company structures, signify a particular challenge in our days. PPC software in particular is marked by complex interdependencies and the difficulty of integrating economic and technical information systems[3]. Great importance has to be assigned to the optimal embedding of the PPC software in company specific PPC tasks. As processes, tasks and objectives of PPC are very distinct for respective companies the use of inflexible standard software may cause problems. This leads to frequent adaptation of the organization to the software and thus not inevitably to an optimal solution. In the best case a company can reach the performance of a competitor using the same PPC software. If on the other hand a company elects a special solution covering its own needs, it has to accept long development times and high costs for made-to-measure systems.

One way to build more flexible PPC systems consists in a first step of modeling the manufacturing processes and in a second one of executing the model as a PPC-system. By the use of reference models a new dimension of flexibility in the process of PPC systems' construction can be reached, at the same time reducing the efforts to adapt the systems. Using this concept each company should be able, with a maximum of flexibility relative to changing situations and the possibility of adapting to changing organizational circumstances, to model its individual manufacturing process and derive functions and data for simulation, planning, scheduling and monitoring. Because of the complexity of the resulting models and because the running production planning and control system shall be automatically derived from the developed model the modeling process must be supported by appropriate software tools.

2. A MODELING APPROACH FOR THE CONSTRUCTION OF PPC-SYSTEMS

2.1 Basics of the Approach

This chapter briefly describes the underlying modeling method MFERT[1]. It covers only the main elements of the method which are necessary to understand the described software tools. For more detailed information see [4]. In order to apply reference models for planning and controlling manufacturing processes a framework is necessary. The proposed framework comprises a modeling method, which provides a process chain based representation of the manufacturing process using a predefined set of construction elements. For the representation of PPC relevant elements of the production process so-called *elements_in_state* are defined. They represent either information, material or resources. The single steps of the manufacturing process are described by *operations* of elements and their attributes. The main elements of the method are:

- *element_in_state*: An element_in_state is an element in a certain state. An element is a static representation of a part of the universe of discourse (i.e. material, resource or information). It is characterized by its attributes.
- *operation*: An operation is a transformation of elements from one state into another. It is characterized by its input and output elements and by the relative time of the inputs and the outputs to a reference element.

Both model elements can be grouped together into categories which contain *elements_in_state* or *operations* of the same class. These categories can be hierarchically organized and are called *element_in_state_category* and *operation_category* respectively. This is based on the object oriented paradigm for data modeling. The model elements *element_in_state* and *operation* cover all informations that describe the production process but they do not cover its structure. To describe this structure a graph based model is used. In this graph there are two kinds of nodes corresponding to the two kinds of data:

- *element_in_state_category_node*: A node of this kind represents elements in state of a certain category and its subcategories. Graphically it is represented by a triangle (△).
- *operation_category_node*: Nodes of this kind represent operations of a certain category and its subcategories. Graphically it is represented by a rectangle (□).

These two types of nodes can be linked together by *directed edges* representing the flow of material in the production process (△–□). Only nodes of different kind may be connected by an edge. The result is a bipartite, directed *graph* which represents a model of the production process or at least a part of it. The *functionality* of the system is realized by allocating functionality to each node of the model. A model of the enterprise can consist of several models of different levels of abstraction. They can be linked together by interfaces representing the flow of data between the levels of abstraction. The interfaces of the lowest level can be linked to the reality[2] or to simulation modules representing the reality. Figure 1 depicts a small example of a rear axle production process which was modeled in a common project with a big german automotive company.

Based on this modeling approach, the object-oriented system OOPUS[3] [4,5] has been developed. It enables a model-oriented construction of manufacturing control systems. The model structure is built by pre-defined modules. From a method repository functionality has to be assigned to them. The

1. MFERT is a german acronym and stands for „modeling of manufacturing systems"

2. I.e to terminals that give information to the production planner or directly to machines on the workshop. And on the other hand to terminals or sensors that provide information of the progress of the production.

3. OOPUS is another german acronym and stands for „Object oriented framework for the generation and integration of individual production planing and control systems"

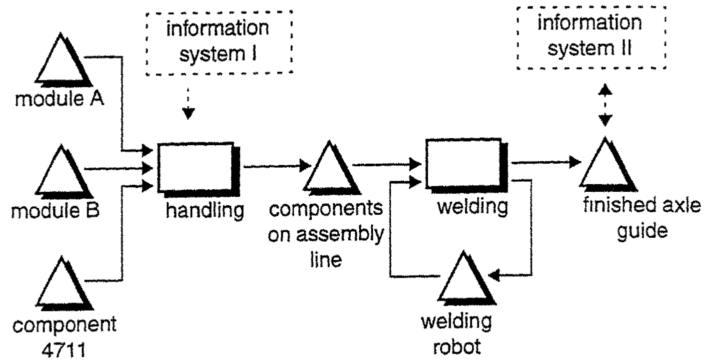


Fig. 1: Partial model of the rear axle production with cojoint information systems.

running system itself is executed decentralized: functionality of partial structures of the model is assigned to agents in the network. This approach allows the scaling of the PPC system. The OOPUS system serves as an integrated framework for the purposeful ***modeling, simulation and generation of PPC-systems*** (see figure 2).

2.2 Reference Models within the Approach

The essential strategic potential of the modeling method explained above can only be fully exhausted when predefined construction sets of elements are provided which include already basic and generic knowledge about production tasks and PPC tasks as well. As the wheel has not to be invented again and again, it is ensured that the creation of company specific PPC systems will result in reduced expenditure compared to a normal individual programming of PPC systems.

We describe such sets of elements as reference models that can be characterized as follows: in reference models different characteristic features of companies, e.g. product structure and manufacturing organization, as well as different types of manufacturing processes such as making, moving, transporting or verifying a product, are taken into account. One of the objectives of applying reference models is the efficient description of the production task (e.g. describing the flow of material) by means of so-called structure models. These structure models for productions tasks have to be completed with company individual restrictions (e.g. calendar of a company, restrictions on quantity). Structure models establish the basis for describing and implementing production planning and control tasks (PPC tasks) with MFERT/OOPUS for the respective production tasks. Figure 2 shows in a schema the proposed concept to create PPC systems with the application of reference models.

The hierarchic modeling approach allows the consideration of different task levels (Material Requirements Planning, Capacity Requirements Planning, Shop Floor Scheduling and Control. By limitation to the main characterizing attributes the approach supports considerably the reduction of complexity. Thus our understanding of reference models goes above the common interpretation of reference models as a rough framework [1].

2.3 Conceptual Aspects and Modeling Examples

In the following chapter a few examples of reference models are depicted. These refer as well to the description of production tasks as to the description of PPC tasks. The characteristic feature ‚product structure' describes a product by specifying its constructive design. This already forms a view of the production task even if the required resources for the manufacturing process and some states of the product are neglected. The constructive design however is of essential importance for the creation of

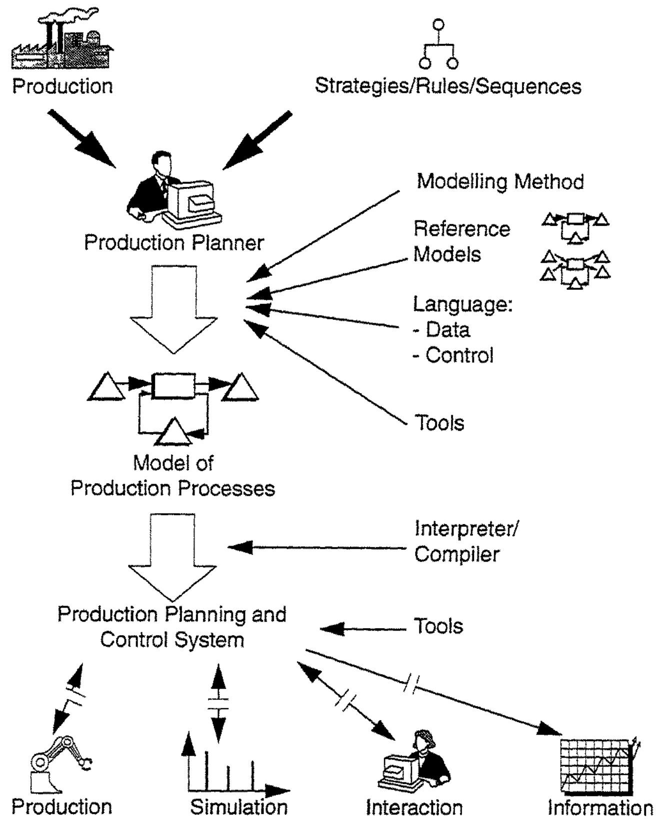


Fig. 2: From production strategies to PPC systems.

models for material requirement planning as at this level statements about demanded quantities of parts within this structure can already be deduced.

With the two basic structures in figure 3 the product structure of one product can be described. This kind of models exceeds mere bill of materials (BOM), as durations are assigned to the manufacturing transformation nodes. Together with the functionality assigned to all the modeling elements a time schedule can be calculated in an early phase of planning the production. In order to model the multi use of parts or products, manufacturing element nodes with several output paths are required. With this enhancement any desired product structures can be modeled in one single (net) model thus being able to consider dependencies among parts used in different products.

The above models may provide a first production schedule and quantities of demanded products resp. parts. In order to schedule the capacity requirements the models have to be extended by representations of resources and operations.

Within capacity requirement planning characteristic features of the production task can be used to deduce typical reference models. For the manufacturing types ‚series production' a basic sequence of manufacturing operations can be defined (e.g. turning on a lathe - milling - drilling - grinding), that is nearly identical for the respective products to machine [6]. Exceptions from this basic manufacturing sequence result on the one hand from leaving out certain manufacturing operations and on the other hand from repeating one or several operations. Has a basic construction element, see figure 4, been defined once, any exceptions concerning the sequence of manufacturing operations in the respective

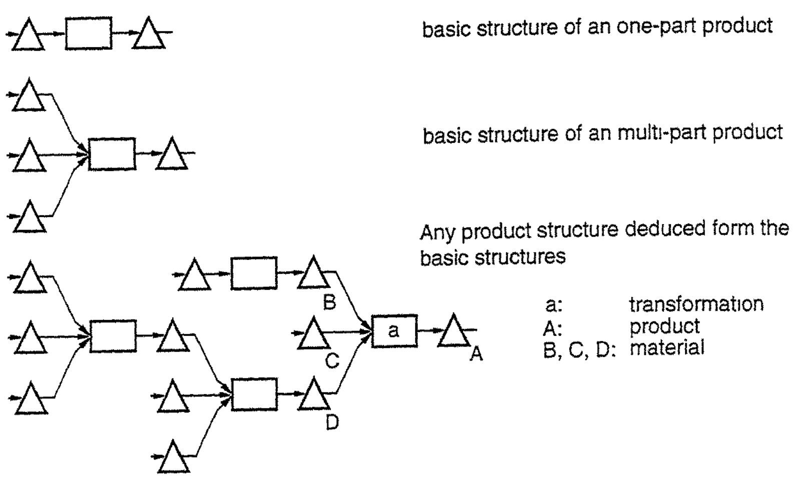


Fig. 3: Construction set for product structures.

context can be modeled (reusability, time saving). Therefore the model, depicted in figure 4 can be looked upon as reference model for series manufacturing, as far as the structure of the model is concerned. As already mentioned above, the respective functionality has to be added (e.g. implementation of priority rules). For further types of manufacturing organizations reference models can be defined as well.

The hierarchic modeling approach allows the specification of manufacturing processes at any detailed level. The level of specification is determined by the respective PPC task. For example reference models can be defined for different kinds of transportation processes or set-up processes (machine tool). The required planning and control functionality is assigned to each node within the reference model. In the case of material requirement planning each node calculates e.g. gross demand, net demand, possibly lot sizes (e.g. in case of series production), secondary demand, scrap allowance and lead time shifts. Nodes at the level of capacity requirement planning own for example priority rules to calculate order sequences and functions to manage capacity demands and free capacity.

The proposed concept allows the realization of all known planning and controlling philosophies (e.g. JIT or OPT). Within the different levels of PPC tasks the respective reference models can be combined according to the companies production tasks and production objectives. The different levels are linked via bidirectional communication interfaces hence realizing control loops in both directions. Interfaces between the model and the ‚real manufacture environment‘ are necessary as well to ensure via model updating the consistency of the PPC model states with the manufacturing reality.

3. CONCLUSIONS

In this paper we have presented an approach for a model based construction of PPC systems with reference models. So PPC systems are more oriented on the individual needs of industrial companies by the simultaneous consideration of strategical demands such as the flexible adaptation to changing market environments. At our institute the proposed method is available in a prototypical tool system for generating company specific PPC systems. Currently the implementation of this prototype is prepared at a well known german automotive concern. A PPC system generated with our tool is installed at our institute and provides good computing results.

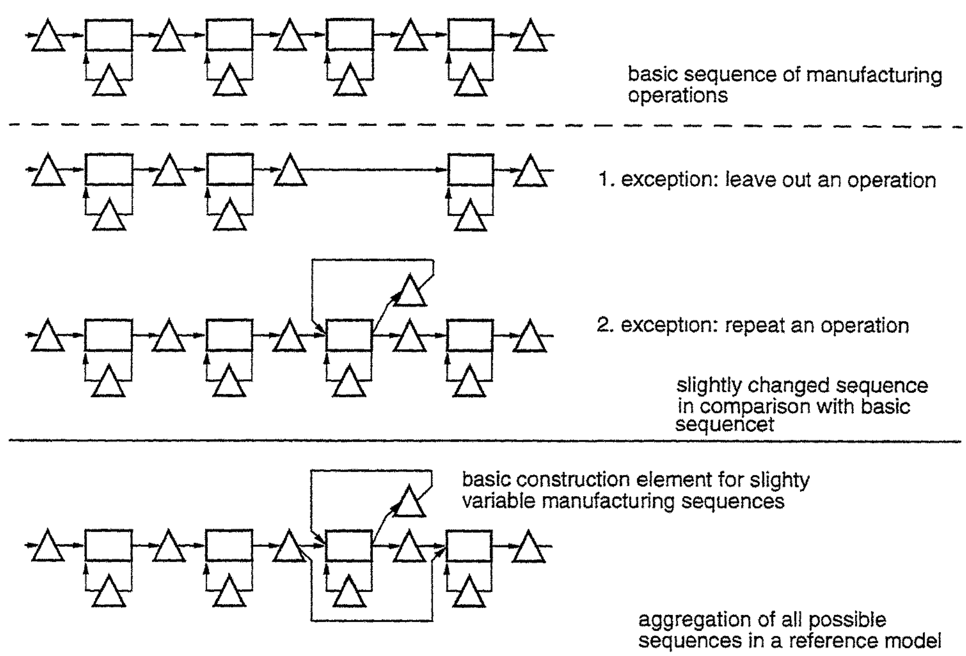


Fig. 4: Reference model for manufacturing types with slightly variable sequencing.

4. REFERENCES

1. Mertins, K.;Rabe, M.;Könner, S.: Reference models for simulation in the planning of factories, IPK Berlin (Ed.), 1995

2. Burmann, R., Manufacturing Management: Principles and Systems, Mc Graw-Hill, Berkshire, UK, 1995

3. Dangelmaier, W.;Langemann, T.;Wenski, R., "Model-oriented Construction of Manufacturing Management Information Systems", In: CESA`96: Symposium on Discrete Events and Manufacturing Systems, IEEE, pp. 346-351, Lille - France, 1996

4. Dangelmaier, W., "Distributed production planning and control in adaptive manufacturing processes - THE MANDATE APPROACH.", In: Improving Manufacturing Performance in a Distributed Enterprise, Advanced Systems and tools, ESPRIT Working Group 9245, pp. 63-70. Edinburgh, 1995.

5. Dangelmaier, W.; Felser, W.; Henkel, S.; Holtkamp, R., "OOPUS - a distributed system for modelling, simulation and production planning and control." In: Improving Manufacturing Performance in a Distributed Enterprise, Advanced Systems and tools, ESPRIT Working Group 9245, pp. 79-88. Edinburgh, 1995.

6. Gaither, N., Production and Operations Management, Fifth Edition, The Dryden Press, Fort Worth et al., 1992

A TASK-CENTRED METHODOLOGY TO SUPPORT AN INTEGRATED AND OPEN COMPUTER AIDED MANUFACTURING SYSTEMS DESIGN ENVIRONMENT

B. WU and R.S. Hull
School of Industrial and Manufacturing Science
Cranfield University
United Kingdom

ABSTRACT

This paper presents the current state of development of an integrated and open computer-aided manufacturing systems design environment to support decision making in manufacturing systems reorganisation and the adoption of new techniques and technologies. In particular it discusses a Task-Centred methodology and its support within the environment. The aim is to provide the designer with a number of integrated computer tools, a systematic approach to manufacturing systems design and an integrated data model. All the elements required to support the designers in executing a particular design task are compiled as a single entity within a computer-based document. The designers can then focus their attention on the tasks and the creation of the task deliverables. A key element is the provision of an open and flexible design framework in which the designers can construct a design methodology to suit the specific requirements of their project as well as their particular needs and preferences.

KEYWORDS

Manufacturing Systems Design, Methodology, Task-Centred Approach

1. INTRODUCTION

The changing competitive global market and its effect on manufacturing enterprises has been widely reported. Reduced production volumes, reduced product life cycles, increased product variety are increasingly becoming the norm for manufacturers. There is therefore a need to reduce product development lead-times, to reduce the costs incurred in such development and to recover these costs over shorter time periods. Rapid product development techniques are seen as one of the means of responding to the changing market conditions and reducing costs. A complimentary approach is to take a systems-wide view of the company, increasing integration and communication between business functions and adopting world class manufacturing and concurrent engineering techniques. Manufacturing systems re-design and business process re-engineering are being advocated as ways of introducing the necessary step changes and of providing foundations for continuous improvement. This paper proposes a computer based environment to support manufacturing systems design projects through the provision of simple tools and techniques and the delivery of on-line methodologies, which can be customised to the particular needs of the design team. This is achieved through the application of a systems design framework, whose fundamental elements are self-contained design tasks.

The paper is divided into four parts. Section two describes the problem domain and outlines why an integrated design environment is required. Section three discusses the principles and concepts behind the task-centred approach and section four describes the task-centred approach within an integrated and open computer aided manufacturing systems design (I/O-CAMSD) environment, applying the concepts of the I/O-CAMSD framework and its constituent elements and outlining the application of a generic framework to construct a specific design approach. Finally, applications and future directions of research, particularly with respect to concurrent engineering, are discussed.

2. THE PROBLEM DOMAIN

The adoption of new technologies and techniques, such as CIM and concurrent engineering, and the restructuring of the organisation and its manufacturing systems can be greatly assisted through the application of methodologies and structured approaches [1]. To reflect the true complexity of the problem of transforming the business Heim and Compton [2] suggest that a holistic approach is required. This will also be complex and hence Kidd [3] proposes that a methodology would need to be supported by a computer in order to be practicable and usable. However, the use of design methodologies is limited in practice, particularly in small to medium sized companies in the United Kingdom [4,5]. The available methodologies tend to be limited to specific manufacturing sub-systems addressed, often not considering the entire factory and manufacturing system in adequate detail. The result is that the methodologies can be considered to be too generic, due to their often high and simplistic level of abstraction of detail, and at the same time too specific in their limited coverage of systems design issues.

Similarly, there has been a limited acceptance of computer tools, such as simulation, clustering algorithms etc., to assist the process of manufacturing systems design (MSD) within such companies [5]. Indeed, the application of computer tools in the past has required experts, their use has been time consuming and has often required the duplicated collection and input of data when several tools are employed. The current computer-aided approaches address data-tool integration but tend not to consider a systems-wide perspective, the design process itself or individual design task related issues. They can therefore be referred to as tool-centred. An integrated design environment is required which will support a number of design tools, ranging from high level problem solving to low level analytical tools, together with a structured approach or methodology. However, the design approach itself needs to be flexible and adaptable to the needs of the individual users, primarily because every MSD case is different. A task-centred approach is therefore proposed as a means of escaping the tool-centred nature of computer environments and of providing the required flexibility.

3. THE TASK-CENTRED APPROACH

The principle behind the task-centred approach concerns the provision of a methodology to assist users in solving problems and to support decision making. An advantage of the task-centred approach, when its functionality with respect to providing design management and methodological support is taken into account, is given by Ehrlenspiel and John [6] who comment on design methodology and design management. They see design methodology as introducing order into the design process, creating a framework in which intuitive processes occur. And hence, they suggest the term *'method aided design'* or *MAD* to describe the process of working in a design framework methodology, that is working intuitively with methodic support.

At an application level, the distinguishing feature of a computer-aided task-centred approach is a common task configuration interface throughout the design cycle. The result, within the context of manufacturing systems design, is that the task descriptions, instructions, detailed help, processes, tools and data of any specific task can be assembled into a single working package, a MSD task document [7]. Such an approach allows the designers to focus their attention on the design task itself and the end deliverables rather than on the application of a specific tool. Indeed, it has been observed that a focus on task definition provides a degree of rigour that has previously been lacking [8], particularly with respect to computer supported MSD. Completely unstructured approaches to MSD depend entirely on the knowledge, experience and insight of the designers. Whilst it is theoretically possible to take into account all possible or relevant objectives and constraints, and hence provide a greater degree of breadth in the design process, such an approach would not provide the necessary degree of rigour [9].

The task-centred approach can be considered to be more natural for the designers, since it is directed towards the attainment of tangible deliverables, and to be easier to manage and plan through the application of conventional project management techniques. The outcome of a design task should enrich a design decision, with the task document providing the opportunity to evaluate a number of scenarios. In support of this concept, an overview of the conceptual architecture of a MSD task document and of the requirements of a computer based design environment have previously been presented [10,11].

4. THE I/O-CAMSD FRAMEWORK

4.1 The Framework Concept

Following a review of industrial practice, available methodologies, frameworks and reference architectures, the I/O-CAMSD framework was developed. The framework essentially supports a structured mechanism for the provision and execution of design methodologies and the communication of systems designs. It is structured to support the principles of problem solving and systems thinking, providing a series of 'building blocks' with which to undertake a MSD project. The framework is comprised of a four stage approach to MSD: project initiation, requirements specification, conceptual modelling and detailed design. At each stage, a number of systems architectures or subsystems can be addressed (see figure 1). Three principle architectures are used to describe the system: Manufacturing, Human and Organisation, and Information and Control architectures. The framework can then be sub-divided into a number of cells or elements designated as Task Frames [11]. Within such a task frame, which can be considered to represent a self-contained package of work, a collection of design tasks exists which addresses a specific subsystem at a particular stage in the design cycle. Hence, it is within these generic frames that sub-problems are solved and a design concept developed.

This concept then acts as a conceptual constraint for task frames of similar or different architectures at the same level in the design process. Design iterations within and between the task frames can be executed, resulting in a consolidation of the concepts developed at the particular design process level. The consolidated designs are then passed down as a concept proposal for the architectures or subsystems of the next stage in the design process. It is through the application of task frames and their associated design tasks that a modular design approach to MSD can be adopted, in terms of the design process and the eventual manufacturing system itself. This modularisation of the design process permits designers to easily generate and develop design ideas and scenarios and to evaluate their impact in order to enrich the decision making process. The complete framework provides a generic approach to MSD. It is through the specification of a particular design approach, selecting appropriate task frames and design tasks and refining and customising their content, that a specific design approach can be defined.

The framework provides an overall structure to the task-centred methodology, with the task frames grouping related design tasks in a coherent and logical manner. It ensures that the design process itself is executed in a coherent and consistent manner. The design task and task frame selection, which forms the basis of the methodology, is derived primarily from the project terms of reference, as defined in the project initiation stage. Where a manufacturing strategy is available for analysis, suggested links and guidelines to appropriate task frames and design tasks, based on the content of the strategy, are provided. The approach derived from the strategy is subject to user validation and guidance is provided in order to assist the users in customising the design process.

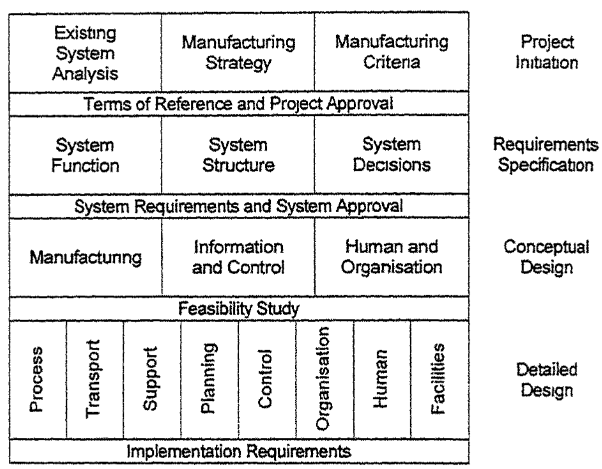


Fig. 1 : The I/O-CAMSD design framework

4.2 The Framework Implementation

The open and integrated nature of the task-centred methodology is best presented through the metaphors used in the prototype implementation of the framework. The computer is being applied in two different ways to support the I/O-CAMSD framework. Firstly, it is being used to provide the tools and data integration and handling required. Secondly, it is being used to deliver the design methodology and associated tools and techniques. These two elements are achieved through the use of Object Linking and Embedding (OLE), Microsoft Visual Basic, and generic system objects which support OLE automation, including the Microsoft suite of office tools. Not only does this provide a rapid development environment for the framework implementation, but it also allows linking to the Microsoft Windows operating system and other business applications. It also has a potentially large user base in manufacturing industries, particularly within small to medium sized enterprises to which the research is primarily targeted.

The computer-based implementation uses three principle constructs to apply the framework. The first is that of a bookshelf which manages the projects. It is here that past design projects and pre-defined projects are available for viewing and customising. The second level, the book level, contains complete MSD projects, comprising the task frames and design tasks.

The design task itself is comprised of three pages. The first page provides an overview and a basic view on the project management status and location of the task within the overall framework. The second page (see figure 2) provides the working area, a guide through the activities within the task and context sensitive help. It guides the user through each of the task requirements, including defining task criteria, selecting task options and setting appropriate design and evaluation techniques, such as specifying a particular algorithm. The final page displays the current working object and results of the task achieved so far. It is within these last two pages and the associated tools that the task deliverables are created. The basic structure is applicable to any general task carried out within the bookshelf. The current task working object is the actual software package being used to apply the MSD tool. In the case of a clustering algorithm, the package is a standard spreadsheet containing the algorithm as a macro.

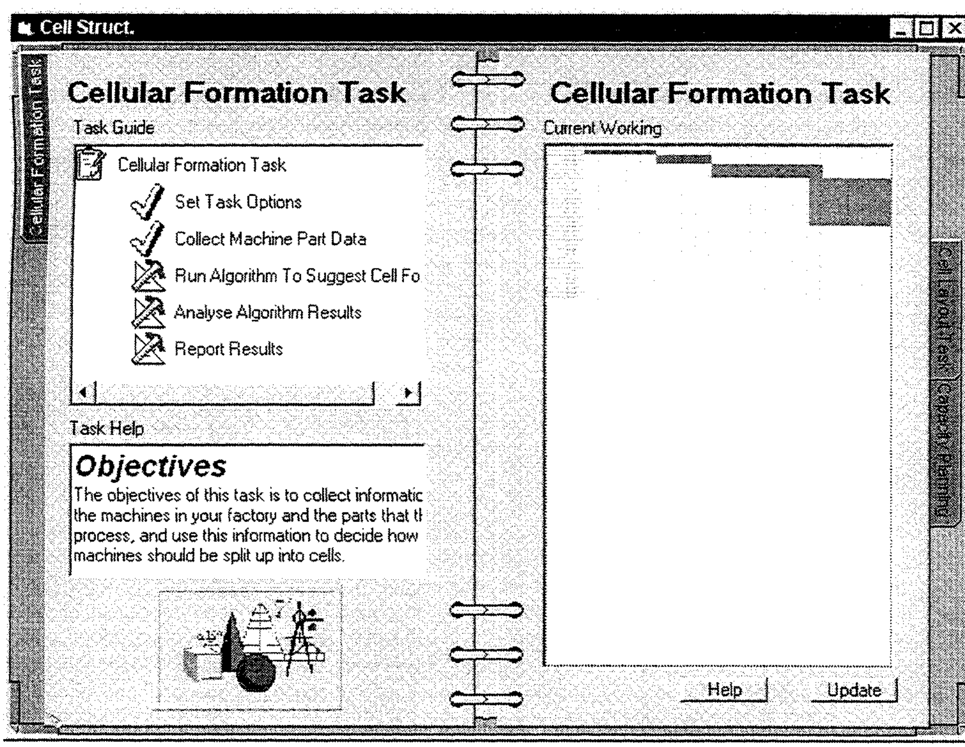


Fig. 2 : Design task document pages 2 and 3

The task-centred approach provides additional functionalities to the I/O-CAMSD framework, primarily those of flexibility and extensibility. The design task itself can be extended through the introduction of additional tools and techniques and associated help and guidelines. For example, a company specific algorithm could be applied within the task by means of defining a macro and its data requirements. In a similar fashion, the framework can be supplemented through the provision of additional design tasks.

5. APPLICATIONS AND FURTHER WORK

Work is currently being undertaken at the implementation level to populate the framework with suitable manufacturing systems design tasks. However, it has been recognised that the application of the framework, and in particular the computer based implementation, has much more wider implications than merely acting as a computer based MSD tool. The application of the standard user interface to configure all the tasks within the adopted methodology, and the underlying software framework to interface and integrate the appropriate and relevant tools, is sufficiently generic that a variety of methodological approaches and philosophies could be supported. An example is that of business process re-engineering, which essentially follows a standard problem cycle and applies a variety of tools and techniques to model, analyse and re-design the business. The modularity of the framework and the framework implementation provide a further axis for development. The project initiation stage, which is currently dominated by the manufacturing strategy analysis as a means to provide the project terms of reference, is not directly linked to the framework and the ensuing methodology. The principle is that the project initiation stage merely provides guidelines for further action and the selection of appropriate task frames and design tasks. In addition, the implementation of the framework is sufficiently generic such that it can be extended or altered to take into account other requirements. Hence, the project initiation stage and the design tasks and task frames can be modified separately and the suggestive logic added at a subsequent time.

An additional application is the consideration of concurrent engineering philosophies and the product introduction process in particular as a specific MSD case. Indeed, there are many similarities in the required approaches. Yu and Yule [12] observe that adopting concurrent engineering philosophies encompasses the combination of three main strategies: a team based strategy, a CAD/CAM central database strategy, and a prototyping strategy. Each has its equivalent when MSD is examined. Teams and MSD task forces are advocated as effective means of implementing MSD [13]. Similarly, access to design and process data, both in terms of the product and the manufacturing system is important [14]. Finally, the ability to examine different scenarios and undertake 'what-if' analysis has been recognised as a requirement for a CAMSD environment and as one of the particular advantages that a computer may provide [8].

Previous attempts have been suggested to couple product design to the manufacturing system within reference architectures, see for example Kovacs *et al* [15]. However, it is perhaps as a support tool for concurrent engineering, rather than as an overall concurrent engineering framework, that a benefit could be immediately attained from the I/O-CAMSD environment. The benefit is primarily achieved by providing methodologies, tools and techniques for the design of the manufacturing system at the product unit, cell and workstation levels of facility design rather than the actual processes themselves. Hence, the target is to achieve reduced time to market, lead-times and costs simply by improving the systems design cycle. If a concurrent engineering framework is considered, then it can be seen that the current computer based task-centred approach outlined in this paper provides some of the fundamental elements of such a framework and could be easily modified to suit such a purpose. The provision of design methodologies, a standard user interface within a windows environment, simple tools and techniques and integration with more complex tools is readily available. The result is an easily accessible means of communicating design ideas, goals, objectives and constraints throughout the design life cycle, from conception to production. Whilst data can be readily drawn from a host of different databases, the primary limitation of the framework at the current state of development is the lack of CAD tools and drafting tools of sufficient quality and wide user base which possess OLE automation capabilities.

6. CONCLUSIONS

The conceptual structure and implementation of the I/O-CAMSD framework and the task-centred methodology described in this paper are based on research work in close collaboration with a number of industrial partners. The paper has presented the key elements of a task-centred methodology for computer aided manufacturing systems design, namely the design task documents the task frames and the I/O-CAMSD framework. Such a structure, which is open and adaptable, provides a general mechanism for task, tool and data integration and for the definition of a specific approach to MSD. Areas in which the generic structure of the framework and its computer based implementation could be applied have also been explored, particularly with respect to concurrent engineering applications.

7. ACKNOWLEDGEMENTS

The work reported here was executed under EPSRC UK Grant ref. GR/J 84656. The authors would also like to thank Edwards High Vacuum International Ltd., Information Engineering Ltd., Kenard Engineering Co. Ltd., LucasVarity and Synchronized Manufacturing Ltd. for their support.

8. REFERENCES

1. Wu, B., Manufacturing Systems Design and Analysis, 2nd Ed, Chapman and Hall, London, 1994.
2. Heim, J.A. and Compton, W.D., Manufacturing systems. Foundations of world class practice, National Academy, Washington D.C., 1992.
3. Kidd, P., Agile Manufacturing, Forging New Frontiers, Chapter 6, p.137, Addison-Wesley, Wokingham, 1994.
4. Devereux, S., Smith, P. and Wood, D., 'A Survey of the Use of Design Methodologies for Implementing Change in manufacturing Companies in the United Kingdom', International Journal of Manufacturing System Design, Vol. 1, No. 1, pp.51-58, 1994.
5. Fritz, S., Schmid, F. and Wu, B., 'A Survey of the Current Practise and Development in Computer Aided Manufacturing Systems Design', Proceedings of The International Conference on Managing Integrated Manufacturing: Organisation, Strategy and Technology, pp.673-689, Hassard, J.S., Forrester, P.L., Hawksley, C. and Tang, N.K.H., University of Keele, Staffordshire, 1993.
6. Ehrlenspiel, K. and John, T., 'Inventing by Design Methodology', Proceedings of the 1987 International Conference on Engineering Design, Vol. 1, p.31, ASME, New York, 1987.
7. Wu, B, 'The Open DSS Structure of Integrated Computer Aided Manufacturing Systems Design', Proceedings of the 14th International Congress on Cybernetics, Namur, Belgium, 1995.
8. Devereux, S. and Wood, D., 'The requirements and application of an integrated computer aided manufacturing systems design environment', Proceedings of the European Workshop on Integrated Manufacturing Systems Engineering, pp.63-68, Grenoble, 1994.
9. Wild, R., Production and Operations Management, chap 6, p.141, 4th Ed., Cassell, London, 1991.
10. Wu, B., 'An overview of the technical requirements for an integrated computer-aided manufacturing systems design environment', International Journal of Manufacturing Systems Design, Vol. 2, No. 1, pp.61-72, 1995.
11. Hull, R.S. and Wu, B., 'Computer Aided Manufacturing Systems Design An Open and Flexible Design Methodology', Advances In Manufacturing Technology X. Proceedings of The Twelfth National Conference On Manufacturing Research, Bramley, A.N., Mileham, A.R., Owen, G.W., pp.126-130, University of Bath, Bath, 1996.
12. Yu, G. and Yule, D.S., 'The Assimilation of Computer-Integrated Manufacturing and Concurrent Engineering', Proceedings of The International Conference on Managing Integrated Manufacturing: Organisation, Strategy and Technology, Hassard, J.S., Forrester, P.L., Hawksley, C. and Tang, N.K.H., pp.401-412, University of Keele, Staffordshire, 1993.
13. Parnaby, J., 'Creating a Competitive Manufacturing Strategy', Production Engineer, July/August, pp.24-28, 1988.
14. Wu, B., 'Information integration - an essential element of integrated CAMSD/CIM environment', Control and Dynamic Systems, Advances in Theory and Applications, C.T. Leondes, Vol. 60, Part 1, pp.123-176, Academic Press, 1994.
15. Kovacs, G.L., Mezgar, I. and Nacsa, J., 'Concurrent Engineering Approach to Support CIM Reference Model-based FMS Design', Computer Integrated Manufacturing Systems, Vol. 7, No. 1, pp.17-27, 1994.

Study of Rapid Intelligent Tooling System Based on RPM Technology

Zhaohui Du, Renji Zhang, Yongnian Yan
The center on Laser Rapid Forming
Department of Mechanical Engineering
Tsinghua University,100084,Beijing,China

Abstract

This paper studies the process of designing and manufacturing tools and dies, which is based on Rapid Prototyping & Manufacturing(RPM) technology. Rapid tooling with short time and good quality can be realized by means of combining RP technique with conversion techniques such as metal casting. The Rapid Intelligent Tooling (RIT) system is a programmable scheme to meet a die need with high reliability and more benefits. RPM and conversion processes ability matrix are constructed for selecting the feasible process approaches according to the criteria of process reliability. Analytic Hierarchy Process and weighted process ability score are introduced to evaluate the priority among factors such as machining cost , time and quality of die. So with these program, an optimal tooling process plan can be decided from a set of feasible paths which satisfy the engineered need well.

Key Words

Rapid Intelligent Tooling, Rapid Prototyping &Manufacturing, Conversion Techniques, Process Reliability Criteria, Analytic Hierarchy Process

1.Introduction

Rapid Prototyping & Manufacturing technology is a kind of integrated technology developed since the mid 1980's. Completely different from traditional shaping technologies, RPM process selectively stack material together from point to line , from line to plane, from plane to body. Then physical 3-D parts can be converted from CAD geometry models directly. Other appellations for this technology are Layer Manufacturing Technology, Material Incress Process or Free Form Fabrication. These appellations represent characters of RPM from different aspects[1]. RPM technology has shown great power in area of prototyping and providing aids for rapid feedback design. Recently, fabrication of functional parts with metal material is becoming an important direction of research and industrial application[2][5]. Although application of RPM in mass production is not economical, this kind technology is very suitable for one case or low volume production. Rapid tooling is one of the best suitable application areas.

There are more and more RP system vendors over the world[3]. With the application of various kinds of RP equipment, a lot of literature reported researches on RPM technical-economic analysis. How to measure and improve the dimensional accuracy and surface finish is being studied[4]. The Center on Laser Rapid Forming of Tsinghua University (China) put forward the concept of multifunctional RPM system (MRPMS), which integrates a few of kinds of RP processes in one platform. And the MRPMS machine has been developed there. But there is no systematic method and software to solve the problem of RPM process planning. Whether or not a process with some kind of RPM and conversion technique can make tool/die with a special order. Which one among the various RP and conversion techniques is the best choice with the consideration of cost , cycle time, physical performance and geometrical form accuracy? These are the requisition and function of Rapid Intelligent Tooling System.

2. Tooling System based on RPM Technology

The Tooling System based on RPM technology includes four main procedures: design/CAD modeling, evaluation, prototyping and tooling.

Design/CAD modeling is a process which computerizes die orders. There are a lot of computer aided designing and drafting tools available. 1996 Rapid Prototyping Direction lists 11 large CAD system vendors who can output STL file, which is a standard file format for RPM technology. CAD modeling with the input of physical model or 2-D photograph, reverse engineering tools provide reliable aids. This process outputs not only die draught, 3-D geometry model, but also machining request(including dimensional tolerance, surface roughness, life-time, etc.)

Evaluation process is to evaluate and modify the design result with simulation tools and virtual prototyping software. For instance, the process analyzes whether or not it is necessary to add support structure while building prototype, the dynamic procedure of metal flow, and the solidification can be simulated before casting. Another important function of this process is precision measurement of prototypes or dies and modifying the design according to the difference between the requirement and result.

For making a specific prototype and die, there are a few of kinds of RPM(such as SLA,LOM,SLS,FDM,3DP,etc.) and conversion techniques(such as investment casting, shell casting, spraying molding, etc.) available for choosing. The outputs, including cost and cycle time for manufacturing, dimensional accuracy, surface finish and life of dies, can be changed from process to process. The next section studies on evaluation criteria for process reliability and multi-object oriented decision making.

3.Process Planning Procedure.

To incorporate multiple factors into the process decision making, a "feed-forward" process planning procedure should be developed (data flowing shown in Fig. 1).

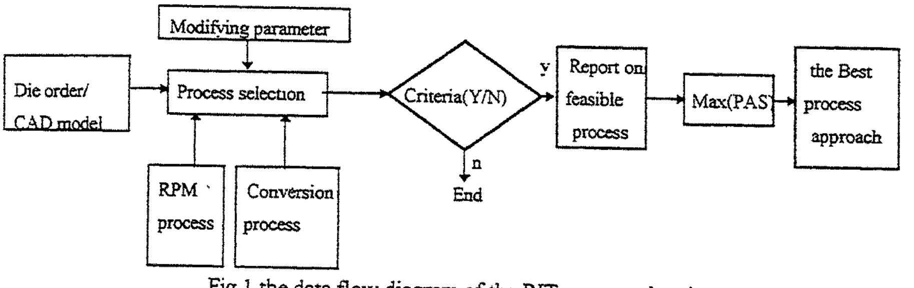


Fig.1 the data flow diagram of the RIT process planning

During process planning, the designer is faced on a set of alternative RPM and conversion techniques to produce a finished die. The die need is input to the process planning scheme. According to the ability of RPM and conversion techniques, feasible process combinations can be chosen. During the scheme of process selection, the reliable criteria should be followed. There are some factors of machining time, cost, part quality and environment effect need to be conscious. And it increases the complexity that the important degree of these factors vary from order to order. The objective of this planning is to select an optimal one from a set of feasible paths which best reflect the manufacturing priorities among production cost, time and quality required. So a "weighted" ability score for evaluating process is introduced. In the following sections, different aspect of the planning scheme will be presented in greater depth.

3.1 Process knowledge base

The capability of available processes can be shown in the process knowledge base. These information is stored in the form of a expandable two-dimensional matrix. Two ability matrixes of RPM and conversion process are built in the RIT system .

It shows a sample of RPM process ability matrix in the Table 1. Only four well-developed techniques including Stereolithography (SLA), Laminated Object Manufacturing(LOM), Selective Laser Sintering(SLS) and Fused Deposition Manufacturing(FDM) have been listed. The items(i.e. material, dimensional tolerance, surface finish, machining cost and time) are specified to determined the ability of each process.

Table 1 RPM process ability matrix

kind of process	kind of material	process tolerance	roughness Ra(μm)	machining cost		machining time	
				synthesis cost	material cost	Scanning Speed	post-process
SLA	resin	IT15	3.7/1.7	$20.0/h $150/kg		500mm/s	1-2 days
LOM	paper	IT16	4.0/2.4	$7.5/h	$40/kg	380mm/s	
SLS	ceramic plastic alloy	IT14	16 / 15	$20.0/h $100/kg		1016mms	1-2 days
FDM	plastic wax	IT15	11 / 5.0	$12.8/h	100/kg	380mm/s	

notes:

1.Machining tolerance value can be converted from Inter -national Tolerance(IT) degree into dimensional tolerance with checking the ISO standard.

2.Surface roughness is varied with the influence of shaping direction(⇒ / ⇑). Average value can be adopted.

3.When calculating the machining cost, three main procedures including preprocessing, part building and post-processing should be considered. Here in order to simplify the computation, a synthesis cost per hour is introduced. So the overall machining cost can determined by the following equation:

Cost=Synthesis cost×time+material cost×part weight

4.Machining time equals to the sum of time spent on the two procedures: part building and post-processing. The gross time for building can be calculated by the following equation:

Building time= volume / (layer thickness × distance between scanning steps × scanning speed)

5.For SLA,FDM,SLS cases, the "volume" in the above equation means the volume of the part need to be built. It can be easily calculated with CAD software or from STL file. For LOM process, the "volume" means the frame volume with subtracting the part volume.

For most cases, the tooling process based on RPM-technology must be with the aids of some kinds of conversion techniques, such as spraying or casting. Presently it has been put forward to study metal casting through using molds made from rapid prototyping techniques. Many process approaches have been developed. In the conversion process knowledge base, four typical techniques including sand die casting, investment casting, wax lost casting and shell casting are specified as examples of available process.The ability matrix of conversion process is shown in Table 2 with similarity to Table 1.

In fact, it is not always practical way that any kind of RP technique can be combined with any kind of conversion process directly. Some paths need transition steps and many of them need improvement of quality. Nowadays there are many research institutes and companies studying on this area all over the world. It can predicted that more kinds of process approach and better performances will be developed. So the ability matrix in Table1 and Table 2 can be expanded and modified.

Table 2 the conversion process ability matrix

kind of process	accuracy	roughness Ra (μm)	min. casting wall thickness	size of casting
sand casting	CT8-13	12.5-25	2.5-5mm	unlimited
investment casting	CT4-6	0.8-6.3	0.3-0.5mm	<10kg
ceramic shell	CT5-8	3.2-6.3	0.5-0.8mm	<1000kg

3.2 The criteria for evaluation of process reliability

In almost each case, not all of the approaches have the ability to finish an engineered die. It is necessary deleting the impossible process path before selecting the optimal one. In order to handle the judgment procedure, criteria are introduced for evaluating process reliability. The approach accepted as a feasible one must follow these criteria. The criteria consist of several rules listed in the following.

Rule 1: the ability index of process accuracy > the minimum of reliable index

Rule 2: the achievable surface roughness < the designed surface roughness

Rule 3: the life of die > the design life of die

Rule 4: the selected process can shape the die freely from the geometry form constrain

Here Rule 1 has be discussed in detail. It is known that machining accuracy shows the concord degree of the part dimensional value between the machined and the designed part. The difference is called tolerance. The factors leading to machining tolerance are multiphase and some of them occur randomly. Traditionally statistics has been applied as a tolerance analytical methodology. A curve with the measured data from a set of part dimensions shows the distribution of the errors. The curve, which is usually a normal distributed one, can be used to analyze the process tolerance. Relevant studies have been taken on RPM area and the normal distributed curves have been gained. According to the regulation of normal distribution, $\pm 3\sigma$ is the system tolerance. To judge whether or not the satisfied product can be manufactured in one process, the process precision capability index (C_i) is used.

$C_i = T/6\sigma$ (where T represents the designed tolerance)

The tooling process discussed in this paper often refers to the combination of RPM and conversion techniques. For this kind of sequential process steps, the overall capability index ($C_i^{overall}$) equals to the product of C_i of every steps.

$$C_i^{overall} = \prod_{j=1}^{n} C_i^j$$

The lower the C_i, the less possibly the selected process meets the design. The minimum C_i (C_i^{min}) represents the process with it should not be chosen as a reliable one which has the ability to manufacture a satisfied part. In some case, the C_i^{min} for metal cutting process is set 0.67

Subsequently, Rule1 can be represented as the following form:

$$C_i^{overall} > C_i^{min}$$

Rule 2, 3, 4 can be handled by a method similar to Rule 1. Rule2 prescribes that a practical process should produce a part with surface finish not worse than designed. In order to improve the part surface finish, the prototype made from RPM techniques often needs a grinding step before shipping it to make a casting mold. The surface roughness of a casting is influenced by multiple factors such as casting process, material, part weight, part geometry form, etc. Spreading coating material on the inside surface of mold cavity is very helpful to improve surface finish.

In the course of tooling, the life of tool/die is considered as an important issue. Rule 3 requires that a feasible process should have the ability to produce die with longer life than designed.

RPM technology is one kind of free form fabrication method. But there exist geometry constraints in casting process. For example, the minimum thickness of casting wall must be larger than the largest pass in which melted metal can fill fully. Other design regulation such as size of casting and passage fillet should be considered as well.

3.3 Decision making

From above discussions, several process paths reliable to finish a special die order have already been picked out. This section introduces how to select the best solution[6][7].

In order to fully evaluate the trade-offs in these different alternatives, a set of quantifiable factors such as machining cost, time, precision, surface roughness and life of die need to be analyzed at the decision-making stage. These factors are quantified and listed in the report of feasible approaches with a form of two dimensional matrix (shown in Fig. 2).

The important degree of each factors may vary among orders depending on the site specific consideration. It is not easy to directly decide which tooling approach is the best. Sometimes designer pay more attention on precision or surface finish of die. In other situation, it is more important to decrease the machining time or lower the cost. Before making a decision these dimensions should be evaluated through a prioritization matrix using the Analytic Hierarchy Process (AHP).

The AHP is an evaluation method based on the theory of fuzzy set, which is depended on decision-maker's estimation on multiple dimensions with experience. The process develops a set of pairwise comparisons to form a prioritization matrix among the different factors based on tooling needs[8]. A example matrix is shown in Fig 3. Based on this matrix, a 5×1 weight vector W can be calculated. For a rank value for each row of the matrix in Fig. 3 is determined through the relationship:

$$
\begin{array}{c} \\ A_1 \\ A_2 \\ F=A_3 \\ \cdots\cdots \\ A_n \end{array}
\begin{array}{c} \begin{array}{ccccc} F_1 & F_2 & F_3 & F_4 & F_5 \end{array} \\ \begin{bmatrix} 0.15 & 45 & 230 & 4 & 40000 \\ 0.25 & 30 & 180 & 7 & 15000 \\ 0.15 & 20 & 300 & 20 & 80000 \\ & & & & \\ 0.40 & 50 & 150 & 8 & 40000 \end{bmatrix} \end{array}
$$

Fig. 2 The report on feasible approaches

$$
X = \begin{array}{c} \\ F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \end{array}
\begin{array}{c} \begin{array}{ccccc} F_1 & F_2 & F_3 & F_4 & F_5 \end{array} \\ \begin{bmatrix} 1 & 1/4 & 1/2 & 1/2 & 2 \\ 4 & 1 & 2 & 2 & 8 \\ 2 & 1/2 & 1 & 1 & 4 \\ 2 & 1/2 & 1 & 1 & 4 \\ 1/2 & 1/8 & 1/4 & 1/4 & 1 \end{bmatrix} \end{array}
$$

Fig. 3 The prioritization matrix among factors

$$R_i = \left(\prod_{j=1}^{k} X_{ij}\right)^{1/k} \text{ (where } X_{ij} \text{ are the elements of the AHP matrix)}$$

The element of W are then computed by a normalization

$$w_i = R_i / \sum_{j=1}^{n} R_j, \quad j = 1,2\ldots n$$

For the matrix in Fig.6, the W is determined to be $[0.105, 0.421, 0.211, 0.211, 0.052]^T$.

The final process ability score(PAS) is used as an index for measuring the priority of process.

$$PAS = \sum_{i=1}^{n} (-1)^{jo(i)} w_i F_i / \Sigma F_j \qquad i=1,2,\ldots n$$

where jo(i) is a function which equals to even or odd number. If factor i (such as machining time or cost) has a negative effect on process capability, then jo(i) equals to odd number. Contrarily, jo(i) equals to even number when the factors are machining accuracy or surface finish or die life.

After calculation of PAS, all kinds of the feasible approaches can be ordered with the increase by PAS value. It is deduced that the tooling process with higher PAS can bring more benefit and better result. The process with highest PAS is the best choice.

4. Case Study

In Fig. 4a it is shown an art miniature, which is a scaled copy from the well-known tower in Beihai Park (in Beijing, China). It is the engineered need to fabricate the plastic injection mold by the means of the RIT technology based on RPM. The tower has very complex geometry form with not only a lot of knaggy planes but also many short and thin lines on its surface. Because it is used as a plaything, the die need not too high accuracy but it must have good surface finish. At the same time, the die is designed with the production capability of 100 thousand cases. With the RIT program, tooling process planning has been carried on. From several alternatives, finally an optimal approach with combination of SLA- and ceramic shell casting technique is determined. The pattern made by SLA is used to fabricate the ceramic shell mold (shown in Fig. 4b) which is for moulding the injection mold (shown in Fig. 4c). The whole process flow is shown in Fig 6.

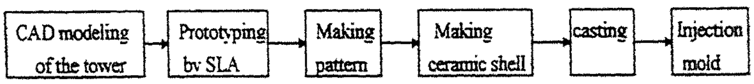


Fig.5 the process flow diagram of the injection mold making

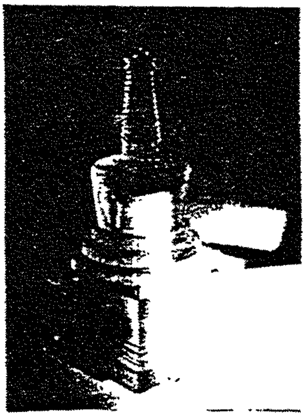


Fig.4a the photograph of the tower miniature

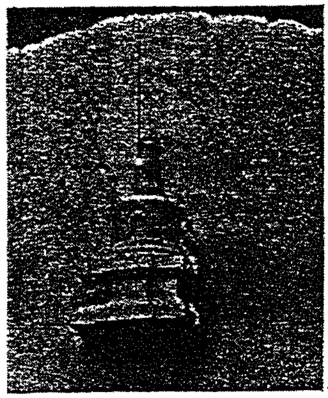


Fig.4b the ceramic shell mold

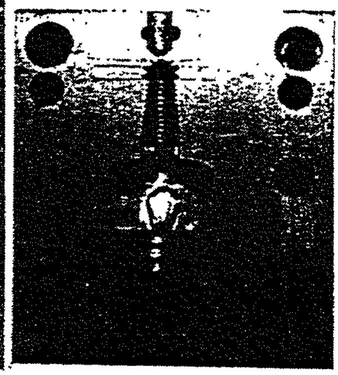


Fig.4c the injection mold

5. Conclusion

Presently studies on directly metal part fabrication with RPM techniques is being advanced. It is a practical and promising way through converting RP model to metal part. This paper emphasizes on the application of this kind of technology in tooling process. The model of Rapid Intelligent Tooling system is constructed. With the combination of RPM and casting techniques, tool/die can be manufactured with low cost, short time and good quality. The abilities of these process approach are stored in the form of two-dimensional matrix, from which the feasible process alternatives can be chosen while following the criteria of evaluation of process reliability. Analytic Hierarchy Process is introduced to the priority weight of factors determined by the tooling needs. The final solution with highest process ability score(PAS) is the best selection. With the programmable procedure, the system improve the accuracy and efficiency for rapid tooling.

6. Reference

[1] J.P.Kruth, Material Incress Manufacturing by Rapid Prototyping Techniques, Annals of the CIRP, vol.40/2/1991, pp603-614

[2]P.M.Dichens, et al., Conversion of RP models to investment castings, Rapid Prototyping Journal,4,1995,pp4-11

[3] Anonymous,1996 Rapid Prototyping Directory, Rapid Prototyping Report,Feb/1996,pp3-6

[4]R.Ippolito,et al.,Benchmarking of Rapid Prototyping Techniques in Terms of Dimensional Accuracy and Surface finish, Annals of the CIRP,vol 44/1/1995,pp157-160

[5]Lena A.K.,G. Sohlenius, Future Direct Manufacturing of Metal Parts with Free-Form Fabrication, Annals of the CIRP, vol.44/1/1994,pp451-454

[6]Nam P Suh, The principles of design, Oxford university press 1990

[7]P.Sheng,M.Srinivasan,Multi-Objective Process Planning in Environmentally Conscious Manufacturing: A Feature-Based Approach, Annals of the CIRP,vol.44/1/1995,pp433-436

[8]Xiao Shaowei et al., Introduction of System Engineering, Tsinghua University Press,1995

THE APPLICATION OF HOT MELT ADHESIVE TO LAMINATED OBJECT MANUFACTURING

Zhang Xianglin , Huang Jin, Li Yan and Huang Shuhuai
Faculty of Mechanical Science & Engineering
Huazhong University of Science & Technology
Wuhan , P.R.China 430074

ABSTRACT

Rapid Prototyping & Manufacturing (RPM) is a fast developing technology. Our research group has made several generations of Laminated Object Manufacturing (LOM) systems (named HRP) and produced a lot of prototypes. To make high precise prototypes, this paper presents some analyses and special experiment results mainly concentrated on the aspects of materials, sheets and hot melt adhesives, and their process and application in LOM. These include: a brief introduction of the characteristics of the newly developed HRP; the special property requirements for substrates and HMAbased on EVA; the ingredients of HMA and their special effects; the coating methods; the bonding in LOM machine of applying HMA to sheets and some suggests for waste removal. Several important data and examples are given.

KEYWORDS

Rapid Prototyping & Manufacturing, Laminated Object Manufacturing, Accuracy, Hot Melt Adhesive , Bonding

1. INTRODUCTION

The new technology of rapid prototyping and manufacturing is a fast developing technology. It is now recognized as a means to greatly reduce the lead time from the conception of a new product to its launch on the market. It gives the opportunity to directly convert CAD files into a physical model in hours rather than the weeks or even months of traditional modelmaking techniques. There are several types in RP such as Stereolithography(SLA), Laminated Object Manufacturing (LOM), Selective Laser Sintering (SLS), Fused Deposition Modeling (FDM). The all apply the principle of making parts layer by layer continuously.

Our research group has made several generations of LOM type RP system. A lot of prototypes have been made on it. The work principle can be illustrated by Fig.1. The LOM system produces models by bonding together a number of laminates which are progressively cut by a CO_2 laser. Special paper with a kind of hot melt adhesives (HMA) coated on the down side is fed from the supply reel and stops on the utmost layer of the part. A special heating device passes over, heating, measuring, compressing and bonding the paper onto the previous layer. The laser mounted above cuts the required internal and external profiles of each successive paper layer according to digitized information. During the cutting process waste material within the cutting area is crosshatched into small pieces (Fig.1b) for easy removal once the model is finished. The accuracy and the reliability depend on several aspects. As far as we know there are few literature dealing with problems from the field of the applied materials and the processing. This is just what we want to discuss.

2. THE PROPERTY OF SUBSTRATES

The sheet is the bone and is easy to be bonded by using adhesives. It weighs from $30\,k/m^2$ ~ $100\,k/m^2$, which depends on the required part precision. The better the consistence of paper thickness, The easier to get high size precision in Z direction. Owing to special driving mechanism aimed to decrease the web tension in the horizontal section, the tension strength is not specially required. The thinnest sheet, which weighs $30\,k/m^2$ can be automatically operated. The strength of the coated sheet increases 87% compared with the raw sheet by our measurement. The up and down faces of sheet need not cover additives, so it allows fluid adhesive easy to penetrate into the sheet. From the photos (Fig. 2 to Fig. 7), we can see the sheet is porous for the HMAs.

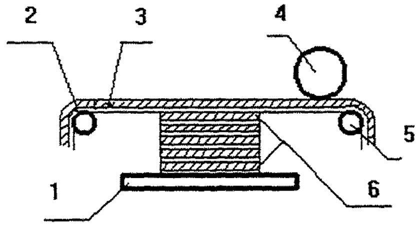


Fig. 1a
1. platform 2. adhesive coat 3. sheet
4. heating roller 5.driving roller 6. part

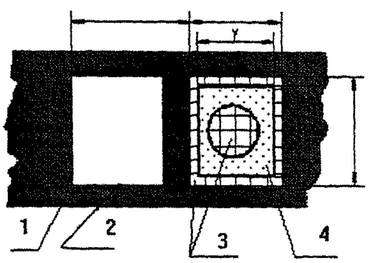


Fig. 1b
1. sheet 2. boundary 3. waste area
4. part area

3. THE RESEARCH OF HMA AND ITS APPLICATION

3.1 The Formulation of hot melt adhesive

The adhesive bonds sheet layers together and it must be structural type. The prototypes made in LOM should be able stand loads to certain extent. The adhesive has following properties: (1) high bonding strength and durability in various unfavorable environments; (2) easy to be coated thinly on the sheet; (3) fast set time of less than 3 ~ 10 seconds; (4) high cohesion in room temperature to keep it in solid state.

Hot melt adhesives have found acceptance and usefulness in many industrial applications. Hot melt adhesives are bonding materials which are solid at room temperature, but soft and fluid at elevated temperatures. A strong adhesive bond is rapidly formed on cooling and hardening. It is thermoplastic type, and no curing time is required. Hot melt adhesive systems offer increased production speeds and lower costs than solvent or waterborne adhesive systems. It is especially suited for continuous operations. Ethylene vinyl acetate (EVA) copolymer is easy to get and the price of it is cheaper, so the EVA based HMA has been used in a wide ranges and is selected to apply in our LOM system.

EVA forms the backbone of the adhesive and provide the cohesion. Some kinds of materials, such as wax, rosin, filler and some other additives should be put together with EVA to form the required adhesive. The ratio of VA to EVA is about 18 ~ 33%. The higher the ratio, the lower the melting point of EVA and the hardness and elasticity. It seems that the ratio does not affect the paper bonding. Another most important parameter is the melt viscosity (MV) which is inversely proportional to the average molecular weight. The ratio of EVA in the adhesives is about 50%, and the wax is less than 40%. Wax has a much lower melting point than EVA. It can supply good creeps when coated onto the sheet or heated in LOM. But when wax surpasses some extents, the peel strength and the hardness will obviously decrease. Rosin can increase the adhesion strength and the creeps, but it takes more time to melt. Some special fillers have to be added to get a clear profile when the laser facula cuts the layers, so that less amount of adhesive is left in the cut trace and the wastes is easy to be separated from the part. We find that it is better for fillers to be less than 12%, otherwise the wetting ability will be obviously affected.

The thermogravimetry experiments show that our adhesive obviously begins to lost weight at $198\,^{\circ}C$, which limits the highest temperature to be applied. Another experiment shows that the ash left after firing at $1000\,^{\circ}C$ are black carbon powders weighing less than 2% of the original weight, so the prototype can be used as patterns in shell casting.

3.2 The coating method

The solvent is not needed for coating of HMA of EVA type. We have tried four kinds of coating methods. (1) Simple knife coating. The sheet may be scared by big particles and the solvent can't always be in high precision. So the method can only be used to try the sheet and the adhesive (2) Roll coating. The method is frequently used to coat solvent adhesives. The device is complicated and the thickness precision does not meet the need of higher parts. (3) Electrostatic spraying. When all the ingredients for adhesive are of powders, highly homogeneous coating can be obtained by this method The EVA powders can only be produced at very low temperature because of its thermoplasticity. (4) Hot extruding. The molten adhesive are transmitted by gauged pump and extruded out through a narrow slot. Both high operation speed and precision can be obtained. The last method is adopted for large scale processing. The thickness error is less 5%. The thickness of coated adhesives must be decided for care. We suggest it not less than 0.015mm and not high than 0.035mm, depending the operation precision of LOM machine.

3.3 The Application in LOM

Because the sheets are porous, there are several types of forces exist between the sheet and adhesive, the molecular force and the mechanical force. The adhesive coated sheet will be heated and pressed to bond together in LOM machine. The following aspects should taking into accounts. (1) The heating power should be high, so the utmost layer quickly gets to the melting point and the operated layers is less effected. (2) The pressure of the hot roller and the laser power should be adjusted according to the real layer thickness so that the high bonding quality can be assured. (3) high precise measuring of temperature and height and some special experiments are required.

3.4 The Consideration for Easy Removal of Wastes

waste removal is a much completed problem. This may be the main drawback of LOM The waste plays the role of supporting, but it has to be bonded to the part Without knowing the shape of the part in advance, the part itself is easy to be hurt when the waste is removing. So we have taken several improving measures from the two aspects of software consideration and strictly control of the processes

3.5 The Comparison of Several Examples

Fig.2 to Fig.7 are photos about sheets, coated adhesives and part cubics taken by a Scanning Electronmicroscope. They are respectively magnified 200,300,500 times that can be showed on the bottoms of every photos. Fig.2 is bare raw sheet with longer fibers which weighing $30\,k\,/\,m^2$. The sheet in Fig.3 is the same as that of Fig.2, but it is coated with a HMA layer on the down face. The penetrating property is very good. The Fig.4 and Fig. 5 are two samples of other LOM researchers. The Fig.5 shows that it is coated by powder spraying method. The Fig.6 and Fig.7 show two profiles of two parts. In Fig.6 the adhesive is squeezed out and it is a little troublesome for the waste removal. In Fig. 7 the sheet fibers are truncated by laser facula.

4. CONCLUSIONS

Material plays a more and more important role in RPM. The above experiment and analyses contributes a lot to the making of qualified prototypes. Nevertheless, there still many problems unsolved and some new problems will expose as our new HRP machines work in large run But the analysis method and experimental principles can be referred.

5. ACKNOWLEDGMENTS

The project is financially supported by the Natural Science Foundation of China. The others members of our research group, Mr. Xiao and Ms. Han also contribute a lot to the research.

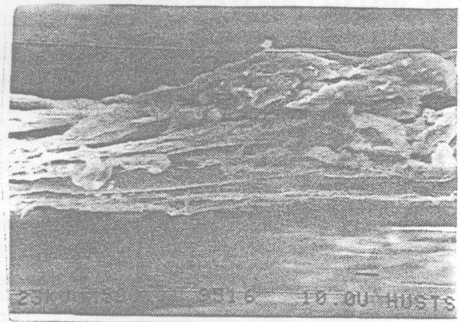


Fig. 2

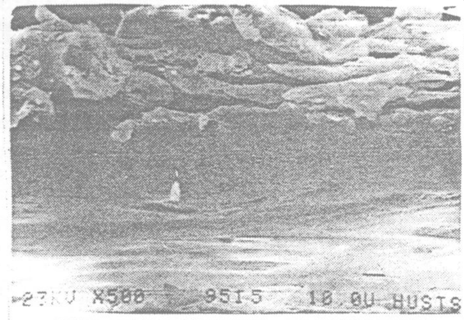


Fig. 3

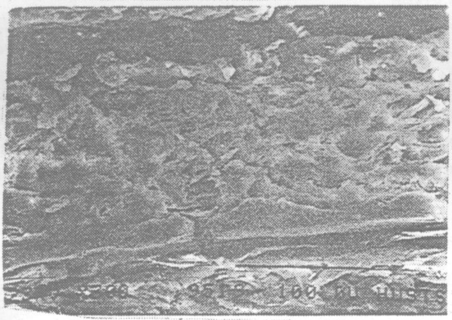


Fig. 4

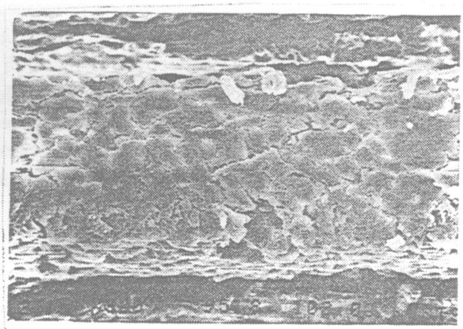


Fig. 5

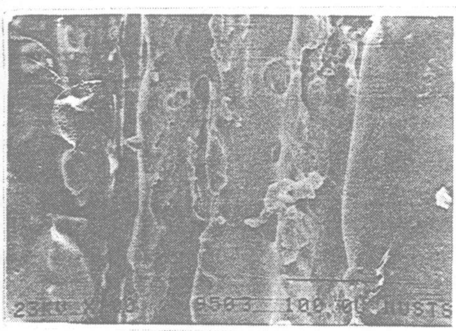


Fig. 6

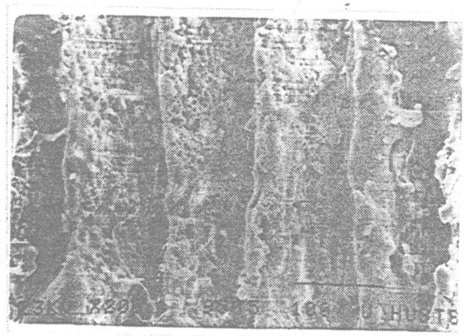


Fig.7

6. REFERENCES

1. Liu Bin et al, How to Improve the Accuracy of Prototype Built by the Laminated Object Manufacturing System, Presented to the International Conference on Manufacturing Automation (ICMA'97), Hong Kong.

2. Shields, J., Adhesives handbook , Butterworths, 1984.

INTERNAL COMPANY USE OF RAPID PROTOTYPING AND 3D-DIGITIZING IN THE ENGINEERING DESIGN PROCESS

C. Birke, C. Beyer, K.-H. Grote
Department of Mechanical Engineering Design
Otto-von-Guericke-University Magdeburg
Magdeburg/Germany

ABSTRACT

Today the trend goes towards shorter product life cycles, which in turn require reduced time spent on the product development. Beside the use of computer hard- and corrosponding CAD-software a number of systems are available that build (Rapid Prototyping) or record (3D-Digitizing) real tangible three-dimensional bodies on basis of different technologies. By using these technologies an effective support of parallel working processes is given.

An important working field at the Department of Mechanical Engineering Design at the Otto-von-Guericke-University Magdeburg is research and further development of the implementation of these technologies in the structured product development process and to give guidelines to the engineer for successful work. Productive procedures and tools for the realization of a computer integrated product development process are developed. This process includes the effective application of technologies to produce the prototypes for supporting simultaneous engineering, presentations, producibility analysis and production.

KEYWORDS

Rapid Prototyping, Three-Dimensional Digitizing, Simultaneous Engineering, Product Development

1. INTRODUCTION

The development of machines and engineering parts is very dynamic and subject to a strong competition pressure. On national and international levels new products (or adaptation of existing products) have to be developed in the shortest possible time. The requirements for high quality and competitive characteristics stay the same or increase. To be able to sell competitive products on the market innovative solutions are needed. Furthermore development times and development costs play an essential role.

As currently up to 75% of the production costs are determined during design, development and production planning, efficient procedures and tools on these fields have to ensure a fast availability of high-quality products. At the present a number of systems are available that build up three-dimensional objects on the basis of different technologies (Rapid Prototyping) or capture three-dimensional bodies.

As a result of the producibility analysis and/or tests of a component, it is very common that the component must be changed on the basis of production, assembly problems or because of the test results subsequently. This happens usually „on stage", either through manual re-work or through changes in the NC-program. The current component is then much different, compared to the stored solid model in the computer. The manual changes have to be added to the solid model in the CAD-system. In most cases this requires a lot of manual work and often there is no possibility to record the real changes on the delivered part (e.g. cast model). This problem leads to differences between the delivered part and the CAD-data. Especially the presence of surface geometries that represent freeform areas in the CAD-data requires a lot of manual work.

2. RAPID PROTOTYPING - FAST AVAILABILITY OF MODELS

Apart from the multitude of different technologies there are a lot of definitions for the classification of the discussed procedures. The terms „Rapid Prototyping" or „Solid Freeform Manufacturing" are used most commonly.

To this class of Rapid Prototyping procedures exclusively those technologies count that build physical, three-dimensional models of a part by additive shape building principles. It means, that the part to be generated has not to be carved out from a solid body, it results from the addition of material or the phase transition of material from fluid or powder to the solid state [1]. The production is accomplished without the use of shapes or tools. In opposition to classic procedures (milling, turning) the additive manufacturing processes can reduce the production period very extensively, although the classic procedures are currently developed more efficient by the usage of computer technologies.

3. SUPPORT OF THE PRODUCT DEVELOPMENT

An important research field at the Department of Mechanical Engineering Design at the Otto-von-Guericke-University Magdeburg is research and further development of the implementation of these technologies in the structured product development process. This work concentrates on the efficient use of technologies for prototyping and product manufacturing to support the product planning and development. The possibilities and restrictions for using these technologies at the early stages of generating solutions and the elaboration of product characteristics are of special interest.

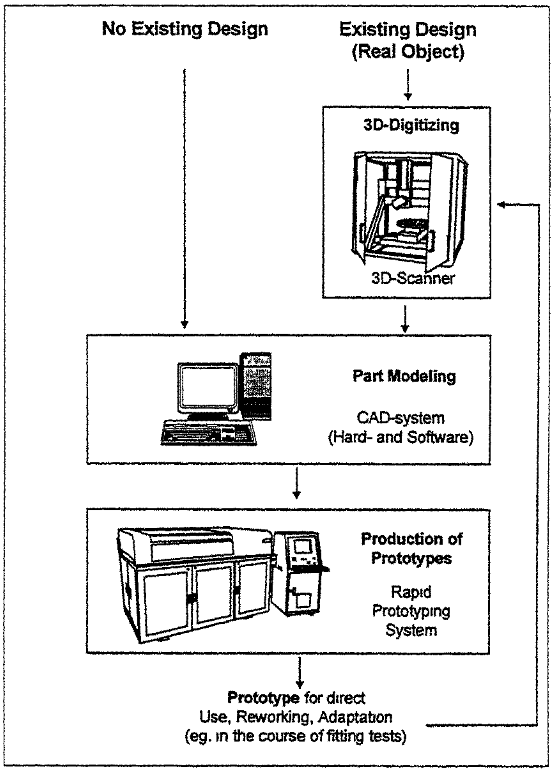


Fig. 1 Process chain (loop) for the Computer Integrated Prototype Development

A closed process chain from solid modeling, rapid prototyping (using the LOM-technology) and the three-dimensional digitizing is realized. It is the aim of this work to get information about the fast and economical generation, verification and re-engineering of product data. The importance of powerful technologies and algorithms on the mentioned field shows products that are or were successful on the market, but quickly had to be adapted to the current market development [2; 3].

4. FORMING A PROCESS CHAIN

The research work at the Otto-von-Guericke-University Magdeburg proceeds continuously. At this time some solutions are presented:

Starting-point is the already mentioned closed process chain (loop, fig. 1) from Solid Modeling, Rapid Prototyping and 3D-Digitizing. Within this chain a pick of parts or part families passes through the single steps to get reliable results of a complex automatic and error-free feedback of the data.

Suitable parts of different complexity are made or selected from already existing models and adapted as verification objects. The different complexity-levels have been determined by the requirements of the engineering design and the computer supported modeling of parts. The next step is the determination of the tolerance fields of the equipment and - if necessary - an adaptation of the part (position, orientation) occurs.

Starting out with objects of simple part geometries the course of the research work continuously increases the complexity of the geometries.

A critical analysis of all interfaces provides indications for an optimal use of the individual components and of their combination. Problems are expected at the data transfer and data transformation for parts with high complexity.

Within a loop every model is being created and compared to the data and for geometry of its starting situation. Besides the comparison of the initial stage (Desired Data/Values) and the edited stage (Real Data) the models - of different complexities- are compared.

Previous investigations have shown results for good arrangements and machine parameters for the prototype manufacturing with high precision, high surface qualities and short manufacturing cycle. A systematic, established evaluation will be published when the statistical research program is completed.

5. TECHNICAL REALIZATION

The research work at the Department of Mechanical Engineering Design of the Otto-von-Guericke-University is facilitated by several state-of-the-art CAD computer workstations, a Rapid Prototyping System with Laminated Object Manufacturing (LOM) procedure, and an optical (non-contact) three-dimensional digitizing device (3D-Scanner) that is needed for the digital capture of existing models.

The preparation of the geometrical data is the same for all the different rapid prototyping systems. Three steps are essential: The data supply (3D-modeling of the object), data preparation and data processing.

To generate the files for the prototyping device several procedures have been tested. In the majority of cases 3D-CAD-software systems are used. The Department of Mechanical Engineering Design works with the programs Pro/ENGINEER and I-DEAS.

a) Paper Block with embedded Model b) Decubing the part c) Finished Model

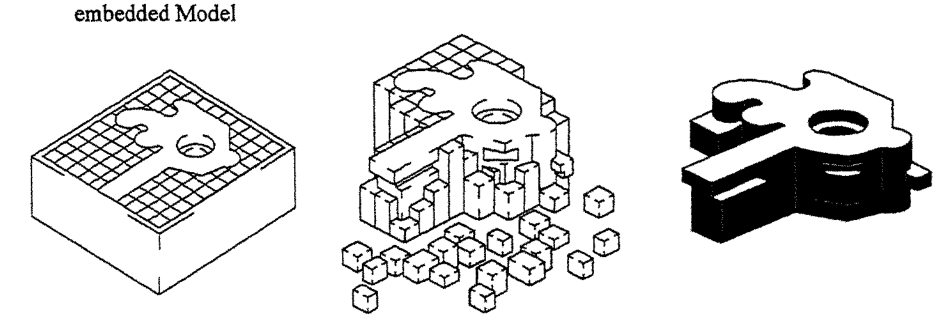


Fig.2 LOM: Decubing the part from the foil block [6]

The Laminated Object Manufacturing (LOM) Rapid Prototyping procedure is briefly explained. Single layers (slices) of geometries are cut, stacked and glued. The result is a stack that forms the desired three-dimensional body (construction part).

A laser beam cuts the two-dimensional cross-sections from foil material. The series of cutting and bonding the single layers on the stack is automatic and computer controlled. The part has characteristics like plywood and allows corresponding follow-up treatment procedures. Figure 2 shows the embedded part in the foil block and the steps to the decubed and finished object.

Rapid Prototyping is the technology of manufacturing models using computer data. Various ways lead to the modeling of 3D-objects. In the majority of cases 3D-CAD-programs are used. Another possibility to get the needed data files for prototyping is to digitize an existing real tangible object (e.g. cast part). The reason for this procedure could be that no 3D-CAD data is available.

The digitized, recorded „data cloud" is a quantity of adjacent points, transformed by software into a „sensible" structure. The results are geometry data in a file form compatible to traditional CAD-systems, which can be processed as traditional CAD-files.

For modeling the objects, a multitude of different soft-and hardware platforms are available with some different surface tracing programs, such as Surfacer, STRIM, ScanCAD surface.

A focus of the research work at the Department for Mechanical Engineering Design is to generate solutions that are suitable for medium-sized enterprises. The factors costs for hard and software as well as the responsible staff are extremely important. That is why the research is concentrated on findings of common rules and strategies independent from the special technical equipment (using not only one software or surface tracing program on a single platform).

The multitude of the systems allow to simulate various conditions in the real industrial operation and to guarantee the transition of the research results into practical applications.

6. PREFERENCES OF RAPID PROTOTYPING TECHNOLOGIES

The use of these technologies supports the parallel adjustments of work processes. A fast availability of prototypes represents a modern help that can increase learning and decision making processes considerably. Corresponding manufactured parts allow - because of immediate and relative easy availability - fit and assembly analyses of constructions already in an early phase of the product development process. These models are used to examine the product design and to establish the manufacturing of the later series production. The models represent a connecting link as information support for the parallel work of the product- and process structuring sequences.

The designer has the opportunity to reflect the results of the systematic development process [7] in form of one or more prototypes tangibly. The existence of prototypes supports the activities of the working area following the engineering designer's work.

Experience has shown that tangible models are better suited to give an imagination of the object than (computer-) graphics or verbal descriptions.

Design models	Function models	Manufacturing models
• CAD model verification (Model main measurements) • Visualizing objects • Test and proof the concept • Marketing objects (Visual objects for decision making)	• Form and fit models • Flow analysis • Examinations stress distribution • Visual components (dummies especially for automotive industries) • Pre-series components	• Plastic mold parts – vacuum casting – metal spraying • Casting – sand-casting – investment casting – die casting

Table 1: General areas of the application of Rapid Prototyping components, after [4]

An early availability of prototypes - for instance by using rapid prototyping procedures - allow that important steps of the production preparation and planning, tool making, FEM-analysis etc. can be realized earlier in the product development process. In the next step, the re-working/changes are fed back into the computers product model via the 3D-digitizer. The advantage is the usually high expenditure for the later „handmade“ modeling work to bring the computer's model up to date.

Besides the application cases that are subject to the examinations, the principles of rapid prototyping allow a multitude of additional possibilities for use (summarized in Table 1). The outlined sequences represent the concept of simultaneous engineering [5] within the product development process.

SUMMARY AND OUTLOOK

In addition to the known advantages of computer supported engineering design, simulation systems, 3D-digitizing and rapid prototyping a further reduction of development time, reduction of costs and increasing of quality can be reached more effective by a largely automated feedback and check of produced parts. The above discussed verification-loop-method offers the comparison/checking of the rapid prototyping results and the data from solid modeling in this time-saving process.

Especially to re-work technical products that are already successful on the market and continuously subject to adjustments of the worldwide demands in the shortest possible time are widely supported by the discussed procedures.

Another effect of the research is the use of these technologies for the optimization of products and for the computer supported documentation of (existing) engineering parts via 3D-Digitizing (re-engineering). The most important result is the computer stored final product model for further utilization.

The need of successful development of innovative, high-quality products for the worldwide market challenge design development engineers to study the current technologies [8]. Besides the knowledge about these processes the practical work in college and through continued education is necessary.

REFERENCES

1. Geiger, M.; Steger, W., „Rapid Prototyping-Verfahren, Anwendungen, Beispiele", Fraunhofer-Institut für Produktionstechnik und Automatisierung (IPA), Stuttgart,1994
2. Grote, K.-H.; Birke, C. „CAD, Rapid Prototyping und 3D-Digitalisierung als Konstruktionshilfen", Konstruktion, Vol. 48, 5(1996), pp. 137 - 142
3. Grote, K.-H.; Beyer, C.: „Computer Supported Product Development Through Integration of CAD, Rapid Prototyping and 3D-Digitizing Helps the Productivity of Former East German Companies", Poceedings of the 1996 ASME Design Engineering Technical Conferences and Computers in Engineering Conference, Irvine, California, 1996
4. Kochan, D., Solid Freeform Manufacturing - Advanced Rapid Prototyping, Elsevier, Amsterdam; London; New York Tokyo: 1993
5. Krause, F.-L.; Beitz,W. „Produktentwicklung mit Simultaneous Engineering", FACTS, Suppl. in Konstruktion 45 (1993)
6. Maschkowski, P. „L.O.M. - Laminated Object Manufacturing", Intelligente Produktionssysteme - Solid Freeform Manufacturing, International Conference at the Technical University Dresden, pp. 97 - 104, Dresden, 1994
7. Pahl,G.; Beitz,W., Konstruktionslehre: Methoden und Anwendung, 3rd edition, Springer,Berlin, 1993
8. Pahl, G.; Grote, K.-H.: „Interdisciplinary design - Knowledge and ability needed" Interdisciplinary Science Reviews, Vol.21, 4(1996), pp. 292 - 303

A computation efficient slicing algorithm for rapid prototyping

S. H. Choi
Dept. of Industrial and Manufacturing Systems Engineering
The University of Hong Kong
Pokfulam Road, Hong Kong.

ABSTRACT

This paper introduces an algorithm for slicing CAD models for rapid prototyping. The algorithm enhances computation efficiency, particularly in terms of slicing speed and the use of computer memory. If a .STL file is too big, the algorithm can split it into a number of temporary .STL sections which are small enough to be accommodated in the computer memory for slicing. The algorithm ensures smooth transition from one section to the next one during the slicing process. In the slicing process, an adjacency list is created. An adjacency list is a reconstruction of the topological relationship of the facets. It eliminates the need to sort the contour intersection points, and hence enhances the overall efficiency of the slicing process.

KEYWORDS

Rapid prototyping, slicing algorithm, .STL format, adjacency list, sorting.

1. INTRODUCTION

A slicing algorithm can be considered as a cutter that is used to cut a computer graphical model into thin layers with a desired thickness. A prototype may be subsequently built from the contour data of these layers on a rapid prototyping machine, which may be SLA, SLS or LOM [1-8], etc.. It may be said the overall efficiency of a rapid prototyping process, from CAD modelling to realization of a prototype, is sometimes largely affected by the effectiveness of the slicing algorithm.

There is some new development of direct slicing of a CAD model. But most rapid prototyping systems use a slicing algorithm that takes a CAD model in .STL format, which defines the model with a series of triangular surface facets [9-10]. An advantage of .STL format is that it allows universal portability of CAD models to commercially available rapid prototyping systems. However, there are geometrical and topological problems when processing a facet model. These include:

(1) Unbound surface due to gaps between facets;
(2) Non-manifold edges as more than two facets being adjacent to a same edge;
(3) Intersections of facets due to self-intersecting solids, or intersections of multiple solids.
(4) A lack of topological relationship due to random scattering of facets;
(5) The large size of .STL files.

The first three problems are generally affected by the stability of the CAD system, while problems (4) and (5) are due to the inherent nature of .STL files. Because of random scattering of facets, it would be necessary to read in the whole .STL file for slicing, which may be too large to load into the computer memory. In these situations, the size of the .STL file becomes a bottleneck of rapid prototyping. This paper introduces an algorithm that enhances the speed of slicing and alleviates the problem of memory limitation.

The algorithm first scans through a .STL file. If the .STL file cannot be loaded into the memory, it is split into a number of temporary .STL sections. The size of each section is adjusted so that it can be accommodated in the computer memory for slicing. The algorithm ensures smooth transition from one section to the next one. In the slicing process, an adjacency list, which is a reconstruction of the topological relationship of the facets, is generated to enhance the slicing speed by eliminating the need to sort the contour points. This slicing algorithm has been implemented on AutoCAD Development System. It is found to be quite effective despite improvements are needed to make it more flexible and user-friendly.

2 THE SPLITING ALGORITHM

2.1 Characteristics of .STL file

A .STL file is a data file which represents the surface of a CAD model with a series of triangulated facets. It consists of similar blocks of data, as follows:

```
facet normal 1 0000000e+00 0 0000000e+00 0.0000000e+00
   outer loop
      vertex 5.0000000e+00 7 0000000e+00 1 4000000e+01
      vertex 5.0000000e+00 5.0000000e+00 1.2000000e+01
      vertex 5 0000000e+00 7.0000000e+00 1.2000000e+01
   endloop
 endfacet
```

Two kinds of data are used to represent a facet, namely the facet normal vector and the 3 vertices in terms of x, y, z co-ordinates. An important characteristic of .STL files is that the triangular facets are unordered in sequence without obvious topological relationship. This characteristic requires the .STL file to be read in its entirety into the computer memory for slicing. It is therefore quite common that a .STL file may be too big to be processed, and hence becomes a bottleneck of the complete prototyping process.

2.2 The .STL file Splitting Algorithm

The unordered nature of the triangular facets, on the other hand, facilitates splitting the .STL file into smaller .STL sections, which may then be processed in the computer memory. Fig. 1 shows the flowchart of the proposed algorithm for splitting a .STL file. This algorithm consists of the following steps:-

a. Scan the .STL file to find the vertical height along Z-axis, and create File 1 and File 2;
b. Define a X-Y cutting plane at 1/3 of the vertical height from the bottom;
c. Read in one facet from the .STL file, and check its vertical position relative to the cutting plane;
d. If the facet is below the cutting plane, output the facet to File 1. If the facet is above the cutting plane, output the facet to File 2. If the facet intersects with the cutting plane, split the facet into three smaller facets and output them to File 1 and File 2 accordingly. Record the intersection points;
e. Repeat steps c and d until all the facets are finished.
f. Link up the intersection points with triangulated facets and output to File 1 and File 2 to cover up the top surface and bottom surface respectively.
g. Using File 2, define a X-Y cutting at 1/2 of vertical height and apply slicing process to obtain the layer contour of the prototype at mid-height. Store this layer contour, and rewind the .STL file and File 1 and File 2.
h. Using the X-Y plane at 1/2 of vertical height, repeat steps c to f. Now the .STL has been split at mid-height into 2 sections, and the layer contour obtained at step g ensures smooth transition from one section to the other.

The proposed algorithm is computationally simple because it mainly involves comparisons of z-coordinates of the facets. Using this algorithm, a large .STL file can be easily split up into a number of small .STL files, which can be fitted into the computer memory for slice processing to construct the prototype.

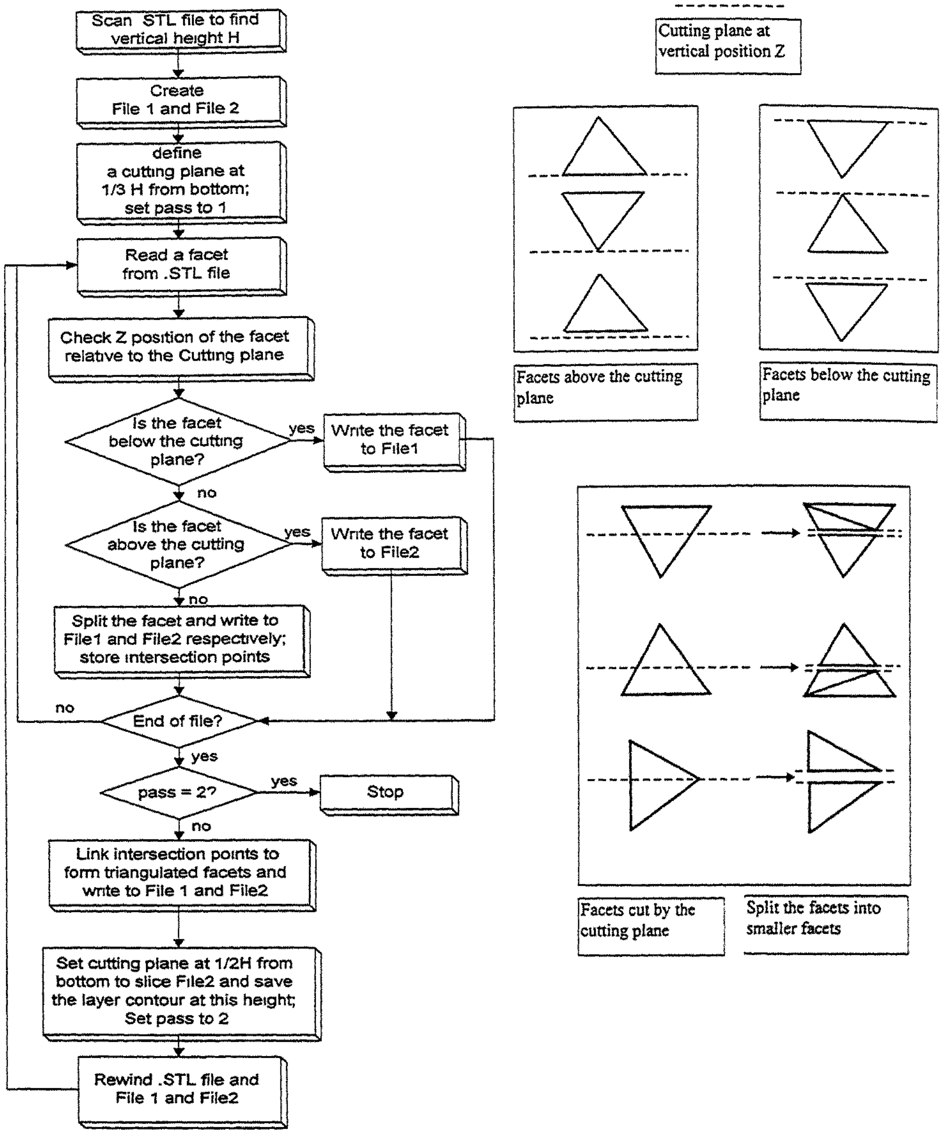


Fig. 1 The flowchart of the proposed algorithm for splitting .STL files

3 THE SLICING PROCESS

The random scattering of facets presents much difficulties in slicing, particularly the speed of sorting the intersection points to form a meaningful contour. To alleviate this problem, a method for reconstruction of topological relationship of facets is proposed to facilitate the slicing process and to eliminate the need of sorting the intersection points.

3.1 The Adjacent List of Facets - Reconstruction of Topological Relationship

Assume that the triangulated surface is bound and that no more than 2 facets share a common edge. Each facet should has 3 edges, and hence 3 adjacent facets. A pair of adjacent facets is defined as two facets which share a common edge. In Fig. 2, the 12 facets of a cube have 12 sets of adjacency information, as shown in Fig 3.

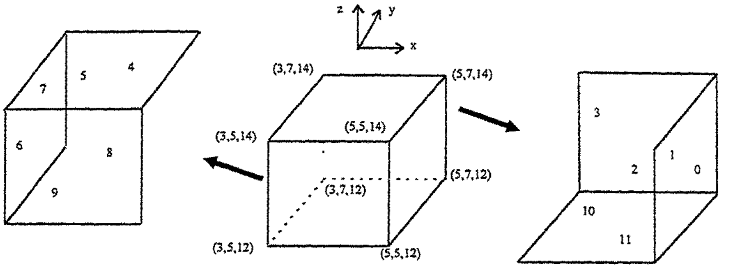


Fig. 2 The facets of a cube

An adjacency list gives the topological relationship of the facets. It is constructed by searching and matching the common edges shared by a pair of adjacent facets.

Facet Number	Adjacency of Edge $E_{0,1}$	Adjacency of Edge $E_{1,2}$	Adjacency of Edge $E_{2,0}$
(0)	1, 1	11, 2	2, 2
(1)	4, 2	8, 0	0, 2
(2)	0, 1	10, 0	3, 1
(3)	7, 0	4, 0	2, 1
(4)	1, 2	3, 0	5, 1
(5)	7, 2	8, 1	4, 1
(6)	9, 2	7, 1	10, 2
(7)	5, 2	3, 2	6, 0
(8)	9, 1	1, 0	5, 0
(9)	6, 2	11, 0	8, 2
(10)	6, 1	2, 0	11, 1
(11)	0, 0	9, 0	10, 1

E_{ij} is the edge formed by vertices V_i and V_j,
Adjacency of Edge $E_{i,j}$ is given as a pair of integers, (m, n), where
m is the number of the adjacent facet;
n is the number of the edge of the adjacent facet common with E_{ij}. n is represented by:
0 to denote edge E_{12} of the adjacent facet
1 to denote edge E_{20} of the adjacent facet
2 to denote edge E_{01} of the adjacent facet.

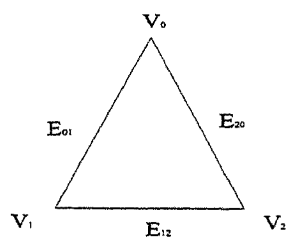


Fig. 3 The adjacency list of the facets of the cube

Referring to Fig. 4, let us consider Facet 0. Its adjacency information is as follows:

Facet 0 1, 1 11, 2 2, 2

The first edge (E_{01}) of Facet 0 matches with edge E_{20} of Facet 1, its second edge (E_{12}) matches with edge E_{01} of Facet 11, while the third edge (E_{20}) matches with edge E_{01} of Facet 2. The adjacency information matches in pair. Hence for Facets 1, 11, and 2:

Facet 1	4, 2	8, 0	*0, 2	*(* Edge E_{20} of Facet 1 matches with edge E_{01} [n=2] of Facet 0);*
Facet 11	# 0, 0	9, 0	10, 1	*(# Edge E_{01} of Facet 11 matches with edge E_{12} [n=0] of Facet 0);*
Facet 2	@ 0, 1	10, 0	3, 1	*(@ Edge E_{01} of Facet 2 matches with edge E_{20} [n=1] of Facet 0).*

3.2 Using The Adjacent List For Slicing

The steps of using the adjacent list for slicing is summarised as follows:

Step 1	Set the cutting plane to the bottom of the prototype;
Step 2	Find the first facet that intersects with the cutting plane;
Step 3	Calculate the points of intersection;
Step 4	From the adjacency information, get the adjacent facet and calculate another intersection;
Step 5	Repeat Step 4 until the adjacent facet loops back to the first one;
Step 6	Search for another set of contour and repeat Steps 2 to 5 until no more facet intersects with the cutting plane;
Step 7	Set the cutting plane upwards by a thickness required and repeat from Step 1 until the cutting plane reaches the top of the object.

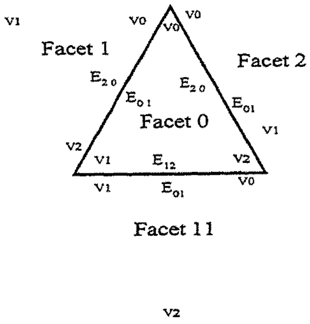


Fig. 4 The adjacent facets of Facet 0

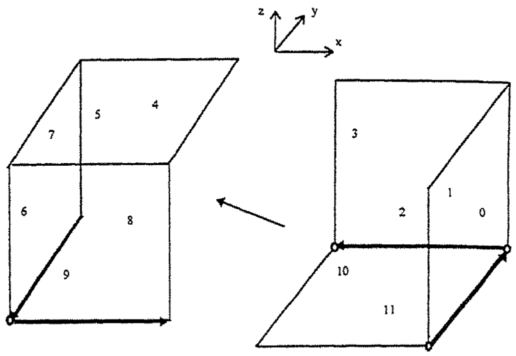


Fig. 5 Slicing at the bottom of the cube

This procedure is illustrated by slicing the cube, as shown in Fig. 5.

Step 1. Set a X-Y cutting plane to the bottom of the prototype

A prototype may be built from the bottom to the top along the Z direction. Therefore, setting a X-Y cutting plane is set at minimum Z.

Step 2. Find the first facet that intersects with the cutting plane

The facet number is assigned according to the sequence in which the facet appear in the .STL file. The facets are done on a first-read-first-processed basis. Hence the first facet of the cube in Fig. 5 that intersects with the cutting plane is Facet 0.

Step 3. Calculate the points of intersection

An edge of a facet is cut by the cutting plane only when the intersection lies on the edge. From the adjacent information, it can be seen that the first point of intersection lies on Edge E_{01} of Facet 0, and the second point of intersection lies on Edge E_{20}.

Step 4. From the adjacency information, get the adjacent facet and calculate another intersection

The adjacent facets of Facet 0 are Facets 1, 11 and 2. The second point of intersection of Facet 0 lies on Edge E_{01} of Facet 2. Hence Facet 2 is the next facet that should be processed to obtain another intersection point on Edge E_{12}. Up to now the slicing process has traversed from Facet 0 to Facet 2.

Step 5. Repeat Step 4 until the adjacent facet loops back to the first one

From the adjacency information of Facet 2, the next intersecting facet is Facet 3. To complete closed contour, Step 4 is repeated until the cutting plane cuts back to Facet 0. The sequence of the facets sliced is:

Facets 0 -> 2 -> 3 -> 7 -> 6 -> 9 -> 8 -> 1 -> 0

Step 6. Search for another set of contour and repeat Steps 2 to 5 until no more facet intersects with the cutting plane

A prototype is in general much more complex than the cube. It may contain cavities or solid features at various locations. In these cases, there should be more than one set of closed contours on a layer. This slicing algorithm search for another set of contour by flagging all facets which have been processed before, and then search for a new facet which intersects with the cutting plane. To complete a layer, Steps 2 to 5 are then repeated until all the facets have been flagged.

Step 7. Set the cutting plane upwards by a thickness required and repeat from Step 1 until the cutting plane reaches the top of the object

Raise the cutting plane by a thickness required and unflag all the facets. Repeat Steps 1 to 6 to obtain a new layer.

4 CONCLUSION

An algorithm which enhances slicing speed and alleviates the limitation of computer memory has been proposed. The algorithm creates an adjacency list to simplify the slicing process. It can also split a large .STL file into smaller sections which may then be accommodated in the computer memory for slicing. It is expected to help enhances the overall efficiency of rapid prototyping processes.

5 REFERENCES

1. Ashley, S., Rapid Prototyping Systems, Mech. Eng. Vol. 113 (4), pp34-43, April 1991.
2. Bagchi, A. and Beaman, J.J., Intelligent design and manufacturing for prototyping, PED-Vol. 50, ASME 1991.
3. Kennerknecht, S. and Sifford, D., New dimensions in rapid prototyping explored for aluminum investment castings, INCAST, Vol IV, No. 3 pp5-10, March 1990.
4. Marks, P., The rapid prototyping revolution.. Better products sooner, Proc. The first international conference on desktop manufacturing, Cambridge, Masschusetts, Oct. 1990.
5. Deckard, C. and Beaman J., Process and Control Issues in Selective Laser Sintering, Proc. Symposium on sensors and controls for manufacturing, pp34-43, PED-33 Nov, 1988.
6. Lindsay, Karen F., Rapid prototyping shapes up as low-cost modeling alternative, Modern Plastics, pp40-43, Aug, 1990.
7. Sachs, E., P. Williams, et al, Three-dimensional printing: rapid tooling and prototypes directly from a CAD model, Proc. Manufacturing International '90, pp131-136, Vol 4, 1990.
8. Gargiulo E.P., Belfiore, D.A., Photopolymer solid imaging process accuracy, pp81-95, PED-Vol. 50, ASME 1991.
9. Wang Y.G., Rapid Photo-machining and the implementation of a slicing algorithm, Chinese National Journal of Electronics and Machinery, Vol 3, 1993.
10. Dolenc A. and Makela I., Slicing procedures for layered manufacturing techniques, Computer-Aided Design, pp119-126, Vol 26, 1994.

RAPID PRODUCT DEVELOPMENT USING A RP-CENTERED INTERACTIVE DESIGN PROCESS

Hsien-Jung Wu
Dept. of Industrial Design, Tunghai University
Taichung, TAIWAN

ABSTRACT

To survive in the global market, companies must get products to market faster than ever before. The ability to quickly reduce time it takes to get new products to market has increased the pressure on all phases of the product development process. Therefore, companies are introducing concurrent engineering into their management practices to handle time compression. The most important technological tools to help concurrent engineering are the use of 3D computer-aided design and the various rapid prototyping (RP) techniques. With the use of RP techniques, problems of evaluating complex design before manufacturing are overcome. These techniques transform a product design into a physical 3D model as a substitute for traditional mock-up used for evaluating design concept and modifying design if required. Consequently, product designers should adjust their design process to respond the design change based on the result of the RP model. This paper proposes a RP-centered interactive design process to improve rapid product development in product design and concurrent engineering systems. This process emphasizes the interaction between product design and rapid prototyping techniques on the design phase of the product development cycle.

KEYWORDS

Rapid prototyping, interactive design process, product development, concurrent engineering

1. INTRODUCTION

In the product development process, designers and manufacturers must find and implement time saving throughout this process without sacrificing quality. To meet the functional requirement of a product and satisfy consumer's need, the concept of concurrent engineering (CE) is introduced to integrate design and other factors in the product development cycle. "Concurrent engineering is the systematic approach to the integrated, concurrent design of products and related processes including manufacture and support"[1]. Concurrent engineering aims to reduce the number of design changes and to detect manufacturability problems in the early design stages. Therefore, a fundamental concern in concurrent engineering is the need to provide the product life-cycle viewpoint in suggesting design alternatives. Research on concurrent engineering issues can be found in [2], [3], [4].

Since the development of product design in a concurrent engineering environment requires the consideration of life cycle factors, designers need to reflect particular attributes of those factors such as manufacturing, marketing, and maintenance. From the viewpoint of product designers, the most essential concern in concurrent engineering is the integration between design and manufacturing. Research on design in the concurrent engineering environment can be found in [5], [6], [7], [8]. Before real molding and manufacturing process, traditionally a mock-up is made to evaluate a product design and its manufacturability, and ensure product quality. However, the operation of making a mock-up is time-consuming and requires well-trained technicians. To solve this problem, rapid prototyping (RP) techniques are developed to reduce the time of product development.

Rapid prototyping is also known as 3D plotting, 3D printing, desktop manufacturing, and free-form fabrication. The RP techniques develop product shapes by building up materials in layers. The layers are slices cut by computer from a software model of the original design. The techniques differ in the way of forming the slices, such as Stereolithography (SLA), Selective laser sintering (SLS), Laminated object manufacturing (LOM), Solid ground curing (SGC), and Fused deposition modeling (FDM)[9]. With the use of RP, by catching design errors earlier, it is possible to fix them with much less time and cost. Most rapid prototyping systems have limited capabilities for functional testing due to the materials they use are no match for the materials that are used in real products. However, recent development of RP technology can directly output prototype parts made out of durable materials that perform similarly to real materials. Description related to RP technology can be found in [10].

2. THE ROLE OF RP IN THE PRODUCT DEVELOPMENT

Rapid prototyping is a group technologies which allow computer-generated data to be quickly turned into physical prototypes as a substitute for traditional mock-up. It has close connection with the application of computer-aided-design (CAD). Rapid prototyping systems depend on CAD data to produce the parts. The rapid prototyping system creates a physical duplicate of the computer-aided design. Therefore, the 3D CAD solid or surface model is the founding point of the RP technology. The CAD date are then converted to the .STL file format through a translator. The .STL file is a standard interface between 3D CAD solid models and the RP systems. These prototypes are used as visual review models for design and aesthetic purpose. In some cases they can be further used as functional prototypes and even sales samples.

In terms of product development process, productivity is achieved by directing a product from concept design to market quickly and economically. Rapid prototyping technology assists this process and automates the creation of a prototype part which communicates more information about the product earlier in the development cycle. Unlike conventional prototyping may take weeks or even months, rapid prototyping is a cost-effective way of constructing prototypes as opposed to conventional mock-ups.

3. DESIGN AND MANUFACTURING INTEGRATION

Many industrial manufacturing problems can be traced back to product design process. Substantial reduction in manufacturing costs can result from revisions at the design stage that efforts can crucially influence the achievement of a product. There is a realization that many manufacturing problems stem form product design are difficult and expensive to manufacture. Therefore, the design process is arranged to create products which fulfill a number of criteria: aesthetic requirements, functional satisfaction, and product quality and reliability. Since the rapid prototypes are used either for aesthetic visualization or functional testing, it can be observed as a tool of integrating design and manufacturing operation. From design drawings (3D CAD files) to prototypes, rapid prototyping technology provides product designers with design advice which responds to manufacturability and marketability concerns.

The concept of design for manufacture (DFM) represents the awareness of the importance of product design as the first manufacturing step in the concurrent engineering environment. The DFM approach preserves communication between all components of the manufacturing system and maintains flexibility to modify design. As a critical step prior to real manufacturing operation, rapid prototyping technology inspires product designers to implement DFM concept through responding the result shown in prototypes where possible errors are caught and corrected in much less time-consuming manner. Consequently, real implementation of rapid product development can be anticipated. In response, a RP-centered interactive design process is required to take the advantage of the time compression provided by the technology.

4. INTERACTIVE DESIGN PROCESS

In the traditional design process, most of time are spent in design rework and compromise among life-cycle factors (manufacture, performance, operability, etc.). As a matter of fact, significant cost benefits can be obtained by reducing design rework. In similar manner, the overall product quality can be completed by reducing compromise solutions. A RP-centered interactive design process is proposed in this paper to accomplish these reductions. This design process can be described as a two-way interactive design process, where two-way means that it consists of the design phase and the RP-response phase. The structure of this RP-centered interactive design process is shown in Figure 1.

In the design phase, product functions are decomposed into several subfunctions depended on the functional requirements. It addresses the conceptual design aspect and design rework of product design. In the RP-response phase, however, steps are executed in a reverse direction. The purpose of this arrangement is to reflect the characteristic of rapid prototyping used in a concurrent engineering environment to achieve time reduction in terms of rapid product development. Basically, design process is defined in the first level as a function that assigns design parameters to particular product requirement.

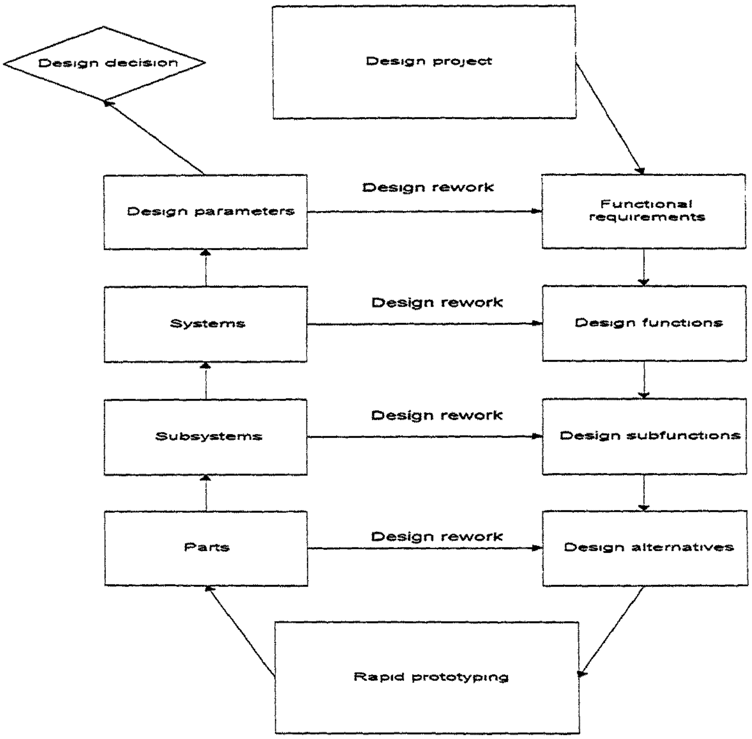


Figure 1. The structure of RP-centered interactive design process

In this design process, conceptual and innovative tasks are finished in the beginning iteration at the right-hand-side design phase in a top-down manner. After the result of rapid prototypes are examined, design rework are then executed (if errors are detected) at different levels based on the composition and in a bottom-up manner at the left-hand-side RP-response phase. The decomposition reflects the review capacity offered by rapid prototyping and different levels may indicate different emphasis such as aesthetic requirement, material features, and manufacturing constraints. Design rework involves the file transfer between RP systems and 3D CAD files. Since design rework is decomposed and done at different levels, the goal of time compression is achieved.

5. CONCLUSIONS

This paper has presented a RP-centered interactive design process. This process tends to utilize the RP technology and reduce the product development cycle time to assure that conceptual idea of the product is properly reflected in the final design decision. Two-way implementation improves the interaction between design and rapid prototyping, and therefore integration between design and manufacturing is enhanced. This paper has advanced the traditional design process to a integrated approach to product design. The advantage of using this interactive design process includes:

1) incorporate the design process with RP technology,

2) decompose product design into various levels to reduce design rework, and

3) achieve rapid product development through the use of RP.

This work aims at assisting product designers to manipulate rapid prototyping technology during their design process. With modifying their routine process by decomposing possible design rework into different levels in the design process, the product development cycle time can be reduced and the goal of rapid product development can be expected

REFERENCE

1. IDA, 1988, The role of concurrent engineering in weapon systems acquisition, IDA Report R-338, December.

2. Kim, K., Cormier, D.R., O'grady, P.J., & Young, R.E., 1995, System for design and concurrent engineering under imprecision, Journal of Intelligent Manufacturing, 6(1), pp 11-27

3. Dowlatshahi, S.,1994, A comparison of approaches to concurrent engineering, International Journal of Advanced Manufacturing Technology, Vol.9, pp.106-113.

4. Syan, C.S. & Menon, U., 1994, Concurrent Engineering - Concepts, Implementation and Practice, Chapman & Hall, Inc., London.

5. Dowlatshahi, S., 1994, Morphological approach to product design in a concurrent engineering environment, International Journal of Advanced Manufacturing Technology, 9(5), pp.324-332.

6. Duda, J., 1994, Modeling of the decision process in a advisory system for technological, International Journal of Advanced Manufacturing Technology, Vol.9, pp.13-19.

7. Harrington, J.V., 1995, Negotiation in a knowledge-based concurrent engineering design environment, Expert Systems, 12(2), pp.139-147.

8. Pourbabai, B. & Pecht, M., 1994, Management of design activities in a concurrent engineering environment, International Journal of Production Research, 32(4), pp.821-832.

9. Majer, 1994, The slice age, Design, February, pp.36-38.

10. Dickens, P.M., 1994, Rapid prototyping of physical parts, in Concurrent Engineering (Syan & Menon eds), pp. 137-150.

11.Baxter, M., 1995, Product Design, Chapman & Hall, London, pp.271-298

A PATH PLANNING ALGORITHM FOR ROBOT MANIPULATORS IN FMS ENVIRONMENT

Ma Yulin Wang Wei* and Fu Yili
Robot&FMS Research Institute
Harbin Institute of Technology
Harbin, P.R.China

ABSTRACT

This paper describes a path planning algorithm for robot manipulators in FMS environment. Not only the motion of the robot but that of the manufacturing cell is considered, the robot path planning is integrated into the manufacturing cell.We present a fast computation method for C-obstacle boundary based on the critical collision joint angle(CCJA) of the robot arms, although the robot arms are substituted by cylinders or rectangular solids, our algorithm does not require that obstacles be expanded and the robot arms be shrunk to line segments ,only the computation of CCJA of every possible collision point is needed. A dynamic step-changeable A* algorithm with a goal-visible-test is employed for path searching, although the path is slightly sub-optimal, a tremendous speed increase is achieved.The planner runs on a Silicon Graphic 4D/70 workstation and has been implemented on a flexible manufacturing cell.

KEYWORDS

Robot Path Planning, C-obstacle, A* Algorithm, Simulation

1. INTRODUCTION

Robot manipulators in FMS environment are used for transporting objects, completing pick-and-place operation,etc.. The presence of obstacles will constrains the movement of a robot, most commercial robots currently have little capability of thinking and decision-making. Inefficient teaching restricts the robot application for factory automation. A relevant research effort is the development of some intelligent capabilities of robots, such as sensing and autonomous motion planning. The ultimate goal is to provide the robot with such control devices that it can accomplish a variety of complex tasks, under conditions known or incompletely known, assigned in terms of desired goals, with little specification of the actions the robot must performs.

Robot motion planning is essential to increase robot autonomy. Many researchers have been involved in exploring rapid and efficient generation of collision-free path.The methods developed include graph searching[1][2][3][4], potential field[5], control algorithm[6] and topologic method[7].

Space limitations preclude a more detailed discussion of the contributions of the authors cited above.Our work was motivated by three observations. First, the robot motion planning is intergrated into the manufacturing cell. Second, we have solved the C-space building problem by using CCJA method.Third, to develop the search process in C-space, our approach makes use of a dynamic step-changeable A* algorithm with a goal-visible-test, and the searched path is modified by a path modifying method.

2. COMPUTATION OF C-OBSTACLE

Many path planning algorithms are based on C-space, from the seminal work of Lozano Perez on

*To whom all correspondence should be addressed.

configuration space[1]. This approach can be grossly describied as being comprised of two phases of during which a data structure is first built that represents the geometric constrains imposed by the obstacles,and then search for a solution.Typically, the process of building the data structure consists of mapping the obstacles from the robot's workspace into the robot's C-space.The C-space is the joint space for robot manipulators.The result of this operation is an explicit representation of all robot configurations that don't result in a collision with some obstacle, that is the robot's free space.

2.1 Description of Obstacles

To compute the C-obstacle conveniently, the enviroment objects are assumed to have verticall sides and horizontal or inclined top/bottom. Actually, many objects can be divided and considered as a composition of several polyhedral objects with vertical sides. An obstacle can be represented by a prism that is a convex hull of the real obstacle.In the other hand, an obstacle can be divided into more prisms, then the difference between the real obstacles and the resulting prisms will be reduced, obtaining a more accurate description, obviously at the price of augmenting the number of prisms and thus the computation time.The description of the obstacle can cause little loss of free space.

All objects are checked to see if they are within the workspace of the robot.Objects within the workspace of the robot are considered as obstacles. Otherwise, they are not obstacles for the robot. For the robot path planning, only obstacles are under consideration.

2.2 Transforming Obstacles to Tori

Most commercial robot manipulators, for example, the PUMA560 and the MOTOMAN-K ,have planar linkages as part of the structure. For robot manipulators, the links can be divided into two groups: the major linkage which positions the end-effector, and the minor linkage which orients the end-effector. For most commercial robot manipulators, the major linkage is a planar 2-DOF linkage, which is in a vertical plane determined by the Z-axis of the robot fixed fram. For articulated robots, including the two just mentioned, we take advantage of the configuration of the main linkage, the first joint determines a horizonal plane, the second and third joint move in a common vertical plane or two parallel vertical planes. According to the features, the obstacles can be approximated with tori centered at the orign of the robot system.This allows us to consider the arm as planar polygon and to operate in a plane, which is determined by the value of the first joint, so path planning in 3D can be accomplished by partitioning the first joint angle.The approximation of obstacles greatly speeds the computation.The degrees of freedom are reduced to only three and the last three joints are considered as free and includes all their possible positions in the polyhedra of the robot hand.This method is a general one and is applicable to most of the commercial robots, by simply changing the corresponding models and their kinematic transformations.

2.3 Critical Collision Joint Angle(CCJA)

A planar linkage is considered because of the observation mentioned above.There can have two kinds of collisions between the polygon obstacle and the rotating link[1]: a vertex of the rotating link with an edge of the obstacle(VE contact), and an edge of the rotating link with a vertex of the obstacle(EV contact). The cases of contact edge to edge or vertex to vertex can be included in EV or VE contacts.Lozano Perez computed the slice projection by VE and EV contacts, our purpose was to compute the boundary of the C-obstacle by VE and EV contacts, the C-obstacle is represented by its boundary not the slice projection.For link K, assume that the previous joint are fixed at certain joint angles, when link K collides with a point, there must be two critical angles: the first angle at which link K contacts the point when link K rotates in a counterclockwise direction, we call the angle lower critical collision joint angle(lower-CCJA), another angle is similar to lower-CCJA, it is formed when link K rotates in a clockwise direction and called upper critical joint angle(upper-CCJA).They are illustrated in Fig.1, angle A is the lower-CCJA and B is the upper-CCJA.

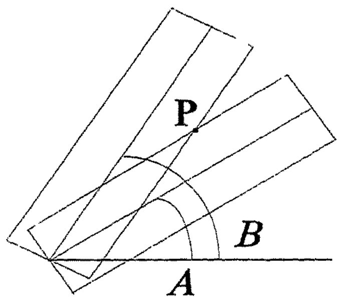


Fig.1: The critical collision joint angles(CCJA).

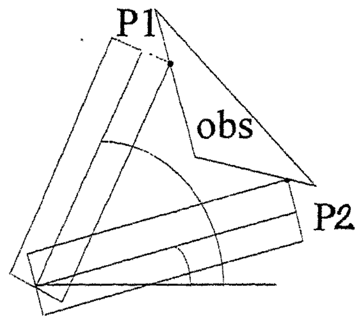


Fig.2: The valid CCJA for an obstacle.

When link K touches a polygonal obstacle(VE or EV contact), there will be a collision point P, the two collision points P1 and P2 as shown in Fig.2.We can compute lower-CCJA and upper-CCJA for a collision point, but only one CCJA is valid for a collision point, because the practical collision can not happen for another CCJA.In Fig.2, for example, the upper-CCJA of the collision point P1 is valid and considered as the upper bound of the C-obstacle of the obstacle, and the lower-CCJA of point P1 is invalid. For collision point P2, lower-CCJA is valid and considered as the lower bound of the C-obstacle.

The link rotates about its joint, the link vertices have circular paths, the intersection points of these paths with obstacle edges are possible collision points.The vertices of the obstacle are also possible collision points. Lozano Perez[1] gave the sufficient condition for a contact between the link and the obstacle.We will see in section 2.4, the test for sufficiency is not necessary in our algorithm.

2.4 The Boundary of C-obstacle

For a planar two-link arm, any point P ∈workspace can correspond either one ,two,or an infinite number of points in C-space.Any C-obstacle has a closed curve boundary, i.e. C-boundary, and any C-boundary is formed by simple closed curves, the joint angles are the configuration parameters.

We describe the links by projecting in the planes perpendicular to the joint axes of the first three joints. A arm projection is a polygon. For an arm ,it has two projections respectively on the plane perpendicular to the first joint axis and on the plane perpendicular to itself joint axis, the projection polygon is descibed as the consecutive vertices.

An obstacle has been transformed into a torus. The rotation of the first joint is divided into angular intervals in $\theta 1$ plane, the vertical plane determined by the Z-axis of the robot's fixed frame may cut through the torus and will generate a rectangular cross-section(RCS) which extends for the angular interval determined by the torus in $\theta 1$ plane. For RCS, all of it or only a part of of it is in the workspace, we must compute the part of RCS in workspace, which is named the true cross-section(TCS), the boundary of TCS consists of line segments and arcs. In the vertical plane, we can calculate the possible collision points of TCS for joint 2 and joint 3. For every possible collision point, we compute the lower-CCJA and upper-CCJA, and get min(lower-CCJA) and max(upper-CCJA) as the lower bound point and the upper bound point of the C-obstacle respectively, calculate all possible lower bound points and upper bound points, the boundary of the C-obstacle consists of the lower bound points and the upper bound points.The procedure consists of the following steps:

(1) Transform an obstacle to a torus.

(2) Get $\theta 1$ interval $\Delta \theta 1$ in $\theta 1$ plane.

(3) Compute RCS in vertical plane and obtain TCS from RCS.

(4) Compute all possible collision points by TCS.

(5) Compute the lower-CCJA and upper-CCJA for every possible collision point.

(6) Get the lower bound point and upper bound point of the C-obstacle boundary by min(lower-CCJA) and max(upper-CCJA) respectively.

(7) Go to step 5, if over the joint limits, the procedure stops.

One example of the algorithm's use are given. Fig.3 shows a wall corner obstacle(Fig.3a) and its C-obstacle(Fig.3b)

(a) (b)

Fig.3: The C-obstacle of a wall corner obstacle

3. PATH PLANNING ALGORITHM

The C-space can be computed explicitly easily with the CCJA method mentioned above, the next course is to search the collision-free path through a grid in the C-space.The A* algorithm[8] is frequently employed for searching the free space, however, even with a bidirectional search, the A* turns out to be rather slow.The alternative presented here is to use dynamic step-changeable A* algorithm with a goal-visible-test.With this method, although the path is slightly sub-optimal, a tremendous speed increase is achieved.

For a node i to be expanded, let the basic step(Bstep) be the distance between node i and its adjacent nodes, let dynamic step(Dstep) be K * Bstep. Frst, the nodes with Dstep will be expanded, if no reasonable node with Dstep exists, then nodes with the distance (K-1) * Bstep should be expanded,the process is continued untill a path is found.

When a node is expanded, a goal-visible-test is made. If a safe path exists from this node to the goal, then end the search, else the search is continued. This extra test will not slow down the process, because once the goal is reachable from current node, tremendous nodes to be expanded will not be expanded.

The algorithm is described broadly as follows:

(1) Let the start configuration S be the root node of the search tree, and put it on OPEN list.

(2) If OPEN list is empty, no solution, exit with failure.

(3) Take the first node i of OPEN list, put it on CLOSED list.

(4) If node i is the goal configuration, get the path from CLOSED list and exit succesfully.

(5) Make a goal-visible-test for node i, if the goal configuration is visible, obtain the path from CLOSED list and exit successfully.

(6) Expand node i with dynamic step Dstep=n * Bstep.

(7) If no cell path with Dstep exists, n=n-1, go to step 6.

(8) Sort OPEN list by the value of the evaluation function f.

(9) Go to step 2.

Where, n is the max step coefficient. f=(1-w)*g(i)+w*h(i), g(i) is the cost of the path from the start configuration to node i, and h(n) is the estimate of the cost of moving from node i to the goal configuration.

The searched path consists of many line segments, a algorithm is employed to shorten the path. For a path point i, a check is made to determine how many path points ahead can be reached with a straight path, the middle path points can be omitted.

4. EXAMPLES

Two examples are given in this section. Fig.4 shows the MOTOMAN-K30S robot and two obstacles(i.e. TCS in vertical plane), the path in C-space is shown in Fig.5.The searched path is shown with solid lines,and the modified path with dashed lines.Where, the max step coefficient is 5, w=0.5 in evaluation function,the number of expanded nodes is 325, the CPU time for generating the collision free path is 40ms using a Silicon Graphics 4D/70 workstation. Fig.6 shows the motion of MOTOMAN-K30S. The second example(shown in Fig .7) shows the motion simulation of the manufacturing cell, which is consists of a MOTOMAN-K30S robot, a rotating working table and a NC lathe.

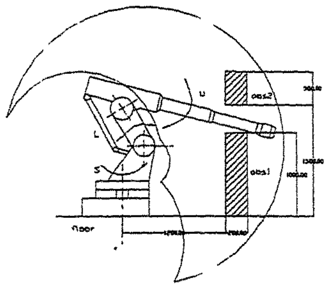


Fig.4 The MOTOMAN-K30S robot and obstacles

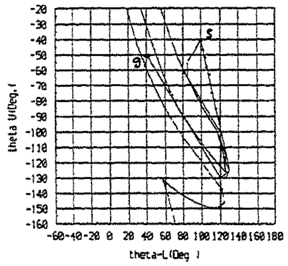


Fig.5 The path in C-space

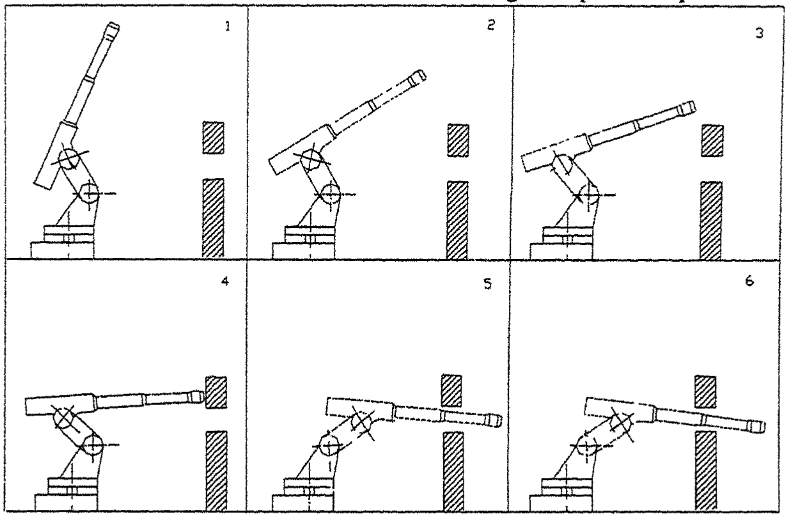


Fig.6 The obstacle avoidance motion of the robot

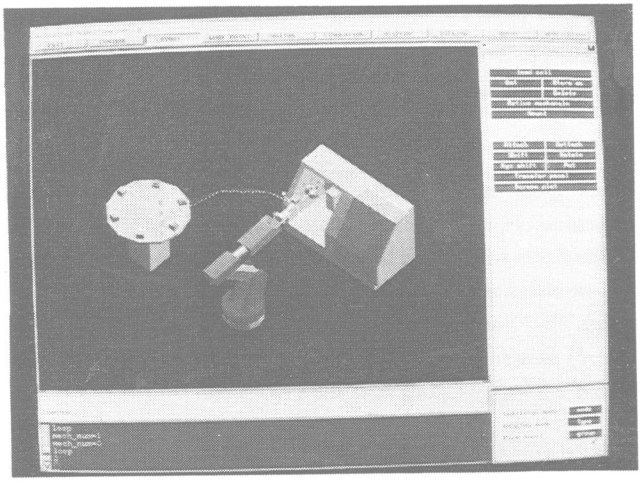


Fig.7 The simulation of the manufacturing cell

5. CONCLUSIONS

Simulations and experiments have shown the algorithm presented here is an efficient method for robot collision-free path planning in a fixed and known enviroment.The computation method for C-obstacle boundary based on CCJA is powerfull.The approximation of obstacles with tori allows us to solve 3D collision-free problem in two projection planes.By employing dynamic step-changeable A* algorithm with a goal -visible-test, a tremendous speedincrease is achieved.

6. REFERENCES

1. Lozano Perez, Spatial Planning: A Configuration Space Approach, IEEE Transactions on Computers,Vol.C-32 No.2 pp.108-120, Feb.1983.

2. Lozano Perez, Asimple Motion-Planning Algorithm for General Robot Manipulators,IEEE J.Robot Automation,Vol.RA-3,pp224-238,June,1987.

3. G.Gini,R.Massa and R.Negretti, Toward Efficient Path-Planning for Articulated Robots, J.of Robotic System ,12(2),pp.93-104,1995.

4. Prabir K. Pal and K.Jayarajan,Fast Path Planning for Robot Manipulators Using Spatial Relations in the Configuration Space, Proc. of the IEEE Int.Conf.on Robotics and Automation,pp669-673,1993.

5. C.W.Warren,J.C.Danos and B.W.Mooring, An approach to Manipulator Path Planning, the Int. J. of Robotics Research, Vol.8,NO.5, pp87-95,Oct. 1989.

6. Zvi Shiller and Steven Dubowsky, Robot Path Planning with Obstacles,Actuator, Gripper,and Payload Constraints,the Int. J. of Robotics Research, Vol.8,No.6,Dec.1989.

7. Ai.Haizhou and Zhang Bo, Simulation Experiment System of Behavior-Based Mobile Robot, Robot,Vol.16,No.2.pp.87-91,March,1994(in Chinese).

8. J.Pearl,Heuristic:Intelligent Search Strategies for Computer Problem Solving,Addison Wesley,1985.

MODEL AND SENSOR BASED CONTROL OF INDUSTRIAL ROBOTS

Ole Madsen, Hans Holm,
Department of Production,
Aalborg University,
Fibigerstræde 16,
DK-9220 Aalborg Øst,
Denmark.

ABSTRACT

This paper presents a system which controls robots that carry out complex production tasks. The control system integrates an off-line programming approach where references to the robots are generated by means of computer models with a sensor based approach, where references are generated on the basis of signals from a sensor, which measures features of the physical world. The concepts of the control system have been industrially implemented in an automatic pipe welding system which is presented briefly in the paper. Experience from this and other implementations of the system concept indicates that the integration of the off-line programming approach with the sensor based approach enables the automation of complex production tasks.

KEYWORDS

Robotics, off-line programming, sensory control.

1. INTRODUCTION

At the Department of Production, Aalborg University, Denmark, a general system, which controls robots that carry out complex production tasks, has been developed. The focus of this work has been on production tasks which has some or all of the following characteristics:

- The product is produced in small batches, even one-of-a-kind.
- The product to be produced has a geometrically complex shape.
- There are relatively large tolerances on the product shape.
- There is a strong and reproducible relationship between the shape of the product, the process control variables, and the result of the process (the quality of the produced product).

There are a number of challenges involved in applying robots to these tasks. First of all, since the batch sizes are small and the robot installations are expensive, the costs and time associated with generating the necessary programs to the robots are critical. Thus, automatic or semi-automatic off-line programming techniques must be used where the robots, the environment, the product to be produced, and the particular process are represented by computer models.

However, the quality of the programs, which are developed on the basis of these models is only good to the extent that the models represent the physical world sufficiently well. Since there are relatively large tolerances on the product shape and since there is a strong relationship between the product shape and process variables to be used, the program must be compensated for the deviations between the modelled product shape and the shape of the physical product.

To carry out the compensations the deviations must be identified by appropriate sensing systems. A number of sensor based systems for control of robots exist today. Some of these sensor based control systems have been used successfully in a number of industrial applications (e.g. ref. 1 and 2). However, in most of the sensor based control systems used in industry, signal conditioning and control vector compensations are based on implicit and/or explicit models embodied in the control system. Normally, these internal models represent standard and often quite simple product shapes.

Such systems cannot be used for processing products that have shapes which are not of a standard product shape type.

Robot control systems have been developed where no programming of the robot is needed (ref. 3 and 4). In these systems a sensor provides the control system with all the necessary geometry references There are, however, a number of problems associated with the application such a control scheme to the manufacturing tasks presented above. First of all, a number of features, which are essential to select appropriate control variables, can only with difficulty be measured (e.g. the plate thickness and material type). Secondly, great efforts must be put on the selection of references for the signal conditioning in order to interpret the sensor measurements correctly. Thirdly, it is difficult with these system to perform collision avoidance checks. Finally, in industrial applications, planning of the manufacturing task is required to provide production scheduling systems with resource and time estimates. To perform these estimations sufficiently accurate, a geometry model based simulation of the manufacturing task is often required.

The control system presented in this paper overcomes the problems mentioned above by integrating elements of an off-line programming system with the sensor based control system. Prior to operation with the control system an off-line planning and simulation of the production task is made. The objectives of the off-line planning and simulation are: 1) to compute time and resource estimates, 2) to perform a collision avoidance check, and 3) to generate the necessary references to the control system. During operation, the control system uses a product shape model from the off-line planning system to provide the control system with references needed for the signal conditioning. This product shape model is constantly updated during operation based on the measurements made by a sensor system so that the product shape model corresponds sufficiently accurate to the shape of the physical product. Control variables are then computed by means of process models which relate the product shape to appropriate control variables. The process models are selected by the off-line planning and simulation system. In this way the system can be used to process different products by changing the product shape and process models. Furthermore, the control system can be used to manufacture products of more complex shapes than the quite simple shapes which can be produced with conventional sensor based control systems.

2. CONTROL SYSTEM ARCHITECTURE.

Figure 1 shows the general architecture of the control system (a more detailed description of the control system can be found in ref. 5). In the figure functional entities of the system are represented by circles, to and from which various signals and data enter and leave (in the figure represented by arrows). The main data entities are represented by pairs of parallel straight line segments. Finally, the figure shows the sources and destinations (terminators) of the data flows. These terminators are represented by boxes and represent other systems which interact with the control system. In the following, first the system terminators are presented, followed by a presentation of the functional entities.

2.1. System Terminators.

As it appears from the figure the system interacts with four terminators. First of with a **Task Manager.** The task manager initiates the control system by sending an identification of the task to be carried out to the control system. This task description contains data related to the modelled product shape (generated by the CAD-system).

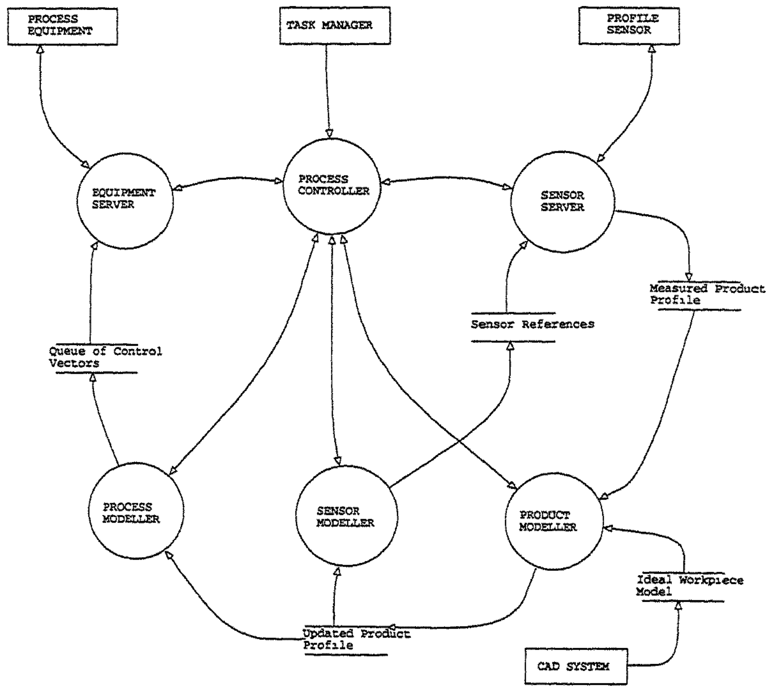


Fig: 1. System outline.

The task of the **CAD-system** is to provide the control system with an ideal model of the shape of the workpiece to be processed.

The task of the **Profile Sensor** is to provide the control system with profile measurements. Different sensor technologies can be used to obtain product profile information. However, most commercially available systems are based on vision systems using structured light principles (as in laser scanners). Figure 2 shows an example of a profile measurement made with a laser scanner.

Fig. 2: Example of a profile measurement.

The **Process Equipment** consists of four basic components which may not all be used to carry out a given production task. These components are: a tool manipulator, a workpiece manipulator, a sensor manipulator, and auxiliary equipment. The tool manipulator is used for manipulating the tool such that desired positions and orientations of the tool relatively to the workpiece can be obtained. The workpiece can be mounted on a workpiece manipulator. The workpiece manipulator can be used for facilitating the control of the process by orienting the workpiece relatively to gravity. The sensor manipulator is used for manipulating the profile sensor so that, during operations, it can measure the product profile ahead of the tool. The auxiliary equipment provides energy and material to the tool.

2.2. Functional Entities.

As it appears from figure 1 the system contains six functional entities: **Process Controller, Equipment Server, Sensor Server, Workpiece Modeller, Sensor Modeller**, and **Process Modeller**.

The main task of the **Process Controller** is to communicate with the Task Manager and to control the sequence in which the functional entities of the control system are triggered.

The **Equipment Server** provides the equipment (i.e. the manipulators and auxiliary equipment) with necessary references. Input to the Equipment Server is a queue of process control vectors computed by the Process Modeller. During operation process control vectors are inserted into the rear of the queue by the Process Modeller. Simultaneously, process control vectors in the head of the queue are read by the Equipment Server and transformed into references to the equipment. The implementation of the Equipment Server requires that the manipulators have appropriate motion interfaces. An example of such a motion interface can be found in ref. 6.

The **Product Modeller** contains a geometry model of the product to be processed. Initially this model comes from the CAD-system and hence represents an ideal product shape. To make sure that the deviations between the ideal geometry model and the geometry of the physical product are sufficiently small, the ideal geometry is updated on the basis of measured product profiles. The Product Modeller computes the curve distance travelled from the start of the process along the processed path to the location where the profile was measured. A similar distance is travelled along the ideal model of the product, and geometrical features corresponding to this location are updated by using the following rules: 1) shape features which can be estimated sufficiently accurate by the ideal model are unchanged, 2) shape features which cannot be estimated sufficiently accurate are either taken directly from the measured product profile or, if this is not possible estimated from a comparison between the ideal model and the measured product profile. The Product Modeller calculates the updated product profile, on request from the Process Controller. During operation, the Process Controller requests for updated product profiles which are located between the present position of the tool and the sensor.

The updated product profile is used by the **Process Modeller** to compute process control vectors by means of process models which relates the product profile to process control vector components (e.g. tool velocity, position orientation etc.). The process control vectors computed by the Process Modeller corresponds to a location on the product between the present location of the tool and the sensor. The process control vectors are inserted into the rear of a queue of process control vectors. Examples of process models can be found in ref. 7 and 8.

The updated product profile is also used by the **Sensor Modeller** to compute sensor references. These references are used by the Sensor Server to perform the signal conditioning.

The task of the **Sensor Server** is to provide the Product Modeller with a measured estimate of the product profile. The Sensor Server generates this estimate on the basis of profile measurements as the ones shown in figure 2. The server acquires profile measurements from the sensor and interprets this information by using the parameters specified in the sensor references. An example of a product profile estimate is shown in the left hand side of figure 3 where the product profile estimate is represented by the points R1,S1,T1,T2,S2, and R2. A number of measured product profile estimates can be combined into a 3D surface model of the product surface, by fitting curves to the points of the product profile estimates. An example of how this can be done is shown in the right hand side of figure 3. Finally, the cross sectional shape of the product can be computed in any plane by computing the intersection between this plane and the curves fitted to the product profile estimates.

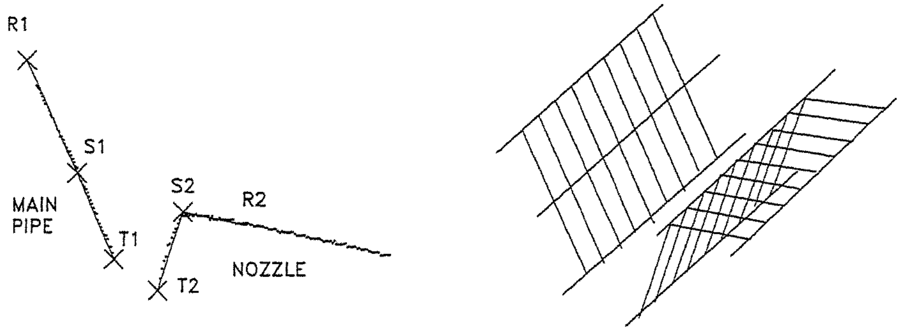


Figure 3. Left: Example of a product profile estimate (R1, S1, T1, T2, S2, R2) computed on the basis of a profile measurement. Right: A 3D surface model of the product is constructed by fitting straight lines to the points of a number of product profile estimates.

3. A SYSTEM INSTANCE.

The principles presented above have been implemented and tested in a control system for arc-welding nozzles perpendicular on large diameter pipes. A detailed description of this system can be found in ref. 9. An example of the products to be welded is shown in figure 4 which also shows the main hardware components of the system. As it appears from the figure the system consists of a profile sensor which is mounted on a robot arm, a little ahead of the welding torch. The workpiece is mounted on a workpiece positioner. Finally, the system consists of a welding machine (not shown in the figure) which provides the electric power to the welding arc and feeds electrode material into the weld pool.

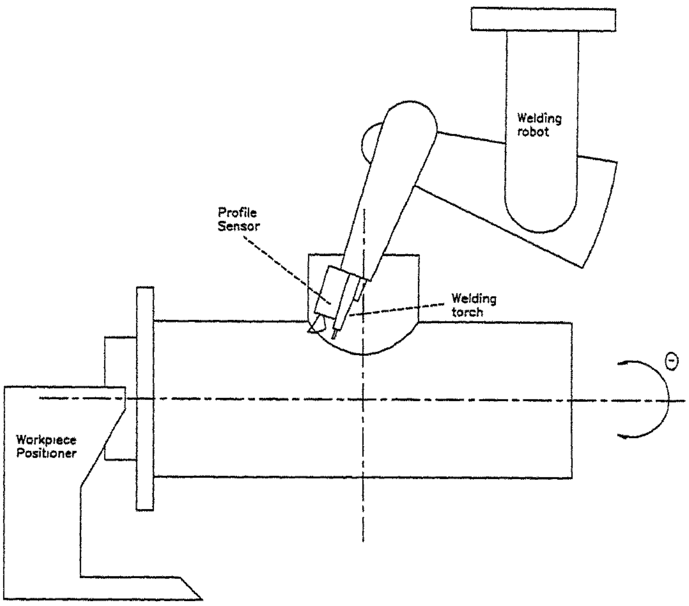


Figure 4. An example of a product to be welded with the control system presented in ref. 9 together with its main hardware components.

Implementation of the system shows that it can be used for performing geometrically complex arc-welding tasks. The system principle has also been tested on other arc-welding seam types (e.g. lap joints). These tests indicate that the system principles can be used to weld seams of varying shapes. Presently, the principles of the system are being implemented in a control system for laser-welding.

4. DISCUSSION

To use the proposed system concept a number of models are needed: models of the equipment and its environment, workpiece models, and process models. Models of the equipment and its environment can be designed in commercially available off-line programming and simulation systems such as PRO/manufacturing, ROBCAD and GRASP, whereas the geometrical models of the workpieces to be manufactured can be designed in CAD-systems as CATIA and PRO/designer. However, reliable quantitative models of the process are rarely available. The lack of process models forms a major obstacle in introducing the proposed system concept in many manufacturing environments. In order to overcome this problem better methods must be made for developing process models.

5. CONCLUSION

In this paper the concept of a robot control system, that carry out complex production tasks has been presented. The proposed control system integrates an off-line programming approach with a sensor based approach.

The control system architecture proposed in this paper has been used in welding applications. Tests with these systems indicate that the control system architecture can be used for controlling robots to carry out complex production tasks. However, it is the assertion of this paper that the principles of the control system can be applied to a range of processes. In order to test this assertion, implementations of the system principles outside the domain of welding is needed.

6. REFERENCES

1. G.E. Cook et. al. (1987) Electric Arc Sensing for Robot Positioning Control, In Robotic Arc welding, Edited by J.D. Lane, IFS (Publication) Ltd, Springer-Verlag, 1987.
2. J. Apples (1987), Application and Economical Aspects of a 3 Dimensional HeNe Laser System, 2nd International Conference on Developments in Automatic and Robotic Welding, London, England, 17-19 Nov., 1987.
3. Ni. Nayak, A. Ray (1993) Intelligent Seam Tracking for Robot Welding, Springer Verlag, Berlin 1993, ISBN 3-540-19826-1.
4. J.P. Huissoon, D.L. Strauss (1994), Automatic Control of a Robotic Gas Metal Arc Welding Cell, 5th World Conference on Robotics Research, Cambridge, Massachusetts, Sep. 27-29, 1994.
5. O. Madsen, H. Holm (1996), Geometry Model and Sensor Based Control of Industrial Robots, 11th IPS Research Seminar, Fuglsø, Denmark, 15-17 April 1996.
6. O. Madsen, H. Holm (1994), A Real-Time Interface for 3D-Tracking in Welding; EURISCON'94; Malaga; Spain; August 22-25; 1994.
7. O. Madsen, H. Holm, J. Boelskifte, I. Hafsteinsson (1995), Model of Root-bead Welding for Off-line Programming and Control, CAPE'95, Beijing, China, 16-18 May 1995.
8. Galopin, M. et. al. (1991) Design of Optimum Adaptation Tables for Robotic Arc welding Using Vision Sensing. In proceedings of EUROJOIN 1; Strasbourg; France; 5-7 November 1991.
9. O. Madsen, H. Holm (1995), Control System Architecture for Robotic Welding of Tubular Joints. 4th International Conference on Trends in Welding Research, Gatlingburg, USA, June 1995

A FORMAL METHOD TO SPECIFY AT AN ABSTRACT LEVEL ROBOTICS-BASED MANUFACTURING SYSTEMS

J.M. Condom, K. Ouriachi
Département informatique
Université de Pau et des Pays de l'ADOUR
Domaine Universitaire - Av. de l'Université - B.P. 576
64012 PAU UNIVERSITÉ Cedex
FRANCE
Tel : 33 5 59 92 31 54
Fax : 33 5 59 80 83 74
Email : {condom,ouriachi}@univ-pau.fr

ABSTRACT

This paper gives some guidelines to make easier the design of robot-based manufacturing systems while contributing to their robustness. We focus on the specification of functionnal requirements using Z formal language. Taking a realistic industry-oriented application, we show how to structure large specifications and specify concurrent aspects considering a true parallelism semantics, to increase system reliability and to improve intellectual control over the system.

KEYWORDS

Robotics-based manufacturing, formal method, Z, reliability, intellectual control.

1. INTRODUCTION

Accurate requirements capture is very important in the design of any system. A mistake at this stage is carried through the entire development process and can be very expensive to correct later [1]. This is particularly the case with safety-critical systems like robot-based manufacturing systems which have highly complex interaction, and the failure of which may not only cause loss of money but also plant damage or injury. Then the use of formal methods to specify such systems seems unavoidable since being based on mathematical concepts they can deliver correctness and enhance software quality. Taking the example of a production cell, we present a specification method aiming at allowing engineers to capture and formalize easily the functionnal requirements.

2. ROBOT-BASED MANUFACTURING SYSTEMS : THE PRODUCTION CELL CASE STUDY

Considering the notion of system as a whole, a robot-based manufacturing system has two components : an environment, i.e. a unit composed of multiple robots cooperating for a production task, and a controller managing the environment. Such a system is reactive and concurrent : reactive because the controller maintain an ongoing interaction with the environment ; concurrent because it is usually composed of several programs running in parallel and cooperating to the manufacturing task. A typical example we use for our application framework is a production cell taken from a real metal processing plant near Karlsruhe and used in a case study aiming at promoting the use of formal methods in industry [7]. This cell is composed of different elements (figure 1) cooperating to the same production task which is to forge metal blanks. The feed belt conveys each blank to a table. This table must rotate to

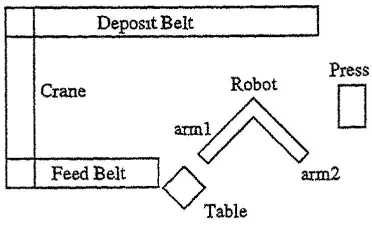


Figure 1: **The production cell**

put the blank in the right position allowing the robot to grip it with its first arm (arm1). Then the robot rotates counter clockwise until its second arm (arm2) aligns to the press. As soon as arm2 has taken the forged metal blank, the robot rotates again to unload arm2 on the deposit belt. Finally arm1 plate is put into the press. The robot is equipped with two orthogonal arms to use the press to maximum capacity. The deposit belt conveys each blank to the crane which picks it up and transports it to the feed belt. The crane acts as a link between both belts that makes it possible to let the system function continuously without the need for an external operator in order to perform demonstrations.

To handle multiple blanks in different parts of the concurrent system, the control software is viewed as a group of cooperating programs, each one managing a component of the environment.

3. THE CONTRIBUTION OF Z TO A RELIABLE AND RAPID PRODUCT DEVELOPMENT

Z is a formal language based on set theory and first order logic developped by the Programming Research Group at Oxford [11]. Its characteristics and the method we developped propose some solutions to specify large real-time systems in a reliable and rapid manner.

3.1 The advantages of Z

Z offers several advantages particularly mentionned by an international survey [4] about large commercial projects using Z. It is reported that Z has often been choosen because it is easily understandable by engineers and it offers a rigorous notation allowing them to clarify their requirements and to express them with precision. It allows a modular and high level approach that can be refined down to a program. This is particularly advantageous for engineers who needn't to deal with unnecessary details at each step of the specification refinement. In addition to those characteristics, some properties expected from the system can be demonstrated like safety and liveness properties confirming the system robustness. All those features not only contribute to the software reliability but also give the engineers a good intellectual control over the system allowing them to specify complex parts of the system easily.

3.2 The specification method

Z wasn't able at the beginning to specify real-time aspects such as concurrency and timing constraints. Since then, works have been undertaken to add those real-time features (e.g. [5],[2],[6]), but concern applications of little size. Our approach investigates the specification of industrial concurrent systems proposing some guidelines to specify large systems easily. First we refer to a true parallelism semantics instead of interleaving semantics commonly adopted and that we consider too restrictive. According to interleaving, only one operation is selected and executed non-deterministically according to a fairness rule [9]. On the contrary, true concurrency involves that all enabled operations are executed in parallel. This semantics contributes to simplify the specification process and to contribute to system robustness. Some guidelines are given below about how to structure large specifications and how to improve intellectual control over the system.

4. STRUCTURING Z SPECIFICATIONS REFERRING TO A TRUE PARALLELISM SEMANTICS

A. MacDonald and D. Carrington [8] specified the production cell at a high level and investigated the issue of structuring Z specifications showing the importance of a good structure within large specifications. We take inspiration from this work adding concurrency features and giving more precision about how the system works. The specification is built up in two steps : the first step is the independent specification of each component with its corresponding control program, we call it the local specification ; the second step is the interaction between components, that is the global specification. The main features of the structuring technique are presented using the robot section of the specification.

4.1 The local specification

Each component and its program is specified in Z by an abstract data type consisting of :

- the state space, a set of states the component may enter ;
- a set of initial states ;
- a number of high level operations allowable in this space.

From a computational point of view, operations can be executed everytime their pre-conditions hold. Therefore suitable pre-conditions and post-conditions must be determined ; the former defines the set of valid states the operations may be applied to, and the latter defines the set of states that may be reached by the system at the end of the operation.

4.2 The global specification

The specification of the system needs an overall state to formalize the interaction between the components. This process called promotion by MacDonald and Carrington [8] includes two phases : the first one, which requires that each abstract data type takes into account an overall state ; the second phase models the relationships between the components. Compared with MacDonald et al.' formalization, the overall state of the cell (*Cell*) includes not only the state spaces of the individual components but also a synchronisation state including all the synchronisation variables required for the communication between components :

$$Cell \mathrel{\widehat=} FBelt \land Table \land Robot \land Press \land DBelt \land Crane \land Synchro$$

Another difference is the declaration required for each operation schema. In an interleaving model, as only one operation is executed at a time, the schema declaration must express that each operation doesn't modify the variables it doesn't manage. Under true concurrency, operations belonging to different abstact data types may have one common state to start running and end their execution in a common state. Hence the only declaration allowed is $\Delta Cell$. The initialization of the cell is a conjunction of the initializations of components and synchronisation variables :

$$Init\_Cell \mathrel{\widehat=} Init\_FBelt \land Init\_Table \land Init\_Robot \land Init\_Press \land Init\_DBelt \land Init\_Crane \land Init\_Synchro$$

Concerning the formulation of interaction requirements, we define a communication by shared variables. Each variable is used as a communication channel between two components. The schema *Synchro* below defines six state variables required for the whole cell, three of them are required for the communication between the robot and the components closest to it : the table, the press and the deposit belt :

Synchro
sync_table_robot : *Sync_Table_Robot*
sync_robot_press : *Sync_Robot_Press*
sync_robot_dbelt : *Sync_Robot_DBelt*
...

Each type is a set of values each one corresponding to a different step of the interaction.

4.3 The global specification of the robot

The robot state is modelled by three state variables : *robot_orientation* representing its position, *arm1_load* and *arm2_load* representing the load of each arm. Four positions are managed :
- the first position (*load_arm*1) aligns aml to the table ;
- the second position (*load_arm*2) aligns arm2 to the press ;
- the third position (*unload_arm*2) aligns arm2 for unloading to the deposit belt ;
- the fourth position (*unload_arm*1) aligns arm1 to the press.

Robot
robot_orientation : *Robot_Orientation*; *arm*1*_load*, *arm*2*_load* : *Component_Load*

where types *Robot_Orientation* and *Component_Load* can take the following values :

Robot_Orientation ::= *load_arm*1 | *load_arm*2 | *unload_arm*2 | *unload_arm*1

Component_Load ::= *loaded* | *unloaded*

Initially the robot is ready to unload the table and its arms are unloaded/

Init_Robot
Robot
robot_orientation = *load_arm*1; *arm*1*_load* = *unloaded*; *arm*2*_load* = *unloaded*

Three operations are required to control the robot : *Load_Arm*, *Unload_Robot*, *Rotate_Robot* In Z, operations on data types are specified by schemas which have two copies of the state variables (Spivey (1992)) : an undecorated set (*Cell*) corresponding to the state of the data type before the operation, and a decorated set (*Cell'*) corresponding to the state after the operation. Those copies are declared by Δ*Cell* :

Δ*Cell*
Cell
Cell'

Then pre-conditions of operations refer to undecorated variables and post-conditions to decorated ones. *Load_Arm* is specified as follows :

$$
\begin{array}{|l}
\textit{Load\_Arm} \\
\Delta \textit{Cell} \\
\hline
((\textit{sync\_table\_robot} = \textit{table\_ready\_to\_unload} \land \textit{robot\_orientation} = \textit{load\_arm}1 \land \\
\quad \textit{load\_arm}1 = \textit{unloaded} \land \\
\qquad\qquad \textit{sync\_table\_robot}' = \textit{arm1\_loaded} \land \\
\qquad\qquad \textit{load\_arm}1' = \textit{loaded} \land \textit{load\_arm}2' = \textit{load\_arm}2) \\
\qquad\qquad\qquad\qquad \lor \\
(\textit{sync\_robot\_press} = \textit{blank\_pressed} \land \textit{robot\_orientation} = \textit{load\_arm}2 \land \\
\quad \textit{load\_arm}2 = \textit{unloaded} \land \\
\qquad\qquad \textit{sync\_robot\_press}' = \textit{arm2\_loaded} \land \\
\qquad\qquad \textit{load\_arm}1' = \textit{load\_arm}1 \land \textit{load\_arm}2' = \textit{loaded})) \\
\textit{robot\_orientation}' = \textit{robot\_orientation}
\end{array}
$$

Load_Arm includes two actions which are arm1 and arm2 loads. The pre-conditions to load arm1 are :
- the table must signal over *sync_table_robot* that it is in the right position to unload ;
- arm1 must be oriented towards the table *load_arm*1 and unloaded ;

The post-conditions of arm1 load are expressed by the lines indented and correspond to arm1 is loaded and arm2 isn't modified.
The same type of pre-conditions and post-conditions are specified for arm2 ; arm2 is loaded when the piece is forged which is signalled by the press (*sync_robot_press* = *blank_pressed*). The last post-condition mentions that the robot orientation is unchanged.

4.4 Remarks on the specification structure and the synchronisation process

The structuring process leads to a two-levels specification and to operations schemas gatherin together actions of the same type ; an operation schema could be specified for each action but it leads to a granularity too fine making, according to us, the specification less readable ; this is particularly the case of rotation operation referring to seven actions.

The synchronization method is both plain and robust. Each component needs only to refer to appropriate shared variables and doesn't require to be aware of the state variables local to the component(s) it interacts with. A direct consequence is that the modification of state variables in a component can be achieved without changing the synchronisation process on condition that, of course, this modification doesn't concern this process too. Another consequence is that the synchronisation process can be easily verified. All those features contribute to a modular and robust approach which is benefit for the specification of large concurrent systems. Another remark which is developped in [3] is that specifying in Z refering to true concurrency should contribute to simplify the specification process and to capture more easily the various states the system may enter. This is a safety factor that must be taken into consideration.

4.4 CONCLUSION

This paper gives a rapid overview of the specification method used to make easier the specification of large real-time systems at an abstract level particularly in the aerea of robot-based automation. This simplification is realized using appropriate structuring techniques and refering to a true parallelism semantics. Further work is the development of a proof method allowing engineers to check system properties easily using an automatic proof environment such as Z/EVES [10].

REFERENCES

[1] J.P. Bowen, "Formal Methods in Safety-Critical Standards", Proc. Software Engineering Standards Symposium, Brighton, UK, 1993.

[2] J.M. Bruel, A. Benzekri, Y. Raynaud, "Z and the specification of Real Time systems", 7th International Conference on : "Putting into practice methods and tools for information system design", Nantes, October 10-12 1995.

[3] J.M. Condom, K. Ouriachi, "On the use of Z to specify robot-based manufacturing systems", Internal Report, Département d'informatique, Université de Pau 1996.

[4] D. Craigen, S. Gerhart, T. Ralston, "An International Survey of Industrial Applications of Formal Methods", NIST GCR 93/626 (Volumes 1 and 2), U.S. National Institute of Standards and Technology, March 1993.

[5] R. Duke, I.J. Hayes, P. King, G.A. Rose, "Protocol specification and verification using Z", Protocol Specification, Testing, and Verification VIII, S. Aggarwal and K.Sabnani editors, pp. 33-46, Elsevier Science Publishers (North-Holland), 1988.

[6] A.S. Evans "Specifying and Verifying Concurrent Systems Using Z", FME'94, LNCS 873, M. Naftalin, T. Denvir, M. Bertran (Eds.), Springer Verlag, Berlin, 1994.

[7] C. lewerentz, T. Lindner (Eds.),Formal Development of Reactive Systems - Case Study production cell, LNCS 891, Springer Verlag, Berlin, January 1995.

[8] A. MacDonald and D. Carrington, "Structuring Z Specifications : Some Choices", ZUM'95 : The Z Formal Specification Notation, LNCS, pp. 203-223, Limerick, Ireland, september 1995.

[9] Z. Manna, A. Pnueli, The Temporal Logic of reactive and Concurrent Systems, Springer-Verlag, New York. 1992.

[10] I. Meisels, M. Saaltink, "The Z/EVES Reference Manual DRAFT", Technical Report TR-96-5493-03a. ORA Canada, 267 Richmond Road, Suite 100, Ottawa, Ontario K1Z 6X3, CANADA,1996.

[11] J.M. Spivey, The Z notation : A reference Manual, Prentice Hall International Series in Computer Science, 2nd edition, 1992.

A SOPHISTICATED ASSEMBLY PLANNING SYSTEM FOR FLEXIBLE ROBOT-BASED MANUFACTURING

H. Mosemann, F. Röhrdanz, and F. M. Wahl
Institute for Robotics and Process Control
Technical University of Braunschweig
Germany

ABSTRACT

Assembly planning and the subsequent execution of the generated plans by robots is one of the key technologies of modern and flexible manufacturing. The choice of an assembly sequence drastically affects the efficiency of the assembly process. In this paper we present our high level assembly planning system ^{High}LAP integrated into a commercial robotic simulation system for flexible robot-based manufacturing. ^{High}LAP takes the CAD descriptions of assembly components and high level assembly specifications using symbolic spatial relationships as input. The assembly planning system generates and evaluates all assembly sequences of a mechanical assembly with minimal user interaction based on *geometrical* and *physical* reasoning. Several evaluation criteria like stability, directionality, assembly pose, manipulability and parallelism are introduced and quantified.

KEYWORDS

Assembly Planning, Assembly Cost Evaluation, Task Planning, Robot-Based Manufacturing

1. INTRODUCTION

Assembly is an interesting area for automation, quality control and flexibility. As the processes in assembly are much more complex than processes in other robotic application areas, intensive research work has been done and still has to be done. The goal in robot assembly is to construct a mechanical product consisting of several parts which can be assembled by robots. A high level assembly planning system generates sequence plans specifying the order in which parts are to be assembled to form the desired product and computes the trajectories to bring the parts together. In addition to such sequence plans, task plans for actually performing each assembly operation must be generated. The *high level assembly sequence plan* is the basis for such *lower level plans*, and taken together they ensure the efficient and flexible assembly of a mechanical product. The main problem in assembly planning is to find a *good* sequence plan; the issues to be considered are very complex. An overview of typical constraints in automated assembly planning can be found in [5]. Numerical evaluation functions are used to decide about the *goodness* of an assembly sequence plan. In the last few years some automated systems have been developed to generate geometrical feasible assembly sequence plans using such evaluation functions (see e. g. [2], [9]). While fulfilling the criteria for *goodness* it is important to search for all valid plans satisfying those criteria. Therefore, it is difficult to find good search heuristics based on geometrical reasoning only. The high level assembly planning system ^{High}LAP introduced in this paper uses for example an assembly hierarchy, an extended cycle finder and physical reasoning to simplify the search for an optimal plan. ^{High}LAP has a modular structure and an open architecture. The system covers all modern aspects of high level assembly planning and introduces new approaches for physical evaluation of sequence plans like the stability analysis. We integrated our high level assembly planning system into the commercial robotic simulation system *IGRIP* (Interactive Graphics Robot Instruction Program [10]) for flexible robot-based manufacturing. Figure 1 depicts the system architecture of ^{High}LAP and the integration into *IGRIP* . ^{High}LAP takes the CAD descriptions of assembly components and high level assembly specifications using symbolic spatial relationships as input. For the specification of *symbolic spatial relations* between the assembly components feature frames can be defined and associated to surface primitives using the *IGRIP* three-tier menu system. The presented system is the first assembly planning system computing and taking into account the range of all stable orientations of an assembly

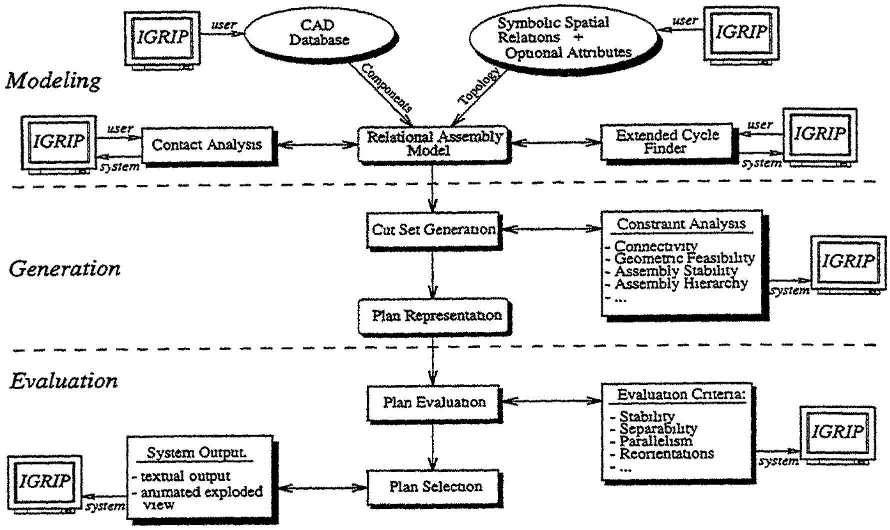
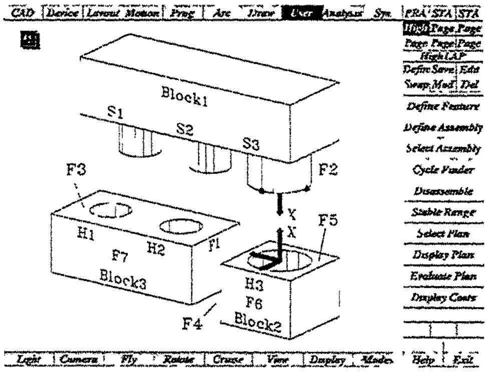


Figure 1: System architecture and user interface based on the commercial simulation system *IGRIP*

considering friction for assembly plan evaluation. For each generated (sub)assembly the range of all stable orientations can be visualized by *IGRIP* and an optimal orientation for the corresponding mating operation is calculated. The generated assembly sequences are stored in an AND/OR-Graph which constitutes a compact representation of all feasible assembly sequences. Each assembly sequence can be visualized by *IGRIP* taking into account the automatically calculated evaluation criteria.

2. MODELING OF ASSEMBLIES

A mechanical assembly is a composition of interconnected components forming a unit. The components consist of rigid bodies bounded by surface primitives (planes, spherical surfaces, cylindrical surfaces). For the specification of *symbolic spatial relations* between the assembly components feature frames of type *FACE, SHAFT, HOLE* can be defined and associated to the surface primitives. In Figure $2a$ a simple example consisting of 3 blocks is shown with some feature frames. Figure $2b$ shows the corresponding feature type specification and in Figure $2c$ the high level assembly specification is given. The principle of symbolic spatial relations bases on [1]; an *extended cycle finder* [4] is used to generate the homogeneous transformations between the components of the assembled product. The key idea of the cycle finder is the reduction of the degrees of freedom defined by symbolic spatial relations by searching cycles consisting of two appropriate relationships between the same components and combining these relationships to a new one with equal or less degrees of freedom. In contrast to the cycle finder published in [6] our system merges all objects sharing a *fix* relationship at an arbitrary reduction step to a new *compound object*. Furthermore, the efficiency of the cycle finder is increased by automatically including relations by a *transitive rule*, if an assembly can't be reduced to *fix* using the above mentioned methods. *HighLAP* automatically detects and classifies the physical contacts of the assembled product (degenerate contacts are handled by planar approximations) and generates an *assembly graph*. The assembly graph associated to an assembly is an attributed simple undirected graph in which each node corresponds to an assembly component and each edge connects a pair of parts which have at least one surface contact. To reason about the feasibility of assembly tasks the assembly representation of *HighLAP* includes a description of attachments which bind one part to another. Attachments may act on surface contacts and eliminate all degrees of freedom for relative motion (for example screw or glue attachments). Therefore, the attachments are associated as optional attributes to the edges of the assembly graph (see [7] for more details about the representation of mechanical assemblies in *HighLAP*). In the following the



(a)

	face	*shaft*	*hole*
Block1	F2	S1,S2,S3	
Block2	F4,F5,F6		H3
Block3	F1,F3,F7		H1,H2

(b)

r_1	S1	*of*	Block1	*fits*	H1	*of*	Block3
r_2	S2	*of*	Block1	*fits*	H2	*of*	Block3
r_3	F2	*of*	Block1	*against*	F3	*of*	Block3
r_4	S3	*of*	Block1	*fits*	H3	*of*	Block2
r_5	F2	*of*	Block1	*against*	F5	*of*	Block2
r_6	F1	*of*	Block3	*against*	F4	*of*	Block2
r_7	F7	*of*	Block3	*coplanar*	F6	*of*	Block2

(c)

Figure 2: (a) Feature frames H3, F5 (coincident with H3) and S3 of an assembly defined under IGRIP, (b) feature type specification in textual form and (c) high-level assembly description using symbolic spatial relations.

product to be assembled is denoted as the *goal assembly* A. It is assumed to consist of n components $C = \{c_1, c_2, \ldots, c_n\}$ and any valid subassembly SA of A is identified by the corresponding set of components $C' \subseteq C$.

3. ASSEMBLY SEQUENCE GENERATION

The generation of valid assembly sequences bases on the well-known disassembly technique of an assembly $A = SA_1 \cup SA_2$ into *two* subassemblies SA_1, SA_2 combined with a cut-set method of the assembly graph. To avoid blind recursion of the cut-set generation, a guided recursive search of connected subgraphs has been implemented in $High_{LAP}$ (see [4] for more details about the guided recursive search). Figure 3a shows an example of a cut in an arbitrary assembly graph. If a cut corresponds to a *feasible assembly operation* it is stored as a hyperarc of an AND/OR-graph which constitutes a compact representation of all feasible assembly sequences. Furthermore, time dependencies and independencies between assembly tasks can be represented. Figure 3b shows the corresponding hyperarc of the valid cut depicted in Figure 3a.

3.1 Constraints for Feasible Assembly Operations

- **Existence of Local and Global Depart Motions:**

$High_{LAP}$ computes the shape of the *local depart space* between SA_1 and SA_2 for translational movements using the edges of the assembly graph with the automatically generated transformations. The analysis of the local depart motion can be done by checking the compatibility of the most restrictive contact with all other contacts. Similar to [3] we use the following types of local depart spaces : *no depart space, half line, line, halfplane, sector of a plane, plane, halfspace, infinite wedge, polyhedral cone, space*. In Figure 4 an example for the depart space of type *polyhedral cone* is shown. For this depart space all valid depart vectors start in the tip of the cone (see Figure 4) and lie completely in the volume defined by the cone. In order to decide whether a set of planar contacts does not completely constrain one subassembly the corresponding depart spaces are iteratively intersected. If the resulting depart space is of type *no depart space* then no local depart motion is possible and the cut into SA_1 and SA_2 is invalid. Otherwise the shape and the position of the depart space is mathematically known. In the following the local depart space of a valid cut into SA_1 and SA_2 is denoted as $lds(SA_1, SA_2)$. The *global feasibility* of depart motions is computed by heuristically choosing a set of directions lying in the local depart space sweeping one subassembly against the rest of the assembly [4].

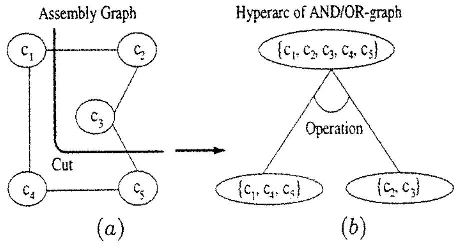


Figure 3: Valid cut in the assembly graph (a) and the resulting hyperarc h=({c_1, c_2, c_3, c_4, c_5}, {c_1, c_4, c_5}, {c_2, c_3}) in the AND/OR-graph (b)

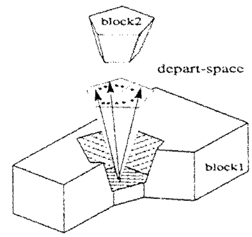


Figure 4: Depart space of type polyhedral cone. The concerned physical contacts which should be established or separated respectively are shown hatched.

- **Stability of Assemblies:**

A (sub)assembly is stable if the rigid components are in static equilibrium under the influence of external and internal forces. External forces arise from gravitation and internal forces from the mutual contact of the objects. For the determination of a stable orientation *HighLAP* takes into account static friction and computes the net forces $\vec{F}_j \in \mathbf{R}^3$ and net torques $\vec{\tau}_j \in \mathbf{R}^3$ acting on the jth component of an assembly [8]. An assembly is said to be potentially stable if there exist contact forces such that the net force and net torque on every body is zero. Whether such contact forces exist can be decided with *linear programming*. *HighLAP* introduces the components of the gravity vector as variables in the linear program and searches among all possible directions of the gravity vector to calculate the set of potentially stable assembly orientations. Therefore, we enumerate all vertices of the convex solution space which is described by the corresponding linear inequalities of the linear program (see [8]).

4. ASSEMBLY SEQUENCE EVALUATION

HighLAP generates an AND/OR-graph representing all possible assembly sequences using the cut-set method with the above mentioned constraints . *HighLAP* assigns the following weights to the nodes and hyperarcs of the AND/OR-graph selecting an optimal assembly plan.

- Stability function: For each AND/OR-graph node *HighLAP* rates the stability of the corresponding assembly A using a stability function. The function bases on the whole set of potentially stable assembly orientations denoted as $so(A) \subseteq \mathbf{R}^3$. Having found all vertices of $so(A)$ the components of the gravity vector are projected onto the unit sphere U with the projection operator Π to mark a stable region on the sphere, denoted as $\Pi(so(A), U)$.

Figure 5 shows an example of an assembly with the corresponding set of potentially stable orientations. *HighLAP* uses the following stability function σ for any valid assembly $A \in \mathcal{P}(C)$ represented by an AND/OR-graph node ($\mathcal{P}(C) = \{c_1, c_2, \ldots, c_n\}$ denotes the set of all subsets of C):

$$\sigma : \mathcal{P}(C) \rightarrow [0, \ldots, 1], \quad A \mapsto \sigma(A) = 1 - \frac{Area[\Pi(so(A), U)]}{4\pi} \tag{1}$$

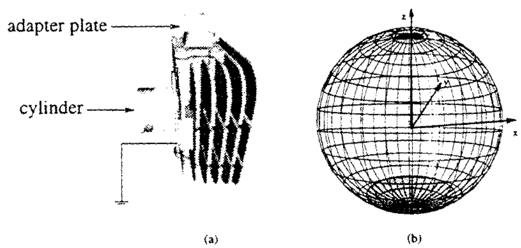


Figure 5: Set of stable orientations considering friction (b) for the $\{cylinder, adapter\, plate\}$ assembly (a)

The following weights for each AND/OR-graph hyperarc $h = (SA_1 \cup SA_2, SA_1, SA_2)$ are applied to rate the difficulty of the corresponding assembly operation:

- Separability function: The shape of the local depart space determines the costs for the separability of the subassemblies SA_1 and SA_2. The costs are increased with the narrowness of the local depart space, e.g. *halfline* costs more than *line*, *line* costs more than *sector of a plane* and so on. The following separability function δ is defined:

$$\delta : \mathcal{P}(C) \times \mathcal{P}(C) \rightarrow [0,\ldots,1], \quad (SA_1, SA_2) \mapsto \delta(SA_1, SA_2) = 1 - \frac{Area[\Pi(lds(SA_1,SA_2),U)]}{4\pi} \quad (2)$$

• Manipulability function: The function κ of a hyperarc takes into account how hard the subassemblies SA_1 and SA_2 are to handle dunng the mating operation with respect to subassembly stability:

$$\kappa : \mathcal{P}(C) \times \mathcal{P}(C) \rightarrow [0,\ldots,1], \quad (SA_1, SA_2) \mapsto \kappa(SA_1, SA_2) = 1/2 * (\sigma(SA_1) + \sigma(SA_2)) \quad (3)$$

• Ratio of the number of components of the subassemblies: In mode (a) ^{High}LAP uses the function ϕ_a to support maximal parallelism during plan execution:

$$\begin{aligned} &\phi_a : \mathcal{P}(C) \times \mathcal{P}(C) \rightarrow [0,\ldots,1[, \quad (SA_1, SA_2) \mapsto \\ &\phi_a(SA_1, SA_2) = \\ &(max(|SA_1|,|SA_2|) - min(|SA_1|,|SA_2|))/(max(|SA_1|,|SA_2|) + min(|SA_1|,|SA_2|)) \end{aligned} \quad (4)$$

In mode (b) small subassemblies SA_1 are added to large subassemblies SA_2 ($|SA_1| \ll |SA_2|$):

$$\phi_b : \mathcal{P}(C) \times \mathcal{P}(C) \rightarrow]0,\ldots,1], \quad (SA_1, SA_2) \mapsto \phi_b(SA_1, SA_2) = 1 - \phi_a(SA_1, SA_2) \quad (5)$$

• Necessity of reorientation: For each hyperarc $h = (SA_1 \cup SA_2, SA_1, SA_2)$ the set of stable orientations $\omega(SA_1, SA_2)$ for assembling the subassemblies SA_1 and SA_2 is calculated by intersecting the corresponding sets of potentially stable assembly orientations:

$$\begin{aligned} &\omega : \mathcal{P}(C) \times \mathcal{P}(C) \rightarrow U, \quad (SA_1, SA_2) \mapsto \\ &\omega(SA_1, SA_2) = \Pi(so(SA_1), U) \cap \Pi(so(SA_2), U) \cap \Pi(so(SA_1 \cup SA_2), U) \end{aligned} \quad (6)$$

If $\omega(SA_1, SA_2) = \emptyset$ then additional fixture costs are assigned for the hyperarc. Otherwise ^{High}LAP chooses a hyperarc orientation corresponding to the center of $\omega(SA_1, SA_2)$.

5. RESULTS

To illustrate the efficiency of ^{High}LAP we describe an assembly planning example of a Yamaha RD80 motorcycle engine. Figure 7 depicts an exploded view of the assembly consisting of 27 components. For ease of illustration the evaluation of the best (dis)assembly sequence for the subassembly {*cr_group, ch, cy, sp, cmb1, cmb2, cmb3, cmb4 , chs1, chs2, chs3, chs4*} is shown in Figure 6. The abbreviations are used as indicated in Figure 7. No stability costs are assigned to the shaded AND/OR-graph nodes representing assemblies which are stable in any orientation. For this planning example the user has specified the initial orientations of the assembly components as indicated in Figure 6. Therefore, three reorientations must be performed (hyperarcs 13, 14, 15) resulting in additional hyperarc costs.

6. CONCLUSIONS

In this paper we presented ^{High}LAP , a sophisticated high level assembly planning system capable to generate stable assembly sequences based on geometrical and physical reasoning. The system takes CAD descriptions of the assembly components and user-friendly high level assembly specifications as input. All valid assembly sequences are generated and evaluated with minimal user interaction. The generated assembly plans are displayed like an animated exploded view. We integrated ^{High}LAP into a commercial robotic simulation system to provide an optimal user interface. In the future we plan the animation of workcells performing the assembly with robots, tooling and fixturing.

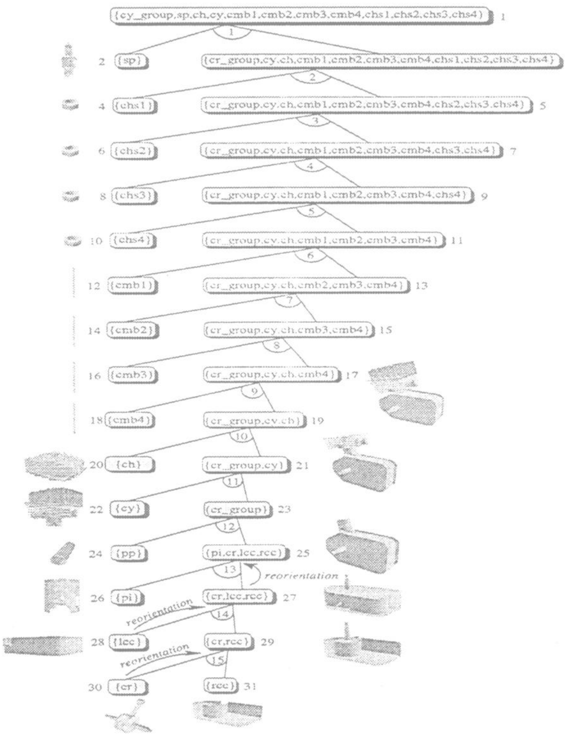


Figure 6: AND/OR-graph of the best assembly sequence for the assembly $\{cr\_group, ch, cy, sp, cmb1, cmb2, cmb3, cmb4, chs1, chs2, chs3, chs4\}$.

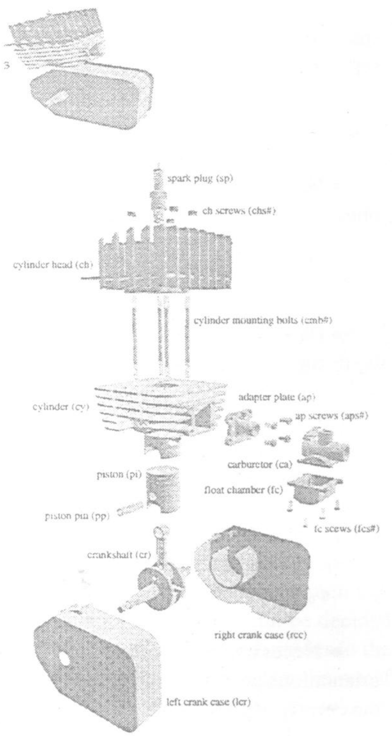


Figure 7: Exploded view of a Yamaha RD80 motorcycle engine.

7. REFERENCES

[1] A. P. Ambler and R. J. Popplestone. "Inferring the Positions of Bodies from Specified Spatial Relationships". Artificial Intelligence, 6:157 – 174, 1975.

[2] L. S. Homem de Mello and S. Lee, editors. Computer-Aided Mechanical Assembly Planning. Kluwer Academic Publishers, 1991.

[3] L. S. Homem de Mello. Task Sequence Planning for Robotic Assembly. PhD thesis, Carnegie Mellon University, May 1989.

[4] R. Gutsche, F. Röhrdanz, and F. M. Wahl. "Assembly Planning Using Symbolic Spatial Relationships". In Graphics And Robotics, W. Strasser and F. M. Wahl, editors,87–104. Springer-Verlag, 1994.

[5] R. E. Jones and R. H. Wilson. "A Survey of Constraints in Automated Assembly Planning". IEEE International Conference on Robotics and Automation, pages 1525–1532, 1996.

[6] R. J. Popplestone. "Specifying Manipulation in Terms of Spatial Relationships". Technical Report 117, Department of Artificial Intelligence, University of Edinburgh, 1979.

[7] F. Röhrdanz, H. Mosemann, and F. M. Wahl. Highlap: "A High Level System for Generating, Representing, and Evaluating Assembly Sequences". IEEE International Joint Symposia on Intelligence and Systems, Rockville Maryland, USA, November 1996.

[8] F. Röhrdanz, H. Mosemann, and F. M. Wahl. "Stability Analysis of Assemblies Considering Friction". Technical Report 5-1996-1, Institute of Robotics and Computer Control, Braunschweig, Germany, May 1996.

[9] B. Romney, C. Godard, M. Goldwasser, and G. Ramkumar. "An Efficient System for Geometric Assembly Sequence Generation and Evaluation". ASME International Computers in Engineering Conference, pages 699–712, 1995.

[10] User Manual and Tutorials, IGRIP Version 3.0. Deneb Robotics, 3285 Lapeer Road West, P.O. Box 214687, Auburn Hills, MI 48321-4687, 1994.

STATE SPACE MODELLEING OF A CLASS OF DISCRETE EVENT SYSTEMS FOR ROBOTIC MANUFACTURING CONTROL.

Hans Holm
Aalborg University, Department of Production
Fibigerstraede 16, 9220 Aalborg East, Denmark.
Telephone : +(45)98154211...2988
E-mail : i9hh@iprod.auc.dk
Fax : +(45)98153030

ABSTRACT

This paper presents outlines for a newly by D. Franke [1] invented arithmetically based method for state space modelling of discrete event systems. The applicability of this method for modelling a shop floor control system, which comprises a scheduler, is demonstrated.

1. INTRODUCTION.

Recently an arithmetic algebra based method has been developed for design of closed loop control of a seemingly wide range of discrete event systems.(See [1]). The closed loop control design by this method is based on an arithmetic algebra state space description of the plant to-be controlled. As it seems to address a wide range of discrete event systems and as this system description method also seems to have the potential of enabling an integrated design of systems composed of sampled continuous systems and discrete event systems it is desirable to test whether it can be used for description of selected important types of systems, which are used when designing multi machine manufacturing control systems. In the paper such a test is reported by presenting an arithmetic algebra state space description of a two robot task sequence control system.

2. MODELLING PRINCIPLE.

2.1 State Modelling.

The main principle of the modelling technique described in [1] is to model the system of interest by a selected set of variables by which the states of interest for the performance of that system can be purposefully described.

2.2 Binary States and Variables.

As the type of system of interest is the discrete event system type, the states of interest of the system are described by presence or absence of an event or by latency or transition of a process. A range of two values of each of the chosen state representing variables is compatible with the characterisation of the system state properties of interest. A variable, which has two values is a binary variable. The chosen state variables are therefore purposefully chosen to be binary. The two possible values of each variable are chosen to be 0 and 1.

2.3 State Modelling for Control.

It is (as in other control system design tasks) wanted to establish a model, by which are described the changes of all selected state variables as a consequence of appearance of an event i.e. a change (binary) in any input receiver variable of the system or a change in any state variable of the system. Announcing appearance of an event a "step", it is wanted to model the value of the selected set of state variables as a function of the values of these same variables in the previous step and the values of the input variables in the previous step. It is hence wanted to establish a model of the following form:

(1): $\mathbf{X}(k+1) = \mathbf{F}(\mathbf{X}(k), \mathbf{u}(k))$, where $X(k)$ is a n-dimensional vector of the state variables, $\mathbf{u}(k)$ is a q-dimensional vector of input variables, $\mathbf{F}(\mathbf{X}(k), \mathbf{u}(k))$ is a n-dimensional vectorial function and k is a scalar counter of steps.

2.4 Binary and Logic Functions.

As all variables are binary they can be interpreted logic variables. The dependence of a logic variable on other logic variables can be represented by a table of truth for the dependant variable as a function of the independent variables. Hence to describe a state variable component $x_i(k+1)$ of the state vector $\mathbf{X}(k+1)$ at step k+1 as a function of the state vector $\mathbf{X}(k)$ and the input vector $\mathbf{u}(k)$ at step k, a table of truth for the state variable component $x_i(k+1)$ as a function of $\mathbf{X}(k)$ and $\mathbf{u}(k)$ can be used. A table of truth completely describes the functional relationship between a dependent variable and a set of independent variables, as this table lists the values of the dependent variable as a function of all possible combinations of the independent variables.

2.5 Arithmetic and Logic Functions.

According to [1] the logic function corresponding to any completely specified table of truth may be represented by a Shegalkin polynomial, which is a polynomial of the form:

$$(2): y = a_0 + \sum_{i=1}^{n} a_i \bullet x_i + \sum_{j=2}^{n}\sum_{i=1}^{j-1} a_{ij} \bullet x_i \bullet x_j + \sum_{k=3}^{n}\sum_{j=2}^{k-1}\sum_{i=1}^{j-1} a_{ijk} \bullet x_i \bullet x_j \bullet x_k + \ldots\ldots$$
$$+ a_{123\ldots n} \bullet x_1 \bullet x_2 \bullet x_3 \bullet \ldots \bullet x_n$$

This polynomial function of the n independent variables x_1 through x_n has 2^n terms. It hence also has 2^n a-coefficients. If these a-coefficients are known the function:

$$(3): y = y(x_1, x_2, x_3, \ldots\ldots, x_n)$$

is known. The table of truth for a Boolean function y of n Boolean independent variables has 2^n different combinations of the n independent variables. Each of these combinations appear in only one row of the table of truth. The table therefore contains 2^n rows.

Inserting the dependent variable as well as the independent variables for each of the 2^n rows of the truth table into equation (2) yields 2^n linear equations with the 2^n a-coefficients as unknowns. As non of the 2^n rows are linear dependent, the 2^n a-coefficients can be determined from these equations. Hence the Boolean (binary) variable y has been determined as an arithmetic function of the Boolean (binary) variables $x_1, x_2, x_3, \ldots\ldots, x_n$.

2.6 Arithmetic Binary State Modelling for Control.

Replacing the binary independent variable y by the binary independent state variable $x_i(k+1)$ at step k+1 and the independent variables x_1 through x_n by the components of the vectors $\mathbf{X}(k)$ and $\mathbf{u}(k)$, an arithmetic function for $x_i(k+1)$ can be derived by the procedure presented in section 2.5 with the components of the vectors $\mathbf{X}(k)$ and $\mathbf{u}(k)$ as independent variables. Repeating this procedure for all $x_i(k+1)$, where $i \in (1,n)$, the desired vectorial function (4) is achieved:

$$(4): \mathbf{X}(k+1) = \mathbf{F}(\mathbf{X}(k), \mathbf{u}(k)).$$

3. A TWO ROBOT TASK SEQUENCE CONTROL SYSTEM.

3.1 System Selection.

The system, for which it has been chosen to derive an equation of the type of equation (1) - a so called state space equation - is a system, which is important for design of control systems for multi machine manufacturing systems.

3.2 Physical System Presentation.

This system is - as shown in figure 1 - a system, which comprises two robots R1 and R2. Each robot is hanging in a gantry of its own. The gantries are shown in figure 1 as the two vertical bars (rectangles). The gantries can roll on two rails, which in figure 1 are shown as two horizontal double lines.

3.3 Manufacturing Plan.

By a manufacturing planning system two sets of tasks - T1 and T2 - have been laid out. At the beginning of the task execution one robot is assigned to each task set. In figure 1 the two task sets - marked T1 and T2 - are represented by two sets of rectangles, where each rectangle represents a subtask - marked Tij. The length of each rectangle represents the expected duration of the task.

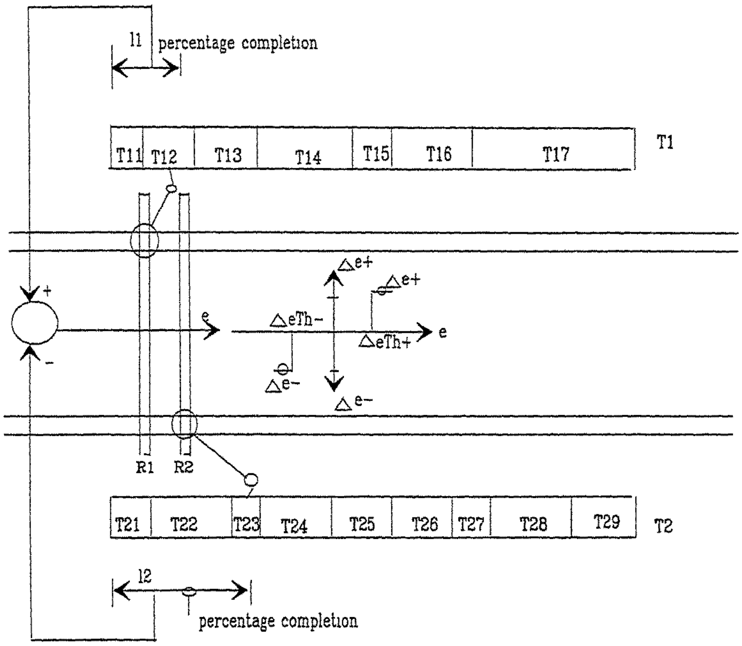


Figure 1.

3.4 Manufacturing Planning Criterion.

As it appears from figure 1 the durations of the two sets of tasks are planned and expected to be equal, for instance because the two sets of tasks are carried out simultaneously on the same workpiece e.g. robotic welding tasks on the same workpiece with two robots on the same gantry.It is thus desired that the two sets of tasks should be started and completed simultaneously.

3.5 System Control Reference.

Based on the production plan and understanding of the task processing a reference for the course of the task processing can be derived. In control engineering terms a reference for the course of the task processing is called a task execution control reference trajectory. For the type of system being treated here typically a task execution control reference trajectory is the reference position of the processing tool as a function of time. In figure 2 the curve T1R(t) is the task execution control reference trajectory for task set T1 of figure 1 and T2R(t) is the task execution control reference trajectory for task set T2 of figure 1.

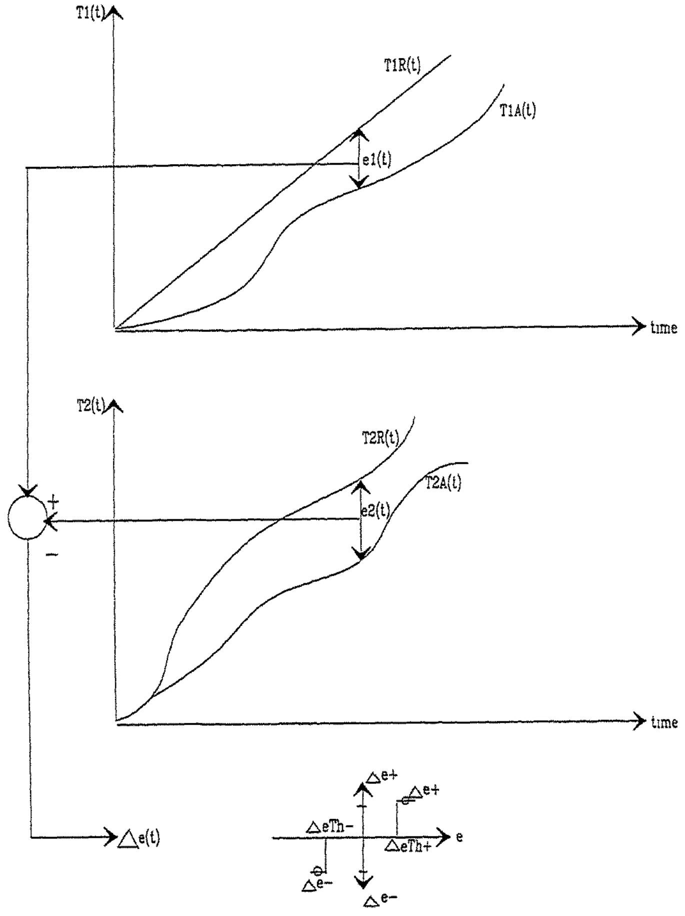


Figure 2

3.6 Processing Disturbances.

Due to (unforeseeable) disturbances of the task processing the durations of the tasks may not be equal to the expected durations.

3.7 Error in Relation to Reference Trajectory.

This means that the actual (or realised) task execution trajectory will deviate from the planned one, which is the task execution control reference trajectory. In figure 2 the actual task execution trajectory for task set T1 is shown as the curve T1A(t) and the actual task execution trajectory for task set T2 is shown as the curve T2A(t). As it is seen from figure 2 the actual task execution trajectory T1A(t) of task set T1 lacks behind the task control reference trajectory T1R(t) by an error e1(t) and analogously does the actual task execution trajectory T2A(t) of task set T2 lack behind the task control reference trajectory T2R(t) by an error e2(t). In the following these errors are called reference trajectory errors.

3.8 Control Goal.

What may be avoided or at least minimised by compensation actions is idle time of one robot, while it waits for a delayed robot to complete a task set. An idle robot may not be capable of moving to another task on another workpiece, while a workpiece on which the two task sets may be performed is still in its processing position. Therefore, if the processing of one task gets ahead of the processing of the other task, the task which is ahead, should release resources to assist on the task, which is behind. Several strategies to achieve this goal can be used. It is not the purpose of this paper to discuss such strategies.

3.9 Control Error.

A strategy has been chosen by which compensating reaction is taken as soon as execution of one of the tasks gets ahead of the execution of the other one by more than a chosen task progress difference threshold. Only when disturbances have arisen, and only when a difference between the reference trajectory errors e1(t) and e2(t) (See figure 2 and section 3.7) has arisen, compensating reactions should be taken. In the following this difference will be called the delay error or the control error. Hence the control error is defined as $\Delta e = e1(t) - e2(t)$. It has been decided, that compensating actions should only be taken if either the control error Δe becomes larger than a threshold Δeth+ or less than a threshold Δeth- (See figures 1 and 2).

3.10 Control Strategy.

As long as the task sets T1 and T2 progress equally, i.e. when Δeth- < Δe < Δeth+ , the robot assigned to a particular task set (T1 or T2) should continue with that task set. If a task set processing state gets ahead of the processing state of the other task set processing state, i.e. if either Δe < Δeth- or Δeth+ < Δe, the robot, which processes the task set, which is ahead, must be transferred to assist on the task set, the processing state of which is behind. A robot should, however, not start processing any other task, until it has completed the subtask, which is currently being processed. A subtask can be processed by one or two robots.

3.11 Selected State Variables.

Only a model for one of the subsystems is developed in the following. The processing velocity of each of the two subsystems can only be influenced by the control system by variation of the number of resources, which is assigned for processing the task set of each of the subsystems. Therefore a binary variable $T1(k) = x_1(k)$, which indicates whether resources are assigned for processing task set one or not is chosen for one state variable. If resources are assigned to task set one T1(k)= 1. If no resources are assigned for processing task set one T1(k) = 0. As mentioned in section 3.10 a robot should not start processing any other task, until it has completed the subtask, which it is currently processing. In order to know whether or not a robot can be released from assignment to a task set, it is necessary to know whether a subtask has been completed or not. A variable s1(k) is set to one, when a subtask of task set one has been completed. When a robot starts processing a subtask s1(k) is set to zero. s1(k) is chosen for state variable 2, hence $s1(k) = x_2(k)$.

3.12 Selected Disturbance Input Variables.

According to section 3.10 one type of action should be taken, if Δeth+ < Δe, another type of action should be taken, if Δeth- < Δe < Δeth+ , and yet another action should be taken, if Δe < Δeth-. In order to distinguish in which one of the three intervals Δe is placed two binary state variables are necessary. *One* of these variables will be called Δe+(k).

If Δeth+ < Δe, then Δe+(k) = 1, else Δe+(k) = 0. *The other one* is called Δe-(k).

If Δe < Δeth-, then Δe-(k) = 1, else Δe-(k) = 0.

Which value Δe+(k) and Δe-(k) will take depends on the development of the disturbances to the subtask processes as presented in section 3.6. Therefore Δe+(k) and Δe-(k) are interpreted disturbance variables. They are given the following disturbance variable names:

$\Delta e+(k) = d_1(k)$ and $\Delta e-(k) = d_2(k)$.

3.13 Table of Truth.

Based on the chosen control strategy - presented in section 3.10 -, the choises of state variables - presented in section 3.11 - and the choises of disturbance variables - presented in section 3.12 - a table of truth for realisation of the control strategy of section 310 has been developed. In table 1 this table of truth is presented. The disturbance combination of the rows R16 through R19 of table 1 are meaningless. The table of truth of table 1 therefore is incompletely specified.

R1 C1	C2	C3	C4	C5	C6	C7
R2	d1(k)	d2(k)	x2(k)	x1(k)	x2(k+1)	x1(k+1)
R3	Δe+(k)	Δe-(k)	s1(k)	T1(k)	s1(k+1)	T1(k+1)
R4	0	0	0	0	0	0
R5	0	0	0	1	1	1
R6	0	0	1	0	0	1
R7	0	0	1	1	0	1
R8	0	1	0	0	0	0
R9	0	1	0	1	1	1
R10	0	1	1	0	0	1
R11	0	1	1	1	0	1
R12	1	0	0	0	0	0
R13	1	0	0	1	1	1
R14	1	0	1	0	1	0
R15	1	0	1	1	1	0
R16	~~1~~ -	~~1~~	~~0~~	~~0~~		
R17	~~1~~	~~1~~	~~0~~	~~1~~		
R18	~~1~~	~~1~~	~~1~~	~~0~~		
R19	~~1~~	~~1~~	~~1~~	~~1~~		

Table 1

3.14 State Space Model.

According to sections 2.4, 2.5 and 2.6 a state space equation for the control of the execution of each of the task sets can be developed by inserting - for each of the rows of the table of truth - the values of that row into equation (2). For a completely specified table of truth this yields 2^n equations for determination of 2^n parameters for the state equation

$x_i(k+1) = \mathbf{F}(\mathbf{X}(k),\mathbf{u}(k))$ for each of the n state variables. In this paper only the state equation for one of the two state variables will be derived for demonstrating the possibility, that such an equation can be set up. For that purpose the equation for $x_1(k+1) = T1(k+1) = f_1(\mathbf{X}(k),\mathbf{D}(k))$ will be set up. Here $\mathbf{D}(k) = (d_1(k),d_2(k))$. For a completely specified table of truth with 4 independent variables and one dependent variable the 2^n equations have been set up and solved for the parameters of the equations. Insertion of the determined set of parameters into equation (2) and replacement of y of equation (2) with $x_1(k+1)$ yields the following state equation for $x_1(k+1)$:

(5): $x_1(k+1) = x_1(k) + x_2(k) - x_1(k)\bullet x_2(k) - x_2(k)\bullet d_1(k)$

or with $x_1(k+1) = T1(k+1)$, $x_1(k) = T1(k)$, $x_2(k) = s1(k)$ and $d_1(k) = \Delta e+(k)$

(6) : $T1(k+1) = T1(k) + s1(k) - T1(k) \bullet s1(k) - s1(k) \bullet \Delta e+(k)$

In order to have a full understanding of this equation it is necessary also to know the corresponding state equation x2 (k+1) = f2 (**X**(k),**D**(k)) in order to understand the influence of s1(k) and Δe-(k) on T1(k+1).

4.CONCLUSION.

The main result of the tests reported in this paper is that the performance of the two presented types of systems, which are systems characteristic for manufacturing systems, <u>can</u> be described by arithmetic state space equations.

5.REFERENCES:

[1]:Dieter Franke: "Sequentielle Systeme. Binäre und Fuzzy Automatisierung mit arithmetischen Polynomen", Braunschweig, Wiesbaden, Vieweg Verlag, 1994.

FUNCTION AND BEHAVIOUR DESCRIPTION IN INTELLIGENT MECHANICAL CAD SYSTEM

C.L. Li, S.T. Tan* and K.W. Chan*
Department of Manufacturing Engineering, City University of Hong Kong, Hong Kong
***Department of Mechanical Engineering, University of Hong Kong, Hong Kong**

ABSTRACT

This paper presents a language that represents the functions and behaviours of a class of mechanical devices whose functions and behaviours can be characterized by the effort and motion involved, and the axis associated with the effort or motion is fixed. The language is based on the *effort and motion state* associated with the axis concerned. Using the language of *effort and motion state*, mechanical functions can be described by a single effort and motion state, a sequence of states or several independent effort and motion states. Examples are given to demonstrate that the language can represent functions such as motion conversion, functions that involves feasible and infeasible behaviours, and that the method can differentiate functions that require energy input from functions that are self-activated. A graph data structure called the specification graph is proposed to represent the design specification of multiple input multiple output and multiple state mechanical device.

KEYWORDS

Functional modeling, Function description, Intelligent CAD

1. INTRODUCTION

With the advance in artificial intelligence, intelligent CAD system will be an indispensable tool in rapid product development. The pre-requisite to the development of an intelligent mechanical CAD system is the ability to represent the functions and behaviours of mechanical devices. A language that is used to describe functions of mechanical device must address the following issues. It must be able to describe a considerable number of functions within the scope of interest. It is desirable that the language is extensible in the sense that it can accommodate new functions that are not yet known when the language is devised. It is common that initial design specification is always incomplete and functional requirements at different levels of details emerge during the design process. It is thus important that the language should support partial or complete description of a function.

The present study concerns with the class of mechanical devices whose functions and behaviours can be characterized by the effort and motion involved, and the axis associated with the effort or motion is fixed. The language for describing effort and motion is first introduced. Examples is given in a later section to illustrate its use in describing mechanical functions and explain how the above issues are addressed.

2. EFFORT AND MOTION OF OBJECT

Consider an object O_A which is constrained to move in a fixed axis X_A. Let $MO(O_A)$ denotes the motion direction of O_A, and $MO(O_A) \in \{+, 0, -\}$ which has the following meaning :

$MO(O_A) = +$ means that O_A moves in the positive direction of X_A.

$MO(O_A) = -$ means that O_A moves in the negative direction of X_A.

$MO(O_A) = 0$ means that O_A has no motion.

Let $\overrightarrow{EI}(O_A)$ denotes the effort input to O_A along X_A. $\overrightarrow{EI}(O_A)$ is specified by a magnitude $|EI(O_A)|$ and a direction $EI(O_A)$, and $EI(O_A) \in \{+, 0, -\}$ which has the following meaning :

$EI(O_A) = +$ means that an effort is applied to O_A in the positive direction of X_A.

$EI(O_A) = -$ means that an effort is applied to O_A in the negative direction of X_A.

$EI(O_A) = 0$ means that no effort is applied to O_A.

Let $EM'(O_A)$ be an ordered pair defined as $EM'(O_A) = (EI(O_A), MO(O_A))$. If $\overrightarrow{EI}(O_A)$ is the only effort applied to O_A, O_A must move in the direction of $EI(O_A)$, and must remain stationary when no effort is applied[1]. Thus, $EM'(O_A) \in \{ (+,+), (-,-), (0,0) \}$.

When considering a kinematic pair formed by two interacting components O_A and O_B, the situation is more complicated. For example, effort applied to O_A may be transmitted to O_B, which output an effort that acts as a source to some other objects. The contact between O_A and O_B may result in a constraint in the motion of O_A and/or O_B.

To describe the possible motion constraints applied to O_A, $MC(O_A) \subseteq \{+,-\}$ is defined :
$MC(O_A) = \{ + \}$ means that motion in the positive direction of X_A is constrained.
$MC(O_A) = \{ - \}$ means that motion in the negative direction of X_A is constrained.
$MC(O_A) = \{+,-\}$ means that motion in both directions of X_A is constrained.
$MC(O_A) = \phi$ means that motion is not constrained.

The *effort and motion* state of an object O_A, denoted as $\overrightarrow{EM}(O_A)$, is defined as the ordered triple $\overrightarrow{EM}(O_A) = (\overrightarrow{EI}(O_A), MO(O_A), MC(O_A))$. If only the direction of the effort is considered, $EM(O_A)$ is defined as follows :

$$EM(O_A) = (EI(O_A), MO(O_A), MC(O_A))$$

The *effort and motion state* of a kinematic pair with components O_A and O_B, denoted as $\overrightarrow{EM}(O_A, O_B)$, is defined as the ordered pair $<\overrightarrow{EM}(O_A), \overrightarrow{EM}(O_B)>$. This notation can be extended to a mechanical device with n components O_i, $i=1,2,\ldots,n$. Its effort and motion state is described by the ordered n-turple $<\overrightarrow{EM}(O_1), \overrightarrow{EM}(O_2), \ldots, \overrightarrow{EM}(O_n)>$.

3. MECHANICAL FUNCTION DESCRIPTION

Many mechanical functions can be described in terms of effort and motion state. When using effort and motion state to describe function associated with a motion axis X_i (without referring to any physical component), the notation $EM(X_i) = <EI(X_i), MO(X_i), MC(X_i)>$ is used.

The table below illustrates few typical examples. Functions can be described by a single effort and motion state, a sequence of states or a several independent effort and motion states. For functions that involve motion conversion from input axis X_i to an output axis X_o, the input axis is specified with non-zero effort $EI(X_i)$ with an associated motion $MO(X_i)$, and $EI(X_i) = MO(X_i)$. This implies that energy is input through X_i. The output is specified with a zero effort input and a non-zero motion. That is, $EI(X_o)=0$ and $MO(X_o)\neq 0$. This implicitly implies that the output motion is being activated by the input motion. To specify a reciprocating output motion, $MO(X_o)$ is set as an alternate sequence of + and -, indicating that its motion direction of X_o changes continuously. Similarly, intermittent output motion is specified by setting $MO(X_o)$ as an alternate sequence of + and 0. For functions that involve infeasible behaviour, such as irreversible motion transmission, ratchet and lock, the infeasible behaviour of a axis X_o can be specified by $MO(X_o)=0$ and $EI(X_o)\neq 0$.This implies motion cannot be activated with an application of effort.

The method can differentiate function that requires energy input from function that is self-activated. As mentioned above, energy input through a axis X_i is specified by setting $EI(X_i)=MO(X_i)\neq 0$. For function that involves self-activated motion, such as self-return, the return action of a axis X_i is specified by $MO(X_i)\neq 0$ and $EI(X_i)=0$. That is, the motion is self-activated without the application of an input effort.

[1] It is assumed that every motion must be activated by an effort. This can be justified by the fact that a minimum effort is required to overcome friction in most mechanical device. Device such as flywheel which stores energy and activates motion is treated as an effort source in our study.

Mechanical Functions	Descriptions using effort and motion states
Motion conversion : constant to reciprocating	If X_i and X_o denotes respectively the input and output axis, $EM(X_i\ ,\ X_o)$ is an infinite sequence : <(+, +, 0) (0, +, 0)> <(+, +, 0) (0, -, 0)> <(+, +, 0) (0, -, 0)> <(+, +, 0) (0, +, 0)> ...
Motion conversion : constant to intermittent	If X_i and X_o denotes respectively the input and output axis, $EM(X_i\ ,\ X_o)$ is an infinite sequence : <(+,+,0) (0,+,0)> <(+,+,0) (0,0,0)> <(+,+,0) (0,+,0)> <(+, +, 0) (0, 0, 0)>) ...
Irreversible motion transmission	Feasible forward transmission from X_i to X_o : $EM(X_i, X_o)$ = <(+,+,0) (0,+,0)>) $EM(X_i, X_o)$ = <(-,-,0) (0,-,0)> Infeasible backward transmission from X_o to X_i : $EM(X_i, X_o)$ = <(0,0,0) (+,0,0)> $EM(X_i, X_o)$ = <(0,0,0) (-,0,0)>
Ratchet	Feasible forward motion : $EM(X_i\ )$ = (+,+,0) Infeasible backward motion: $EM(X_i)$ = ((-,0,0)
Clutch : When clutch is engaged, input motion is transmitted to output. When clutch is disengaged, motion is not transmitted to output.	Engaged : $EM(X_i, X_o)$ = <(+,+,0) (0,+,0)> $EM(X_i, X_o)$ = <(-,-,0) (0,-,0)> Disengaged : $EM(X_i, X_o)$ = <(+,+,0) (0,0,0)> $EM(X_i, X_o)$ = <(-,-,0) (0,0,0)>
Lock : When the door is locked, it cannot be pushed open. When it is unlocked, it can be opened.	If X_d denotes the door, Locked : $EM(X_d)$ = (+, 0, 0) Unlocked : $EM(X_d)$ = (+, +, 0)
Self-return : If an effort is applied the switch, it move from the neutral to a displayed position. When effort continues to be applied, switch remains at the displayed position. When effort is released, switch returns to the neutral position.	If X_i denotes a self-return switch, $EM(X_i)$ = (+, +, 0) (+, 0, 0) (0, -, 0)

A common method of function description is to devise a set of a pre-defined vocabulary[1,2,3] (such as constant to reciprocating, clutch, lock, self-return etc.) from which each mechanical function is assigned a unique name. Knowledge about each function is pre-defined in the knowledge base using methods such as transformation rule[4] or hierarchical function tree[5]. A limitation of such approach is that if new function is to be added, the knowledge base must be updated. Therefore, it is difficult to accommodate new functions that are not yet known when the knowledge base is built. Using the language of effort and motion state, any new function can be described if that function can be characterized by the effort and motion involved.

The method naturally supports partial or incomplete description of function. For example, the lock function description shown above only specifies the locked and unlocked state of the locking function. To be more specific about the lock function, the method of locking (i.e. the transition from the unlocked state to the locked state) can be specified. For an automatic door lock, the locked state is usually achieved by pushing the door to the closed position. For an manual lock, not only the door has to be positioned to the locked position, but also the door knob has to be rotated to the locked position manually. To differentiate the difference between the two types of locks, more effort and motion states can be used to specify the method of locking in the function description. An example in the next section illustrates the details.

4. ELEMENTAL FUNCTION AND MULTIPLE STATE DEVICE

The term **elemental function** is used to refer to the basic element of a function that can be described by a single effort and motion state. As noted from the previous section, the function of a device may be described by several elemental functions. A multiple state device is a device where its function changes at different state. A state is a collection of functions. The functions can be describe by a single elemental function, a sequences of elemental functions or several independent elemental functions. For multiple inputs multiple outputs device, an elemental function may involve one or more of the inputs and outputs simultaneously.

The following table illustrates the functional description of a simple door lock. The mechanical function associated with each state is underlined and the corresponding elemental functions used to describe each function is given.

State and functions	Elemental functions of the door X_D and handle X_H
State : Locked Function : Door being Locked in both opening and closing directions.	$EM(X_D) = (+,0,\phi)$ $EM(X_D) = (-,0,\phi)$
State : Opening Functions : Handle being pushed-to-move to the opened position. Handle being pushed-to-maintain at the opened position and the door being pushed-to-move to the opened position. Handle self-returns to the closed position.	$EM(X_H) = (+,+,\phi)$ $EM(X_D, X_H) = <(+,+,\phi)\ (+,0,\phi)>$ $EM(X_H) = (0,-,\phi)$
State : Unlocked Function : Door is free-to-move.	$EM(X_D) = (+,+,\phi)$ $EM(X_D) = (-,-,\phi)$
State : Closing Function : Door being pushed-to-move to the closed position. The handle remains stationary at the closed position.	$EM(X_D\ , X_H) = <(-,-,\phi)\ (0,0,\phi)>$

5. DESIGN SPECIFICATION

As noted from the previous example, the function in each state of a mechanical device can be described in terms of the elemental functions. With suitable arrangement of the elemental functions, a functional specification can be formed. The method of specifying the functions of a multiple input, multiple output and multiple state kinematic device is based on the observation that for each elemental function, it is associated with a single configuration or a transition from one configuration to another. The specification specifies the configurations, the transition between configurations and the associated elemental function. The specification is represented using a digraph called the specification graph.

Let S denotes the specification graph

$$S = (C,F)$$

where

C is the vertex set which is the set of all configuration c_j.

F is the arc set which is the set of all elemental functions.

For a specification with n inputs/outputs, each input/output is represented as a motion axis X_i, i=1,2,...,n. C is the set of all configurations c_j that are associated with one or more elemental functions. Each configuration c_j is an ordered n-turple $c_j = (x_1, x_2,...,x_n)$, where x_i is called a position label which represents the value of the configuration parameter of X_i at the jth configuration. For each X_i, the set of all x_i is denoted by $\overline{X_i}$. That is, $\overline{X_i} = \{ x_i \mid c_j \in C \}$. An arc $f \in F$, directed from vertices c_k to c_l, specifies an elemental function that involves configuration change from c_k to c_l. The elemental function f is specified by an effort and motion state of the motion axes $EM(X_1, X_2, ..., X_n)$. If an elemental function does not involved configuration change, the corresponding arc starts and ends at the same vertex.

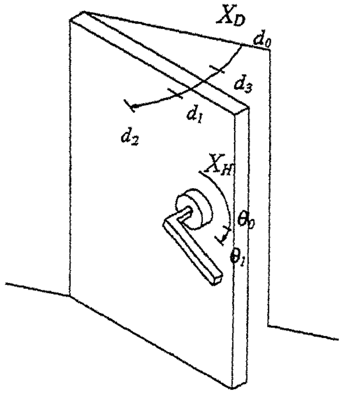
Figure 1 shows the design specification of an automatic door lock. The elemental function is specified through the effort and motion state of axes X_D and X_H, which represents the motion of the door and the handle respectively. In the specification graph shown in Figure 1(b), the Locked state involves two elemental functions f_1 and f_2, which are loops incident at a single configuration c_1. The Opening state involves elemental functions f_3, f_4 and f_5, which corresponds to a path from c_1, via c_2 and c_3 to c_4. The Unlocked state involves elemental functions f_6 and f_7, and the Closing state involves a single elemental function f_8.

6. CONCLUSION AND FURTHER WORK

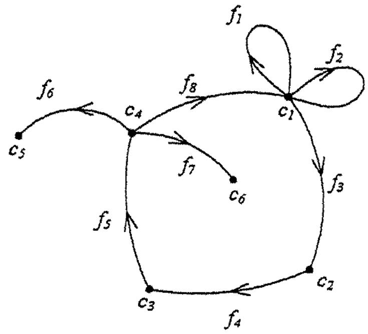
A Language that can be used to describe the functions and behaviours of a class of mechanical devices has been developed. Examples have been given to show that the language is able to describe most of the common mechanical functions. It is extensible in the sense that it can accommodate new functions and it supports partial description of a function. Based on this language, a graph data structure known as the specification graph is proposed to represent the functional specification of mechanical devices with multiple inputs, multiple outputs and multiple operation states. n.

7. REFERENCES

1. C.L. Li, S.T. Tan and K.W. Chan, "A qualitative and heuristic approach to the conceptual design of mechanisms" Engineering Applications of Artificial Intelligence, Vol 9 No 1, pp 17-32, Feb 1996
2. J.A. Collins, B.T. Hagan and H.M. Bratt, "The Failure-Experience Matrix - A Useful Design Tool", Transactions of the ASME, Journal of Engineering For Industry, Vol 98, pp 1074-1079, 1976
3. K. Lai and W.R.D. Wilson, "FDL - A Language for Function Description and Rationalization in Mechanical Design", Transactions of the ASME, Journal of Mechanisms, Transmissions, and Automation in Design, Vol 111, pp 117-123, 1989
4. S.P. Hoover and J.R. Rinderle, "A Synthesis Strategy for Mechanical Devices", Research in Engineering Design, Vol 1, pp 87-103, 1989
5. F.M. Hashim, N.P. Juster and A. de Pennington, "A Functional Approach to Redesign", Engineering with Computers, Vol 10, pp125-139, 1994.



(a) Motion axes X_D , X_H and the position labels



(b) The specification graph

f_1: $<(\overrightarrow{F_{D1}}, 0, 0)\ (0, 0, 0)>$
f_2: $<(\overrightarrow{F_{D2}}, 0, 0)\ (0, 0, 0)>$
f_3: $<(0, 0, 0)\ (\overrightarrow{\Gamma_\theta}, +, 0)>$
f_4: $<(\overleftarrow{F_{D1}}, +, 0)\ (\overrightarrow{\Gamma_\theta}, 0, 0)>$
f_5: $<(0, 0, 0)\ (0, +, 0)>$
f_6: $<(\overrightarrow{F_{D1}}, +, 0)\ (0, 0, 0)>$
f_7: $<(\overrightarrow{F_{D2}}, -, 0)\ (0, 0, 0)>$
f_8: $<(\overrightarrow{F_{D2}}, -, 0)\ (0, 0, 0)>$

(c) Effort and motion states $EM(X_D, X_H)$ of the elemental functions

$c_1 = (d_0, \theta_0)$	$F_{D1} = +$
$c_2 = (d_0, \theta_1)$	$F_{D2} = -$
$c_3 = (d_1, \theta_1)$	$\Gamma_\theta = +$
$c_4 = (d_1, \theta_0)$	
$c_5 = (d_2, \theta_0)$	
$c_6 = (d_3, \theta_0)$	

(d) The configurations and input effort directions

Figure 1. The design specifcation of an automatic door lock

Dimension-driven Parameterized Design of Free Form Shapes

Yu Kai-Ming and Zhu Wei-Ming
Department of Manufacturing Engineering
The Hong Kong Polytechnic University
Hong Kong

ABSTRACT

The paper develops a theoretical basis for dimension-driven parameterized design of free form shapes. Dimension-driven parameterized design of free form shape is to obtain a new free form shape similar to the old one but with a different set of dimension parameter instantiation. In other words, a new member of the free form shape's family will be designed. However, research on dimension-driven parameterized design of free form shapes is an unattempted area. Also, there is still no commercial CAD system that can handle dimension-driven free form shapes successfully. In this research, we exploit a new concept of similar shape. Dimension-driven variation of similar free form shapes is then explained and illustrated with an example. Inverse design of similar shape is also studied.

KEYWORDS

Dimension-driven; Parameterized design; Free form shapes

1. INTRODUCTION

Dimension-driven parameterized design of free form shape is to obtain a new free form shape that is similar to the old one but with different set of dimension parameter values. In other words, a new member of the free form shape's family will be designed. This technology is very useful in variational design. However, there is still no commercial CAD system that can handle dimension-driven free form shapes successfully.

In dimension-driven free form shape research, the first problem to be solved is a proper definition of shapes. This includes how to describe shapes, how to define similar free form shapes and how to preserve shapes in free form geometry variation. Some related works on free form shapes have been done in recent years. Su and Jones studied a shape-preservation theorem [1,2]. Kantorowitz extended the theorem to cases of hyperconvex and loops polygons [3]. Goodman investigated on shape preserving properties on generalizing Ball's work [4]. Piegl discussed the use of shape handles to modify the shape of rational B-spline curves [5]. Although these works covered free form shape preservation, research on dimension-driven parameterized design of free form shapes is still an unattempted area. The paper thus proposes a theoretical basis for dimension-driven parameterized design of free form shapes.

We first revise the definition of similar polygonal shape and then investigate the different types of similarity. Some possible ways to obtain similar polygons are also shown. Next, similar free form curves are defined in relation to similar polygons. Dimension-driven variation of similar free form shapes is then explained and an example is given. Inverse design of similar shape is also studied. Finally, conclusions and future work of this research are discussed.

2. SIMILAR POLYGONAL SHAPE

2.1. Definition of Similar Polygons

We define two kinds of similarity for polygons. One is *metric similar* and the other is *shape similar*.

The definition of metric similar is derived from that of similar triangles. In Euclidean geometry, two triangles that have equal corresponding angles or proportional corresponding side lengths are considered similar. Since polygons can always be triangulated, they can be considered similar if all the corresponding triangles in the tessellation are similar. These similar polygons are called metric similar, that is, they have equal corresponding angles and proportional corresponding side lengths. An example of two metric similar pentagons is shown in Figure 1. The definition of metric similar is, however, too restrictive. Instead, the quantitative shape control restriction is relaxed, such that shape similar only constrains the shape qualitatively. For a pair of shape similar polygons, the corresponding angles only need to be of the same type, say both acute or both convex, and corresponding side lengths need not have the same proportional relationship. Obviously, metric similar can be considered as a special case of shape similar. The definition of shape similar for polygons is based on their corresponding angles (α) and ratios of their corresponding side lengths (r). Nine types of similar polygons are listed in Table 1.

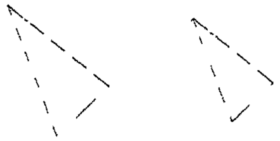
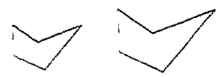
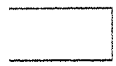
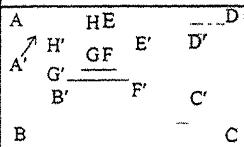
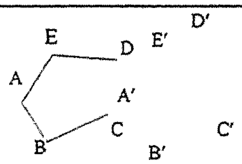
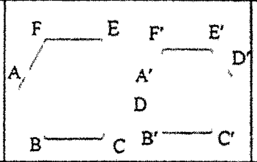
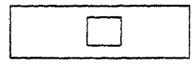
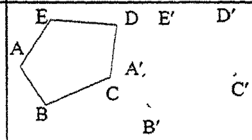


Fig. 1: Metric similar shapes.

Table 1: Nine types of similar polygons.

Type	Corresponding Angles	Corresponding Side Lengths	Mathematical Expression	Example	Remarks
I	All equal	All have the same proportional constant.	$\alpha_i = \alpha_i'$ $r_i = const$ $(\forall i = 1,2,\dots,n)$		
II	Some but not all are equal	All have the same proportional constant.	$\alpha_i = \alpha_i' \wedge \alpha_j \neq \alpha_j'$ $(\forall i, j = 1,2,\dots,n,$ where $i \neq j$ $r_k = const.$ $(\forall k = 1,2,\dots,n.)$		$\angle A = \angle A'$
III	None equal.	All have the same proportional constant.	$\alpha_i \neq \alpha_i'$ $r_i = const$ $(\forall i = 1,2,\dots,n.)$		
IV	All equal.	Some but not all share the same proportional constant	$\alpha_i = \alpha_i'$ $(\forall i = 1,2,\dots,n.)$ $r_j = const. \wedge r_k \neq const.$ $(\forall j, k = 1,2,\dots,n,$ where $j \neq k$		$\frac{AB}{A'B'} = \frac{CD}{C'D'}$ $\frac{EF}{E'F'} = \frac{GH}{G'H'}$
V	Some but not all equal	Some but not all share the same proportional constant.	$\alpha_i = \alpha_i' \wedge \alpha_j \neq \alpha_j'$ $(\forall i, j = 1,2,\dots,n,$ where $i \neq j)$ $r_k \neq const. \wedge r_l = const.$ $(\forall k, l = 1,2,\dots,n,$ where $k \neq l$		$\angle A = \angle A'$ $\angle B = \angle B'$ $\frac{AB}{A'B'} = \frac{CD}{C'D'}$
VI	None equal.	Some but not all share the same proportional constant.	$\alpha_i \neq \alpha_i'$ $(\forall i = 1,2,\dots,n.)$ $r_j = const. \wedge r_k \neq const.$ $(\forall j, k = 1,2,\dots,n,$ where $j \neq k$		$\frac{BC}{B'C'} = \frac{EF}{E'F'}$

VII	All equal	None has the same proportional constant.	$\alpha_i = \alpha_i'$ $r_i \neq$ const. $(\forall i = 1,2,\dots,n.)$		
VIII	Some but not all equal	None has the same proportional constant.	$\alpha_i = \alpha_i' \wedge \alpha_j \neq \alpha_j'$ $(\forall i, j = 1,2,\dots,n,$ where $i \neq j)$ $r_k \neq$ const. $(\forall k = 1,2,\dots,n.)$		∠A=∠A′ ∠B=∠B′
IX	None equal.	None has the same proportional constant.	$\alpha_i \neq \alpha_i'$ $r_i \neq$ const $(\forall i = 1,2,\dots,n.)$		

Among the nine types of shape similarities, we can say that type I is the tightest while type IX is the loosest. Also, for similarities that have different corresponding angle values, the polygons are similar only if the angles have the same type.

2.2. Similar Polygon Transformation

Having defined similar polygons, the next question to ask is how to obtain them. Some possible approaches such as *uniform offset, differential offset, side rotation* and *local vertex movement* will thus be evaluated.

Scaling. Scaling is used to increase or decrease the size of a geometry. If s_x and s_y are scaling factors in the X and Y-directions respectively, then for any two points, we have

$$r_i = \frac{\|\mathbf{P}_i'\mathbf{P}_{i+1}'\|}{\|\mathbf{P}_i\mathbf{P}_{i+1}\|} = \frac{\sqrt{s_x^2(x_i - x_{i+1})^2 + s_y^2(y_i - y_{i+1})^2}}{\sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2}} \tag{1}$$

For any three points,

$$\begin{aligned} \mathbf{P}_{i-1}\mathbf{P}_i \cdot \mathbf{P}_i\mathbf{P}_{i+1} &= (x_i - x_{i-1})(x_{i+1} - x_i) + (y_i - y_{i-1})(y_{i+1} - y_i) \\ &= \|\mathbf{P}_{i-1}\mathbf{P}_i\| \cdot \|\mathbf{P}_i\mathbf{P}_{i+1}\| \cos\alpha \\ \mathbf{P}_{i-1}'\mathbf{P}_i' \cdot \mathbf{P}_i'\mathbf{P}_{i+1}' &= s_x^2(x_i - x_{i-1})(x_{i+1} - x_i) + s_y^2(y_i - y_{i-1})(y_{i+1} - y_i) \\ &= \|\mathbf{P}_{i-1}'\mathbf{P}_i'\| \cdot \|\mathbf{P}_i'\mathbf{P}_{i+1}'\| \cos\alpha' \end{aligned} \tag{2}$$

Thus, for uniform scaling where $s_x = s_y$, $r_i =$ const. and $\alpha_i' = \alpha_i$. Differential scaling occurs when $s_x \neq s_y$, that is, $r_i \neq$ const. and $\alpha_i' \neq \alpha_i$.

Offset. Offset is a special case of translation in which polygonal sides are moved parallel to the original ones. New vertices are obtained by intersecting the corresponding sides.

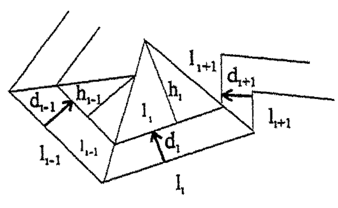


Fig. 2: Offset.

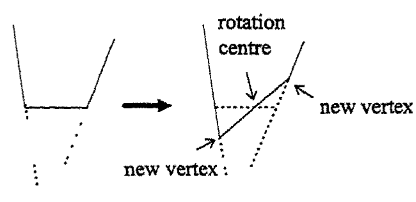


Fig. 3: Side rotation.

For each corresponding side (see Fig. 2),

$$r_i = \frac{l_i'}{l_i} = \frac{h_i}{h_i + d_i} \quad (3)$$

where l is the side length, d is the offset distance and h is the height. Since h_i is not a constant, $r_i \neq$const. in general for uniform and differential offset. However, both uniform and differential offset the new sides being generated are parallel to the original ones, therefore $\alpha_i = \alpha_i'$.

Side rotation. Since rotation is a rigid motion, a rotation applied to the whole polygon will not change its size and shape. Instead, like offset, rotation is applied to individual polygonal side with respect to a certain user defined centre point. New vertices are obtained by intersecting the corresponding sides (see Fig. 3). Obviously, it would cause both $r_i \neq$ const. and $\alpha_i' \neq \alpha_i$.

Local vertex movement. Local vertex movement means to move a single vertex or a group of vertices in the polygon (see Fig. 4). Obviously, through proper pushing or pulling, all nine similarity types can be obtained.

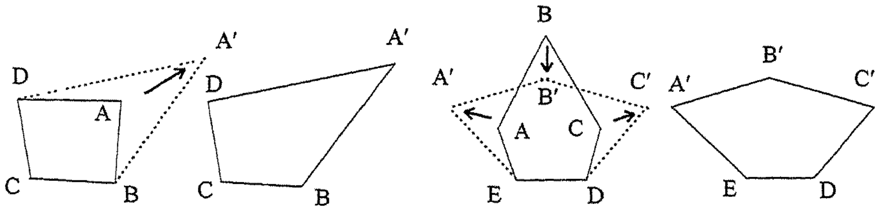


Fig. 4: Local vertex movement.

The above result is summarized in Table 2. It can be seen that only differential scaling, side rotation and local vertex movement allow for shape similar variation (i.e., $\alpha_i' \neq \alpha_i$). However, from the dot product formula, it is difficult for the user to control the angular variation through the scaling factors for differential scaling. In side rotation, only angular dimensions can be controlled but not linear nor radial dimensions. In local vertex movement, the user controls the vertex coordinates rather than the angle acuteness or convexity. Better user interface should thus be devised to allow for full dimension-driven shape similar transform.

Table 2: Approaches of similar polygon transformation and their achievable similarities.

Transformation	Similarity Obtainable	Limitation
Uniform scaling	I	$\alpha_i = \alpha_i'$, $r_i =$ const
Differential scaling	IX	Difficult to control α_i
Uniform offset	VII	$\alpha_i = \alpha_i$
Differential offset	VII	$\alpha_i = \alpha_i$
Side rotation	IX	Control angular dimensions only
Local vertex moving	I to IX	Coordinates movement

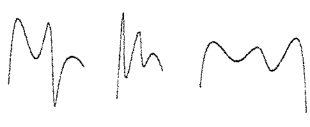
2.3. Similar 2D Free Form Curves

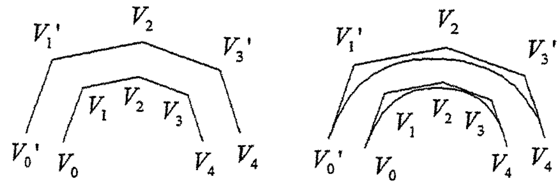
The main purpose of this research is to study the shape-preserving similar transform of free form objects. Unlike polygons, free form curves cannot be characterized by angles and side length ratios. However, from differential geometry, the shape of a free form curve is characterized by its *curvature*. Shape similar free form curves are thus defined as free form curves with the same sequence of curvature types. The curvature type can be positive, negative or zero. The sequence is a discrete distribution where absolute curvature value is not required. Examples shown in Figure 5 are three shape similar free form curves having the same sequence of curvature type: +, -, +, -, +.

From Kantorowitz's shape preservation theorem [3] and properties of NURBS (non-uniform rational B-spline) curves, a free form curve, in general, follows the shape of its control polygon. Therefore, we can conclude that similar control polygon results in similar free form curves (see Fig. 6). For instance, the expression of a similar transformed NURBS curve is,

$$\mathbf{P}'(t,\mathbf{V}) = \mathbf{P}(t,\mathbf{V}) + \Delta\mathbf{P}(t,\Delta\mathbf{V}) = \sum_{i=0}^{k}(\mathbf{V}_i + \Delta\mathbf{V}_i)\mathrm{R}_{i,k}(t) \tag{4}$$

where $\mathbf{P}(t,\mathbf{V})$ is the original curve, $\Delta\mathbf{P}(t,\Delta\mathbf{V})$ is the variational curve, $\mathbf{V}$ is the control polygon and $\mathrm{R}_{i,k}(t)$ are basis functions. For shape similar transform, we need only to maintain the control vertex labels in the control polygon to be invariant, i.e., convex remains convex or acute remains acute. The result also applies to free form curves with repeated control vertices or collinear segments.



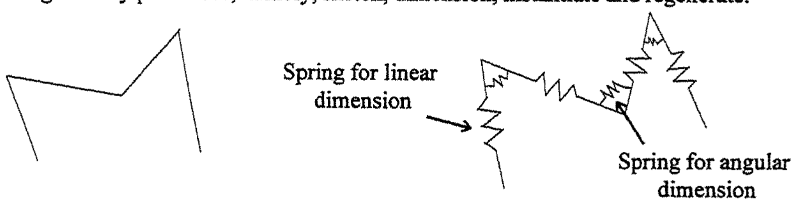


(a) (b)

Fig. 5:Shape similar free form curves. Fig. 6:Similar free form shapes from similar control polygons.

3. DIMENSION-DRIVEN VARIATION OF SIMILAR FREE FORM SHAPES

In dimension-driven parameterized design of free form shapes, a new free form shape is obtained by substituting a new set of dimension values to the old one. The dimensions we give to the free form shape are only critical dimensions characterizing the key design intent (e.g., shape or function). The size will vary while the shape have to remain similar. This is analogous to have some springs fixed on the polygon angles and sides (see Fig. 7). As a result, the relaxed shape similar is more desirable than the restrictive metric similar. Figure 8 shows a flow chart for dimension-driven parameterized design of free form shapes. It follows the steps of a typical dimension-driven variational geometry procedure, namely, sketch, dimension, instantiate and regenerate.



(a) (b)

Fig. 7: Control polygon (a) and springs fixed on it (b).

An example of a vase profile is shown in Figure 8 to demonstrate the procedure. Figure 9(a) is the initial shape. We first input the control polygon and its free form profile is generated. Next, some dimension parameters, the height and width of the vase, the length of segment DE, and the values of angles C and D are specified. The acuteness of the angles and the lengths are constrained in this example. New dimension values are then assigned. Since dimensions can be treated as metric constraints, a constraint satisfaction algorithm (numerical or geometrical reasoning) follows. If successful, a new vase similar in shape but different in size will be constructed (see Fig. 9(b)).

4. INVERSE DESIGN OF SIMILAR SHAPE

In practice, the user may define a free form curve by inputting a number of data points. The free form curve will then fit or pass through all the data points. In this case, the control polygon must be calculated first from the data points. In general, free form curves can be compactly written in the matrix form as **P**=**RV**, where **R** is the matrix of basis functions. The control polygon can be evaluated as

$$\mathbf{V} = (\mathbf{R}^T\mathbf{R})^{-1}\mathbf{R}^T\mathbf{P} = \mathbf{R}^+\mathbf{P} \tag{5}$$

where $\mathbf{R}^+$ is the pseudo-inverse of **R**. The procedure mentioned can then be used to obtain curves of similar shape.

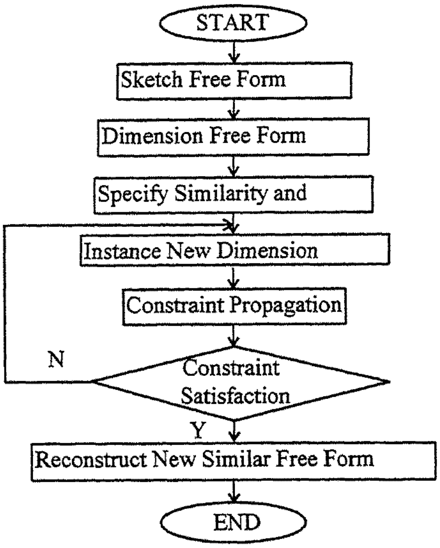


Fig. 8: Dimension-based variation of similar free form shapes.

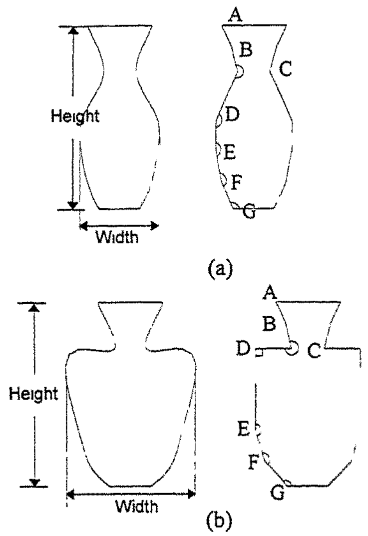


Fig. 9: An example of dimension-based variation of similar free form shapes.

5. CONCLUSIONS AND FURTHER WORKS

This paper has developed a theoretical basis for dimension-driven parameterized design of free form shapes. The theory is suitable to characterize 2D free form curves. In addition to designing customer products and engineering parts with free form features, dimension-driven free form variation is also useful in other areas. For instance, garment design, modeling of human organs, font design, deformable object modelling and reverse engineering are all potential applications. For future development, the theory can be extended to 3D curves and surfaces.

6. REFERENCES

1. Su, B.Q. and Liu, D.Y., Computational Geometry, Curve and Surface Modeling, Academic Press, Boston, 1989.
2. Jones, A.K., Shape control of curves and surfaces through constrained optimization, in Farin, G.E.(ed.), Geometric Modeling: Algorithms and New Trends, SIAM, pp265-279, 1987.
3. Kantorowitz, E. and Schechner, Y., Managing the shape of planar splines by their control polygons, Computer-aided Design, 25, 6, Jun., pp355-364, 1993.
4. Goodman, T.N.T. and Said, H.B., Shape preserving properties of the generalized Ball basis, Computer Aided Geometric Design, 8, pp115-121, 1991.
5. Piegl, L., Modifying the shape of rational B-splines, Part 1: curves, Computer-aided Design, 21, 8, Oct., pp509-518, 1989.

PETRI NET AIDED TOLERANCING FOR PRISMATIC COMPONENTS

P.Eshwaraiah, A.L.Sarat Babu, K. Ravi Raju, N.K.Tewari
Department of Mechanical Engineering
Indian Institute of Technology, Delhi
New Delhi, India.

ABSTRACT

In this paper, an attempt has been made to include tolerancing in computer aided process planning for prismatic parts. It consists of Feature Based Tolerance Modelling System (FBTMS), Tolerance Constraints Formulation (TCF), and Petri Net Aided Process Sequencing modules. The purpose of a FBTMS for CAPP is to enable the designer to represent the tolerance information. Based on the representation, tolerances are classified as inter-feature tolerances and intra-feature tolerances. In TCF, tolerance constraints such as dimensional, geometrical and technological constraints are formulated. By considering the above tolerance constraints, operation sequence is generated using the Petri net approach. A case study is taken to demonstrate the proposed methodology.

1. INTRODUCTION

Computer Aided Process Planning (CAPP) is today recognised as the "missing link" between CAD and CAM. Computer Aided Process Planning for prismatic parts is of paramount importance as they constitute approximately 54% of the total component spectrum. It is more difficult to develop a generative CAPP system for non-rotational parts due to their variety of overall structure and the complexity of part shapes and part features.

Generative CAPP systems based on features are being contemplated by researchers and practitioners. Features represent a collection of geometrical entities in an intelligent form and hence provide information at a higher conceptual level than is available in traditional CAD models. The use of such groups of geometry coupled with the necessary manufacturing related information is seen as a practical means of integrating design and manufacturing.

Geometric dimensioning and tolerancing of a work piece sets important constraints on the selection and sequencing of its manufacturing operations. A new trend is becoming apparent, in the field of computer-aided process planning, which considers tolerance implications in feature based modelling and in process sequencing. The review of some basic contributions confirm the need for developing computerized tolerancing modules for geometric modellers.

Finding the optimal sequence of operation based on tolerance constraints for a given part, is becoming an important aspect in CAPP. In this scenario, an attempt has been made on feature based tolerancing and tolerance constraints formulation; and by considering the tolerance constraints, operation sequence are generated using the Petri net approach.

2. BACKGROUND

Solid modeller based investigation of dimensioning and tolerancing representational issues was initiated by **Requicha** [7]. He proposed a tolerancing theory based on the "variational class" concept. **Kulkarni and Pande** [4] developed a system called, The Tolerance Modelling System (TMS). TMS has been designed and implemented to facilitate solid model based representation and validation of conventional and geometric tolerances and enable their use in analysis and decision making tasks. **Roy and Liu** [9] discussed the requirements of a new CAD data model and proposed a feature-based representation scheme based on the hybrid CSG/B-Rep data structure. This hybrid structure exploits the advantages of both CSG and B-Rep models in representing tolerance information. A relational graph structure (face-based data model of B-Rep) of the object is maintained at each hierarchical level of object construction for associating tolerance and other attributes.

Halevi and Weill [3] used the constraint based operation sequence and have considered the tolerance constraints like dimensional, geometrical, technological and economical constraints. They used the matrix approach for finding the operation sequence. **Chang-Xue Feng and Kusiak** [1] proposed object-oriented classification scheme for machining constraints. Algorithms are developed to check the related machining constraints for different cases of part topology. **Richard et al.** [8] identified the constraints that govern the topological configuration of the target part. They demonstrated a methodology that represents, reasons and plans with multiple knowledge types and multiple planning strategies existing in the domain of process planning. **Prabhu et al.** [6] presented a tree structured approach and recursive algorithm for generation of alternative feasible process plans. However, tree structure based recursive algorithm proposed by them inherently requires large storage space to generate alternative plans.

The aspect of tolerance consideration while sequencing has not been adequately dealt with in the works reviewed above, and very few researchers have reported work in this direction.

3. FEATURE BASED TOLERANCE MODELLING SYSTEM (FBTMS)

All current CAD databases contain information related with geometrical shapes only, while many important engineering specifications, dimensions and tolerances are not associated with geometry Purpose of these CAD databases is only to pictorially represent the object Pictures generated by CAD databases are important for human understanding but except for very limited applications, these pure geometrical representation are of little use to computer aided engineering applications. Lack of complete information from CAD database requires that additional information related with manufacturing applications is separately appended to CAD databases. Unavailability of complete information from single database presents the most severe bottleneck in way of integrating CAD and CAM systems without any human intervention [2].

To this end, this paper proposes a system based on the premise that explicit representation of manufacturing information in the CAD database is needed for process planning.

The architecture of the present modelling system is shown in the figure 1. Part is defined in terms of its geometrical shape and manufacturing features to be machined. Each feature corresponds to a cavity volume to be removed by machining. Since the modelling is feature-based which is higher level information, instead of lines, points and curves which is lower level information in the process planner point of view. This simplifies the planning greatly; however, the flexibility of modelling is limited. The tolerance information is represented in the feature based model. Figure 2 illustrates the content and classification of the geometric tolerance information. Both single feature tolerances i.e. intra-feature tolerances and related feature tolerances i.e. inter-feature tolerances are represented in feature based model. Feature's information is extracted from the modelled part. This modelling data is written to a model file for further working of process planning activity.

4. TOLERANCE CONSTRAINTS FORMULATION (TCF)

4.1 Geometric tolerance constraints

The geometric dimensioning and tolerancing of a workpiece sets important constraints. Different categories of constraints can be classified in the following way:

- dimensioning with a datum as constraint;
- geometric tolerances constraints;
- technological constraints;

4.1.1 Dimensional constraints

The operation surface is dimensionally referred with respect to another surface i.e. datum surface, so it is logical to execute the datum surface before executing the operation surface. The type of surfac e may be cylindrical or flat. Four types of relationships are identified:

- flat-flat;
- flat-cylindrical;
- cylindrical-flat;
- cylindrical-cylindrical.

Figure 3 shows these four possibilities. In Figure 6, B1, B2 are raw material surfaces, F1, F2 are flat surfaces and C1, C2 are cylindrical surfaces. It is logical that raw material surface B1 is constraint to surface F1. Similarly cylindrical, surface C1 is constraint to surface C2 which is also cylindrical surface because of dimensional relationship. Surface F2 is constraint by hole C1 which in turn, is constrained by the raw material surface B2. However, by applying the rule **Plane before hole** [3], it is decided to transfer the tolerance to another flat surface i.e. raw material surface B2. Surface F2 is therefore constrained by raw material surface B2. Hole C1 is constrained by raw material surface B2, but hole C1 is first defined in relation to F2 as a consequence of the tolerance transfer as seen earlier and is constrained by surface F2. The following algorithm identifies these constraints.

Algorithm 1: For dimensional constraints

If (operation surface A is flat and datum surface B is flat)
  Then Surface B is constraint to surface A
If (operation surface A is cylindrical and datum surface B is cylindrical)
  Then Surface B is constraint to surface A
If (operation surface A is flat and datum surface B is cylindrical)
  Then look for other related flat surfaces for surface B and transfer tolerance to that related surface. The new datum surface is constraint to surface A. (Plane before hole)
If (operation surface A is cylindrical and datum surface B is flat)
  Then If (surface A act as datum for other flat surface)
    Then transfer the tolerance to that flat surface, and the new datum surface is constraint to surface A. (Because already machined surfaces has priority over non machined surfaces)
    Else Surface B is constraint to surface A

4.1.2 Geometrical Constraints

Geometrical relationships sets data references as constraints. For example Figure 4 illustrates a geometrical constraints where the tolerance of coaxiality of hole (2) is referred to shaft (1), taken as a datum and therefore constraint. In the same way if the operation surface is perpendicular to another surface i.e. datum surface and therefore constraint.

The algorithm for the geometrical constraints is as follows:

If (Surface B is in geometrical reference to Surface A)
  Then
  Surface B is constraint to operation surface A

4.1.3 Technological constraints

In order to execute sequence of operation properly technological constraints are important. As an example two rules are given below can used for finding the technological constraints.

Rule 1: If the surface A is coaxial with surface B and both are counter boring or counter sinking operations, then go for piloted countersinking tool and surface B is used for guiding the pilot tool. So surface B is constraint to surface A.

Rule 2: If for some processes, an operation must be performed before others can proceed. For example, rough boring comes before semi-finish boring, and semi-boring comes before finish boring. That means for semi-finish boring, rough boring is constraint and for finish boring, semi-finish boring is constraint.

4.1.4 Process sequence based on geometric tolerancing constraints:

Dimensions represent geometric constraints. Tolerances are imposed by the functions(s) of the parts to be designed and thus represent functional constraints [10]. The geometric dimensioning and tolerancing of a workpiece sets important constraints on the selection and sequencing of operations. Out of 14 tolerance characteristics, flatness, circularity, cylindricity, and straightness are related to operation selection. The other 10 geometric tolerances are related to operation sequence.

Two surfaces with a strict parallelism tolerance must be machined in the same set-up in order to reach the tolerance requirement. Two holes with a close concentricity tolerance should be drilled in one set-up. These type of constraints are used to lump selected operations together into groups to simplify the operation sequence problem. A set of this category is formalized and stored in the knowledge base for the sequence of operations.

5. PETRI NET AIDED PROCESS SEQUENCE BY CONSIDERING TOLERANCE CONSTRAINTS

5.1 Petri Nets

A petriNet is a directed bipartite graph which has places, transitions and relations between them (described by directed arcs). The places are represented by small circles, transitions by lines. A token is represented as a dot in the circle indicating the place. PetriNet is executed by firing transitions. Each transition contains one or more input places and output places. A transition can only be fired if all its input places are activated. A place is said to be activated if it contains a number of tokens equal to it's weight. On firing a transition, one tokens are removed from input places, and tokens are added to output places. The total set of tokens of a PetriNet at any instant is called as marking of the PetriNet. it describes the state of execution of the PetriNet. The behaviour of PetriNets can be simulated by firing transitions and creating new markings.

5.2 Working of the process sequencing module

We know that for any activity to take place certain pre-conditions are there and for the activity is associated with time. When the activity is complete the status of the system changes and hence certain post-conditions are found. A petri net model is analogous to the above facts and its input and output places denote the pre and post conditions while the transition denotes the activity and has a certain time attached to it.

The working of the process sequencing module is illustrated in the Figure 5. Input to this module is precedence relationships(tolerance constraints) framed by TCF module. System findouts enable processes based on precedence relationships. An enabled process is one which has no constraints, which means that this process can be executed. An enabled transition fires by removing tokens from each of its input places the number of tokens equal to the weight of the corresponding incoming arcs and depositing in each of its output places the number of tokens equal to the corresponding weights of output arcs. Processes which are executed are free from constraint. Then the system looks for another process to execute, The procedure is repeated until all the processes are executed. The result of module is the operation sequence.

6. CASE STUDY

Figure 6 shows an example part chosen for case study. In figure 6, the raw material surfaces are represented as B1 to B4 and operation surfaces are represented as 1 to 9. The working of various modules is described in following sections.

6.1 FBTMS module

The part is modelled in UNIGRAPHICS solid modeller using the existing library of features. This part model is input to the FBTMS, the system identifies the each feature on the part. Then the system prompts for intra feature tolerances for selected operation surface. User has to select intra feature tolerance type and its value. Then the system prompts for inter feature tolerances for selected operation surface. User has to select tolerance type, its value and datum surface. This is repeated for all operation surfaces.

The geometrical and tolerance data of each feature is written the model file in a predetermined format. The format of the data file is suitable to read the tolerance constraint formulation module. The sample of model file is shown here:

(make-instance (gensym) of CBORE_HOLE
(id 1)
(cbore_dia 20.0000)
(dia 14.0000)
(c_bore_depth 5.0000)
(hole_depth 25.0000)
(tip_angle 0.0000)
(layer 1)
(parent nil)
(children nil)
(make-instance (gensym) of CBORE_HOLE
(id 2)
(cbore_dia 20.0000)
(dia 14.0000)
(c_bore_depth 5.0000)
(hole_depth 25.0000)
(tip_angle 0.0000)
(layer 1)
(parent nil)
(children nil)

Tolerance information

surface: 1a
surface type: cylindrical
surface finish Ra 6.3
diametrical u_tolerance 1.000, l_tolerance 0.000
concentric with 4a, tolerance value 0.050
surface: 1b
surface type: cylindrical
surface finish Ra6.3
diametrical u_tolerance 1.000, l_tolerance 0.000
concentric with 4b, tolerance value 0.050

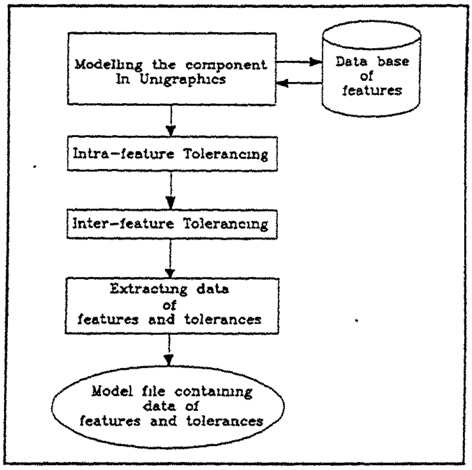


Figure 1: Architecture of FBTMS

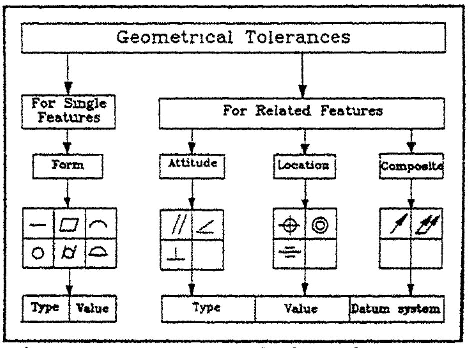


Figure 2: Classification of Geometric tolerances

6.2 Tolerance constraints formulation module

Input of this module is model files generated by the FBTMS module. This file consists of geometric feature information and tolerance information of various features. The system identifies the different classes of constraints based on the tolerance information. Then the system generates the output in a model file which can be read by process sequence module. The output of this module is as follows:

surface: 1aF
dimensional constraints :
geometrical constraints : 4aF
technological constraints : 4aF
surface: 1bF
dimensional constraints :
geometrical constraints : 4bF
technological constraints : 4bF 4aF

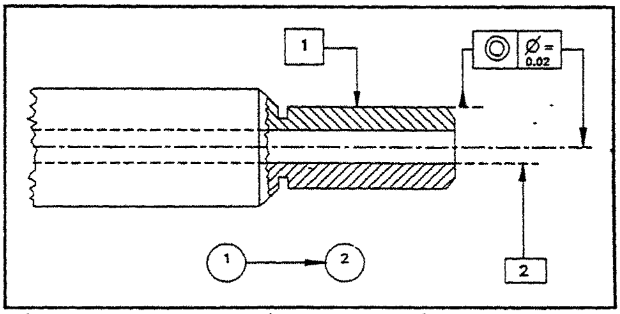


Figure 4: Geometric Constraint

6.3 Process sequencing module

Input of this module is output generated by the tolerance constraints formulation module. It generates the operation sequence. A typical PetriNet for the example component is shown in figure 7. Operation sequence can be obtained by executing Petri Net model. The generated sequence for the example component is as follows:

Operation Sequence

Operation 1:	3F	
Operation 2:	5F	
Operation 3:	7F	
Operation 4:	8F	9F
Operation 5:	4aF	
Operation 6:	1aF	4bF
Operation 7:	1bF	2aF
Operation 8:	2bF	
Operation 9:	6aF	
Operation 10:	6bF	

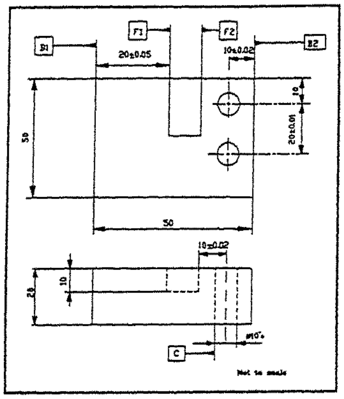


Figure 3: Types of dimensional constraints

7. CONCLUSIONS

An attempt has been made to represent a tolerance information in feature based modelling system. The technological details (tolerance information) are classified as intra-feature and inter-feature tolerances. Intra-feature tolerances are attributed to the feature geometry. Inter-feature details establish the relationships between various form features and primitive features. Geometric dimensioning and tolerancing represent geometric constraints. The geometric dimensioning and tolerancing of a workpiece sets important constraints on the selection and sequencing of operations. Various types of tolerance constraints are identified. An effective generation of operation sequencing is proposed using petri net. Tolerance constraints are considered as input to the operation sequencing module.

REFERENCES

[1] Chang, X. F., and Kusiak, A., 1995, "Constraint-based design of parts", Computer-Aided Design, 27(5), 343-352.

[2] Clark, A. L., and South, N. E., 1987, "Feature Based Design of Mechanical parts", Proceedings of AUTOFACT, 1-69.

[3] Halevi, G., and Weill, R. D., 1995, "Principles of process planning: A logical approach", Chapman & Hall.

[4] Kulkarni, V. S., and Pande, S. S., 1996, "Representation of feature relationship tolerances in solid models", International J of Production Research, 34(7), 1975-1994.

[5] Ngoi, B. K. A., and Ong, C. T. A., 1993, "A complete tolerance charting system", International Journal of Production Research, 31(2), 453-469.

[6] Prabhu, P., Elhence, S., Wang, H. and Wysk, R., 1991, "An operation network generator for Computer Aided Process Planning", Journal of Manufacturing Systems, 9(4), 283-291.

[7] Requicha, A. A. G., 1983, "Toward a theory of Geometric tolerancing", International Journal of Robotics Research, 2(4), 45-60.

[8] Richard, J. Mayer, Umesh Hari, Chuan, J. SU., and John Yen, 1995, "Plan generation strategies for a knowledge - based automated process planning system", International Journal of Computer Integrated Manufacturing, 8(6), 399-410.

[9] Roy, U., Liu, C. R., and Woo, T. C., 1988, "Feature based representational scheme of solid modeller for providing dimensioning and tolerancing information", Robotics and Computer Integrated Manufacturing, 4(3/4), 335-345.

[10] Salomons, O. W., Houten, F. J. A. M. van and Kals, H. J. J., 1993, "Review of research in feature-based design", Journal of Manufacturing Systems, 12(2), 113-131.

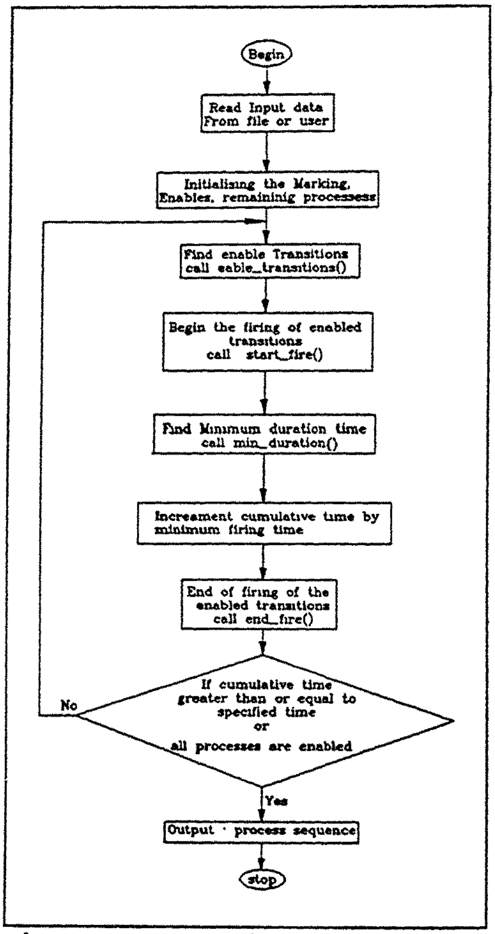


Figure 5: Working of process sequencing module

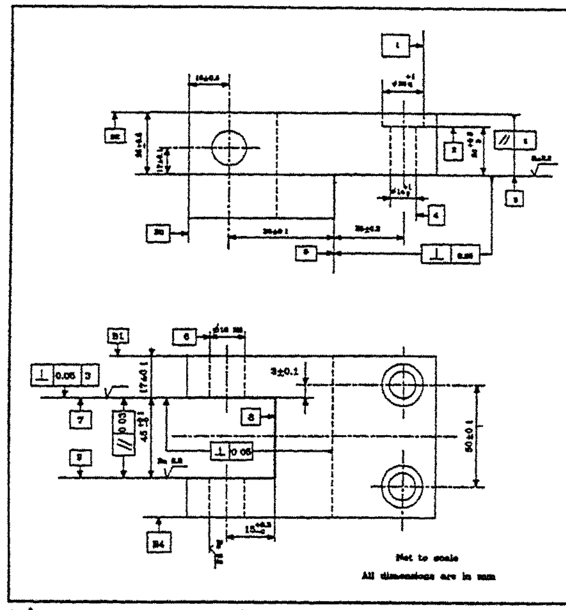


Figure 6: Example component

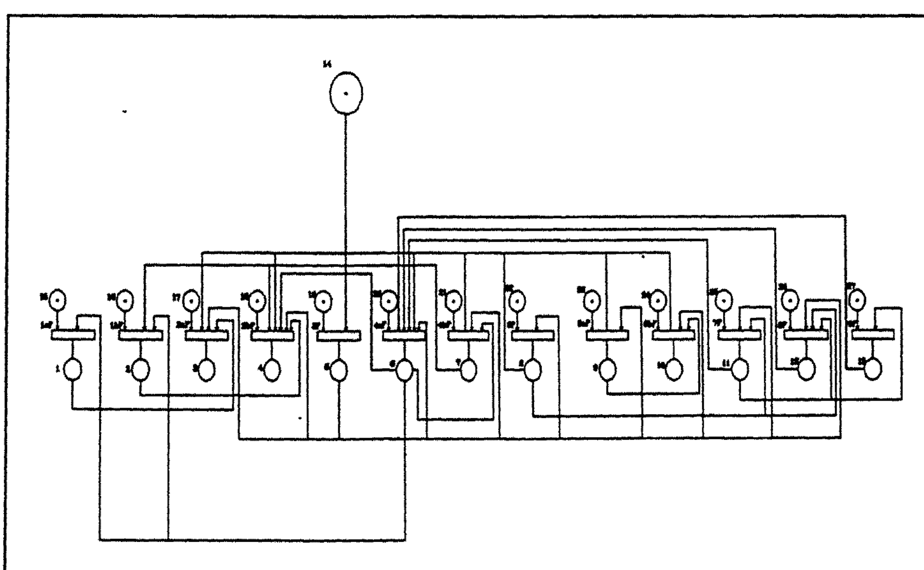


Figure 7: A typical Petri Net for the example component

AUTOMATIC MESH GENERATION FOR SIMPLIFICATION OF FINITE ELEMENT ANALYSIS

Junko Fukuda
Faculty of Engineering,
Kyushu Kyoritsu University
Kitakyushu 807, Japan

ABSTRACT

An automatic mesh generation system has been developed for easy implementation in order to prepare large amounts of input data with respect to the formation or generation of nodes of the finite elements which are needed for a finite element analysis. The mesh generation method by means of random number is capable of adapting the arbitrarily shaped and multiply connected planar domain. Then the method is possible to effectively adapt to transformation of a shape of domain through the process of the shape optimization. By using minimum necessary information for generating mesh of triangular elements in the arbitrarily shaped of a given domain, large amounts of input data of finite element analysis can be prepared automatically. This system simplifies the user's task while saving manpower in carrying out the finite element analysis.

KEYWORDS

Mesh Generation, Finite Element Analysis, Optimization

1. INTRODUCTION

A design system of optimizing contour shape to minimize stress concentration in a flat plate was previously proposed[1]. It was shown that this system was highly effective in reducing the stress concentration by adapting the optimized shape around the discontinuous part of ship structure and hatch opening. The Shape Optimization Method to minimize the maximum stress around discontinuous parts of plate without inserting a thick reinforcement plate has been presented earlier. The finite element method (FEM) using the constant strain triangles is applied in view of stability of numerical analysis for calculating the displacement and stress of the structure. The improved values of the stresses at nodes on the external boundary of the domain are needed in the process of the shape optimization, but the irregular and dispersed values of FEM results include the errors which are influenced sensitively by the mesh pattern. The accuracy of stresses at a node on the external boundary of the domain which are approximated by simple averages of stresses of the adjoined elements having the node in common, is generally known as having lower than those of internal ones. We have proposed a smoothing method furnishing more accurate stresses at a node on the boundary of the domain and applied this smoothing method to the optimization system for improving the accuracy of the stress distribution obtained by FEM[2]. Further, it is proposed the method of automatic mesh generation which is possible to effectively adapt to transformation of a shape of the domain through the process of optimization. Consideration is taken to avoid the formation of elements having acute angled corners and to control the size or distribution density of finite elements in order to avoid the loss of numerical accuracy of result of the analysis[3]. The Mesh Generation Method by means of random numbers is capable of generating mesh of triangular elements in the arbitrarily shaped and multiply connected planar domain. The design system of optimizing contour shape has applied 3 methods, (a) Mesh Generation, (b) Stress Smoothing and (c) Shape Optimization[4]. This paper presents an automatic mesh generation system which has applied the above-mentioned mesh generation method. This system has been developed for easy implementation of preparing large amounts of input data with respect to the formation or generation of nodes of the finite elements which are needed for the finite element analysis. By using minimum necessary information for adapting the arbitrary shape of a given domain, large amounts of input data of finite element analysis can be prepared automatically. Moreover, it has a function of pre-processing system using a personal computer and image scanner which produces input data for the above-mentioned automatic mesh generator by interactive mode[5]. This system also has a function of projecting and modifying the pattern of generated mesh by automatic mesh generator as required[6].

It is confirmed that this system simplifies the user's task while saving manpower in carrying out the finite element analysis.

2. CONCEPT FOR AUTOMATIC MESH GENERATION

In this system emphasis is laid on the following points with an aim to avoid the loss of numerical accuracy in the analysis, to improve practicability and so on.

1) To handle the arbitrarily shaped and multiply connected planar domains.
2) To generate mesh in the whole domain with the least possible information.
3) To control triangular element density in any part of the domain.
4) To distribute triangular elements smoothly all over the domain.
5) To economize the computing time by using the geometrical method of triangular element mesh generation, in case of simple domain.
6) To use data generated by this system directly as the data of a substructure method.
7) To modify the generated mesh pattern through interactive mode while one is checking it on CRT.
8) To produce the shaped information data of the domain by using an image scanner.

3. INPUT DATA FOR AUTOMATIC MESH GENERATOR

It is possible to specify the following input data in the process to automatically generate triangular mesh in the domain of an arbitrary form based on the system.

1) The external form of the domain is defined by connecting nodes to it with straight lines. This domain however may be multiply connected. Accordingly, necessary nodes for that must be specified.
2) The whole domain is divided into an appropriate number of sub-blocks composed of connected straight lines and each sub-block, which may be multiply connected, is specified by a set of nodes on respective boundary lines.
3) When it is necessary to generate nodes automatically by equally dividing the intervals between the nodes on the boundary line of the domain and the boundary line between sub-blocks, the number divisible into equal parts between those nodes is specified. For instance, the figures in the parenthesis of Fig.1 show the number divisible into equal parts.
4) When specification of nodes is required to be arrayed in regular sequence on the specified curve AB, the node numbers at both ends of curve the AB are specified, but for the rest of the nodes, only coordinates of nodes are specified. The node numbers on curve AB are automatically given in the process. For instance, ■ marks without number in Fig.1 show the nodes corresponding to them.
5) Certain nodes within the domain are specified when it is necessary. For instance, ※ marks in Fig.1 show the nodes corresponding to them.

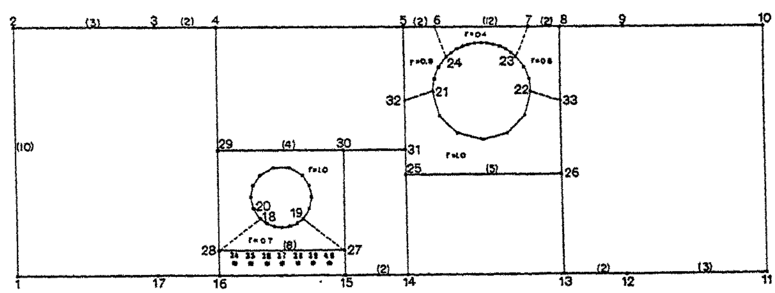


Fig.1 : Input data to generate mesh automatically in plane domain with 2 circular holes.

6) Any of the 3 methods of mesh generation, (a) Regular Mesh Generation(RMG), (b) Semi-Regular Mesh Generation(SRMG), (c) Irregular Mesh Generation(IMG), may be specified for each sub-block. Also when none of them are specified, any of them can be selected automatically according to the external form of sub-block or the state of domain, whether it is simply connected or not. For each sub-block where Irregular Mesh Generation is specified, the density factor r must be specified, if necessary, by dividing sub-block further into small domains. The sub-block with a circular hole in Fig.1 comes under them.

4. OUTLINE OF AUTOMATIC MESH GENERATOR

Interim nodes are generated, according to specified number divisible into equal parts, on every segment line, where the intervals of nodes are specified to be equally divided.. Also with such polygonal sub-block as shown Fig.2, it is assumed as quadrilateral sub-block in a broad sense by specifying appropriate 4 principal nodes denoted by mark ◉ in the figure on the boundary line of sub-block. Interim nodes are generated on the segment of a line whose partitioning number is not specified so that the average distance between them may be equal to the distance between nodes on a set of confronting sides of each quadrilateral sub-block. Then interim nodes are generated on the external boundary of the domain and the boundary lines of the sub-blocks will be determined. As the next step, nodes specified by input data and generated automatically are given new consecutive number. The next step is made to generate interior points and form triangular elements within each sub-block. The above-mentioned 3 methods are possible for that purpose. a) The method of Regular Mesh Generation is applied for the sub-block of which the form is a convex quadrilateral and the numbers of nodes on the respective confronting sides are equal. Fig.3 shows mesh patterns generated by using this method. b) The method of Semi-Regular Mesh Generation has a function of varying elements distribution smoothly in the sub-block by using difference between numbers of nodes on the confronting side of it. But the form of sub-block should be a convex quadrilateral. Fig.4 shows mesh patterns generated by this method. c) The method of Irregular mesh generation by using random number is applied for the sub-block whose form may be a multiply connected concave polygon. It has a function of controlling triangular element density in any part of the sub-block by specifying several density factors. Later this method is described in detail. Fig.5 shows mesh patterns generated by this method.

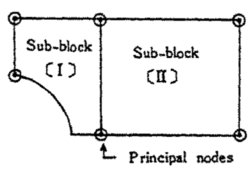


Fig.2 : Principal nodes.

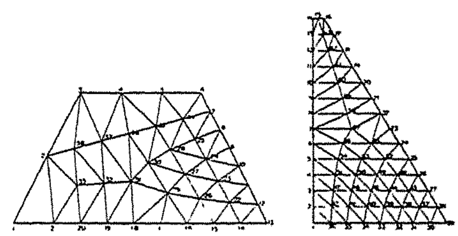


Fig.4 : Mesh patterns generated by method SRMG.

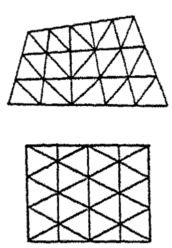


Fig.3 : Mesh patterns generated by method RMG.

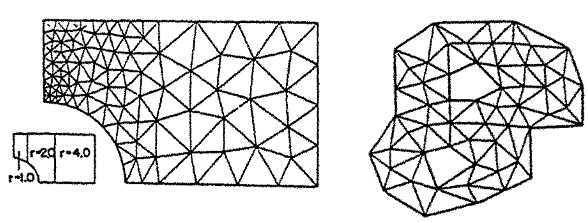


Fig.5 : Mesh patterns generated by method IMG.

5. METHOD OF MESH GENERATION BY USING RANDOM NUMBER(IMG)

This method is illustrated by generating mesh in the multiply connected domain as shown Fig. 6. The nodes on the boundary line of the domain are defined.

5.1 Generation of Interior Nodes

The following is the process showing how to generate nodes in the multiply connected concave domain.

1) In order to avoid the formation of triangular elements having acute interior angles near the boundary of the domain, a polygon is made first by drawing parallel lines inside the perimeter of the domain and at a designated interval d, as shown by dotted lines in Fig. 6(a). Nodes are not generated in the belt zones between the solid line and the dotted line. In case the value of d is not provided by input data, the average distance between the nodes on the boundary of domain will be calculated automatically and $\sqrt{3}\ /2$ times of it will be given as the value of d.

2) The lowest nodes on the boundary of the domain is assumed as P and the far left hand side as Q. A square rectilinear grid of designated gauge r is superimposed over the circumscribing rectangle from lower left to upper and from left to right row. This circumscribing rectangle is superimposed over the domain while two sides of the rectangle pass through P and Q on the boundary line of domain, as shown in Fig. 6(a).

3) An attempt is made to generate one interior node in each square by uniform random number. A generated point will be retained as an interior node if it lies in the domain, if it lies outside the belt zones mentioned above and if nodes on the boundary of domain, fixed interior nodes and previously generated nodes lie outside the hole of the radius r centered at the point. Several consecutive attempts are made to generate a point to satisfy the above-mentioned conditions in any square. If no point is successfully generated after several attempts, then no interim node is generated in that square and the next square is considered. This process is repeated until all squares have been tested.

4) In order to make uniform distribution of interior nodes by adding new points where interior node density thus generated is rough, the process is repeated by shifting the position of the circumscribing rectangle as much as r/2 toward vertical and horizontal direction.

The attempt made so far is to generate interior nodes so that interior nodal density is uniform if possible, but it is desirable to increase the nodal density in the regions close to where stress is concentrated. In such case, it is possible to designate some region in the domain and a nodal density factor r within each region, as shown Fig. 7. By repeating the process for each region, interior nodal density is controlled all over the domain. Interior nodes will be given nodal number in the order of nodes generated.

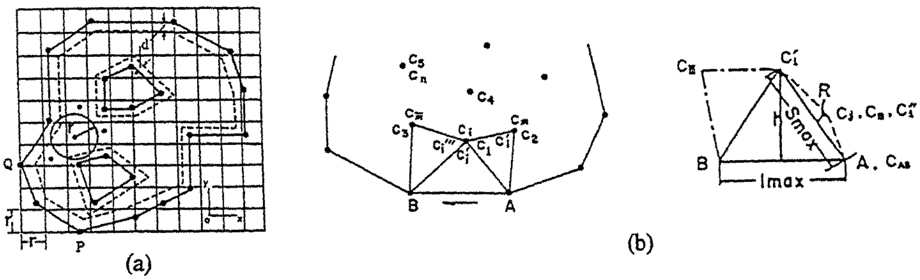


Fig.6 : Method of mesh generation using random numbers.

5.2 Generation of Mesh

The following is the process showing how the algorithm interconnects these points to form a triangle which have been generated by the method described above.

1) The boundary nodes (nodes on the boundary line of the domain) and interior nodes are assumed as C_k.
2) Any two consecutive boundary nodes A and B are chosen from boundary nodes first and the line segment AB is given the direction as shown in Fig. 6(b). The problem now is to determine a node C_k as a opposite vertex for base side AB from C_k so that the triangle $\triangle ABC_k$ is geometrically optimal. The prescribed number n ($n \leqq 5$) points from the least length $|AC_k| + |BC_k|$ are selected from the points which lie on the right hand side of $\overline{AB}$. They are assumed as $C_1, C_2, \dots C_n$.
3) It is checked to see if $\triangle ABC_k$($k = 1,2,\dots,n$) overlaps any previously generated triangle or any fixed boundary segment. If this is the case, C_k is discarded from further consideration as an opposite vertex for base AB. The points selected from C_k by this way are expressed as C_i ($i = 1,2, \dots, n'$).
4) It is checked if $\triangle ABC_i$ ($i = 1,2,\dots,n'$) contains any point C_k ($k = 1,2,\dots, i-1$). If such is the case, then C_i is also rejected from further consideration and the selected points are expressed as C_i'.
For the determination of node C_i' which should be selected to form a new triangle, consideration is taken on the following closeness factor which measures the deviation of a triangle from an optimal triangle. By handling the following process from a) to f) for each $\triangle ABC_i'$, calculate β_i' as is mentioned later.
a) Let l_{max} be the length of the longest side of $\triangle ABC_i'$ and let h be the distance between that side and the opposite vertex. The value of γ_i' is obtained from the following equation:

$$\gamma_i' = h / l_{max} \tag{1}$$

If $\triangle ABCi'$ is equilateral, it is easy to show that the value of γ_i' is $\sqrt{3} / 2$. This value is called form factor. Form factor for optimal triangle is assumed as α and it's value is given by input data. The value of $|\gamma_i' - \alpha|$ is used as a measure of deviation of that triangle from an optimal triangle. In order to estimate the consequence of selecting a new element, the next step is made for the triangle which will have AC_i' or BC_i' as base sides in a later stage of element generation.
b) Using AC_i' as a base, select from C_i ($i = 1,2, \cdots n'$) points C_j which lies on the right hand side of $\overline{AC_i'}$ and take up a point C_m which minimizes the length $| AC_j + C_i'C_j |$. The value of δ_A is calculated by the following equation.

$$\delta_A = (\overline{AC_m} + \overline{C_i'C_m}) / \overline{AC_i'} \tag{2}$$

c) Using BC_i' as a base, select from C_i ($i = 1,2, \cdots n'$) points C_j which lies on the left hand side of $\overline{BC_i'}$ and take up a point $C_{\bar{m}}$ which minimizes the length $| BC_j + C_i'C_j |$. The value δ_B is calculated by the following equation.

$$\delta_B = (\overline{BC_{\bar{m}}} + \overline{C_i'C_{\bar{m}}}) / \overline{BC_i'} \tag{3}$$

d) Let the triangle corresponding to Min(δ_A, δ_B) be $\triangle C_{AB}C_i'C_i''$. In the case of $\delta_A \leqq \delta_B$, node A corresponds to node C_{AB} and C_i'' to C_m. If $\delta_A > \delta_B$, node B to C_{AB} and $C_{\bar{m}}$ to C_i''. This will be to select the triangle $\triangle C_{AB}C_i'C_i''$ which will have the most acute interior angles.
e) Calculate the value of form factor γ_i'' for $\triangle C_{AB}C_i'C_i''$ by the following equation.

$$\gamma_i'' = R / S_{max} \tag{4}$$

Where S is the length of the longest side of $\triangle C_{AB}C_i'C_i''$ and R is the distance between that side and the opposite vertex.
f) Calculate

$$\beta_i' = (\gamma_i'' - \alpha)^2 + (\gamma_i' - \alpha)^2 \tag{5}$$

Here α will be assumed to be as $\sqrt{3} / 2$ for an equilateral triangle.

5) Let a node C_i' corresponding to Min (β_i') be C_i'''. This node C_i''' is determined as the optimal opposite vertex for side AB and $\triangle ABC_i'''$ is defined as the new triangular element.

6) By repeating this process on the assumption that one side of the newly generated triangle is to correspond to the above-mentioned side AB, the correspondence between all the interior nodes and all triangular elements is determined.

6.EXAMPLE OF APPLICATION

Fig.7(b) shows mesh pattern which are generated by specifying several density factors in a domain as shown in Fig.7(a) using IMG method. Fig.8(b) shows mesh pattern which is generated in the discontinuous shaped plate by input data as shown in Fig.8(a) while the mesh generator selects automatically any of the 3 method of mesh generation for each sub-block.

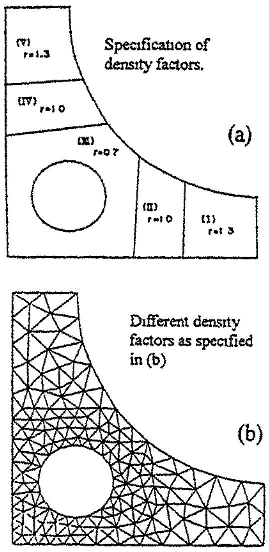


Fig.7 : Example of partitioning and controlling density factors and mesh pattern.

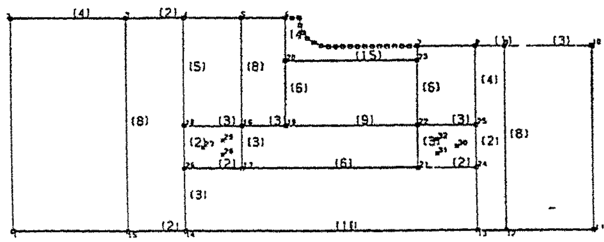


Fig.8(a) : Input data required to generate mesh for the discontinuous shaped plate.

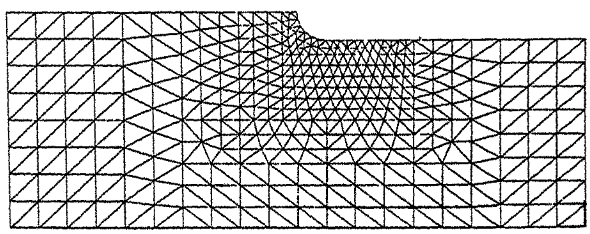


Fig.8(b) : Mesh pattern generated by input data as shown in Fig.8(a).

7.REFERENCES

1. Fukuda, J., 'Design System of Optimizing Contour Shape to Minimize Stress Concentration in Flat Plate', Computer Applications in Production and Engineering, Sun, Q., Tang, Z. and Zhang, Y., pp.102-110,Chapman & Hall, London, 1995.
2. Fukuda, J., 'Accuracy of Stress Smoothing Methods in Finite Element Analysis', The West-Japan Society of Naval Architects, No.87, March, pp239-262, 1994.
3. Suhara, J. and Fukuda, J., 'Automatic Mesh Generation for Finite Element Analysis', Advances in Computational Methods in Structural Mechanics and Design, Oden, J.T., Clough, R.W. and Yamamoto, Y., pp.607-624, The university of Alabama Press, Alabama, 1972.
4. Fukuda, J., 'A Method for Optimizing Contour Shape to Minimize Stress Concentration in Flat Plates', Information Processing Society of Japan, vol.36, No.8, pp.1760-1777, 1995.
5. Fukuda, J., Katsuki, H., Kaneko, K. and Hirano, Y., 'Pre-Processing System for Automatic Mesh Generator Using a Personal Computer and an Image Scanner', The West-Japan Society of Naval Architects, No.88, August, pp.195-211, 1994.
6. Fukuda, J., 'Automatic Finite Element Mesh Generation and Modification by means of Interactive Graphics', Information Processing Society of Japan, vol.20, No.6, pp.494-500, 1979.

DESIGNING AT A DISTANCE

A.J. Medland
School of Mechanical Engineering
University of Bath
Bath
United Kingdom BA1 7AY
telephone: 01225 826158
e-mail: a.j.medland@bath.ac.uk

ABSTRACT

With the move towards global working has come the incentive to undertake remote and distributed design. Whilst the problems of manufacture can be controlled and organised by forming the activities into processes or assemblies with well defined interfaces into the product plan or configuration, such decomposition cannot be readily achieved in the area of conceptual design. Here the activities of concept generation and evaluation are highly interrelated, with a high degree of interdependence on all downstream processes.

The issues of design working are further complicated by the personal aspects of skill, culture and experience. It is therefore very difficult to compare methods of design working and how they would be influenced by distributed working.

Work to date has indicated that an holistic view needs to be taken of the design right from the stage of generating the product specification In this all stake holders need to be represented. Such a process should then lead to a series of testable propositions that can be used to bound and interface individually undertaken designing activities. Through such a process the problems of conflict can be both identified and resolved.

The current research has led to the proposal of a conceptual system running within a constraint resolution environment. Here the specification is reformed into constraint rules which can be addressed and resolved by selected design variables. These variables in their turn are owned by individual designers, whilst the constraint rules apply globally. The truth of individual rules, sub-problems (comprised of clusters of rules) through to the complete design can be checked. This approach relies upon allowing remote activities to proceed independently as long as all rules remain true. Only when conflicts occur are the necessary designers brought together.

Such an approach is currently under construction and will be evaluated on the design of mechanical problems to address problems in the machinery industry.

KEYWORDS

Design Theory, Constraint Resolution, Remote Working.

1. BACKGROUND

The increase in global enterprises has brought with it the gradual dispersion of a company's various business activities between a number of different countries. These are chosen on a wide variety of bases which include the regional markets to be penetrated, the availability of labour (both skilled and appropriate), financial incentives, and economic barriers. Some of these depend on cultural suitability as well as on the traditional regions in which particular skills are available (through education and the existence of similar industrial enterprises).

This process, as seen in a number of world-based enterprises, has resulted in highly dispersed production system, both at a component and complete product level. In the aircraft industry components and sub-assemblies right through to wings, tails, and body sections are often built in different countries and transported to the parent factory for assembly. Conversely many automobile manufacturers supply similar components (sourced both locally and from abroad) to various assembly plants which are tailored to produce vehicles for a local market.

All such globally distributed activities concentrate on decomposing and manipulating discrete sub-operations or component assemblies, within an overall management process. The success of this approach is dependent upon the degree to which these sub-operations and assemblies can be undertaken independently. Although all assemblies must be related to others in order to provide the desired function and form, much of the necessary interaction can be reduced by the careful ordering of events and by strictly pre-defining the interface relationships between mating parts and spaces. Thus the various sub-activities can proceed with the minimum of interaction with the other remote processes.

The problems of such manufacturing organisations are thus seen to be addressed through the creation of an efficient management and control system. Each remotely related element must therefore be well defined, not only in terms of its physical interfaces into the product, but also in terms of how the creation of this element interfaces with the management plans. The risks of some kind of failure, from not delivering on time through to the not meeting and holding specification, need to be considered, and contingency plans prepared if such risks are high.

Although contingency planning can take place in manufacturing, and, possibly, in the later stages of detail designing, it is unclear at present how such events can be catered for during the activities of conceptual design.

2. REMOTE DESIGN WORKING

Companies have already embarked upon programmes of remote (or distributed) design working. Information on how such programmes are working is difficult to come by because much of it is confidential, both in terms of the practices and of the product information being conveyed. Even that information which is available is difficult to assess, as there are no agreed methods for evaluating or comparing design methods. Design methods, with and without remote working, cannot be easily compared, on account of the differences observed between each individual designing case [1]. The approach taken on each occasion is dependent upon the design objectives, current experience and knowledge, available technology, skills, management, personalities, and cultures.

Whilst it is not possible to compare directly methods of design working, it is possible to predict that remote design working will become a necessary way of working and will, in its turn, create a whole new set of problems. Remote design working has to come into practice, not only for the reason previously given for global working, but also because certain aspects of design are highly specialised and cannot be maintained by many businesses on a continuous basis. Instead the need is often bought in, on demand, and may therefore be located anywhere in the world. This will bring with it not only the problems of communications, security, and confidentiality, but also the social problems of interaction, commitment to conflicting goals (from those of individuals through to those of the business and the product), managing diverse skills remotely (through "non-eye-to-eye" contact) over different social structures and time zones.

3. CONCEPTUAL DESIGN

The author is currently involved in a number of studies in design and remote working. These range from the supervision of student project work at a foreign university, to collaborative research into the exchange of design information and manufacturing rules with other universities. This is leading to an understanding of the needs and form of communications required in the early stages of design.

Many years of experience as a practising designer and observations of others at work have shown that although the early stages may appear to an observer to be chaotic, with the designer leaping from one idea to another, they are a vital ingredient in the ultimate success of a project. It is here that the ambiguities of the specification are resolved and boundaries set on the work to be done and changes allowed. It is the "free-roaming" activity that defines the design space for the future work.

This free roaming process is often incorporated within formal activities of brain-storming, lateral thinking, Synetics etc. These not only create the boundaries and regions of possible solution (which are obviously closely coupled), but also produce commitment within the team to a product goal. This is not easy to reproduce in a remote working situation, and an alternative structure, based on constraint resolution processes, is therefore being considered.

4. A CONSTRAINT APPROACH TO DESIGN

The underlying structure of design can be seen to be that of decision making. Although the process of design is often represented as a sequence of interrelated processes it is the decision structure which determines what processes occur and in what order[2].

The decisions arise from the need to meet the overall objectives of the design. A network of decisions therefore leads back from the end goal through intermediate decisions to an array of known states (figure 1). For each decision a necessary action or process can be chosen to allow the decision to be validated (figure 2)..

In attempting to formalise this process, and without stifling creativity, a constraint resolution approach has been proposed. Here all requirements and decisions are transformed into a series of constraint rules that meet or support the original specification. Each constraint rule is formed as a proposition whose truth can be tested. Singly or in clusters the truth of an aspect of the design can be tested, or the truth of the total design can be established by evaluating all constraint rules[3] .

The decision making structure within such a constraint resolution process can then be seen as a total truth maintenance system. Should any decision be untrue, the end goal can never successfully be achieved. Within such a system a true solution can then be sought by suitable changes to the design variables. This can be performed manually or undertaken automatically within a computer environment by direct search techniques.

As constraint rules can be formed to allow the testing and resolution of both equality and inequality relationships, the truth of geometric, mathematical and set rules can be handled. The approach can thus cope with mixtures of geometric, analysis, management and manufacturing decisions, within the same structure.

The process can be further refined by allowing those who are responsible for a chosen aspect to retain 'ownership' (and thereby control) of the relevant design parameters. variables. Only the owner can then freely manipulate or change such design variables[4]. However as the final design must be true, from every point of view, the constraint rules are considered to be globally applied and accessible by everybody.

5. PROPOSED STRUCTURE

The proposed design structure imposes an additional level of resolution within the design process. Here the details of the specification are first transformed into a series of testable propositions (or constraint rules) that can be used to verify (or otherwise) the appropriateness of an subsequent design activity (figure 3). The inclusion of new or the modification of existing rules (to meet changes in the specification) can take place throughout the designing activity with the consequences, to the current stage of design being observed.

The creation of the testable propositions needs to be performed with all parties concerned and a collective agreement reached. If any requirement in the specification is left either unagreed or untestable then agreement can not be assured and a conflict may arise. Once the basic constraint rules have been formed then if sub-systems can be defined together with their agreed interface rules.

When the basic and interface rules exist the sub-activities can be essentially pursued independently with individual designers applying their own set of 'owned' design variables. The individual designing activities can proceed independently as long as none of the rules fail. An analysis activity may for example show that properties of strength will not be met, whilst the creation of geometric features could show that parts do not assemble.

In all cases of violation of rules the designer is required to take some corrective action. This may be simply to change the sub-design until the violation is removed and an appropriate design is established. The designer can contact the 'owner' of an offending part and negotiate for appropriate changes to be made (even in both parts) which will result in a true state being re-established for both. Should the conflict exist back with the testable form of the specification, then negotiations can be entered into to change, re-interpret or remove the offending aspect of the specification.

Whilst some aspects of change can be freely entered into (such as locally owned activities) others will be restricted dependent on their importance to the overall design and to the level of commitment entered into by other aspects of the design. Whilst all changes should be considered, those that result in major effects on function and market or require major design and manufacturing changes to be considered, should be only embarked upon after due consideration.

6. DEMONSTRATION OF APPROACHES

Two different approaches have been investigated. The first divides the problem up into functional cells (or sub-problems) whilst the second assigns ownership to variables and thereby provides a means for dynamic problem subdivision.

In both approaches the specification is re-cast into a series of testable propositions before further design activities are undertaken. Such actions are carried out as a continuous problem refinement activity throughout the product design.

6.1 Functional Cells

Within the functional cells approach the design is broken down into sub-problems in which the interrelationships can be readily understood. Within the design of a tracking and handling mechanism, for use on a production line, the design must be able to simultaneously satisfy the requirements of matching the specified line conditions, undertake the assembly operation, enter and leave the restricted region of the operating zone. These individual constraint groups can be separately defined, modelled and then resolved (figure 4).

The interdependency of these cells however, requires that they must all be concurrently true if an acceptable solution is to be found. They are thus formed into a network (as shown in Figure 4) and a procedure evoked to 'fire' each cell as sufficient information becomes available for its resolution. Once fired one cell will then trigger another. Such firing continues until no more cells are left to be resolved or an untrue state appears in a resolved cell. The system then backtracks allowing the 'free variables' of the previously resolved cells to be changed in order to seek an alternative solution that can make all cells true. Thus in the production line assembly problem, the line parameters, entry and exit points as well as the assembly operation itselfcan all be manipulated in the search for an acceptable state.

Within a remote designing environment it is envisaged that each cell should be 'owned' and handled by a group based within a single location. When conflict occurs in the network, each group is then in a position to discuss, modify and reconfigure their cell in the light of the problem identified within the network. Discussion thus occurs firstly at a high level between groups to identify how the conflict arose between cells and then at an interior cell level in an attempt to find a new cell solution that will allow all cells to be made true.

6.2 Variable Ownership

Within the second approach investigated, the creation of testable propositions proceeds as before. These are considered to be applied globally to allow the problem to be resolved. The design variables are however considered to be owned by individual members of the team, which are assigned as they arise. Such an assignment takes place within a hierarchical 'management' structure, with the ownership and responsibility being passed down the structure [4]. Thus in conditions of conflict the system can back track up the hierarchy to find an owner at which the conflict can be addressed and resolved.

The ownership of design variables allows the system to identify, for each designer, the variables that can be used in the solution of the sub-task. If for instance, the designer is directed to work on a specified component, then the design variables necessary for such a task are assigned. The system can then identify all of the rules within the testable proposition list that will be influenced by a change in these variables. These are then collectively assigned to the design task and have to be resolved if the final design is to be acceptable.

The designer is then free to work on a solution to the problem, in isolation from all other members as long as other variables are not being reassigned that make the selected rules untrue. The designer is thus presented with fixed values (taken at the start of the session) for the other design variables. Should values, influencing the current design, start to change, then this occurrence will be signalled to the designer. Remote communications can then be opened between the two designers and the effect of these changes identified and resolved.

Should an individual designer be left with an untrue design, even after direct communications between those involved, the 'violation' is automatically reported up the hierarchical management structure, to that point at which all the conflicting variables are owned by a single person. These can then be solved by accepting one design and forcing the other to comply, with a modified set of rules or both designs can be brought together and all the related rules solved simultaneously.

In such an approach designers can be assigned to address differing problems that are not simply sub-problem or component related. Reassignment of variables can allow modifications to be investigated to allow, say, the dynamic characteristics of a complete mechanism to be improved without allowing the form of the assembly being changed.

7 CONCLUSIONS

The process of remote design working severely changes the form of the interrelationships between members of the design team. No longer is there an automatic and informal means for understanding individual responsibilities and what is being undertaken by others.

The proposed approach as investigated in this paper sets out to replace this informal structure with a procedure which whilst providing an awareness and control, does not over-constrain the designers in their way of working. The method requires that firstly all aspects of the outline specification, by agreement, are turned into testable propositions. This occurs at and early stage of the design and can be refined as the design proceeds.

Sub-groups of these rules need to be resolved for different aspects of the design. In order that the final design is acceptable, the a situation must be found in which all sub-groups (or elements of the design) are concurrently true.

Two approaches to problem decomposition are proposed and presented. Firstly the decomposition to the problem into functional cells and the other into dynamic cell through the ownership of design variables. Whilst the second method does provide greater flexibility (and can be formed to represent the structure of the first) it presents the design team with a very much more complex environment within which to work. The benefits of flexibility may then be outweighed by the confusion that could arise due to the dynamic reassignment of variables.

Work into this approach for remote working is still at an early stage with a suitable environment still under development. As this become available the benefits and problems of each approach should become clearer.

8. ACKNOWLEDGEMENTS

The work described in this paper was carried out under project supported by the LINK Programme in the Design of High Speed Machinery. The financial support and the help of all those involved is gratefully acknowledged.

9, REFERENCES

1. L. Blessing, A. Chakrabarti & K Wallace, A Design Research Methodology'. ICED 95 Praha 1995, pp 50-55.
2. A. J. Medland 'The Computer Based Design Process', 2nd edition. Chapman & Hall 1992
3. A.J.Medland, G. Mullineux, A.H.Rentor & B.I. Twyman 'A Software Environment for the Design of Manufacturing Machinery'. Proc. 31st Int MATADOR conf., Manchester 1995 pp 449-454.
4. A.J. Medland, 'The Use of Constraint Techniques in the Resolution of Design Conflict, Control and Management in a Mechanism Designing Environment' Journal of Engineering Design. Vol. 6, No. 3, 1995 pp231-238.

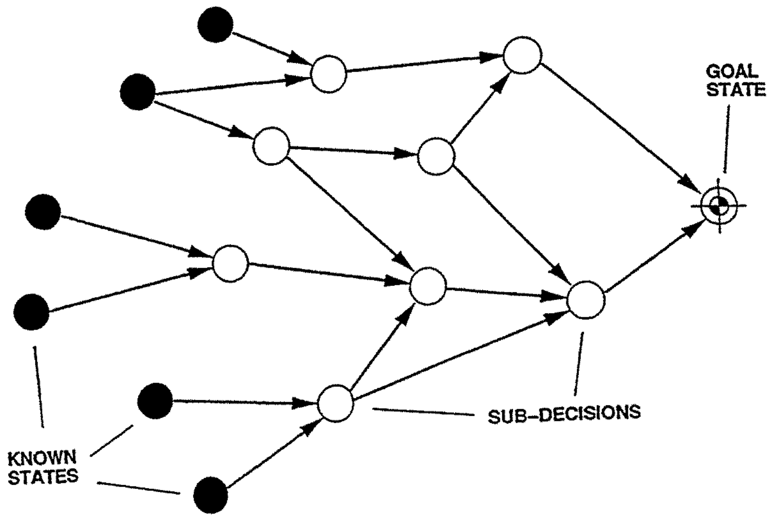


figure 1 Network of decision

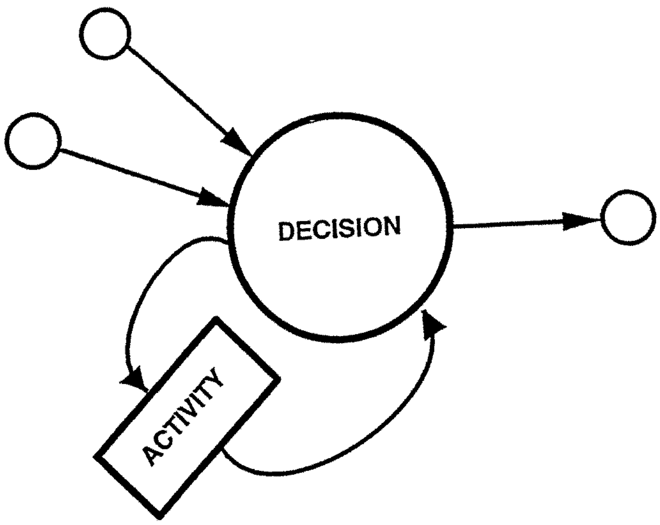


figure 2. Decision method activities

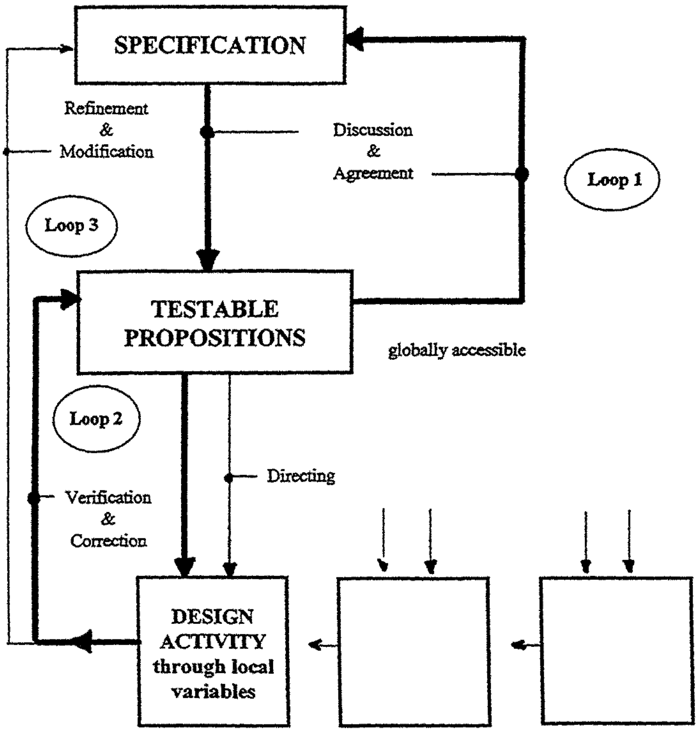


figure 3. Proposed design structure

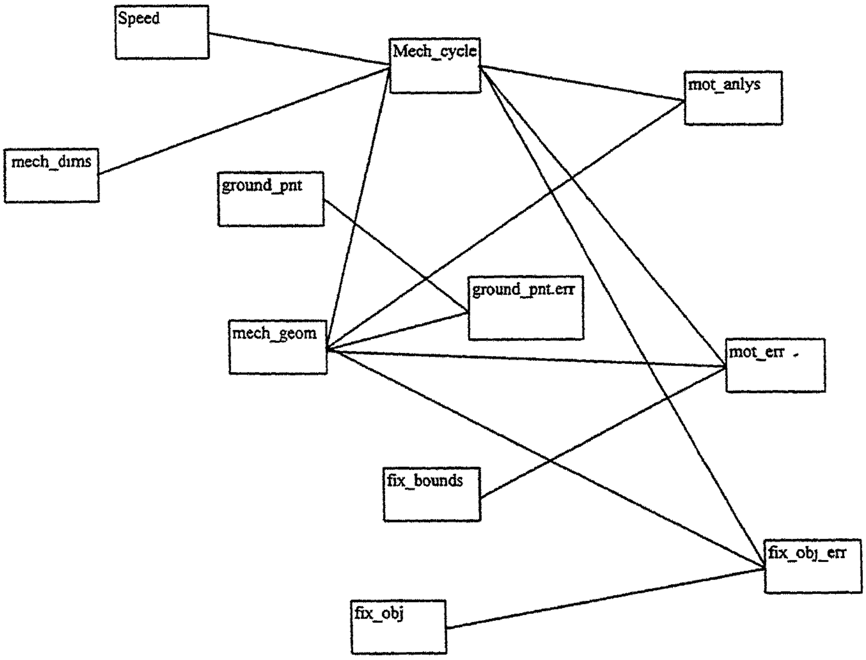


figure 4. Cellular decomposition of a Tracking and Handling problem

A TOLERANCE MODEL WITH SURFACE IRREGULARITY CONSIDERATION

C.K. Lui, W. S. Sze and S.T. Tan
Department of Mechanical Engineering
The University of Hong Kong
Hong Kong

ABSTRACT

This paper presents a tolerance model which can handle the effect of surface irregularity in tolerance specifications. The model is variational based in which a particular model variable is assigned to the extent of surface irregularity and the rest of the model variables are the measures of different kinds of nominal variations. Tolerance specifications are represented by a collection of constraints which enclosed a region in the space spanned by the model variables. The applicability of the model in tolerance representation is illustrated by examples representing tolerance specifications of size, orientation and form in the model space. Theory related to the model variable of surface irregularity and application of the model in tolerance justification are also discussed.

KEYWORDS

Tolerance Model, Surface Irregularity, Tolerance Justification

1. INTRODUCTION

The specification of tolerances is important in both the manufacturing cost and the functionality of an engineering product. Tight tolerances can maintain functionality but incur an excessive manufacturing cost and may be unattainable by the available manufacturing process. Loose tolerances give rise to lower manufacturing costs but may affect functionality.

In order to obtain a cost effective tolerance specification with the variation of functionality kept in an acceptable range, tolerance analysis and tolerance synthesis should be performed. Tolerance analysis studies the effect of a tolerance specification on the functionality of the product. Tolerance synthesis distributes tolerance values to a tolerance specification by optimising the manufacturing cost. Both these functions can only be carried out if there is a mathematical foundation for tolerance representation.

In most current tolerance researches, tolerances are usually modelled by the elements of nominal geometry variations, i.e. only size and orientation errors of the face features are included in the tolerance models. As the effect of surface irregularity is ignored, the formulation of tolerancing problems based on these models may not fully reflect the nature of the geometric deviation of a manufactured part. On the other hand, in tolerance standards such as ASME Y14.5 [1], size tolerance and orientation tolerance of a feature also control the form of the feature if form tolerance is not specified. If the effect of surface irregularity is ignored, the representation of these tolerances may be incompatible with the standards.

This paper proposes a tolerance model which supports the solving of tolerancing problems in a more realistic standpoint by taking the effect of surface irregularity into account. As the development of the tolerance model is based on the concept of variational modelling with one of the model variables assigned to the extent of surface irregularity, a brief review of the related materials is given in Section 2. The details of the tolerance model and the associated examples are respectively presented in Section 3 and Section 4. The application of the model in tolerance justification is discussed in Section 5. Further work of applying the model to more complicated tolerancing problems is discussed in Section 6.

2. RELATED WORK

2.1 Theories of Tolerance Representation

Various tolerance representation theories have been proposed by different researchers. Requicha and Chan [2] starting from the viewpoint of geometric modelling formulated an offset boundary approach for tolerance representation in which tolerance zones are constructed by subtracting the expanded and the shrunk versions of the nominal feature. In this theory, mathematical rules are used for constructing different types of tolerances. The values and the corresponding datums of different types of tolerances are viewed as attributes of the features.

Jayaraman and Srinivasan [3, 4] developed the virtual boundary requirement (VBR) approach for tolerance representation. In this theory, functional requirements such as the ability to be assemblied and material bulk are captured in a specific form of tolerances designated as VBRs. These VBRs can be converted into other form of tolerances designated as conditional tolerances (CTs). Different types of tolerance can then be derived in the form of CTs.

Variational modelling is the methodology for the flexible design of geometry [5]. Geometric model of an object can be defined and altered by solving the set of simultaneous constraint equations of model variables. Turner [6, 7] applied this concept in tolerance representation by modelling the nominal deviations of an object with a variational model. In his theory, model variables are associated to the surfaces of an object and each model variable is a measure of the degree of variation of some geometric property. The geometric variations should be carefully selected in such a way that the basis vectors of the model variables form a vector space. Tolerances are represented as functions of the model variables. Tolerance limits indirectly impose an in-tolerance region in the vector space.

2.2 Theory Related to Surface Irregularity

In a recent paper [8], the authors have proposed a model based method to derive a tolerance measure of surface irregularity from the digitised profile data of some machined surfaces. In the method, the surface height, z, which is measured from the mean line of the profile, is considered as a random process with respect to the traverse length, x, such that the height of a point along the traverse direction can be written as a random function $z(x)$. The structure function of the profile data is fitted to an empirical model by means of nonlinear regression to obtain the model parameters. The tolerance measure of surface irregularity, T_I, which can be determined from the model parameters, is defined as the statistical extreme value of the positive maxima of the random function $z(x)$. If $z(x)$ is assumed to be a Gaussian process, its statistical extreme value of the negative minima will be equal to that of the positive maxima and the extent of surface irregularity will be equal to $2T_I$. Based on the knowledge of T_I, the worst case value and the statistical characteristics of the extent of surface irregularity can be evaluated

3. THE TOLERANCE MODEL

In computer aided design (CAD) systems, objects are represented by their nominal geometry in a solid model. These objects are the error-free versions of the real manufactured instances. As each manufactured object will deviate from its nominal geometry, the basic concept of tolerance representation by variational modelling is to identify the possible variations introduced by the manufacturing process and to provide a mathematical description of the varied object by introducing independent model variables. Although the objects represented by the variational model are still an idealisation of the manufactured counterparts, the model can be used as a platform for the evaluation of tolerancing problems if it can closely approximate the real situation.

Depending on the type of geometric elements selected to undergo variations, there are various ways to construct a variational model for tolerance representation. According to [7] and [9], applying variations to the dimension parameters of constructive solid geometry-based (CSG) primitives wil restrict the possible variation modes of the object. In feature-based systems, the variational coverag of a model is dependent on the choice of features used and the sequence of feature formin operations. If variations are introduced to the vertex coordinates of an object, problems will arise i polyhedral models with faces containing more than three vertices since three vertices alone can full determine a planar surface. To allow the surfaces of an object to vary is a favourable strategy for th construction of a variational model because: 1) part functions are more closely related to the surface of an object rather than the vertices, 2) the model can represent the tolerance standards mor naturally, and 3) the model can be enhanced to include curved surfaces.

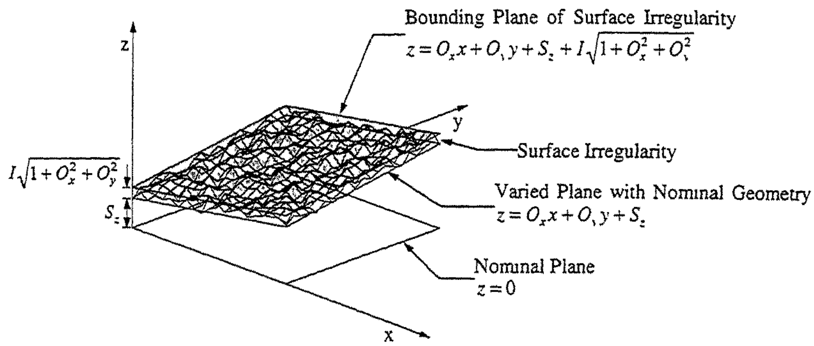


Figure 1: Associating model variables with a planar surface

As an object can always be treated as a collection of intersecting surfaces that form the bounding faces, a variational model can be automatically constructed from the information inside a CAD system. In our proposed method, variations are introduced to each surface by associating a set of independent model variables . In the set, one of the model variables is assigned to measure the extent of surface irregularity while the others are assigned to control the nominal geometric variations. As illustrated in figure 1, model variables are associated with a planar surface which is defined in its local coordinate system with one of its vertices positioned at the origin and its surface normal parallel to the z-axis. The equation of the nominal planar surface can then be written as:

$$z=0 \tag{1}$$

Orientation variations and position variation are introduced to the surface by associating the model variables of nominal variations O_x, O_y and S_z with the surface. In the variational model, the equation of the planar surface subjected to the nominal variations can be written as:

$$z=O_x x+O_y y+S_z \tag{2}$$

In the above equation, O_x and O_y are the model variables which control the orientation variations of the planar surface about the y-axis and the x-axis respectively. The model variable S_z controls the position variation of the planar surface in the direction of the outward normal. The effect of surface irregularity is introduced to the variational model of the planar surface by associating a surface irregularity model variable, I. The variable I controls the extent of surface irregularity in the direction of the outward normal of the nominally varied surface. With the variable I, an offsetted surface can be constructed from the nominally varied surface to represent the bounding surface of surface irregularity. The equation of the bounding surface of surface irregularity in the variational model can be written as:

$$z = O_x x + O_y y + S_z + I\sqrt{1+O_x^2+O_y^2} \quad (3)$$

When a variational model is defined by associating model variables with all the surfaces in the nominal model, the set of model variables forms a Cartesian space with its dimension equal to the number of model variables in the set. Each point in the space corresponds to an instance of the varied object. Measures of tolerance are defined as functions (tolerance functions) of the model variables. Tolerance limits are imposed on the tolerance functions such that each tolerance specification defines a group of constraint inequalities which enclose a region in the model space. A varied object instance is said to be in-tolerance if the point corresponding to the instance is either on the boundary or inside the region, i.e. the values of the model variables of the instance satisfy all the tolerance constraints.

4. ILLUSTRATIVE EXAMPLES

In this subsection, examples are used to illustrate the application of the tolerance model with surface irregularity consideration for the representation of some tolerance specifications. The examples are simplified (only three model variables are used) so as to provide easy visualisation of the in-tolerance regions in a three dimensional space.

In the following examples, a rectangular block is used as the target object which is subjected to tolerance specifications. For illustration purpose, we assume variations only occur in one of the block faces and the other faces are all perfect. In order to limit the number of model variables to three, variations of the variable face are further assumed to be two dimensional. The target rectangular block is shown in figure 2. In the figure, the rectangular block is shown in such a view that the width of the block is parallel to the view vector. L and H are respectively the dimensions designating the length and height of the block. The upper face of the block is subjected to two dimensional variations, i.e. the orientation variation about the axis parallel to the length of the block is assumed to be zero. As all the other faces are assumed to be perfect, the dimension L will remain constant while the dimension H will vary.

Figure 3 shows the variational model of the rectangular block. In the figure, M_1 is the model variable controlling the position variation of the upper face and M_2 is the model variable controlling the orientation variation. M_3 is the model variable approximating the true extent of surface irregularity, I. The approximation is valid since $M_3 = I\sqrt{1+(M_2/L)^2} \approx I$ if $L >> M_2$.

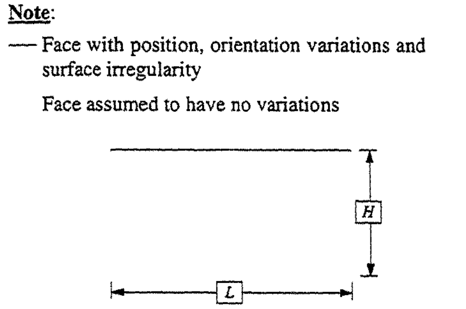


Figure 2: The two dimenional view of the rectangular block used in the examples

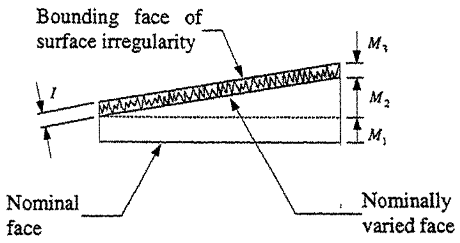


Figure 3: Variational model of the varied planar face

4.1 Size Tolerance

Size tolerance (or limits of size) is the tolerance specification which applies solely to individual features of size [1]. It prescribe the extent within which variation of geometric form, as well as size, are allowed. A size feature is either a cylinder or spherical surface, or a set of two opposite elements or opposite parallel surfaces, associated with a size dimension. On two parallel planar surfaces, size limits control flatness and straightness of the elements of both surfaces and also the parallelism relationship between these surfaces [10].

Figure 4: A size tolerance

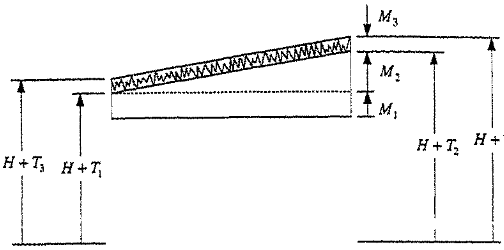


Figure 5: Tolerance functions for size tolerance representation

As shown in figure 4, a size tolerance of $\pm T_S$ is used to control the variation of the height of the target rectangular block. In order to represent the size tolerance in terms of the model variables, eight tolerance variables are introduced. In figure 5, T_1 and T_2 measure the variations of the top left and the top right vertices of the nominally varied face relative to the bottom face. T_3 and T_4 measure the variations of the top left and the top right vertices of the bounding face of surface irregularity relative to the bottom face. As there is no datum specification in size tolerance, measurements can be taken from either the top or the bottom face. The other four tolerances variables, T_1', T_2', T_3' and T_4' are theoretically necessary to measure the variations of the bottom left and the bottom right vertices relative to the normally varied face and the bounding face of surface irregularity respectively. However, these can be neglected based on the following argument:

$$T_1' = \frac{(H+T_1)}{\sqrt{1+(M_2/L)^2}} - H \approx T_1 \text{ if } L >> M_2. \text{ And similarly, } T_2' \approx T_2,\ T_3' \approx T_3 \text{ and } T_4' \approx T_4.$$

The four tolerance variables T_1, T_2, T_3 and T_4 can be written in terms of the model variables as shown below:

$$\begin{aligned} T_1 &= M_1 \\ T_2 &= M_1 + M_2 \\ T_3 &= M_1 + M_3 \\ T_4 &= M_1 + M_2 + M_3 \end{aligned} \qquad (4)$$

As these tolerance functions should always be within the tolerance limits and the model variable M_3 is an approximation to the extent of surface irregularity I which should always be greater than zero, the following constraints can be established:

$$\begin{aligned} &-T_S \le T_1 \le T_S, \quad -T_S \le T_2 \le T_S, \\ &-T_S \le T_3 \le T_S, \quad -T_S \le T_4 \le T_S, \\ &0 \le M_3 \end{aligned} \qquad (5)$$

The set of constraints can be reduced to the following form:

$$-T_S \le T_1, \quad T_3 \le T_S,$$
$$-T_S \le T_2, \quad T_4 \le T_S,$$
$$0 \le M_3 \tag{6}$$

These constraints enclose a region in the space spanned by the model variables as shown in figure 6, which is the in-tolerance region of the size tolerance specification.

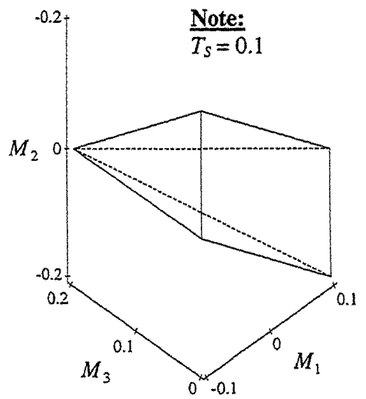


Figure 6: In-tolerance region of a size tolerance

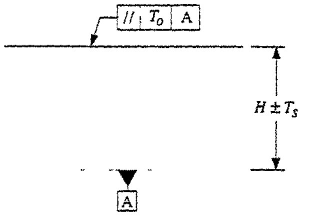


Figure 7: Specification with an additional parallelism tolerance

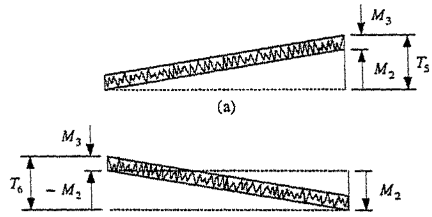


Figure 8: Tolerance functions for parallelism tolerance representation

4.2 Parallelism Tolerance

Parallelism tolerance is the tolerance specification used to control the parallelism of one feature relative to another. As a member of the orientation tolerance family, the parallelism tolerance of a feature should be specified relative to one or more datum features. When a parallelism tolerance of a feature surface is designated to be related to a datum plane, the tolerance zone is defined by two parallel planes parallel to the datum, within which the surface of the designated feature must lie. On plane surfaces, parallelism tolerance control flatness if a flatness tolerance is not specified [1].

Although the size tolerance already constrains the orientation variation of the top face relative to the bottom face of the target block in a $2T_S$ region (figure 6), a parallelism tolerance of size T_O can be used to further control the parallelism of the top face relative to the bottom face as shown in figure 7. As parallelism tolerance does not control the relative distance between the designated feature and the datum feature, the representation of this tolerance should be independent of the model variable M_1. To represent a parallelism tolerance, two new tolerance functions are introduced. As shown in figure 8(a), the tolerance function T_5 measures the difference in vertical positions of the top right vertex of the bounding face of surface irregularity and the top left vertex of the nominally varied face. On the other hand, in figure 8(b), the tolerance function T_6 measures the difference in vertical positions of the top left vertex of the bounding face of surface irregularity and the top right vertex of the nominally varied face. Therefore, T_5 and T_6 can be expressed in terms of the model variables as shown below:

$$T_5 = M_2 + M_3$$
$$T_6 = -M_2 + M_3 \qquad (7)$$

As the magnitude of these tolerance functions should be less than T_O, the constraints imposed by the additional parallelism tolerance are as follow:

$$T_5 \le T_O, \quad T_6 \le T_O, \qquad (8)$$

These additional constraints tighten the tolerance specification of the target block and the in-tolerance region of the tolerance specification is shown in figure 9.

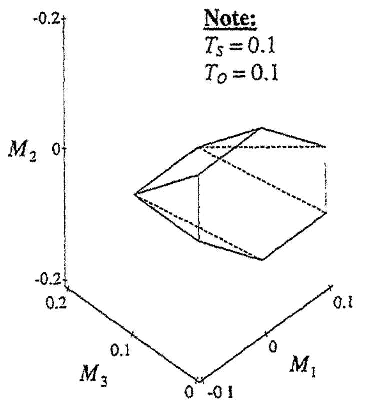


Figure 9: In-tolerance region of a size and a parallelism tolerance

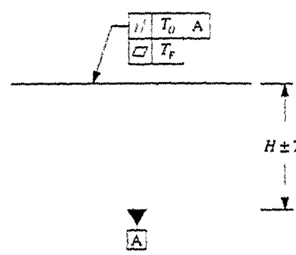


Figure 10: Specification with an additional flatness tolerance

4.3 Flatness Tolerance

Flatness tolerance is a tolerance specification used to control the form of a flat surface. Similar to other form tolerances, flatness tolerance controls individual feature and does not relate to datums. When a flatness tolerance is specified on a surface feature, the tolerance zone is defined by two parallel planes within which the surface must lie [1].

In figure 10, the form variation of the upper face of the target block is controlled by applying a flatness tolerance of size T_F. A new tolerance function T_7 is introduced to provide a measure for the distance between two parallel planes within which the upper surface of the block must lie. That is:

$$T_7 = M_3 \qquad (9)$$

And the flatness tolerance imposes a tolerance constraint to the tolerance specification which is shown as follows:

$$T_7 \le T_F \qquad (10)$$

The in-tolerance region of the tolerance specification of the target block with an added flatness tolerance is shown in figure 11.

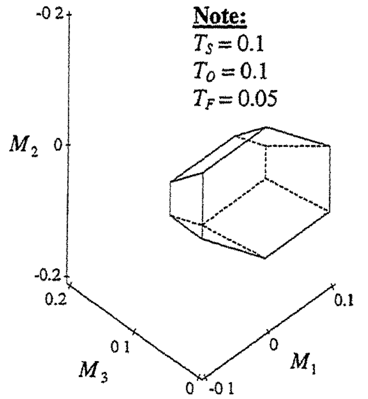


Figure 11: In-tolerance region of a size, a parallelism and a flatness tolerance

5. TOLERANCE JUSTIFICATION OF AN INDIVDUAL COMPONENT

A properly assigned tolerance specification should always maintain the variation of component functionality in a desirable range. However, this specification may not be a cost effective specification which is attainable by the available manufacturing processes. To justify a tolerance specification of an individual component in terms of cost effectiveness and attainability, we can compare the yields with the optimum yields of the component corresponding to the available manufacturing processes. (The optimum yield of a manufacturing process can be determined from the information of unit component cost, unit tolerance cost and the number of component to be manufactured). For a cost effective tolerance specification, there should be at least one available manufacturing process with its yield approximately equal to the optimum yield. On the other hand, if all the simulated yields of the available manufacturing processes are much smaller than the optimum yields, the tolerance specification can be justified to be unattainable.

The yield of a component corresponding to a manufacturing process is approximated by means of statistical simulations. In the simulations, the model variables of nominal deviations (M_1 and M_2) are assumed to be random variables which follow the normal distributions with means equal to zero and standard deviations equal to the worst case values of the model variables in the size tolerance specification divided by a constant. i.e. $\langle M_1 \rangle = \langle M_2 \rangle = 0$, $\sqrt{\langle M_1^2 \rangle} = T_S / k$ and $\sqrt{\langle M_2^2 \rangle} = 2T_S / k$, where $\langle \; \rangle$ is the mean function and k is a precision constant which is inversely proportional to the setup precision and the measurement errors of a manufacturing process. The model variable of surface irregularity (I) is assumed to be a random variable which can be characterised by means of the method outline in Subsection 2.2 such that the distribution of the random variable can be obtained from the model parameters of a manufacturing process [8]. With these assumptions, varied component instances of a manufacturing process can be simulated by sampling the model variables from their distributions. The model representation of tolerance specifications (such as equations (6), (8) and (10)) are used as the basis to determine whether a varied component instance is within the specification. The yield of a component corresponding to a manufacturing process can be determined by simulating a large number of varied component instances and counting the number of the in-tolerance component instances.

6. FURTHER WORK

The root of tolerancing problems is to cost effectively allocate attainable tolerances with the proper types and sizes to all the components in an assembly such that the variations of functionality of all the components are constrained within desirable ranges. In this paper, we have presented a model for tolerance representation in which surface irregularity is included as a model element. This model on one hand can provide the representation of tolerances in a more realistic standpoint; on the other hand, it can correlate tolerances to manufacturing processes. The further works of our research is to extend the tolerance model for solving more complicated tolerancing problems by: 1) developing an assembly model which can represent the interrelationships among various components in an assembly, 2) identifying and establishing the relationship between component functionality and model variables in the tolerance model, 3) developing a cost model of tolerances, and 4) developing the methodology for the non-linear optimisation of manufacturing cost subjected to the constraints of tolerances and functionality requirement.

7. CONCLUSION

In this paper, a tolerance model with surface irregularity consideration is proposed. The model is variational based with one of the model variables assigned to the extent of surface irregularity. The applicability of the model to tolerance representations is demonstrated with examples of size, orientation and form tolerances. The statistical characteristic of the model variable of surface irregularity can be correlated to the manufacturing processes by means of surface irregularity characterisation. Application of the model in tolerance justification of an individual component is discussed. The model can be further extended to solve more complicated tolerancing problems which will be the future issues in our research.

8. ACKNOWLEGMENTS

The authors would like to thank the Department of Mechanical Engineering of the University of Hong Kong for the support of the project.

9. REFERENCES

1. ASME National Standard ASME Y14.5M-1994, Dimensioning and Tolerancing, The American Society of Mechanical Engineers, New York, 1995.
2. Requicha, A.A.G., and Chan, S.C., 'Representation of Geometric Features, Tolerances, and Attributes in Solid Modelers Based on Constructive Geometry', IEEE Journal of Robotics and Automation, vol.RA-2, pp.156-166, 1986.
3. Jayaraman, R., and Srinivasan, V., 'Geometric Tolerancing: I. Virtual Boundary Requirements', IBM Journal of Research and Development, vol.33, pp.90-104, 1989.
4. Srinivasan, V., and Jayaraman, R., 'Geometric Tolerancing: II. Conditional Tolerances', IBM Journal of Research and Development, vol.33, pp.105-125, 1989.
5. Light, R., and Gossard D., 'Modification of Geometric Models through Variational Geometry', Computer-Aided Design, vol.14, pp.209-214, 1982.
6. Turner, J.U., 'A Feasibility Space Approach for Automated Tolerancing', Journal of Engineering for Industry, vol.115, pp.341-346, 1993.
7. Turner, J.U., 'Tolerances in Computer-aided Geometric Design', Ph.D. Thesis, Rensselaer Polytechnic Institute, 1987.
8. Lui, C.K., Tan, S.T. and Sze, W.S., 'Tolerance Modeling of Engineering of Engineering Surface: A 2D Approach by means of Fractal Superposition', Proceedings of the 1996 ASME Design Engineering Technical Conferences and Computers in Engineering Conference, ASME, 1996.
9. Gupta, S. and Turner, J.U., 'Variational Solid Modeling for Tolerance Analysis', IEEE Computer Graphics and Applications, pp.64-74, May, 1993.
10. Meadows, J.D., Geometric Dimensioning and Tolerancing: Applications and Techniques for Use in Design, Manufacturing, and Inspection, Marcel Dekker Inc., New York, 1995.

ON THE ASSEMBLABILITY OF A PRODUCT

Hung-Yao Hsu and Grier C. I. Lin
Centre for Advanced Manufacturing Research
University of South Australia
The Levels, Pooraka,
South Australia, SA5095
Australia

ABSTRACT

In this study the assemblability of a product is investigated on the basis of the accessibility of each component. The accessibility of each component is defined as the probability of having a collision-free path to approach its destination. The motivation for doing this research is due to the understanding that the number of components and their relative arrangement within a product have a significant influence on the assembly operations of the product. Consequently, if there were a quantitative measurement of this influence, the optimal product configuration (ie the arrangement of components within the product) could be determined during the conceptual design stage and the best method of joining two mating components could be decided accordingly. As a result of this study, traditional Design for Assembly (DFA) methods can be reinforced by the provision of the product assemblability.

KEYWORDS

Assemblability, Accessibility, Product Configuration, Design For Assembly Method

1. INTRODUCTION

Due to the pressures created by a highly competitive market, many efforts have been explored by researchers and industrial companies to shorten the time-to-market of products with competitive price, satisfactory quality and reasonable profit. Concurrent engineering (CE) concept is one of the promising approaches to addressing this issue. The essential idea of CE is to implement all the relevant tasks simultaneously (or nearly simultaneously) by multi-disciplinary team-work. CE is a strategic concept, leading to a systematic approach of the integration of design, production, and related processes dealing with all aspects such as maintenance and disposal at the end of the product life. According to the concept of CE there are two levels of concurrence which need to be considered. The first is the organisation concurrence that ensures all the relevant tasks are conducted simultaneously towards meeting the objectives of a project. The second level is the individual concurrence that ensures individuals conduct their duty by taking all the other aspects of the project into consideration.

Design for Assembly (DFA) has been a valuable individual concurrence tool for a designer to foresee potential assembly problems of a product and resolve them during the early design stage. In traditional DFA, there is no quantitative measurement of the influence of the arrangement of components on the assembly operations (ie the assemblability of a product). Therefore the designers must rely on their experience to determine the product configuration (ie the arrangement of components within the product). Consequently, if there were a quantitative measurement of this influence, an optimal arrangement could be chosen and the best method of joining two mating components could be decided accordingly during the conceptual design stage, and a better rearrangement of components can be achieved during the redesign stage.

In this study the assemblability of a product is investigated on the basis of the accessibility of each component. Firstly, a generalised accessibility for a component can be derived on the basis of the accessibility proposed by the authors[1]. An algorithm for calculating the assemblability of a product is then developed. The whole idea of this new development is illustrated by an example. This paper is concluded with a summary about the contribution of this study and the future research directions.

2. GENERALISED ACCESSIBILITY FOR A COMPONENT

In this section, a review about the previous study[1] will be given followed by the development of a generalised accessibility for a component.

2.1 Two Fundamental Assembly Modes and the Associated Accessibility

Stack-up assembly and distributive assembly are two fundamental assembly modes in general assembly operations. They and their associated accessibility (in bits) are described as follows. In the following discussion, the shape of each component is described by its bounding box, which is the smallest rectangular prism that can contain the part completely.

The ***stack-up assembly*** mode means all the components are piled up in such a way that later one is attached above to the previous one. In other words, the assembly direction coincides with the components' allocation direction (as illustrated in Figure 1).

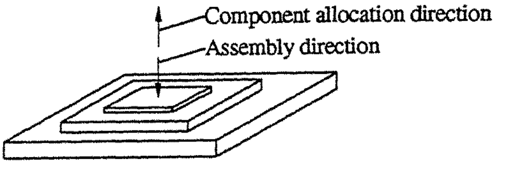


Figure 1 The stack-up assembly

The accessibility for the attaching component in each pair of mating components is defined by the following formula:

$$I = \log_2(\text{base area of the attaching component} \; / \; \text{common area between two components}) \tag{1}$$

where the numerator is the base area of its bounding box; and the denominator means the area enclosed by the mating features.

In the ***distributive assembly*** mode the assembly direction is perpendicular to the component allocation directions and is illustrated in Figure 2. All the components are assembled from top and distributed in both X and Y axial directions.

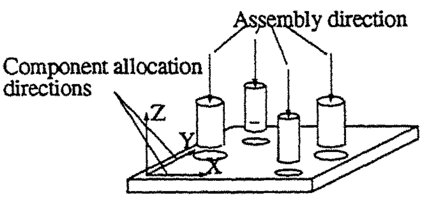


Figure 2 The distributive assembly

In order to define the accessibility, a better understanding about the configuration space is essential. The following information about configuration space is cited from Lozano-Pérez[2]:

The *configuration* of a polyhedron is a set of independent parameters that characterise the position of every point in the object and is defined relative to an initial configuration. In this initial configuration, by convention, a fixed vertex of the polyhedron coincides with the origin of the global coordinate frame. For a polyhedron A, this vertex is called the *reference vertex of* A, or rv_A. The number of parameters required to specify the configuration of a k-dimensional polyhedron, A, relative to its initial configuration, is d, where $d = k + k! \,/\, [2!\,(k-2)!]$[3]. Thus, the configuration of A can be regarded as a point $x \in R^d$; this d-dimensional space of configurations of A is denoted $Cspace_A$. A in configuration x is $(A)_x$; A in its initial configuration is $(A)_0$. Not all possible configurations in $Cspace_A$ represent legal configurations of A; in particular, configurations of A where $A \cap B_j \neq \emptyset$ (B_j is any obstruction) are illegal because they would cause collisions (ie not safe). These illegal configurations are the result of a mapping of the B_j into $Cspace_A$. This mapping exploits two fundamental properties of objects: 1) their *rigidity*, which allows their configurations to be characterised by a few parameters and 2) their *solidity*, which requires that a point not be inside more than one object.

The $Cspace_A$ obstacle due to B is defined as $CO_A(B) \equiv \{x \in Cspace_A \mid (A)_x \cap B_j \neq \emptyset\}$, denoted $CO_A(B)$. Thus, if $x \in CO_A(B)$ then $(A)_x$ intersects B, therefore x is not safe. Conversely, any configuration $x \notin CO_A(B_j)$ (for all objects B_j) is safe. If A is a convex polygon with fixed orientation, the presence of another convex polygon B constrains the configuration of A, in this case simply the position of rv_A, to be outside of $CO_A(B)$, a large convex polygon, shown as the region in Figure 3 (c).

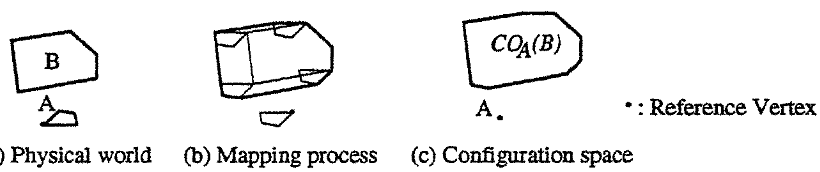


Figure 3 The $Cspace_A$ obstacle due to B, for fixed orientation of A

The configuration space (*Cspace*) obstacle is used to define the accessibility of a component with respect to (wrt) other components. The underlying idea can be illustrated by the example in Figure 3 (a) where the locations of both A and B are their respective destinations in the assembly. When examining the influence of object B on the approach of object A, the situation can be referred to the $Cspace_A$ obstacle due to B as shown in Figure 3 (c). Two lines can be drawn as shown in Figure 4 to define the *forbidden angle* where object A will collide with object B. The *accessible region* is the region other than the area enclosed by the forbidden angle. Here it is assumed that the object approaches its destination along a straight line path without changing its orientation. Consequently, the accessibility for object A to approach its destination under the presence of object B is defined as follows:

$$I = \log_2(360° / (360° - \text{forbidden angle})) \qquad (2)$$

If there are other objects present, the forbidden angles between object A and other components are determined respectively. All the forbidden angles are summed up to obtain the ***net forbidden angle*** of object A. The accessibility is determined by using the net forbidden angle in Equation (2). These two fundamental assembly modes consider only the idealised situation. Further extension is required for the calculation of the assemblability of a product. This will be discussed in the next section.

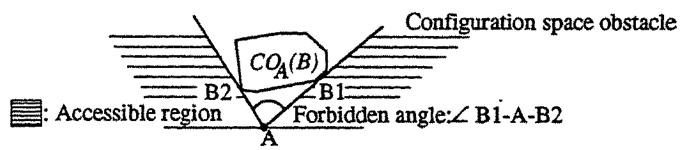


Figure 4 The forbidden angle due to $CO_A(B)$

2.2 Generalised Assemblability for a Component

The assemblability of a product should take at least the following three issues into consideration: (a) component shape, (b) component allocation and (c) assembly plane. The shape of a component means its geometric shape. The allocation is the position and orientation of a component. The assembly plane is the plane where components are attached to the base part.

The shape of a component will affect its own accessibility when the mating features can not be located easily by the assembly agent. This is the major concern in the stack-up assembly mode and is characterised by the Equation (1). Besides, the shape and orientation of one component in conjunction with the shape and orientation of other components also affect the shape of its *Cspace* obstacles and thus its accessibility. The forbidden angle of each component depends on the shape and the relative position of his *Cspace* obstacles. Therefore the accessibility of one component is a function of the shape, position and orientation of all the component within a product. However the Equation (2) does not consider the influence of the height of components. The profile of the assembly plane that will affect the approaching direction of a component to its destination is not considered in the Equation (2) either. In order to resolve these limitations, a general form of the Equation (2) is needed.

Before discussing the general form, the influence of height should be considered. Referring to Figure 2, the assembly plane is XY plane (hereafter principal plane) while YZ plane and ZX plane are planes perpendicular to the assembly plane (hereafter secondary planes). In the secondary plane the same mapping procedure as in *Cspace* can be used to decide the influence of the height. The theoretical accessible direction in the secondary plane is 180 degrees because the component is assumed to be assembled to the base part externally As shown in Figure 5, the forbidden angle due to the presence of component B in the *Cspace* can be decided by the same ideas of Figure 4. It is a function of the shortest distance between the reference vertex of component A and component B (refer to Figure 5), and the height of the obstacle

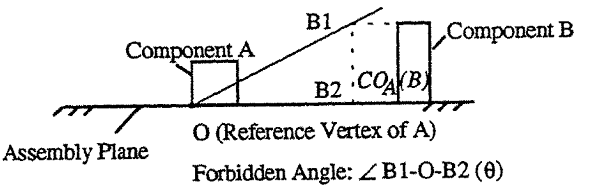


Figure 5 The forbidden angle in the secondary plane

For a component to access its destination along straight line path with fixed orientation, the ideal accessible volume (IAV) can be represented as a unit hemisphere (with the reference vertex as its origin) in *Cspace*, as shown in Figure 6(a). Every point on the surface is the projection of an accessible direction. When there is another component around, a forbidden volume can be defined as shown in Figure 6(b) where ϕ and θ are forbidden angles on the principal plane and the secondary plane respectively. The forbidden volume (FV) is equal to $(\Delta\phi\sin\theta) / 3$. The net accessible volume (NAV) is defined as the difference between IAV and FV, denoted by IAV \ FV and is shown in Figure 6(c). If there are many obstacles, the NAV can be decided by the difference between IAV and the FV induced by each obstacle. The general form of Equation (2) for a component can be defined as follows:

$$I = \log_2(2\pi / (2\pi - \sum_{j=1}^{n} \Delta\phi_j \sin(\theta_j))), \quad (3)$$

where j is the number of the obstacles;

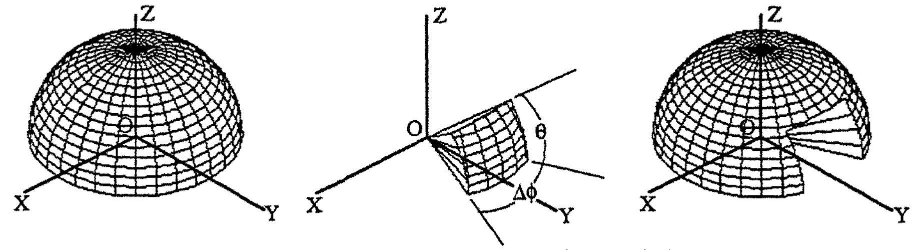
$\Delta\phi_j$ is the forbidden angle induced by the component j in the principal plane and

θj is the forbidden angle induced by the component j in the secondary plane.

After deriving Equation (3), the generalised accessibility for a component is defined as combination of Equations (1) and (3), and has the form as follows.

$$I = \log_2((\frac{\text{base area of the attaching component}}{\text{common area between two components}})(2\pi / (2\pi - \sum_{j=1}^{n} \Delta\phi_j \sin(\theta_j)))) \quad (4)$$

This generalised form can be used to determine the accessibility of each component within a product. In the next section how to decide the assemblability of a product on the basis of the accessibility of components will be discussed.



(a) Ideal accessible volume (IAV) (b) Forbidden volume (FV) (c) Net accessible volume (NAV)

Figure 6 The net accessible volume of a component

3. AN ALGORITHM FOR DECIDING THE ASSEMBLABILITY OF A PRODUCT

The assemblability of a product is defined as the sum of the accessibility of its components in each assembly plane. A product usually has many assembly planes on its base part to which other components are attached. Therefore an algorithm is required to allow the designer to decide the assemblability of a product effectively. The algorithm can be described as follows:

a) identify the assembly directions and the assembly planes associated with the base part;
b) identify the assembly modes associated with each assembly plane;
c) calculate the accessibility of each component in an assembly plane by Equation (4);
d) repeat the c) step until all the assembly planes have been gone through;
e) sum up the accessibility of each component to have the assemblability of a product.

4. A WORKING EXAMPLE

A printed circuit board (PCB) has two components of different size as shown in Figure 7. The length, width and height dimensions of component A (CompA) are 225, 135 and 50 (in mm) respectively. The length, width and height dimensions of component B (CompB) are 120, 105 and 70 (in mm) respectively. The assemblability of the PCB wrt each position of comp B is calculated according to the algorithm of Section 3, and some findings of interest are highlighted below.

The assembly mode is the distributive mode. There is only one assembly plane with one assembly direction. In the principal plane, the forbidden angles ($\Delta\phi$'s) for both CompA and CompB are 57 deg. wrt position 1 of CompB and 71 deg. wrt position 2 of CompB, as shown in Figure 7. The forbidden angles (θ's) for CompA in the secondary assembly plane are 22.26 deg. wrt positions 1 of CompB and 18.43 deg. wrt positions 2 of CompB. The accessibility of CompA wrt position 1 of CompB is 0.063 bits. The accessibility of CompA wrt position 2 of CompB is 0.065 bits. The forbidden angles (θ's) for CompB in the secondary plane are 16.30 deg. and 13.39 deg. at positions 1 and 2 respectively. The accessibility of CompB at position 1 is 0.046 bits. The accessibility of CompB at position 2 is 0.047 bits. The assemblability of this PCB wrt position 1 of CompB is 0.109 bits and it is 0.112 wrt position 2 of CompB. A smaller assemblability implies that the associated product configuration is easier to be accessed than the one with larger assemblability.

Some interesting findings are summarised as follows (refer to Figure 7):

a) The influence of one component on the other component is mutually consistent in the principal plane; therefore, angle $\angle$ B1-Oa-Ob1 is equal to angle $\angle$ A1-Ob1-Oa.
b) The shape of the *Cspace* obstacle of CompA is independent of the position of CompB, ie the shape of $CO_A(B_1)$ is the same as the shape of $CO_A(B_2)$.

c) For any two components, the forbidden angle of one component can be constructed by drawing two parallel lines from its reference vertex wrt the sides of the existing forbidden angle of the other component in the principal plane. As shown in Figure 7, the forbidden angle ∠ B2-Oa-B3 of CompA can be constructed by drawing two lines Oa-B2 and Oa-B3 from Oa. These two lines are parallel to the edges Ob2-A1 and Ob2-A2 of the forbidden angle ∠ A1-Ob2-A2 of CompB at position 2 respectively.

d) The assemblability of a product provides the designer with a quantitative measurement of the influence of component arrangement on assembly operations.

e) The position, shape and orientation of components affect the assemblability of a product.

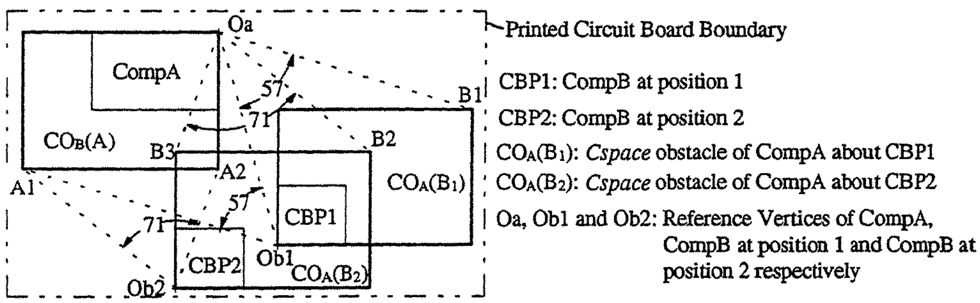


Figure 7 A deliberate printed circuit board

5. CONCLUSION

This paper explores the development of the assemblability of a product on the basis of the accessibility of components. The underlying ideas are based on the previous study[1], information content of axiom design theory[4], DFA guidelines and configuration space obstacle.

The achievements of this study are summarised as follows:

a) A generalised accessibility for a component is defined.
b) An algorithm for determining the assemblability of a product is also proposed.
c) Assemblability of a product gives designers a better understanding of the influence of the arrangement of components on assembly operations. Hence the allocation of components within a product can be planned in such a way that it will facilitate assembly operations and also fulfil the functional requirements during the design stage.
d) This study will also benefit the product redesign by comparing the assemblability of all alternatives.
e) Traditional DFA evaluation methods can be reinforced by the provision of assemblability.
f) Research in this area can be further developed on the basis of this study.

Potential applications of the assemblability such as design for disassembly, assembly planning, optimisation of the product configuration and analytical DFA evaluation method will be explored in conjunction with further developments in the nearest future.

6. REFERENCES

1. Hsu, H.Y., Lin, G.C.I. and Kao, Y.C., "On the Allocation of Components within a Product", to be appear in the Proceedings of the 1st International Conference on Engineering Design and Automation, Bangkok, 1997.
2. LOZANO-PÉREZ, T., "Spatial Planning: A Configuration Space Approach", IEEE Transaction on Computers, C-32, 2, pp.108-120, 1983.
3. BOTTEMA, O. and ROTH, B., Theoretical Kinematics, pp.10, North-Holland, Amsterdam, 1979.
4. N. P. SUH, The Principles of Design, (Oxford University, New York, 1990), Chap. 5.

SOFTWARE PROTOTYPING AND OBJECT-ORIENTED SPECIFICATION AS A METHODOLOGY FOR CUSTOMER-ORIENTED DEVELOPMENT OF SOFTWARE SYSTEMS - AN EXAMPLE FROM THE AUTOMOTIVE INDUSTRY

K. Mertins, R. Albrecht, S. Beck, O. Bahns, B. La Pierre
Division Systems Planning
Fraunhofer Institute for Production Systems and Design Technology, IPK Berlin
Berlin, Germany

ABSTRACT

The development of complex software systems for the manufacturing industry has recently become under the strong influence of cost and time savings while equally having to produce high quality, consistent and convincing customer solutions. More so, the key word „reusability"- whether it refers to products obtained at the various phases of the software development process or actual software components - is often associated with solutions to reduce these restrictive factors and preserve investments in information technology.

During a project with the automotive industry, IPK Berlin specified a new production management system for an Asian car manufacturer using an object-oriented specification method and extended it with the parallel development of a system prototype, creating the foundation for a combinative methodology, the „Prototyping and Object-Oriented Specification Method (PTOOS)".

KEYWORDS

Automotive Industry, Production Control, Prototyping, Specification, Object-oriented Software Development, PTOOS

1 INTRODUCTION

An Asian car manufacturer needed a software system which would control and manage the various phases of the automobile production process. The initial task of the software system was to plan, by means of complex algorithms, the specific sequence in which customer ordered cars were to be produced. This included a wide range of technical, managerial and physical parameters and criteria, which needed to be met to reach an efficient production plan. Secondly, it was demanded that the software should monitor the production process in order to react to unforeseeable situations, such as machinery breakdowns, part shortages, etc. It was therefore targeted for the system to be able to dynamically re-sequence the current order (change the order of the cars already being produced within certain limits), as to correct these unwanted situations. The main technological process stages within the problem domain of the project are shown in Figure 1.

During the first discussions with the customer, it became obvious that the complexity of specifying (and implementing) such a software system would be very high. It was soon recognized, that a close customer-developer relationship would need to evolve in order to complete such a software package successfully. The decision was made to use an advanced, object-oriented specification method to cover the problem domain thoroughly. It would allow for an evolutionary approach to system development, i.e. use the results from previous phases and integrate them into the succeeding phase. This led to the decision, in addition to this approach, to use a *Prototyping* strategy to software development, allowing for even closer customer-orientation and a better understanding of the problem at hand for both the developer and customer teams. The prototype of the future control system could illustrate the relevant functionality in sufficient detail during the various phases of development and also serve as a basis for system testing, evaluation and experimentation and for training. As the project proceeds, the evolutionary approach also supported by this strategy would ensure that the prototype could evolve to a software solution „ready for coding".

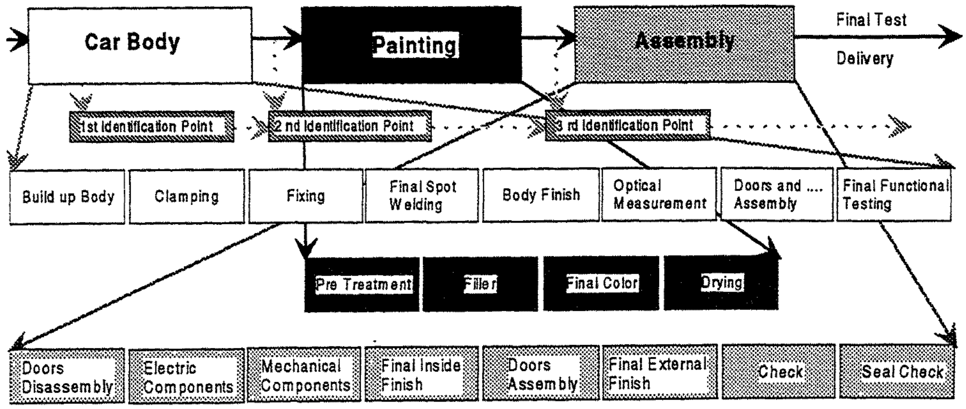


Figure 1: Car builders areas to control

2 THE PROTOTYPING AND OBJECT-ORIENTED SPECIFICATION METHOD (PTOOS)

In this section, the method used during the specification and prototyping phases of the production management project will be introduced.

2.1 Object-oriented Specification (OOS)

In the object-oriented specification method, classical specification steps are abstracted to object-like classifications. These abstractions of the problem domain are textual descriptions of the occurring objects, attributes, services, functions and modules. Appropriate graphical representations of the user interface are designed using CASE tools or, preferably, so-called *integrated development environments (IDEs).*

An OBJECT can be an order for a product, an operation, a tool, a storage, etc. Each of these OBJECTs are described by ATTRIBUTES, such as „identification number", „status" or „customer number". A SERVICE description provides the processing within an OBJECT, using data of the OBJECT or requesting data from other OBJECTs. The SERVICE is a textual description of logic processes modifying the ATTRIBUTES of the OBJECT. The FUNCTIONS transform the SERVICES into implementation-oriented descriptions of algorithms. Mostly one SERVICE consists of several FUNCTIONS.

The interface „window masks" (or FORMS) are usually designed within a programming language's *integrated development environment* (IDE) or a CASE tool. The definitions of all the above named components are specified in the *specification glossaries,* namely the *object glossary, attribute glossary, service glossary, function glossary, module glossary* and the *form glossary* (see Figure 2).

The GUI component of the object-oriented specification method, the textual descriptions of objects, functions and services as well as the IDE approach allows for the construction of a specification-based prototype *(specification prototype).*

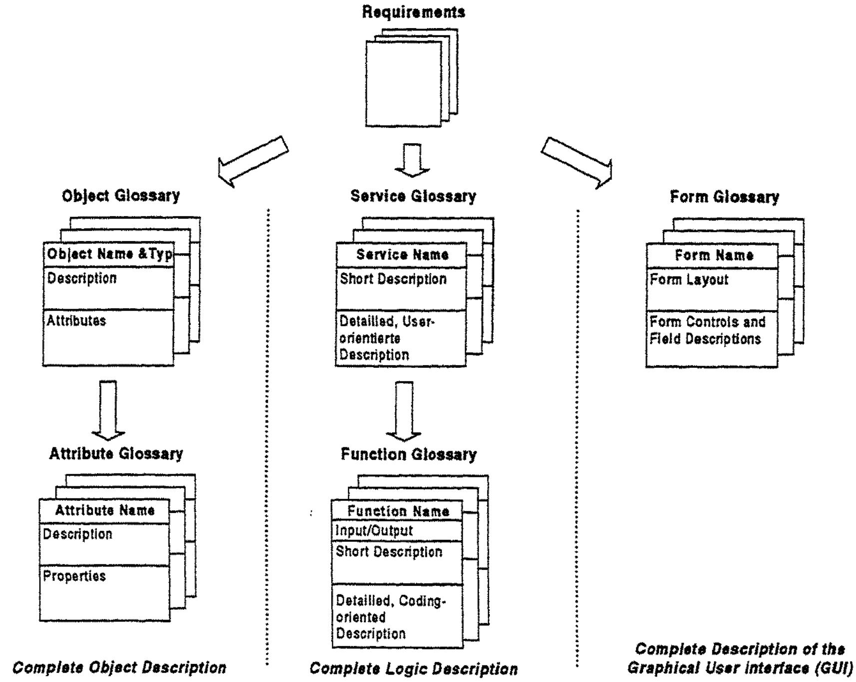


Figure 2: Components of the OOS method

2.2 Prototyping

Software prototyping as an approach to the evolutionary system development of core functionality has the following advantages:

- working versions of the future software system are produced early in the development process, visualizing functionality and potentials to its users
- continuous and intensive confrontation with a real, although restrictive software system, gives the user a clearer understanding of what he wants, which inevitably leads to a more comprehensive and precise specification
- within a short period of time, different versions of a system design can be evaluated by means of modal prototypes to aid decision-making
- it provides a communication basis for discussions among all groups involved in the development process, especially between developers and users
- a creative and customer-oriented cooperation between developers and users is encouraged.

Returning our attention to the results obtained at the specification level described above, the window masks from the form glossary, essentially representing the objects defined by the specification, are „hooked up" as defined by the services. The forms expose functions by means of controls (buttons, property sheets, etc.) on these masks. These controls do not actually perform any functionality, they merely load and display the textual description from the specification documents (glossaries). These masks are then grouped into the specification-defined modules, which describe a certain scenario within the specified software system. This process involves iterative steps, so that the prototype grows as the specification is completed.

The product of the specification and prototyping phase is a running version of the future software system boasting a broad variety of features, but little actual functionality, called a horizontal prototype. A vertical component may be implemented for specific scenarios, depending on the customers requests.

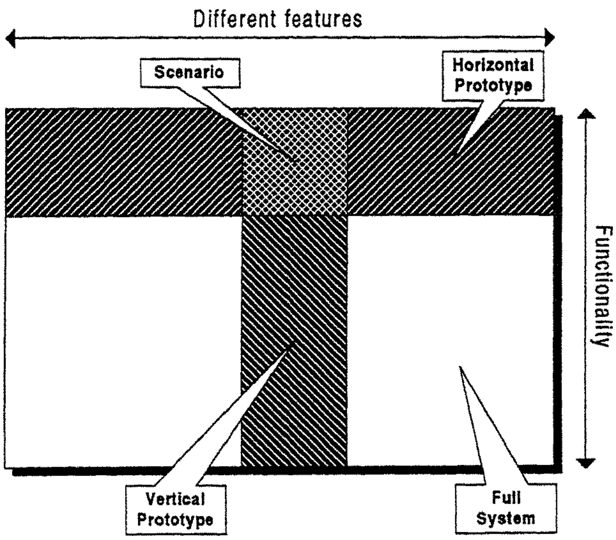


Figure 3: Aspects of prototyping

2.3 The PTOOS System Development Process

Taking the two methods described above and combining them into the new PTOOS method results in the following development steps, also illustrated in Figure 4.

2.3.1 Requirements Analysis, Architecture Design and Work Packages

The development process begins with the analysis of the requirements in which, through interviews and discussion with the customer, a draft version of the requirements is formulated. In the next step, the attempt is made to describe the functionality of the future system by designing the system architecture. As a result, so-called *work packages* are identified which distinguish self-contained components of the system architecture as a whole. This approach allows for the organization of specialized teams in which its members can deal with their work package more intensely.

2.3.2 Glossaries

The teams now begin the compilation of the specification glossaries for each defined work package. At this stage, discussions with the customers continue to take place, in case the requirements documents are insufficient to completely describe the objects, attributes, services, and functions of the system. Team members also concurrently begin with the design of the graphical user interface using the chosen development platform's IDE according to the requirements and the specification described in the glossaries.

2.3.3 Specification Prototype

Once enough information has been gathered in the glossaries to fully specify one component of a work package, implementation of a software prototype based on the specification, a specification prototype, can be launched. This is done by „hooking up" the forms defined in the form glossary according to the functions and data flow anticipated for in the future system. This process means adding controls and fields to represent data, adding command buttons to trigger services and functions and creating links to the specifications documents. Links to documents replace missing code of the core functionality, since the goal here is not to implement the final, full-fledged software system.

Building the specification prototype is the core activity of the PTOOS method. A first version of the future software system can be presented, based on an object-oriented specification, giving a first impression to its future users.

2.3.4 Evaluation of the Prototype

While using this prototype version of the system, he or she can decide whether it complies with the requirements. If this is not the case, the requirements may need to be modified. If so, the documents of the requirements analysis may need to be revised. If the requirements are adequate, it may be necessary to revise the glossaries instead. This procedure allows for a review of the requirements and a revision of the specification.

If, however, the prototype complies with the stated requirements, another component of the work package can be specified and prototyping-wise implemented. This cycle is repeated until the specification is complete. Concurrently, other work packages are compiled the same way by other teams. The specification prototype matures to a pre-release running version of the whole future system, which is now ready to be coded in the following implementation phase of the software development process.

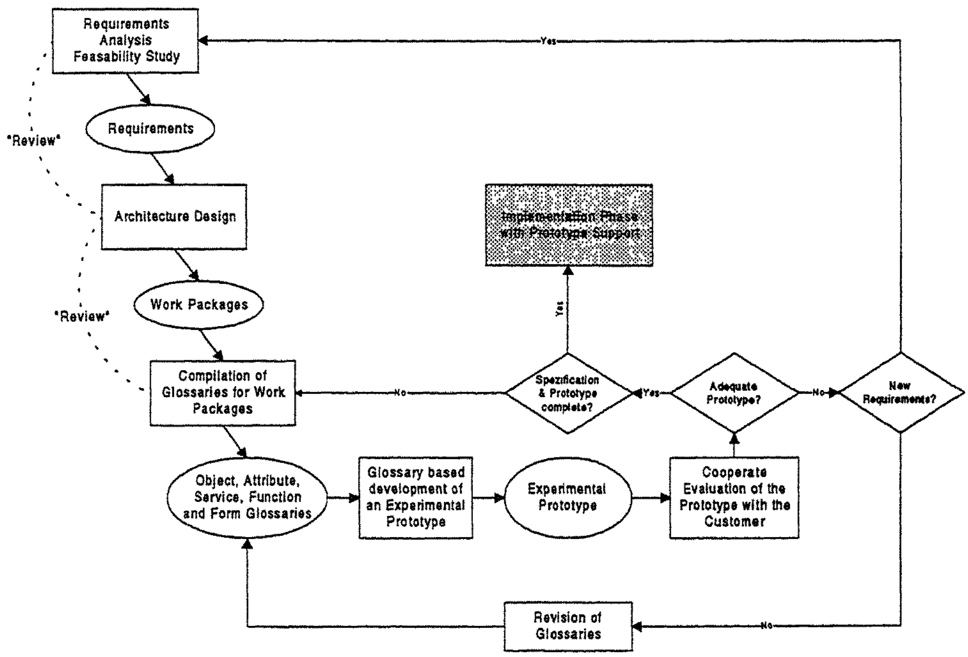


Figure 4: System development process based on the PTOOS method

3 CONCLUSION

Combining the prototyping approach with the object-oriented specification strategy into the PTOOS method has the benefits, that

- a methodical system development process is performed by consistent and continuous specification & realization support,
- a close and clear customer - developer cooperation is supported,
- restrictive system development factors such as cost and time are reduced by a modular "reusability" approach,
- a flexible system development is supported by providing means for the evaluation of alternatives, and
- the complexity of the development process is reduced by an IT-supported specification mapping and context-related retrieval potentiality.

The application of this approach positively contributed to the cooperative design and development of the production management system for Asian automobile manufacturer. The car manufacturer's team recognized the usefulness of „visually" following the progress of specification and used the prototype for orientation within this complex software system. The possibility of navigating the specification documents via the system's graphical user interface allowed the developers to use the prototype as a reference for the implementation of the real software system.

4 REFERENCES

1. Mertins, K.; Rabe, M.; Albrecht, R.; Beck, S.; Bahns, O.; La Pierre, B.; Rieger, P.; Sauer, O.: „Gaining certainty while planning factories and appropriate order control systems - a case study". Proceedings Seminar CAD/CAM'95, p. 8B1 - 8B20, Bandung, 1995.
2. Mertins, K.; Rabe, M.; Albrecht, R.; Beck, S.; Bahns, O.; La Pierre, B.: „Applications of customer-supplier relations on decentralized short term production planning and control in automotive industries.", Conference Proceedings of the 4th International Conference on Advanced Manufacturing Systems and Technology (AMST), Udine, September 1996.
3. Wilksch, S.: Wissensbasierte Spezifikation von Systemen zum fertigungsnahen Auftragsmanagement., Produktionstechnik Berlin, Forschungsgebiete für die Praxis 144, Editor: Spur, G., Carl Hanser Verlag, München, 1994.
4. Budde, R.; Kautz, K.; Kuhlenkamp, K.; Züllighoven, H.: Prototyping - An Approach to Evolutionary System Development., Springer Verlag, Berlin, 1992.
5. Floyd, C.: „A systematic look at prototyping", In: Budde, R.; Kautz, K.; Mathiassen, L.; Züllighoven, H.: Approaches to prototyping, Springer Verlag, Berlin, 1984.
6. Kilberth, K.; Gryczan, G.; Züllighoven, H.: Objektorientierte Anwendungsentwicklung: Konzepte, Strategien, Erfahrungen, Vieweg Verlag, Braunschweig/Wiesbaden, 1993.

IMPROVEMENT OF THE PLANNING QUALITY IN HIGHLY NETWORKED PROCESSING INDUSTRY PRODUCTION BY MEANS OF SIMULATION AND COOPERATIVE LOCAL CONTROL STATIONS

Dipl.-Ing. J. Pirron
Dipl.-Ing. J. Arnold
Dipl.-Ing. O. Reisch
Fraunhofer Institute Manufacturing Engineering and Automation (IPA)

ABSTRACT

The demands of the market for short delivery periods makes planning and control of multi-stage, customer order-oriented production difficult and complex. This is true especially for the semi-finished products industry, the semiconductor industry, the food processing industry, the paper industry, as well as parts of the chemicals industry. So far, neither rough scheduling PPC (Production Planning and Control)-systems nor disjoint local control stations have been able to manage the enormous coordination effort for planning multi-stage linked production. Thus, the objective is to develop an integrated system from globally coordinated planning and short-term production control with distributed cooperative local control stations. The expected benefits of such a system are reduced inventories and lead times (by about 30%) and an increase of the delivery reliability (by about 60%). In the past few years, the Fraunhofer Institute for Manufacturing Engineering and Automation (IPA) has developed a system architecture which meets these demands. A prototype is currently being used successfully in industrial processes.

1. INTRODUCTION

The logistic performance measures, like delivery periods, delivery reliability, throughput, inventory, and costs increasingly determine an enterprise's potential in the market. In many basic industries [1, 2], the simultaneous optimization of all these performance measures is very difficult. The pressure to build more and more complex plants which are only profitable with high utilization rates and large batches cause extreme problems for production control in linked productions fraught with conflicting restrictions. The manufacturing of semi-finished steel products serves as a good example for that. All kinds of different production technologies and their specific restrictions require a new batch composition for each process stage. (Depending on the company, the term batch or campaign composition is used.)

In order to reduce set-up costs and increase throughput, large batches are desired. In addition, the technology requires a certain sequence of orders within the batches. This leads to the situation that, depending on the production stage, the waiting time of a customer order for a suitable "slot" can be very long within this complex weave of technological restrictions and economic criteria. The proceeding of a customer order from the steel factory to the ready-to-deliver coil resembles a steeplechase rather than a smooth flow. Also, the plants' process capability often does not improve at the same rate as the customer demand for quality rises. Resulting materials quality shortage increases the stochastic character of the process and simultaneously reduces its predictability.

Huge buffer stocks of mostly allocated materials are built up between individual production stages as safety stocks in case of materials shortage, as well as to resolve conflicts of sequence. In the case of predominantly customer order-oriented production, large inventories of work in process (WIP) cause long lead times. Under these conditions, short delivery periods and high delivery reliability are wishful thinking rather than reality. For industries with oversupply it is more and more these values that have a major impact on an enterprise's position on the market.

The IPA in Stuttgart has successfully introduced several production planning systems for multi-stage processes. A conventional system could not have been satisfactory in any of the cases. Neither rough scheduling PPC (Production Planning and Control)-systems nor disjoint local control stations were able to cope with the enormous coordination effort necessary for the planning of multi-stage productions [3, 4]. The complexity of production control becomes obvious when considering a steel factory with an extreme variation of batch sizes:

- between 20,000 and 30,000 material units (slabs, coils, paletts, packages) with 10 to 15 work cycles each (about half a million work cycles have to be coordinated within 2 to 3 weeks)
- all orders are customer orders
- approximately 30 plants with 2 to 15 restrictions each
- work in process between 150 to 250 million marks
- 70 to 90% of the orders are not delivered on time

up to 200,000 tons average inventory of finished products

An additional disadvantage is lacking adaptability. Especially the oversimplification or incompleteness of order planning components call into question the use of standard systems in this essential part of a PPC-system. Accordingly, the user acceptance is low.
In most cases, only expensive external software experts are able to counteract these shortcomings. Thus, the introduction times are long, and the costs for small and midsize companies are intolerably high. Besides that, special adjustments make the systems even more inflexible and thus unsuitable for the growing dynamics of market, business processes, and manufacturing structures. But it is especially control loops in the shape of continuous improvement processes already implemented in many enterprises which increasingly require a high adaptability and agility of EDP-systems.

2. SYSTEM ARCHITECTURE

The planning system developed by the IPA supports planners from the realisation of company goals to the optimisation of local targets with respect to an optimal control and transparency of customer orders.
An existing rough planning system determines delivery dates for customer orders on the basis of a simple static capacity check. It is important that a request is dealt with immediately and that the distribution gives reliable information about delivery date and volume quickly to the customer.
In order to guarantee an on-time delivery of orders, fine-tuned scheduling of the individual work cycles is necessary. In contrast with existing systems, the global planning [5] carries out a general streamlining of all work cycles of customer orders under consideration of important initial conditions. The conflict between short lead times and optimisation of plants regarding technological restrictions and economic batch sizes is resolved globally. The results are optimal scheduling of promised delivery dates, as well as realistic guidelines for the individual production stages. The perfect moment for materials usage (start of production) is set. The determined 'time windows', i.e. the local planner's disposition time, considers the transport times between work cycles, so that no inconsistencies concerning availability of materials occur in the plants. Among others, the following initial conditions are taken into account:

- shift schedule of plants
- campaign dates
- planned and time-consuming set-up
- capacity use in manufacturing
- existing plant programs
- current inventory conditions of the plants (work status of started customer orders)
- important technological restrictions specific of plants (e.g. maximal changes in width)
- transport times

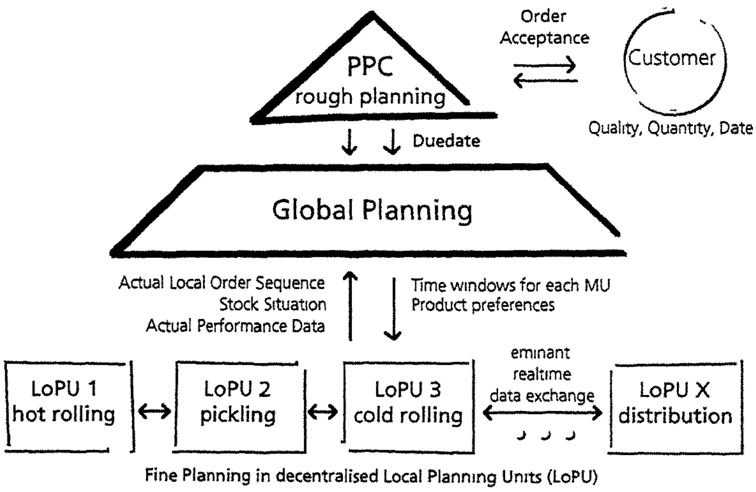


Figure 1: System Architecture of Global Planning and Decentralised Cooperative Local Control Stations

Whereas the optimisation potentials of the global planning are located in medium-term areas, the local planning units focus on short-term production control which is carried out by means of optimal batches or order assignment sequences of the plants. Plant programs have to be set up under consideration of all restrictions specific of plants and the given 'time windows'.
Nowadays, job shop scheduling for following shifts based on the plant's order inventory. In contrast to that, the new system architecture supports a discrete event real time coordination with the upstream and downstream plants. As soon as upstream and downstream production stages include an order in their job shop schedule, the set production dates, material availability and demand are known globally. If individual 'time windows' cannot be kept, conflicts may arise. By means of a standardised evaluation system, the cooperative local control station system supports the solution of such conflicts with consideration of global objectives.
With the help of data from the shop floor disposition, existing local control stations are updated. Thus, real time supervision of the realisation of plant programs in the manufacturing process is not possible. For purposes of reactive control, the local control stations that are integrated in this architecture are connected online with the shop floor disposition and production data capturing systems. The stochastic influences of production are visualised on the surface of local control stations and their results immediately become transparent. With the help of an integrated alarm system, schedule conflicts caused by discrepancies between desired and actual performances can be recognized without delay. Efficient graphic user surfaces inform the planner about the results his decisions had on upstream and downstream plants, as well as changes of customer orders in production. The planner is thus able to make the right adjustments, while again taking into account global guidelines and consequences for upstream and downstream stages of production.

3. GLOBAL PLANNING FOR AN OVERALL JOB CONTROL

Via an interface - developed by the IPA in collaboration with the customer - with the overlaid PPC-system, the latest order pool including delivery dates, work schedules, and current work progress is being integrated daily in the global planning data base. Exact process times per material unit and work cycle for a specific customer order are needed for realistic scheduling. If not available otherwise, they can be determined by means of a process times model.
The essential scheduling is carried out by an enterprise-oriented simulation model based on the simulator SIMPLE++ which was created by the IPA in cooperation with the AESOP GmbH/LLC Stuttgart. After all the pertaining data has been loaded into the model via interface with the global planning data base (SQL), backwards scheduling, based on the promised delivery dates, determines the latest possible order processing dates. With regard to these dates, realistic deadlines are then set in the forward simulation.

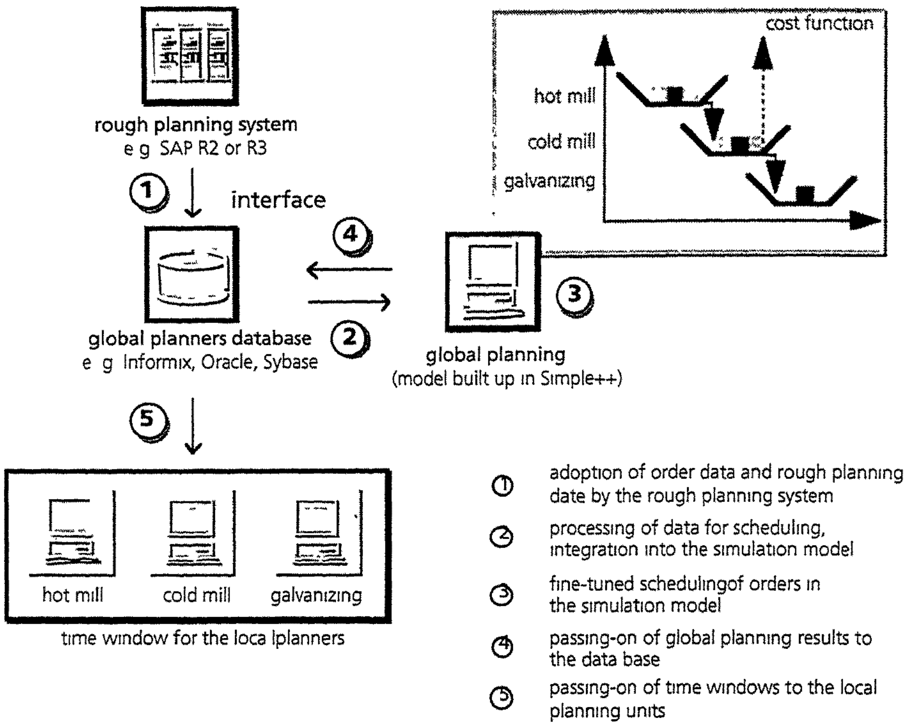


Figure 2: Course of Global Planning

With the help of these global planning scheduling guidelines, the availability date at the plant is set in such a way that the transitional time is added to the production date of the upstream plant. The access to the plant's incoming stores is the beginning of the disposition window for the decentralised planner. The end of the disposition window is determined by the global planning production date at the plant. Between theses two times, the decentralised planner can take care of the order, while taking into account all the restrictions without jeopardising the planned manufacturing process of the order.

4. SCHEDULING BY MEANS OF A HEURISTIC METHOD BASED ON DISCRETE EVENT SIMULATION

The optimisation of the order processing sequences within the global planning building blocks is based on user-oriented heuristics. The IPA has a huge object library with standardised control moduls, as well as strategies specific of industry and technology. In comparison with other scheduling methods and systems with fixed lead times, the model-based simulation with heuristics has the advantage of calculating lead times dynamically, based on realistic processing times for each order at each production level. Waiting times and transitional periods result from the current state of inventory in front of the plants and the available transport capacity.

Unlike MRP-based systems that carry out a simple capacity requirements planning, discrete event simulation in production control make it possible to apply the system's set of rules not only to the line's global view, but it also can be controlled via order parameters. This creates the possibility to set the rules for a plant and to determine the sequence of individual orders. So, for example, orders of class A customers can be prioritised. However, this rule can be limited, so that it is only valid if the state of inventory of all succeeding plants allows this.

5. BUILDING THE SCHEDULING MODELS

In contrast to standard systems, here the user gets involved already in the planning of the system architecture. This enables the planner or logistics expert to make short-term adjustments in case of changes in the organisational structure. If changes in the production structure occur, like the model's integration in and adaptation to new plants or factories, the building and updating of the model is carried out by a logistics expert who is familiar with the problems that exist in manufacturing. Expensive and time-consuming coordination cycles between company and EDP are not necessary. The possibility to model processes and planning logic in a user-friendly manner was created through user-defined building blocks in SIMPLE++ which were developed by the IPA specifically for this purpose. The fact that the system is composed of user-defined building blocks, as well as the global system's object-oriented technologies, make it possible to integrate organisational changes of manufacturing into the system within minutes. This strenghens the company's agility in the continuous improvement process.

6. NETWORKED LOCAL SHOP FLOOR CONTROL STATIONS

The time windows' for each work cycle of all material units which were determined by the global planning and which are not in conflict with each other have to be transformed into feasible plant programs. The first step is a computer-controlled method for the creation of an automatic plant program, the second one is a graphic-interactive local control station surface for the manual modification of plant programs. Both tools are modelled specifically for plants.

In order to delegate the task of modelling to the logistics expert, simulation is used. SIMPLE++ serves as the evaluation system within an integrated problem-solving method consisting of genetic algorithms and simulation. This method provides the planner with a fine-tuned sequencing that considers all restrictions of one plant. Again, one important advantage of the system developed by the IPA is its ability-after a project-accomanying introduction of the overall system - to transfer the expert knowledge on to the decentralised logistics specialists. They contribute directly to the optimisation process with their experience and knowledge. Thus, timely and informationally critical arrangements with experts, as is the case with the global planning, are not needed.

In the local control station, the latest work progress, as well as the planned plant programs, are presented online. Efficient functions, like resequencing of work cycles with 'Drag & Drop' supports the generator of plant programs in the building and modification of programs. All relevant data of a material unit are shown, and important numbers, like coffins, are illustrated graphically.

Global planning and local planning are coordinated in such a way that usually production dates can be found within the leeway of 'time windows'. Process stochastics, like plant or material failures, and special plant restrictions not shown in the global planning can have the effect that material units have to be manufactured outside of their 'time windows'. Work cycles, tightly networked due to reasons of inventory and processing minimisation, cause conflicts with upstream and downstream plants. Here, the notion of cooperative local control stations offers an efficient and - regarding delivery reliability and the minimisation of costs - optimal solution. Based on the principle of model-view-controller (MVC), the local control stations are networked online in a broker-server-architecture [6, 7, 8]. That way, each conflict in scheduling with its specific impacts on deadlines, especially customer delivery dates, as well as production and set-up costs which can be influenced by planning, can be identified. An internal cost model, which is integrated into the local control station network, is the key to an optimal solution concerning costs and delivery dates. An alarm system warns early of critical shifts within the system of scheduling and material availability.

7. SUMMARY: BENEFITS AND OUTLOOK

Due to the networked state of decentralised entities that was accomplished with the help of a global planning simulation, and due to the provision of all decentralised logistics experts with up-to-date information, the whole planning system changes from an accumulation of disjoint manufacturing areas to a customer-oriented cosmos. It used to be the planners who were informed best about the problems of their plants. Today, however, there is an openness about planning information that incorporates the entire production network. The transparency of upstream production stages widens the planning horizon of each planner. The replanning of a specific plant causes a number of consistence checks within the new architecture, so that the planner knows in advance the effects the replanning of each single order has on the entire system.

In addition, the early recognition of conflicts and bottlenecks and solutions simultaneously suggested by the planning system create the necessary foundation of trust for a stronger networking of the production. The above described transparency and the planners' willingness to respond to customer concerns has an impact on the management ratios of the considered enterprise.

A stronger networking of the local planning entities has the effect that the planner does not consider exclusively his own inventories. This kind of planning that exceeds one particular area is expected to result in a 30% reduction of the work in progress. Since the production in the described enterprise is

customer-oriented, the reduction of total inventories is connected with the reduction of lead times (by about 30%). In the case of inventories to the value of 150 to 250 million marks, it is expected that this reduction of inventory will save several million marks each year.
It is important to make sure that the utilisation of the plants and the total throughput is not affected negatively by the above described measures. The consideration of campaigns and crucial technological restrictions within the work cycle scheduling (that goes beyond single production stages) might even improve the composition of batches and reduce the percentage of set-up time and costs. At the same time, early identification of conflicts and common objectives increases the reliability of customer order deliveries. In the past, very few customer orders have reached the finished products storage on time. Now, delivery reliability of up to 90% is possible. This, in return, will change the distribution, so that hardly any safety buffers between the date of the customer order and the finished products storage will be needed. Inventories can be further decreased, and this will have very positive effects on the capital tie-up costs at the end of the value added process. Due to the improvement of these performance measures, the enterprise will become more flexible and agile on the market and economically more efficient in production. The described optimisation of the logistics parameter, together with the open architecture, are an incentive for the IPA to further develop this architecture. At present, the main focus of the project is the development of heterogeneous and international production networks. The fast adaptability of these networks is one of the enterprises' major goals. Compared to other PPC-projects,the realisation of the planning system can be carried out with considerably less effort. This was possible due to continuous use of object orientation and user-friendly modelling of the simulation within both global and local planning. The implementation of local planning in the shape of a data base solution in a client-server-environment in the system described above could have created the necessary discrete event-oriented technologies only at high costs. The IPA has great experience in the field of object-oriented technologies, especially in distributed discrete event applications, which helps to avoid mistakes that occur already during the conceptualising of new systems.

8. REFERENCES

1. Große-Oetringhaus, W. F., Fertigungstypologie unter dem Gesichtspunkt der Fertigungsablaufplanung, Duncker & Humblot, Berlin, 1974.
2. Overfeld, J.: Produktionsplanung bei mehrstufiger Kampagnenfertigung: Untersuchung zur Losgrößen- und Ablaufplanung bei divergierenden Fertigungsprozessen, Reihe Schriften zur Produktion, Ed.: Witte, T.; Rieper, B., Diss., Fachbereich Wirtschaftswissenschaften, Universität Osnabrück, Verlag Peter Lang, Frankfurt am Main, 1990.
3. Haupt, R.;Nöfer, E.: Produktionsplanung und -steuerung dezentraler Einheiten, in: Corsten, H. (Hrsg.): Handbuch Produktionsmanagement, Wiesbaden, 1994.
4. Zäpfel, G.: Dezentrale PPS-Systeme - konzepte und theoretische Fundierung, in: Zägfel, G. (Hrsg.): Neuere Konzepte der Produktionsplanung und -steuerung, Linz, 1989
5. Arnold, J.;Schuler, K.: Simulation als Betriebsprognosesystem in der Halbzeugfertigung, Tagungsband CAT '93, Stuttgart, 1993.
6. Shan, Y. P.: An Event-Driven Model-View-Controller Framework for Smalltalk, OOPSLA'89 Proceedings, 1989.
7. Krasner, Pope: A cookbook for using the Model-View-Controller user interface paradigm in Smalltalk-80. Journal of Object-Oriented Programming, 1988.
8. Pree, W.: Design Patterns for Object-Oriented Software Developement, Addison-Wesley, 1994.

AN ALTERNATIVE CONCEPT TO MRPII FOR MASS CUSTOMIZATION BASED ON THE OBJECT-ORIENTED PARADIGM

Claus Rautenstrauch
Department of Information Science
University of Konstanz
Germany

ABSTRACT

By *mass customization* (MC) we mean a synthesis of mass production and the production of customer-specific ("customized") products. Companies which employ MC are in the position to produce large numbers of customized products in a short period of time. New concepts of manufacturing planning and control (MPC) are needed for this. Since with MC coarse planning in the usual sense is no longer possible, procedures for ad hoc generation of bills of materials and work schedules as well as scheduling are necessary. These are discussed in detail in this contribution, and an object model for a MPC expansion is presented. At the close of the presentation there will be a brief look at an actual implementation of the presented concepts.

KEYWORDS

Mass customization, manufacturing planning and control (MPC), time and capacity scheduling, manufacturing resource planning (MRP II)

1 MASS CUSTOMIZATION

1.1 Definition and Demarcation From Mass Production

By *mass customization* (MC) we mean a synthesis of mass production and the production of customer-specific ("customized") products. In the past companies had a choice between producing low-cost mass produced products or taking orders to produce customized products at relatively high prices. Contemporary MC aims at combining the positive qualities of both alternatives: (mass) production of individual goods using mass production processes with the highest product quality and shortest possible delivery times [1]. Suppose we take the customer's viewpoint: In the ideal case products manufactured using MC should not differ from mass produced (MP) goods, whether in price, quality or delivery time. The only distinguishing characteristic should be the perfect fulfillment of individual demands.

The driving force of the production process is the client with his specific demands. MC thus comprehends *self-customizing*, in which the customer himself configures the desired product from the producer's offering. Beyond this new product ideas are also worked out on the basis of active analyses of user needs. Examples of products produced by MC processes are specialized tools, vehicles and vehicle parts, truck body works, electronic pagers, windows and outdoor clothing or also information products such as product-specific operating instructions [1, 2]. However, not all products are suitable for MC. On the one side, customers can easily be overtaxed by an excessively broad product offering. On the other hand, with certain products there are limits imposed by the simplicity or particular complexity of the product's design.

1.2 Organizational Foundations

Though at first glance the concept of MC seems simple, it is difficult to realize in practice. Pine [1, 3] describes very graphically a few examples from the Japanese automobile industry in which MC efforts were unsuccessful. The reason is that MC cannot be developed by a process of evolution from existing concepts like MP or *continuous improvement.* Instead, radical redesign of a company's organization structures and the structuring of its business processes is required.

Basically MC necessitates *organic* organization. The foundations of MC are dynamic networks consisting of relatively autonomous work groups. These interact with other work groups independently of individual product specifications. In this sort of organization the relationships between the individual work groups and their composition typically undergo a process of continuous change. The workers in such networks are not only involved in carrying out assignments, but are also active in operational planning and quality control. This demands high qualifications and the readiness on the part of employees to assume partial responsibility in the execution of tasks.

In an organic organization managers take on the roles of coordinators. Together with work groups they plan the respectively most effective form of cooperation, as well as supporting and monitoring the course of production. Not only the role of employees and management, but also that of technology changes with MC. With conventional organization technology is chiefly employed in automating and accelerating work processes. It therefore has a *substitutive* character in regard to the employment of human capital. In an organic organization, by contrast, technology has a primarily *supportive* function. This is because in this case human creative and problem-solving potentials will be systematically exploited ("Robots don't make suggestions").

2 OBJECT-ORIENTED MPC APPROACH FOR MC

2.1 Limits of Manufacturing Resource Planning (MRP II) Systems for MC

In contrast to MP, for the planning of production processes with MC neither basic data in the form of electronic bills of materials nor work schedules for the parts to be produced are available as a basis for planning. Since planning procedures for MPC systems based on the MRP II concept start with these basic data, it is obvious that MPC systems cannot be employed for MC. Of course it is conceivable that each part theoretically manufacturable by MC could be stored as a variant in an MRP II system. However, the great range of possible variants leads as a rule to an explosive growth in the size of bills of materials- and work scheduling databases. The systems thus tend to become unmanageable.

A further barrier to the employment of MRP II systems is the multi-stage planning concept underlying these systems [4]. Here in production overall planning precedes current fine-tuned planning, which with MC, however, cannot realistically provide advantages. Coarse planning is therefore useless with MC.

Nevertheless, even with MC the MRP II concept need not be completely superseded by other concepts. MC does not mean that companies produce just any products. Rather, for products belonging to specific product categories, desired variants can be produced in the smallest desired quantities. This implies that specific parts and assemblies are produced as *base elements* which are identical even with custom manufactured parts. The production of base elements can, as in the past, be planned and carried out in the conventional manner. The functions for MC, that is, the ad hoc generation of bills of materials and work scheduling, as well as real time scheduling, can be grounded on the database of the MRP II system [2]. This includes drawing on master part file bills of materials, work schedules and operations. Since it can be assumed that base elements are the chief causes of inventories, inventory management problems can be solved by conventional MRP II systems. Thereby MPC expansion for MC focuses on short-term fine-tuned planning.

2.2 MC Functions

2.2.1 Bill of Materials Generation

Since with MC the end products are specified by customers, neither bills of materials (BOMs) nor work schedules are available in advance. The first task is therefore to create these data ad hoc from the customer's part specifications.

The concept of BOM generation presented here, which suitably supports MC requirements, is based on the object-oriented paradigm. First, objects of the *part* class will be dealt with. Besides the usual attributes like (specific) part number, designation, type, etc., a part can *possess* a number of additive attributes. The possession of such attributes is specific; an attribute can only be possessed by a specific object. If a part object occurs several times in the real world, the object must be given a quantity attribute. Further, a part object can be *labeled* with any desired number of attributes, whereby this label need not be specific. An object which possesses an attribute is always labeled with it. Further, a number of values can be assigned attributes of a part.

The labeling and possession of attributes has the following background: The specification of a customized part occurs through the definition of attributes and values which the part should have. If the customized part belongs, e.g., to a bicycle, this can be described with the attributes: *frame type, frame height, frame color, mud guard color, number of gears, type of gearshift*, etc. The customized bicycle is then completely described by the concrete values: *city bike frame, 91 cm, green, black, 5, hub gearshift*, etc. For the configuration of customized products, *configura-*

tion tools can be used which are based on electronic product catalogues [2]. Such systems support users by showing all available components and meaningful combinations of them and excluding physically impossible configurations. Often advisory systems are also integrated into such configurators. It is thereby ensured that actually realizable products are specified and the subsequent generation of bills of materials and work schedules can dispense with time-consuming consistency tests.

The object specified by the customer is instanced as a new object. This new object does not *possess* any attributes, but it is *first labeled* with all attributes. Then all instances of parts are sought which are labeled with them. The connection between labeling and the possession of attributes as well as assigned values is as follows (the examples given in parentheses refer to figures 1, 2 and 3):

♦ If a part is labeled with an attribute which is assigned no appropriate value, this means that at a deeper level of the production structure for this part a part appears which possesses the appropriate attribute (e.g., A is labeled with attribute 1, which D possesses). If several parts are labeled with the same attribute and appropriate value, this means that they appear in the same BOM (e.g., parts H and I of assembly D). The connection between the parts can therefore be adopted from the standard BOM of the part from the database of the MRP II system. Parts without labels are standard parts which do not influence the customizing of products (e.g., part F). A part which possesses an attribute may be labeled with other attributes (e.g., part D possesses attribute 1, but is also labeled with attributes 3 and 5). This means that at deeper steps taken by this part a part appears which possesses this attribute (e.g., part K, which possesses attribute 3).

♦ If a part possesses an attribute, this part determines, on the basis of the assigned value, the further structure of the BOM. In this part's BOM only parts may appear which were assigned the corresponding value. In this way the structure of customized variants is precisely described.

♦ If a part is labeled with an attribute it does not possess and a value is also assigned to the attribute, the part belongs to a BOM whose top element is the possessor of the corresponding attribute (e.g., parts H, I, J and K belong to assembly D).

The algorithm for the ad hoc construction of a BOM is now presented and illustrated on the basis of a simple example. The starting point is the specification of a part X, which was specified with the individual values A_{1i}, A_{2j}, A_{3k}, A_{4m}, and A_{5n}, with A_{pq} as the value q of attribute p.

BOMs are represented here as BOM trees, and quantity coefficients are omitted, since they are adopted without change from the MRP II database. Each node is divided into four domains: above left is the part designation, below left the value for which the node is valid, above right the attributes with which the part is labeled, and below right the attributes which the part possesses. The BOM can be labeled using simple algorithms and is therefore not dealt with in detail here. The algorithm for the ad hoc generation of bills of materials is as follows:

1. In the first step all parts are selected from the part class which are labeled with the attributes of the specified part and to which no value is assigned (parts A, C and E). Further, all parts are selected to which values are assigned which appear in the specification of the end product and which possess all the attributes of the assigned values. In the example these are parts D, G and M. The values are to be stored in the part class as variant designations.

2. Parts labeled with the same attribute are interconnected, in accord with the one-level BOMs of the MRP II system, and the base elements are supplemented (in the example these are B and F). Figure 1 shows the previously generated interim BOM.

3. Now only parts are lacking which depend on attributes and values which are already present in the interim BOM. Therefore all parts are now selected to which values are assigned for which a possessor is already present in the interim BOM. Additionally, parts of this sort are also selected which possess attributes with corresponding values and to which additional values are assigned which are already present in the interim BOM (parts H, I, J and K). Part D shows, with its labels (attributes 3 and 5), that at deeper levels additional individual parts appear.

4. On the basis of BOM information from the MRP II database, the nodes are then connected with each other. Figure 2 shows the expansion of an interim bill of parts through supplementation of part D's subtree.

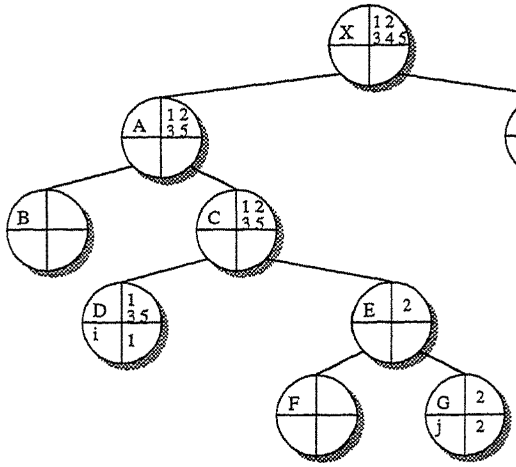


Fig. 1: Interim BOM

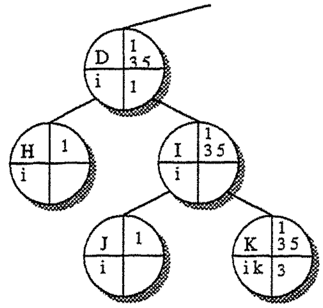


Fig. 2: Expansion with assembly D

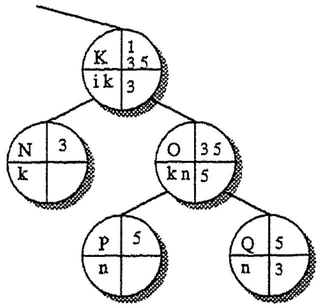


Fig. 3: Expansion of the interim BOM with assemblies K and Q

The example shows on the basis of K's label that a further individual attribute is present on a deeper level. Therefore steps 3 and 4 must be repeated until no additional parts are selected in step 3. Figure 3 shows the BOM tree for part K after twice-repeating steps 3 and 4.

2.2.2. Creating Work Schedules

In conventional MPC systems, work schedules which represent internally produced parts are respectively assigned to the nodes of a BOM. Thereby it is determined which manufacturing steps are to be performed in order to prepare the part represented by the nodes.

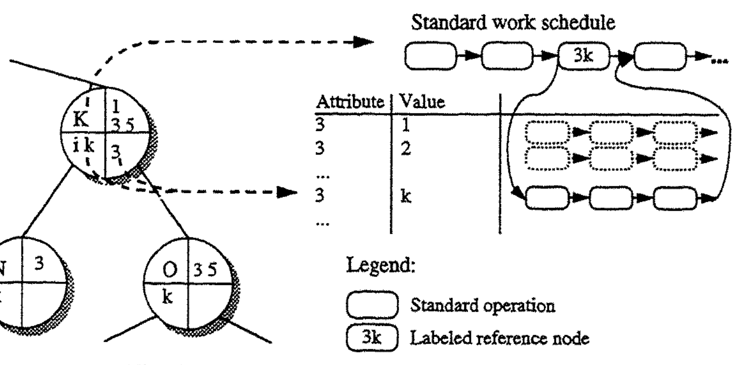


Fig. 4: Bills of Materials and Work Schedules

Work schedules are represented in simplified form here as chains of operations in which each node represents an operation. For MC the concept is expanded in that standard work schedules can contain *reference nodes* which indicate the variant operations or operation chains which depend on the value of an attribute. Each reference node is labeled with the attributes which are to be supplemented for the variant-specific operation chains. A work schedule for a part to be concretely produced is created by first selecting the standard work schedule for the part. Then the respective variant operation chain is substituted for the reference nodes for all attribute-value combinations which are noted in the BOM for the respective parts. Figure 4 shows the connection between bills of materials and work schedules.

2.2.3 Scheduling of Operations

Since with MC there is no overall planning, the operations must immediately be scheduled in detail. Operations in MRP II systems are typically assigned to resources or respectively resource groups. However, this concept will be deviated from here, since it is too inflexible for MC.

An operation defines an indivisible sequence of activities which are to be carried out by a work group. BOMs define sequence conditions and quantities in which parts are to be manufactured. With the assignment of work schedules to network nodes it is additionally specified which activities must be carried out for the production of individual parts. Work groups are modelled in the system as a group of individual resources, whereby skills needed to carry out specific activities are paired with each individual resource (employees, machines, facilities, means of transportation, etc.). These skills are acquired by work groups in the organizational selection of work groups. For each operation, master data show how much time is required for the individual activities.

In scheduling, each part studies the work group object to determine which work group can contribute the greatest share of the capacity need for the execution of the operations of its work schedule. Then a shop job for a respective part is assigned to this work group and the necessary capacity is reserved. A *shop job* is defined here as work instructions for a work group which state not only the activities to be carried out, but also the time- and material need. For the determination of the starting dates consideration must be taken of the capacity which has already been reserved through other shop jobs. Scheduling can be done either forward or backward.

The advantage of this procedure is that a reorientation of work groups has no influence on the course of the scheduling process. However, after the reorientation of the work groups rescheduling is necessary. The heuristics are to consistently assign shop jobs to the work groups which can employ the greatest share of the needed capacity. This is justified by the premise that work groups are always selected with the aim to accomplish the greatest possible amount of related parts of jobs. If it turns out in scheduling that the promised delivery dates are frequently missed or individual work groups are greatly over- or under-utilized, then the work groups must be reconfigured. Since the basic data structures – work schedules and operations – are not fundamentally changed by MC, simple rescheduling can be carried out as in the past with a conventional job scheduler.

2.3 Object Model

The relationships presented in section 2.2 are based on the object model shown in figure 5. As a metamodel the notation of the extended entity relationship model (EERM) was chosen [5]. This is because it starts from standard entity relationship model (ERM) and simultaneously supports object-oriented concepts. The notation is to be interpreted as follows:

- *Entity types* (in object-oriented terminology *classes*) are represented as rectangles, *relationship types* as diamonds and *attributes* as rectangles with rounded corners. This corresponds to the notation in the standard ERM. Lines connecting classes with relationship types are labeled with (min-max) cardinalities indicating the minimal and maximal numbers of elements which a list or group can contain (* = unlimited). Lines of recursive relationship types are labeled with the designation for the interpretation of the direction (in figure 5 this is the relationship type "direct").

- Attributes can be *object-valued*, i.e., themselves consist of a *group* or *list* of objects of another class with which the attribute is connected by the □– symbol. Each group or list is assigned a (min-max) cardinal number. If the □– symbol has an arrow point on the end, then the class that the arrow points to is dependent on the attribute, i.e., an object of an independent attribute exists only if this appears in an object-valued attribute. Object-valued attributes are used if relationships between entity types are permanently valid, otherwise relationship types are employed.

♦ The triangles connect *input-* and *output classes*. Input classes are connected with the basis line and output classes with the corners of the triangles. Input classes transmit their attributes to the output classes, whereby output classes can have additional attributes. Thereby output classes are *specializations* of output classes. The equal signs in the triangles symbolize that the entity quantities of the output classes are disjunct.

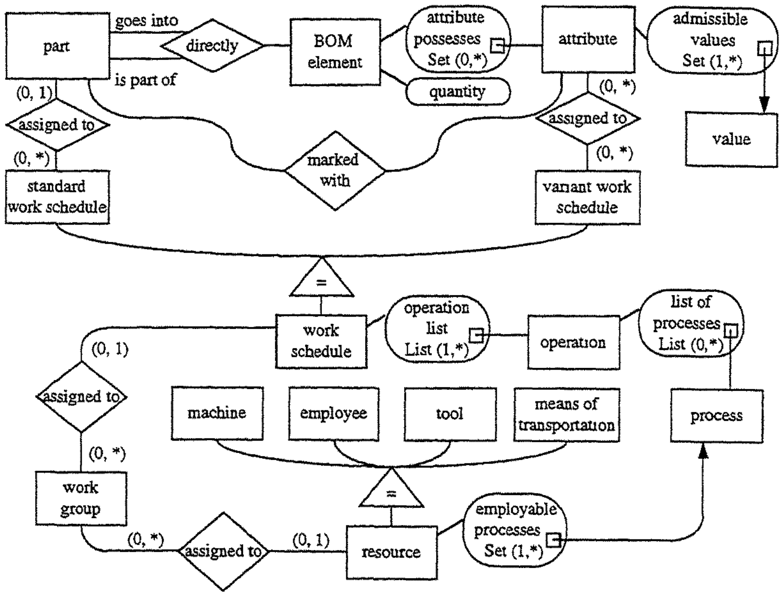


Fig. 5: Object Model

3 PROSPECTS

The procedures presented here have been realized in greatly simplified form by a praxis project in which individualized technical documents were created. The simplifications occurred in two areas: On the one side, the BOM are only one-level, and, on the other, the materials (text blocks, illustrations and tables) as well as the resources (special programs) are present only in electronic form. Nevertheless, on the basis of this system not only the general functional competence of the concept, but also its excellent performance can be demonstrated.

An essential point to be dealt with in the future is how to support the training of optimal work groups. Production system efficiency depends specifically on putting together the most effective work groups. Since the order situation and the work group composition interact, procedures must be developed to support this harmonization.

REFERENCES

1. Pine II, B. J., Victor, B. and Boynton, A. C., „Making Mass Customization Work“, Harvard Business Review, 36, 5, pp. 108-119, 1993
2. Moad, J., „Let Customers have it their Way“, Datamation, 41, 6, pp. 34-39, 1993
3. Pine II, B. J.: Mass Customization – The new Frontier in Business Competition, Harvard Business Press, Boston, 1993
4. Isaa, T. N., Czajkiewicz, Z. J., „MRP II: Manufacturing Resource Planning: System Development, Implementation and its Impact on Productivity at the Factory Level“, Modern Production Management Systems, Kusiak, A. (ed.), pp. 711-727, North Holland, Amsterdam et al., 1987
5. Engels, G., Gogolla, M., Hohenstein, U., Hülsmann, K., Löhr-Richter, P., Saake, G., Ehrich, H.-D., „Conceptual modelling of database applications using an extended ER model“, Data & Knowledge Engineering, 9, 1, pp. 157-204, 1992/93

SIMULATION FOR EVALUATION OF SCHEDULING RULES IN FLEXIBLE MANUFACTURING SYSTEMS

F. T. S. Chan[1] and R. W. L. Ip[2]

[1] Department of Industrial and Manufacturing Systems Engineering, The University of Hong Kong, Hong Kong
[2] Department of Manufacturing Engineering, City University of Hong Kong, Hong Kong

ABSTRACT

This paper explores the operational problems of Flexible Manufacturing Systems(FMS) through simulation. The purpose of this study is to evaluate various combinations of scheduling rules in the FMS system, i.e. given a set of dispatching rules applied at the workstations, the effect of different machine selection rules to the system performance are analysed. Various measures, such as makespan, lead time, machine utilisation, net profit, time delay and inventory level at local input buffer are examined.

KEYWORDS

Simulation, Scheduling, Flexible Manufacturing Systems

1. INTRODUCTION

In recent years due to the high competitive market, it has become necessary for manufacturing systems to have quick response times and high flexibility. The largest single event having a positive impact on manufacturing improvements has been the introduction of FMS. Owing to its highly automated nature, a typical FMS has a high investment cost. Hence it becomes necessary to select the best configuration at the design stage, and more importantly to identify the most efficient scheduling rule at the operating stage itself. The scheduling decision rules can have a substantial effect on the system performance.

An FMS, through a careful implementation, enables an organisation to respond rapidly and economically, in an integrated manner, to significant changes in its operating environment. Some of the advantages of FMS include: improved capital/equipment utilisation, reduced work-in-progress(WIP) and set up, substantially reduced lead times and increased the profit margin.

The high investment cost of FMS justifies the use of computer simulation support. For example, in the operation phase of the installed manufacturing system, it is required to maintain high system performance by predicting the system behaviour under any feasible production schedule which can meet the daily production requirements and by selecting the most effective production schedule among the alternatives prior to its implementation. However, the overall system performance is usually contributed from several criteria measurements, such as machine utilisation, net profit, makespan etc. Hence, the selected production schedule should not be aimed at just one single objective, say, just aimed for a high profit return; this may cause an unacceptable long makespan. This consideration may dictate the use of simulation technique in selecting the best scheduling rule during the operation phase of FMS. In this paper, an attempt is made to explore the operational problems of FMS via simulation.

2. OVERVIEW OF SCHEDULING AND SIMULATION MODELLING

When considering the operational problems of FMS, such as scheduling and loading strategies, simulation methodology seems to be useful to address these issues. Many authors(1,2) have used various criteria for the generation of optimum schedules, such as number of tardy jobs, number of completed jobs in process inventory and machine utilisation. Montazeri and Wassenhove(3) stress the need for simulation prior to actually setting up the FMS. They use a user-oriented discrete event simulator to mimic the operation of a real life FMS. Stecke and Solberg(4)have carried out a detailed simulation of a real life system. They have tested various alternatives and evolved loading and control methods which significantly improved the systems production rate. O'Keefe and Haddock(5) discuss the advantages of data driven generic simulators for FMS.

There are many commercially available simulation languages and packages such as SLAM, SIMSCRIPT, WITNESS, SIMFACTORY, SIMAN, ARENA. A set of important attributes while selecting simulation software for manufacturing is given by Law and McCourse(6). Generally, different authors give different statements of the functions of a simulation package tool, depending mainly on how detailed this statement is. In summary, the following advantages of simulation methodology for modelling FMS can be highlighted:

- Simulation can reduce the risk of installing an FMS which may not provide sufficient flexibility.

- A simulation model can represent important characteristics of an FMS more realistically. It may incorporate the complex interactions which may exist between various variables, for example, loading strategy at workstations and policy of machine selection for alternative routings.

- Alternative FMS designs can be evaluated easily in a controlled environment.

- A computer simulation model's ability to address directly the measures of performance typically used in FMS evaluation helps to calculate the same measures of system performance for hypothetical FMS configurations as used in judging the real systems.

3. DEVELOPMENT OF THE MODEL

3.1 The Simulation Model

The objectives of the present simulation model include:

- To model a typical FMS system in an easy-to-understand package and to develop a simulation model in a PC-based environment so that the model can be used as a pedagogical tool while introducing FMS.

- To study the effect of various dispatching rules at the workstation and different machine selection rules for alternative routings.

The first step in any FMS simulation study is to establish the system configuration and then to develop a simulation model. In this regard, keeping the objectives of the simulation model, SIMFACTORY has been chosen as a programming vehicle.

A simulation model is developed to address scheduling rules in a typical FMS environment as shown in Figure 1. The logic of this simulation approach is presented in Figure 2. This is a typical prototype of FMS with seven jobs and six workstations. The same situation could have been very difficult to model through an analytical model. Through simulation one can address various "what if" scenarios. A scheduling rule is used to select a job to be processed from a set of jobs waiting for service. These rules can also be used to introduce workpieces into the system, to route parts in the system and also to assign parts to facilities. Scheduling runs may be static or dynamic. Because of the large number of scheduling rules, it is not obvious which scheduling rule to select in a given environment. Thus there has been a tendency to select a simple rule randomly (e.g. FIFO). This rule is then hardwired into the automated FMS system. Studies(5), however, have shown that the selection of the scheduling rules can have a significant impact on system performance. Hence, in recent years, substantial research and study has been carried out in analysing these scheduling rules. The purpose of the present simulation is to analyse various combinations of scheduling rules in the FMS system, i.e. given a set of dispatching rules applied at the workstations, the effect of different machine selection rules to the system are studied.

3.2 Assumptions

The assumptions are as follows:

- There is flexible alternative operation routing for products.
- There is only one operation at a time on a machine.
- One machine of each type exists at a workstation.
- Set up times are constant.
- There is an uninterrupted operation on the machine once the processing starts.
- The machines are continuously available for production.
- The same pallets are used for different jobs.
- The job list is present at the beginning of simulation. No new jobs are added.
- Raw materials, tools, jigs, fixtures, etc. are present and released immediately when required.
- Cost data are provided.

The operation routings and processing times of the FMS are shown in Table I. It may be noted that this particular configuration of products and machines has multiple routings. The model was developed and run on a PC-AT system. A combination of five rules at the workstation and three rules for machine selection was simulated. A set of rules used is given in Table II. All the other data are presented in Table III.

Referring to Table IV, the carrying cost reflects the cost of having parts sitting in the factory. This applies to raw materials and to WIP parts. It reflects the time value of money: the longer a part remains in the process, the higher is its cost to produce. It also reflects the cost of stockroom and queue space; it may be expensive to keep large amount of WIP within your process. Furthermore, each final product has a sale price. This is the market value of the product, i.e. this is the amount you are paid for the product. The developed simulation model compares this value to the cumulative cost incurred to produce the part. This comparison is one measure of the net gain or loss in the production of the part.

From Table V, the job allocation rate is the cost to apply to the job handled by the workstation. For instance, a station works on a single part for 10 minutes. If the job allocation rate is $15/minute, then the job will leave the station with $150 added to its cost. This provides a powerful way to allocate the cost of operating the station to each of the jobs which is handled during the process. The busy rate is the amount per minute which the job costs while it is busy. For workstation, this is the time spent performing machining operation. The idle rate is the amount per minute the workstation spent in nonproductive states. These two rates are used to compute the total cost for operating the station. The time spent in busy or idle states is multiplied by the correct rate, and added to obtain the total operating cost.

It may be noted that the results reported are for the illustrative FMS configuration only. However, the model developed is general enough to handle a case of m jobs and n machines with alternate routings, random machine failure, resource allocation, etc.

4. SIMULATION RESULTS AND DISCUSSION

For the given configuration and data, the FMS was simulated with different scheduling rules. Table VI summarises the results for all the rules.

Refer to Table VI, those values with (*) are the corresponding best performance measure. As a result, the best combined scheduling rules for each individual performance measure is summarised in Table VII. It can be observed that the SQL/SIPT combined rule outperforms all the other rules in obtaining the shortest makespan and the highest average machine utilisation. On the other hand, the SQL/LTPT is the best in achieving the maximum profit. However, the question raised here is: "Are they the best overall performers?". Some managers prefer to achieve a very short makespan of producing a batch, so that they could launch another new batch for other customers. In that case, they may decide to adopt the SQL/SIPT rule, but it has to be stressed here that the profit gain($60649) is the bottom fourth among the fifteen combined scheduling rules. On the other hand, if the managers prefer to receive a maximum profit for the batch, they may implement the SQL/LTPT rule. However, the corresponding makespan(5950 minutes) is also the bottom fourth among all the other rules, this may not be accepted by the production strategy of the company. This consideration leads to the introduction of Analytic Hierarchy Process(AHP) technique to compile the overall best performed rules(7).

The approach of AHP involves the structuring of any complex problem into different hierarchy levels with a view to accomplishing the stated objective of a problem. Using a method for scaling of weights of the elements in each of hierarchy levels with respect to an element of the next higher level, a matrix of pairwise comparisons of elements is constructed where the entries indicate the strengths with which one element dominates another. This scaling formulation can be translated into a largest eigenvalue problem which results in a normalised and unique vector of weights for the elements for each level of hierarchy. These weights are then used to determine the single composite vector of weights reflecting the relative importance among entities at the lowest levels of the hierarchy that enables the accomplishment of the objective of the problem(Figure 3).

AHP provides a comprehensive structure representing a combination of one's rational, irrational and intuitive judgments in the decision process. It is easy to apply, facilitates understanding of the problem, and provides functions to evaluate inconsistencies in the decision maker's judgments. However, development of AHP model is beyond the scope of the present study, hence it is suggested for further investigation.

5. CONCLUSION

For the given configuration and data, the FMS was simulated with fifteen different combined scheduling rules. From Table VII, it is clear that the combination of rules plays an important role while evaluating the system performance. Thus it must be emphasised that, in an FMS environment, a single rule aimed at a single objective is not desirable. Depending on the scheduler's objective (i.e. different people may have different perspective in assigning the weight on each evaluation criterion), performance of various combined rules may vary on attributes such as makespan, utilisation, profit, etc.

The work reported here is only a pilot study carried out on the FMS configured as a part of pedagogical material being developed to understand the strength of simulation as a design tool for FMS. The major emphasis is on presenting the details of the simulation model developed and the results obtained. The motivation for this approach stemmed from the fact that very few studies have been reported in this direction, in particular, different machine selection rules combined with various dispatching rules at workstation. Using the developed simulation model, a feasible combined scheduling rule can be found to give an optimum performance measure.

In the proposed AHP modelling, the weight assigned on each evaluation criterion is very subjective. It needs further investigation upon the significance of various weights on the performance of scheduling rules.

6. REFERENCES

1. Denzler, D. R. and Boe, W. J. "Experimental investigation of flexible manufacturing systems scheduling decision rules", *International Journal of Production Research*, Vol. 25, No. 7, 1987, pp. 979-94.
2. Gupta, Y. P. and Goyal, S. K. "Flexibility tradeoffs in a random flexible manufacturing system, a simulation study", *International Journal of Production Research*, Vol. 30, No. 3, 1992, pp. 525-57.
3. Montazeri, M. and Wassenhove, L. N. "Analysis of scheduling rules for an FMS", *International Journal of Production Research*, Vol. 28, No. 4, 1990, pp. 785-802.
4. Stecke, K. E. and Solberg, J. J. "Loading and control policies for a flexible manufacturing system", *International Journal of Production Research*, Vol. 19, No. 5, 1981, pp. 650-55.
5. O'Keefe, R. M. and Haddock, J. "Data-driven generic simulators for flexible manufacturing systems", *International Journal of Production Research*, Vol. 29, No. 9, 1991, pp. 1795-1810.
6. Law, A. M. and McComas, M. G. "How to select simulation software for manufacturing applications", *Industrial Engineering*, Vol. 7, No. 7, 1992, pp. 29-33.
7. Saaty, T. L. "How to make a decision: the analytic hierarchy process", *European Journal of Operational Research*, Vol. 48, No. 1, 1990, pp. 9-26.

Job Type	Operation 1	Operation 2	Operation 3	Operation 4	Operation 5
1	M4	M1/M2	M4	M1	M6
	(30)	(50)	(110)	(30)	(10)
2	M3	M1	M2	M3	M6
	(90)	(70)	(120)	(40)	(10)
3	M5	M4/M1	M5/M2	M4	M6
	(80)	(110)	(140)	(50)	(10)
4	M2	M3/M4	M2/M1	M5	M6
	(50)	(150)	(70)	(60)	(10)
5	M3	M1	M4	M2	M6
	(40)	(70)	(20)	(70)	(10)
6	M2/M5	M1/M3	M4/M5	M2	M6
	(20)	(50)	(240)	(80)	(10)
7	M3/M1	M4/M5	M3	M4	M6
	(60)	(60)	(70)	(90)	(10)

Note: M1/M2 indicates alternative routing of machine 1 or machine 2.
(x) indicates machining time is x minutes.

Table I. An operation sequence of the FMS: alternative routing is possible for some operations

Workstation	Machine Selection
FIFO	RAN
SIPT	SQL
LIPT	LULIB
STPT	
LTPT	

Key:

FIFO	:	First In First Out
SIPT	:	Shortest Imminent Processing Time
LIPT	:	Longest Imminent Processing Time
STPT	:	Shortest Total Processing Time
LTPT	:	Longest Total Processing Time
RAN	:	Random
SQL	:	Shortest Queue Length
LULIB		Lowest Utilisation of Local Input Buffer

Table II. Scheduling rules used in the FMS simulation model

Information	Input data
Capacity of local input buffer	10 jobs
Capacity of local output buffer	10 jobs
Machine setup time when perform different operation	10 minutes
Interarrival time of a batch	50 minutes
Batch size	7 jobs (job 1 to job 7)
Maximum number of batch created	10 batches

Table III. Data for the FMS simulation model

Job Type	Raw material/piece ($)	Carrying cost/minute ($)	Sale price/final product ($)
1	100	0.1	3,000
2	100	0.1	4,400
3	100	0.1	8,000
4	100	0.1	6,000
5	100	0.1	2,000
6	100	0.1	9,000
7	100	0.1	3,600

Table IV. Cost data for each job

Workstation Number	Job allocation rate ($/min)	Busy rate ($/min)	Idle rate ($/min)
1	10.	1.	0.1
2	15.	1.	0.1
3	12.	1.	0.1
4	5.	1.	0.1
5	10.	1.	0.1
6	5.	1.	0.1

Table V. Data for computing the total operating cost

Performance Measure	Dispatching rule at workstation	Machine Selection Rule RAN	SQL	LULIB
Makespan (min.)	FIFO	5280.	5360.	5260.
	SIPT	5640.	*** 5200.**	5670.
	LIPT	5440.	5580.	5470.
	STPT	5670.	5460.	6080.
	LTPT	6240.	5950.	5990.
Average lead time (min.)	FIFO	3548.	3521.	3566.
	SIPT	2889.	2807.	2825.
	LIPT	2914.	3063.	2931.
	STPT	2694.	*** 2586.**	2683.
	LTPT	3310.	3318.	3332.
Average machine utilisation (%)	FIFO	80.32	79.1	80.6
	SIPT	75.18	*** 81.54**	74.76
	LIPT	77.96	76.	77.52
	STPT	74.78	77.66	69.74
	LTPT	67.94	70.84	70.76
Net profit ($)	FIFO	51406.	56700.	56277.
	SIPT	71248.	60649.	63521.
	LIPT	77450.	73265.	77947.
	STPT	81379.	79775.	80590.
	LTPT	76662.	*** 82428.**	73782.
Average delay at local input buffer (min.)	FIFO	620.95	608.16	621.6
	SIPT	478.60	482.85	499.3
	LIPT	499.16	526.8	*** 409.4**
	STPT	499.72	459.6	504.2
	LTPT	536.58	598.9	564.5
Average WIP at local input buffer (no. of part)	FIFO	6.76	6.54	6.78
	SIPT	4.84	5.28	5.02
	LIPT	5.24	5.38	*** 4.34**
	STPT	4.98	4.78	4.62
	LTPT	4.94	5.62	5.38

Table VI. A summary of the simulation results

Performance Measure	Combination of scheduling rules *Machine Selection Rule/Dispatching rule at workstation*
Makespan	SQL/SIPT
Average machine utilisation	SQL/SIPT
Average lead time	SQL/STPT
Net profit	SQL/LTPT
Average delay at local input buffer	LULIB/LIPT
Average WIP at local input buffer	LULIB/LIPT

Table VII. Summary of the best combined scheduling rule for each individual performance measure

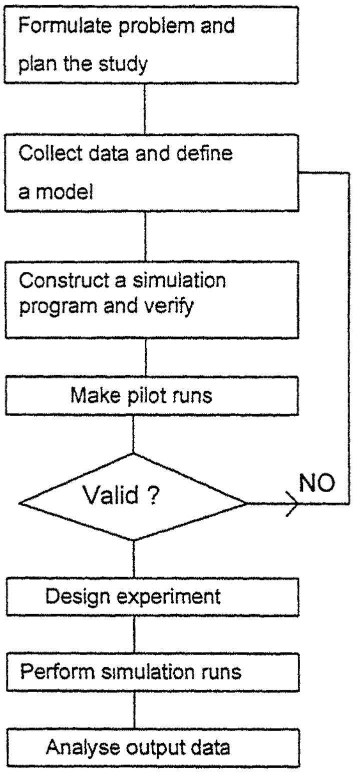


Figure 2 Simulation study approach

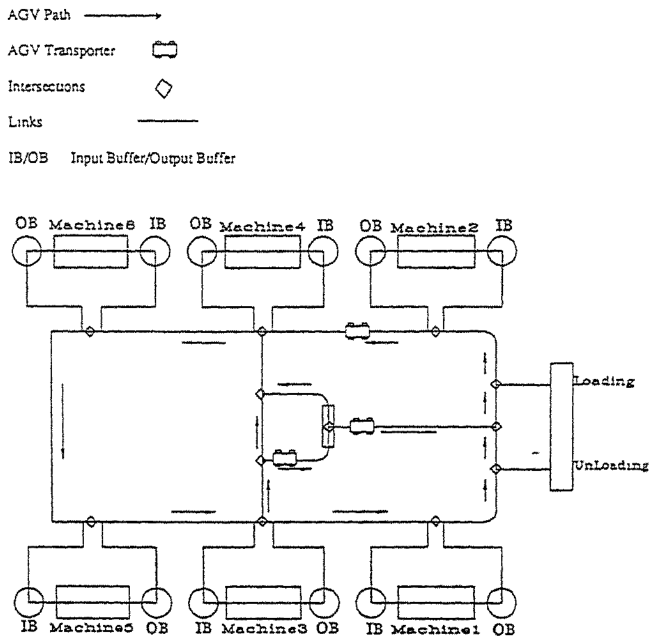


Figure 1 Configuration of the flexible manufacturing system

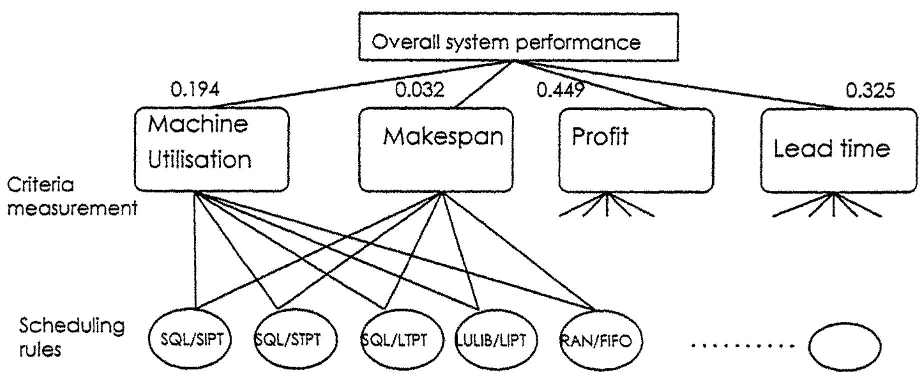


Figure 3 AHP model with different scheduling rules applied to FMS

SHOP-FLOOR SCHEDULING AND DATA CAPTURING SYSTEM FOR AUTONOMOUS MANUFACTURING ISLAND

F.W. Lam, Y.J. Ma, W.B. Lee, and S. Zhang*

Department of Manufacturing Engineering, The Hong Kong Polytechnic University, Hong Kong
*Department of Mechanical Engineering, Tongi University, Shanghai, P. R. China

ABSTRACT

In this paper, a form of decentralized manufacturing system known as Autonomous Manufacturing Island (AMI) is discussed. An AMI is a highly autonomous work unit within an enterprise which has its own production planning and control intelligence made possible by the automation of manufacturing information flow through out the factory. The Shop-floor Scheduling and Production Control System and Shop-floor Data Capturing System are the main supporting modules for the implementation of AMI. The above modules serve as an important link between the shop floor and the office and business system.

KEYWORDS

Autonomous Manufacturing Island, Shop-Floor Scheduling and Production Control system, Shop-Floor Data Capturing system, Information Flow.

1. INTRODUCTION

Technology innovation always is a key point of factory development [1]. It does not necessitate mean highly sophisticated automated production lines, robots and large-scale computers. The appropriate adoption of automation is an important solution as automation is the tool of controlling and improving the manufacturing process but not the end-goals. It is widely recognized that for a manufacturing system to be adaptive to rapid external changes, its sub-systems or units should have a high degree of self-autonomy (i.e. with self-governing and control functions) rather than a high degree of automation designed according to some master plans. One of the examples of such automation is the concept of Autonomous Manufacturing Island (AMI) proposed by Zhang [2]. It is a decentralized manufacturing system with autonomy. AMI consists of certain CNC machines, a team of skillful workers and a workshop-oriented computer networking, and acts as a small production plant in a large manufacturing organization. In the following section, the concept of AMI is explained and the supporting systems for implementing AMI are discussed.

2. THE CONCEPT OF AUTONOMOUS MANUFACTURING ISLAND

AMI is a decentralized manufacturing system with a high degree of autonomy. It consists of a mix of CNC machines and conventional machines, a team of skillful workers and a workshop-oriented computer networking, as a small division in a factory. The concept and characteristics of AMI are shown in Figure 1.

In AMI, by means of personal computer and network [3], all of the manufacturing processes, production preparations work, work organizations are integrated in one system [4]. The boundaries between different jobs and departments, the technical and the clerical staff are removed. The organization thus formed can be called task-specific.

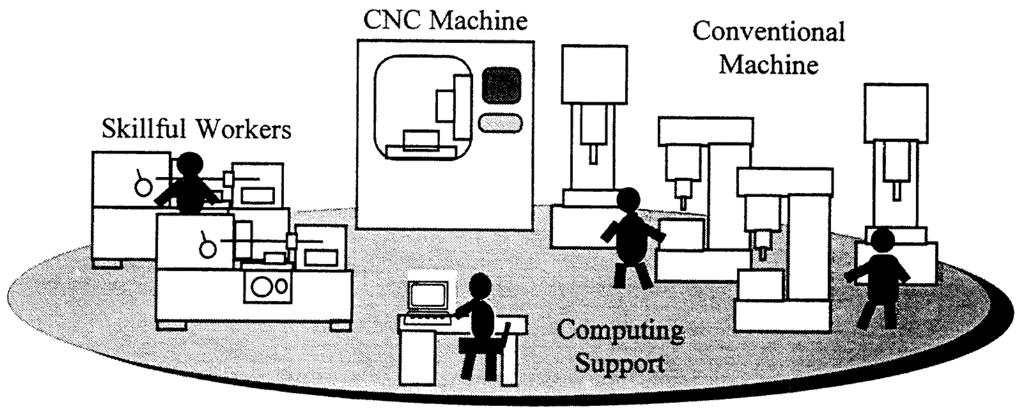


Figure 1: The concept of AMI

The traditional function-oriented organization of enterprise leads to an increasing complexity combined with ineffective utilization of resources and barrier of information flow. AMI uses the know-how of all the staff to create process-oriented production design. Transparent objectives are the basis. These are the pre-conditions for concentrating the enterprise resources on the value-added processes. The structures based on AMI, as shown in Figure 2, with low complexity would allow optimal usage of resources and quicker response to the market demands.

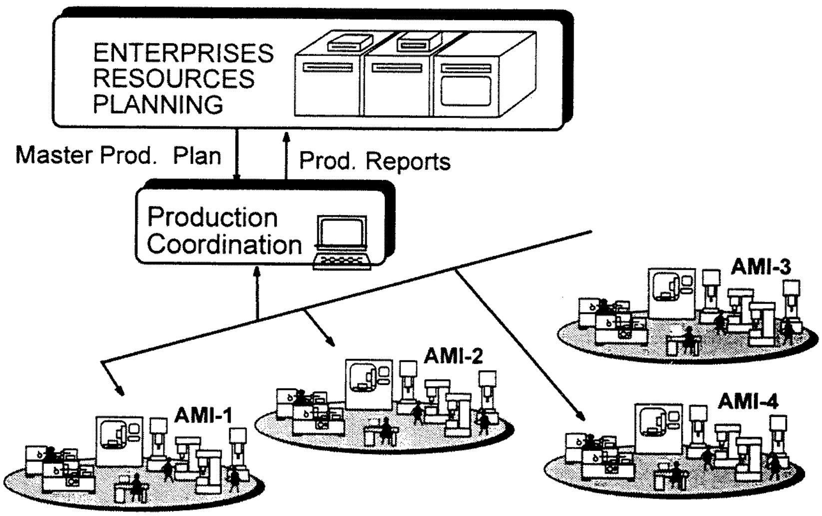


Figure 2: Decentralized production system based on AMI

In order to achieve a high level of planning reliability and responsiveness, especially in case of disturbances, an improved structuring of planning processes is necessary.

3. ESSENTIALS OF SHOP-FLOOR SCHEDULING AND DATA CAPTURING SYSTEM

Conventional production scheduling is characterized by early fixing of planning data and rigid passing through of planning results within the processing chain. Central planning does not provide any feedback of machine-specific data or fine "in-process" planning to meet changing situations [5, 6]. AMI requires planning and control of processing tasks and available capacities. Based on each specific task, the most suitable AMI with available capacity and delivery is chosen for the job. Within AMI the principle of closed-loop control enables an event-oriented planning closely related to the shop floor, as shown in Figure 3. The AMI has its own job scheduling and control system, production data capturing devices and simulation tools to form a closed-loop control circuit in the enterprise planning and control.

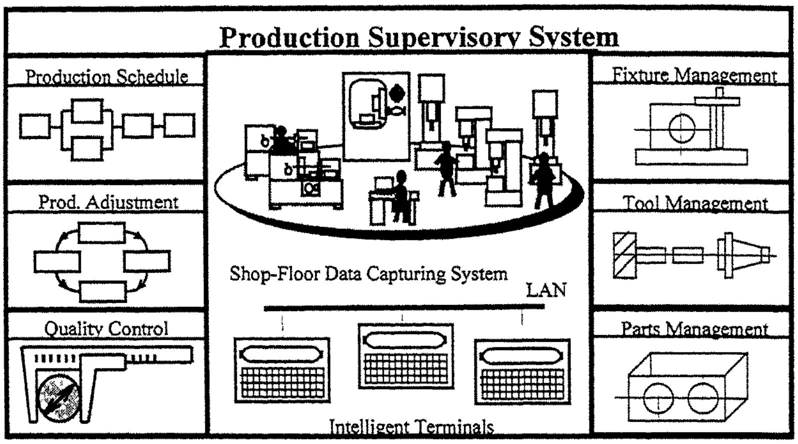


Figure 3: Closed-loop Control System for Autonomous Manufacturing Island

The ability to automate manufacturing information flow forms the basic backbone for the formation of AMI [7, 8]. The architecture and the information flow of the Shop-Floor Scheduling and Production Control System and the Shop-Floor Data Capturing System are shown in Figure 4.

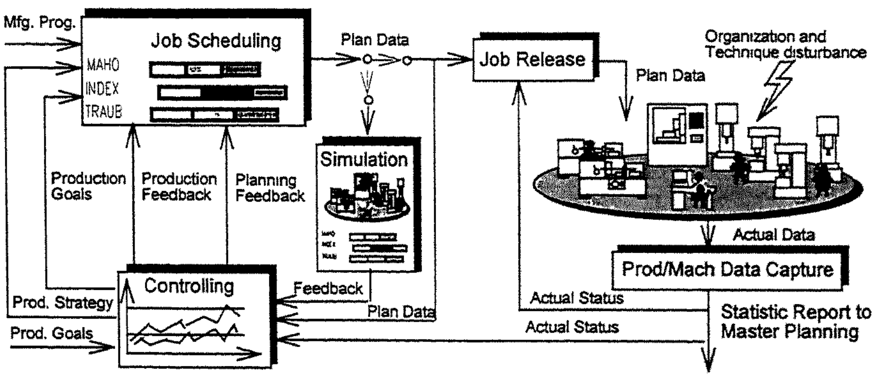


Figure 4: Planning and control function of AMI

3.1 Shop-Floor Scheduling and Production Control System

The followings are the main objectives of Shop-Floor Scheduling and Production Control System for AMI:

- To ensure the products will be manufactured with pre-defined quality and delivered to the customers within the specified time frame.
- To improve the flexibility of rapid production.
- To shorten the production lead time.
- To reduce the stock holding level.
- To improve the productivity of major production equipment.

In order to fulfil the above mentioned objectives, the in-house developed Shop-Floor Scheduling and Production Control System includes the following features:

a) Data Management
   - Working calendar: To identify the working days, working shifts and the holidays of the production plant.
   - Resources information: To process the relevant data about the capacities of the production equipment, tooling information and different types of raw materials.
   - Staff information: To manage the information about the staff working in the production plant especially their level of skills.
   - Parts information: To update the records of the parts database; the information can also be converted from the output of Computer-Aided Process Planning (CAPP) which is another in-house developed module for AMI.

b) Production Scheduling
   - Job orders: To generate a complete list of job orders with detail requirements for the period concerned.
   - Machine allocation: To allocate the various job orders to the appropriate machines automatically by the system with the considerations of process plans, machine loading and the availability of various production resources. Different results can be simulated graphically and the allocation can be adjusted manually if necessary.
   - Production sequencing: The best result will be chosen based on the delivery date and the job numbers for individual machine will be assigned automatically.
   - Production schedules: According to the result of production sequencing, a master production schedule for the period concerned will be compiled and stored in the central database.

c) Production Control
   - Tuning of production schedule: The daily production schedule can be fine-tuned and revised if necessary.
   - Processing of job tasks: To issue the new job tasks to the appropriate Intelligent Terminals according to the production schedule and update the task lists upon the completion of the job.
   - Production statistics and report: To generate the statistics and compile a report about the number of finished goods, work-in-progress, defectives, and other production information such as completion times, defective codes, machine breakdown, etc. All the data are collected by the Shop-Floor Data Capturing System.

d) Financial Auditing
   - Salary control and report
   - Prime cost control and report

3.2 Shop-Floor Data Capturing System

The Shop-Floor Data Capturing System mainly describes the following situations:

a) Job-specific Data
- Production capabilities
- Machine loading
- Machine status

b) Production-specific Data
- Defective
- Machine fault
- Machine waiting due to shortage of material

c) Machine-specific Data
- Machine status
- Loading/unloading
- Interruption

For the monitoring of every CNC machine an Intelligent Terminal is required. Based on the above mentioned criteria, the following sub-modules are developed for data capturing purposes:

a) Current job
- Job content: The operator can retrieve information about the current job.
- Starting time: The exact time has to be recorded and added to the database.

b) Input by barcode reading
- Barcode of job card: To identify whether the content of the current job card matches the master production schedule.
- Barcode of staff ID: To check whether the worker is operating the right machine and record down any changes.

c) Part program management
- Download of part programs: To download the previous prepared part programs from the Computer-Aided Manufacture (CAM) System to the Intelligent Terminal.
- Modification of part programs: To modify the part programs if necessary.
- Transfer of part programs: To transfer the part programs from the Intelligent Terminal to the CNC machine.

d) Production report
- Finished components: To report the total number of QC passed components.
- Defectives: To report the number of defective products and the corresponding defect codes.
- Scrap materials: To report the amount of materials being scraped.
- Returned materials: To report the amount of materials not being used and returned to the store.
- Completion time: To report the total time spent for the current production.

e) Machine report
- Machine waiting: To report the time wasted due to the shortage of materials, toolings, fixtures, and other faults, etc.
- Machine fault: To report the down time and the corresponding reasons.

f) Enquiry
- Machine status: To check whether the machine is operating, standing-by or under maintenance, etc.
- Production status: To check the number of finished components, defectives, etc.

4. CONCLUSIONS

The Shop-Floor Scheduling and Production Control System and the Shop-Floor Data Capturing System are playing an important role and tightly integrated with CAD, CAM and CAPP systems in AMI. They provide the important links between the business planning, management systems and the individual machine control systems. The modular approach makes it easier to link the shop floor to the management system through modern networking technology. The aim of AMI is to optimize the utilization of production resources. Its starting point is a clear perception that a new work organization is needed to optimize the utilization of resources through the automation of information flow, reformation of work organization on the shop floor and improvement of personal qualification in continuation. Implementation of AMI could realize with less investment, but gain remarkable benefits.

REFERENCES

1. National Research Council, "Information Technology for Manufacturing – A Research Agenda", National Academy Press, 1995.

2. S. Zhang, "Advanced Manufacturing Technology", Shanghai Mechanical Engineering Society, March 1993.

3. J. R. Pimentel, Communication Networks for Manufacturing, Prentice Hall, 1990.

4. A. Kusiak, Intelligent Manufacturing Systems, Prentice Hall, 1990.

5. J. Rickel, "Issues in the Design of Scheduling Systems", Expert Systems and Intelligent Manufacturing, Elsevier Sciences, New York, 1988.

6. S. Becker and R. Parr, "Scheduling Manufacturing Systems", Handbook of Design, Manufacturing and Automation, Wiley-Interscience, 1994.

7. A. Bauer, R. Bowden, J. Browne, J. Duggan and G. Lyons. Shop Floor Control System, London: Chapman & Hall, 1991.

8. S. A. Spiewak, "Automatic Supervision of Machine Tools", Automatic Supervision in Manufacturing, Springer-Verlag, 1994.

RAPID PROTOTYPING OF GOLF CLUBS

R Jones and S R Mitchell
Department of Manufacturing Engineering
Loughborough University
United Kingdom

ABSTRACT

The design and manufacture of golf equipment is a major and extremely competitive industry with club head design following traditional methods for the design of sculptured products in that the hand crafting of prototypes is a significant part of the design process. Manufacturers are taking advantage of modern technology by using computer aided design (CAD) methods, but they still require prototype models for design appreciation and play testing. Iron clubs are generally solid and prototypes can be CNC machined. However, modern 'woods' are generally made of metal and are hollow. It is virtually impossible to machine these shapes and tooling must be produced to enable wax production for investment casting. Rapid prototyping methods enable waxes to be produced directly giving significant savings in tooling and time. This paper considers the factors involved in the design and prototyping of club heads and their effect on the use of feature based CAD methods for modelling. Information from CAD models has been used to manufacture club heads, by CNC machining, stereolithography, laminated object manufacture, fused deposition modelling and selective laser sintering.

KEYWORDS

Sculptured Surfaces, Feature Based Design, Golf Clubs, Rapid Prototyping

1. INTRODUCTION

The demand for golf continues to increase and the business is estimated to be worth around $80 million per year. As with other sports enthusiasts, golfers are continually looking for enhanced performance from their equipment. To a certain extent the Rules of Golf (Ref. 1) are restrictive, but manufacturers continually attempt to develop new aesthetically and functionally enhanced designs, taking advantage of new manufacturing techniques and materials. Since the golf industry is fashion conscious manufacturers must get their products and innovations to market quickly if they are to maintain or expand their market share. At present over 80 per cent of the world's golf club heads are made in South East Asia (Taiwan in particular). This makes prototyping a particularly lengthy operation for Western countries because of the communication difficulties involved. Without using CAD/CAM tools a new design concept could take up to 2 years to come to the market

Traditionally heads have been hand crafted by clubmakers and professional golfers. Few, if any, drawings were used so that the prototype club head model effectively embodied the design specification. However, in recent years computer based methods have been introduced to reduce lead times and increase design flexibility (Ref. 2 and 3).The design process requires two forms of prototype:

(i) for design shape appreciation;
(ii) for practical play test evaluation.

For design appreciation the prototypes need not be produced in metal, and resin or wax is frequently used instead. The most important requirement is that the prototype is available quickly to be handled by the designers. The play test prototype, however, needs to be produced in metal and have equivalent characteristics to a production model. At present these models take considerable time to produce and this study has been conducted to evaluate the suitability of rapid prototyping methods for producing golf club head prototypes.

2. CLUB HEAD DESIGN CRITERIA

The rules (Ref. 1) governing club head shape are generally concerned with maintaining a traditional appearance. Club heads tend to have two essential features, an angled (lofted) face to project the ball into the air and a cylindrical hosel for attaching the shaft. Loft angle, head weight, shaft length and shaft 'lie' angle, vary through a set of clubs to give different trajectories. Clubs are mostly sold in matched sets, based on the principle that each club should 'feel' the same so that a player uses the same swing with any club in their set. The clubs are usually matched by a process known as swing weighting (i.e. combining shaft length and head mass to match a predetermined static lateral moment). The validity of this technique is often questioned, but the result is that club heads need to be manufactured to a weight tolerance of ± 2 grammes to achieve an acceptable swing weight tolerance on headweights generally between 170 and 300 grammes, depending on the club.

Although the position of the centre of gravity and mass distribution is important the head shape accuracy is not as significant. Essentially, the club must only look like the original prototype, and have the same inertia characteristics. Engraved (and sometimes embossed) logos are also important design features. Aesthetically they help establish brand identity, and functionally they have a significant effect on the head inertia characteristics, often subtracting as much as 4 grammes from the total head mass. They may be placed anywhere on the head, apart from the hitting face area where there are restrictions. Figure 1 shows a graph of logos positions on iron club heads and their placement frequency. The allowable face markings or grooves have an additional affect on ball spin after the club impact and are strictly controlled in terms of both size and shape. To achieve the production accuracy required to meet the regulations the face markings are normally machined directly on the wax tooling masters.

Metal wood heads are hollow since they must have an equivalent mass to their wooden predecessors to achieve the same swing weight. Investment cast heads for woods were first developed in 1976, based on a two piece construction which allows the core to be removed from both the investment wax and the cast head. The sole plate is subsequently welded in position and ground to give an acceptable join. Typical wall thicknesses are from 2.5 mm on the face and less than 1.0 mm on the 'crown' for a 'driver'. Figure 2 shows a typical thickness variation on the club crown or upper surface. Over recent years there has been a tendency to increase head sizes to give increased face area and moment of inertia. Driver head sizes have increased from 175 cc to 280 cc. while maintaining the same head mass (around 200 grammes for a driver). To achieve this some section thicknesses are now extremely thin (as little as 0.5 mm on the crown). Vacuum casting methods need to be employed for these section thicknesses and high strength head materials such as titanium are also being used.

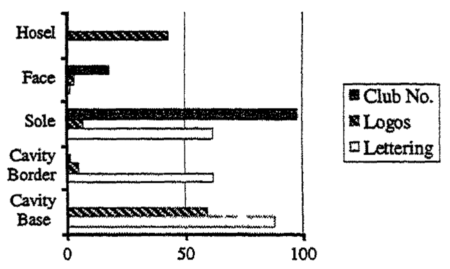


Figure 1. Position and Frequency of Logos on Iron Heads

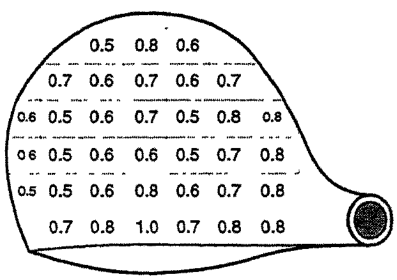


Figure 2. Thickness Variation an Example Wood Club Crown

3. CAD DESIGN OF CLUB HEADS

To adequately model the predominantly three dimensional sculpted shape of wood or iron club heads it is necessary to employ 3D surface modelling software. Several commercial CAD packages offer suitable functionality, and the system used for this work was DUCT, Delcam International Ltd.'s surface model based CAD/CAM software. To increase design efficiency a feature based approach to design has been developed (Ref. 4). The club head is modelled using a set of extended surfaces, which are subsequently blended and trimmed. The triangular mesh output files from DUCT were post-processed into valid STL files using Delcam's Trifix utility. The result is a well defined three dimensional computer model. Two club head models were created for this exercise:

- A solid cavity back 5-iron having a loft angle of 31 degrees, a lie angle of 60 degrees, and a volume of 31002 cubic mm.
- A hollow wood (driver) having a loft angle of 10.5 degrees, a lie angle of 56 degrees and an external volume of 180,000 cubic mm.

Examples STL meshes for the two clubs are shown in Figures 3 and 4 (although for illustration purposes these are coarser than the actual meshes used). The hollow wood had a circular aperture of 30 mm diameter cut into the base to mimic the sole plate. The section thickness of the wood varied between 2.5 mm at the centre of the face to 1.1 mm at the top of the crown. It should be noted that the sizes of these club heads had not been optimised, they were simply created as typical examples. Even though the logos and face markings are considered significant contributors to the club head design, they were omitted from the head prototyping trials to separate the issues of general shape from ornamental detail reproduction. Manufacturer's trade marks and player's signatures are frequently used, and they need to be modelled in fine detail. The logos can be modelled using Delcam's ArtCAM software and subsequently wrapped onto the club surfaces. The simple face markings can be modelled using DUCT and again wrapped onto the club face. However, to model the logos on a typical iron club head, even at a substandard resolution, requires 30,000 facets to be added to the original 20,000 facet model resulting in an STL file size increase from 7 to 13 mb.

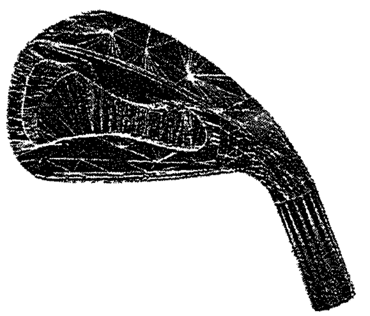


Figure 3. STL Mesh for an Iron Head

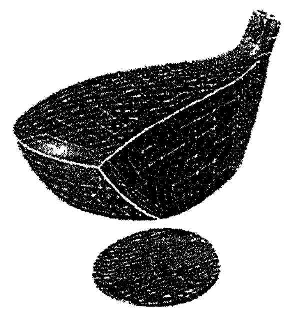


Figure 4. STL Mesh for a Hollow Wood

In order to assess the suitability of RP methods for the aesthetic and functional requirements of logos and face markings, several separate test models were developed. Three examples are discussed in this paper :

(i) The Slazenger Panther trade mark.
(ii) The Seve Ballesteros signature.
(iii) A set of grooved face markings.

Two processes, stereolithography and wax jet printing were selected for manufacturing logos. The section thicknesses for this work are extremely thin (0.5 mm) and it was considered that the other processes would have difficulty in achieving this level of detail. In each case the test model is based on a shell with 5 external surfaces having curvature values similar to a golf wood head. Each logo was wrapped on to each of the 5 surfaces at various orientations to demonstrate the effect of layer build direction on accuracy. Engraved and embossed versions of both logos (to a depth and height of 1 mm respectively) were modelled side by side on each surface to highlight any differences between the two. Figure 5 shows results for the Slazenger Panther logo.

The STL file sizes for the models were as follows:

Cavity Back Club	7.0	mbyte
Hollow Wood	13.0	mbyte
Slazenger Logo	7.5	mbyte
Balesteros Signature	10.0	mbyte
Face Grooves	2.5	mbyte

File transfer to the RP subcontractors was achieved by Internet file transfer, DAT tape cartridge, or segmentation and compression for multiple 3.5" HD floppy disks.

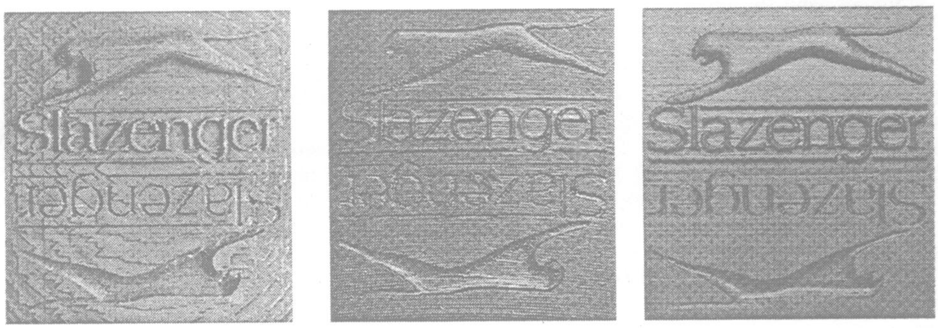


Figure 5. Slazenger Logo Results for Different Build Orientations (90°, 45° and 0°)

4. RAPID PROTOTYPE MANUFACTURE

The majority of prototype golf club heads are produced by milling and with CAD designs becoming available CNC machining is increasingly used. Evaluating sculptured part accuracy produced by different processes is often difficult because of the problems in establishing suitable measurement datums. For this work the benchmark process is CNC milling. The external surfaces of the iron and wood present little difficulty, but it is virtually impossible to completely machine a hollow wood. If wax prototype models are required for investment casting, then wax injection tooling must be manufactured at an additional cost.

The RP processes considered for evaluation are shown in Table 1.

	Solid 5-Iron	Hollow Wood	Logos etc
CNC Machining			X
Stereolithography (STL)			
Laminated Object Manufacturer (LOM)		X	X
Selective Laser Sintering (SLS)		X	X
Fused Deposition Modelling (FDM)			X
Wax Jet Printing	X	X	

Table 1 Rapid Prototyping Process Evaluation Trials

The factors used for evaluation were:

- cost
- lead time
- manufacturing time
- section thickness
- surface finish
- weight
- lie and loft angles
- offset
- hosel diameter

Table 2 gives details of the values of principal club characteristics for the 5-iron club by the four RP processes. The LOM and STL values were obtained from 6 samples and the SLS and FDM from 1 sample.

The surfaces of the models produced by LOM were sealed to enable them to be used as "waxes" for investment casting. This was achieved using 4 casts of epoxy paint spray with the models sanded down between coats. The mass value of the model increased to an average of 34.3 grammes. and the hosel diameter increased to 13.9 mm. The iron models were built in various orientations to give minimum build time (hosel horizontal) and also to give desirable surface finish (face horizontal and hosel vertical).

	Loft (°)	**Lie** (°)	**Offset** (mm)	**Weight** (g)	**Hosel Diameter** (mm)
LOM (average)	31.0	60.0	4.6	31.7	13.1
(range)	(1.5)	0	(1.0)	(2.8)	(0.3)
STL (average)	30.5	60.0	3.8	16.6	14.0
(range)	(1.5)	(1.0)	(1.2)	(0.9)	(0.6)
SLS 31.0	60.0	5.2		35.1	13.3
FDM 31.0	59.0	4.8		36.7	14.0

Table 2. Measured Club Specification Parameters (Irons)

The wood models were produced by STL and FDM methods with the club in the address position and with the hosel vertical. The significant feature of the club manufactured in the address position was the poor surface finish produced on the crown of the club, which is the most visible part of the club. The hosel vertical orientation produced a more satisfactory result, although a poor finish was achieved on the underside of the toe, which is in a more acceptable region. The thinnest part of the club is at the centre of the crown and the approximate designed value was 1.1 mm. The centre of the face is the thickest section was designed at 2.5 mm. Table 3 gives measured values for the RP woods.

	Crown Thickness (mm)	**Face Thickness** (mm)	**Hosel Diameter** (mm)
SLS (address position)	1.38	2.84	11.55/11.80
SLS (hosel vertical)	1.60	3.01	11.50/11.85
FDM (hosel vertical)	0.85	3.12	11.55/11.78

Table 3. Measured Club Specification Parameters (Woods)

The thin section on the crown of the club caused some processing problems for the STL method. The surface skin thickness was supposed to be 0.3 mm but during processing the liquid polymer was trapped in the honeycomb and could not be drained. This part of the model is therefore solid following the U.V. hardening process.

5. DISCUSSION

The models were produced by various universities and companies on a grace and favour arrangement with charges only being levied for consumables. The work therefore was integrated into the normal work programme when possible. It is difficult to be specific about lead times and processing times since if commercial rates were charged priorities would change.

However, the shortest machine build time quoted for the iron club was for the LOM models, around 4.5 hours (the longest was 7 hours). This does not compare favourably with CNC machining of the same model in resin where times of less than 60 min have been achieved. Removal from the blank and hand finishing are still required, but times for this process compare with RP finishing.

The time to produce the hollow wood model was significantly longer than the iron with 12 hours being the shortest for the FDM process. Rapid prototyping of this shape shows some promise since the metal model can only be made by casting into moulds. It was envisaged that there would be difficulties in achieving the crown thickness, however, this may prove acceptable down to thicknesses of 1.0 mm. The significant differentiator is cost and RP methods will have to compete against cheap wax injection tooling from South East Asia. A tool set can be produced for as little as £1000 including engraving. This price is still difficult for RP manufacturers to compete with, but the advantage of reduced lead time may offset cost factors.

Generally the shape and dimensions of the models were considered acceptable. There was some concern regarding the weight variation for the iron models. An 8 per cent weight variation on a cast steel head would result in 16 gms, which is in excess of the required tolerance. Further work needs to be undertaken to use the models in the investment casting process to determine the effects of the weight variation noted.

Because of the fine detail required for logos and text the build direction was of considerable significance. Surfaces with small rates of curvature gave poor results since the slice height can be a significant proportion of the logo height. Text or a signature will have fine details at various orientations and elements running parallel to the slice can easily be lost. Wax jet printed examples showed a considerable improvement for both raised and indented parts. The build direction is of less importance, however, the process time increases considerably.

6. ACKNOWLEDGEMENTS

The authors would like to thank the EPSRC, Dunlop Slazenger International Limited and Delcam International Limited for their support, and also the companies and universities who processed the models in the interests of this research.

7. REFERENCES

1. The Rules of Golf, The Royal and Ancient Golf Club of St Andrews, 1992.

2. Jones R, Computer Based Methods for the Design and Manufacture of Golf Clubs, Science and Golf, (Proc. of 1st World Scientific Congress of Golf), London, pp 280-285, E&F.N. Spon., 1990.

3. Mitchell S R, Newman S T, Hinde C J and Jones R, A Design System for Iron Golf Clubs, Science and Golf 2 (Proc. of 2nd World Scientific Congress of Golf), London, pp 390-395, E&F.N. Spon., 1994.

4. Jones R, Mitchell S R and Newman S T, Feature Based Design Systems for the Design and Manufacture of Sculptured Products, Int. J. Prod. Res., Vol. 31, No. 6, pp 1441-1452, 1993.

AN APPROACH ON DECISION SUPPORT FOR SELECTION OF PROCESS CHAINS OF THE TOOL AND DIE MAKING INDUSTRY UNDER SPECIAL CONSIDERATION OF RAPID PROTOTYPING TECHNOLOGIES

H. Dürr, U. Kaschka
Department of Mechanical Engineering and Process Technology
Technical University of Chemnitz-Zwickau
Germany

ABSTRACT

The paper describes a methodology for the decision support to select and compare alternative process chains of tool and die making industry. The presented methodology allows the comparison according to criteria such as quality, costs and time and supports the selection of appropriate technological strategies.

First of all it is necessary to create a comprehensive knowledge base with technologies and products of tool and die making industry, focussing especially on processes of Rapid Prototyping. The implementation of these data into the company's data structure is enabled by a STEP-oriented process chain model. Suitable process chains in consideration of technical and economic aspects can be assigned to each product. Furthermore it is possible to plan the technological strategy for a single part or part family.

Using industrial representatives the methodology of decision support is going to be tested and programmed for a computer-based decision support and offer calculation system.

KEYWORDS

Rapid Prototyping, Tool and Die Making Industry, Decision Support

1. INTRODUCTION

The current situation in tool and die making industry is characterized by an increased number of variants, shorter cycles of product innovation and within minimum limits of time. To secure market shares, it is vital to reduce the time from the initial product idea to the launch of the product. Furthermore, flexibility and keeping the terms of delivery are important in order to stay competitive.

Based on these demands, there can be observed a trend towards application of innovative manufacturing methods. In this case, technologies such as Rapid Prototyping, High Speed Cutting (HSC) and 5-axes-milling are mostly emphasized. Particularly the technologies of Rapid Prototyping (also named as "generative manufacturing processes" or "layer manufacturing") contain a high potential to reduce the entire cycle of product development by building models and prototypes within shorter periods as well as reduction of costs of product's modification by using physical models in the early stages of design. Additionally, the so called "rapid tooling"-technologies are the main topic of many research works. These technologies should be enabled the direct production of tools and dies in the series material by means of layer manufacturing methods in the future.

Contrarily to this development potential, user often have insufficient knowledge regarding opportunities and limits of innovative technologies. Permanent new developments and different degrees of industrial maturity of offered technologies complicate decision making. Suitable practical tools for a direct comparison with conventional manufacturing processes are missed. Producers of machines and service bureaus are mostly unable to provide an objective overview, because they will always tend to sell "their own" solution.

This topic has already been dealt with a number of research activities /1-6/. These concentrate mainly on evaluation of single manufacturing techniques without investigating the whole process chain from CAD data preparation up to a possible post processing. However, time and cost savings from single

building procedure can be lost again by ignoring the usage conditions of preceding and succeeding process-steps especially in the field of Rapid Prototyping technologies.
In other research approaches, the conventional methods (e.g. the NC-milling) being still relevant for the production of prototypes, tools and dies are neglected.
In general, there is a demand for decision support systems for assistance in complex decision problems in design and process planning focussing on tool and die making industry, which can be integrated into the existing electronic-data-processing environment of a company without any problems.

For that reason the most important objective was to develop methods for decision support in the selection of process chains of the tool and die making industry and to offer practical solutions for all potential users.

2. PROCESS CHAINS IN THE TOOL AND DIE MAKING INDUSTRY

Recent development in the field of generative technologies and HSC extends the opportunities of producing models, tools and dies. Figure 1 provides an overview of the most important process chains.

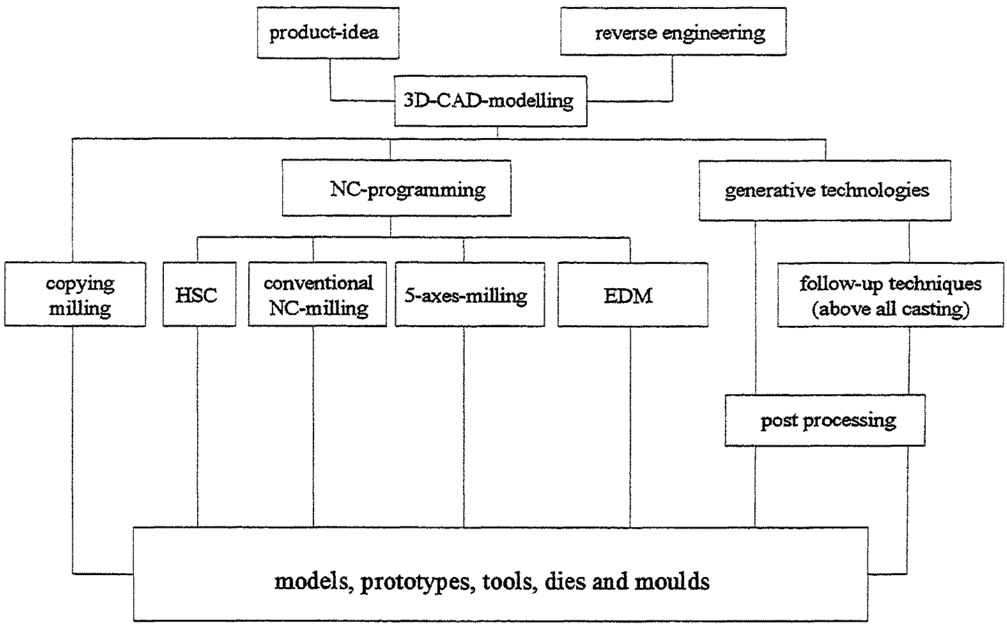


Fig. 1 : Process chains in the tool and die making industry

The selection of the most appropriate process chain for a specific part must include the different time, cost and quality priorities of the enterprise. Additionally there must be considered aspects of business management (operational organisation), manufacturing technology (available machines and devices) as well as information processing (available CAD- and other software systems). In order to achieve an economic result, advantages and disadvantages of techniques and process chains have to be evaluated carefully. Growing information diversity, increasing complexity of demands and partly contradictory objectives require systematic instruments for decision support.

3. DECISION SUPPORT METHODOLOGY

First of all there is a most general question: To which extent is the methodological support in decision problems appropriate from the viewpoint of the user? In this connection it must be taken into account that the application of complex evaluation systems stands for an important cost factor as well.
In many cases, particularly in small and medium-sized enterprises the decision is based on intuition of experienced experts. This method incorporates risks in more and more complex situations and increasing large data stocks as a result of introduction of new technologies.
On the other side there are strictly determining methods of selection and evaluation. Their results are always definite rankings of the evaluated alternatives. This fact, however, leads to problems of acceptance for the user, who had been given solutions whose generation cannot be reproduced sufficiently.
In consideration of all these facts a kind of decision support seems to be useful which offers methods suitable to the problems, on the one hand, and provides the responsible person a certain range of decision making, on the other hand.

A two-step approach combining selection of technological strategy and evaluation of process chains by cost - benefit - analysis are fulfilled these requirements. Figure 2 shows a classification of both methods. Figure 3 illustrates the integration within the flow of business information.

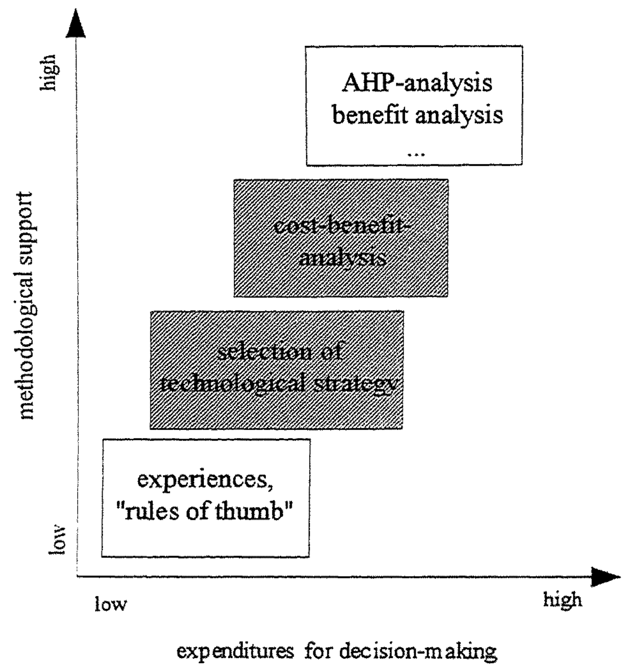


Fig. 2 : Classification of used methods

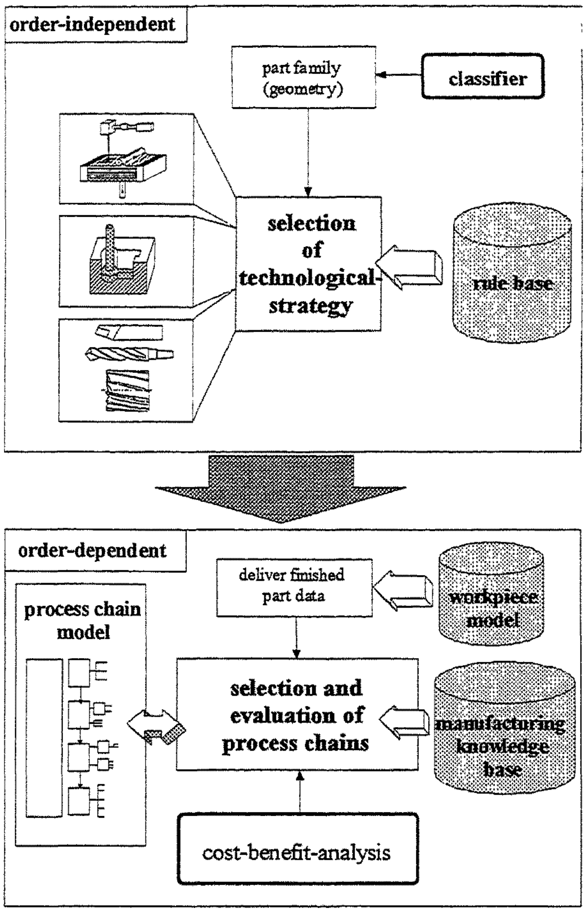


Fig. 3 : Decision support inside the business data flow

3.1 First Step: Selection of Technological Strategy

The selection of technological strategy is a method for rough planning. Technologies and process chains to be used are planned quickly and easily.

Based on the results of analyses and an intensive evaluation of process chains there has been developed a set of rules which allows to allocate technological strategies to single parts or a number of parts (part families). Within the project, technological strategies are understood as the classification of manufacturing methods according table 1.

Technological strategy	Examples of manufacturing methods
conventional	NC-milling, EDM, copying milling
unconventional	High-speed cutting (HSC)
generative	Rapid Prototyping technologies such as SLS, SLA, FDM, LOM
combined methods	NC-milling and laser generating, NC-milling and eroding

Tab. 1 : Technological strategies

To develop a set of rules the criteria of significant severity must be found out and quantified For tool and die making industry this means particularly of the attribute "geometrical complexity". For instance, the more complex the part geometry, the higher is suitability of generative manufacturing processes to produce especially models and prototypes.
All of components can be subdivided into classes of complexity by means of a classifier The existence of complicated form elements (e.g. ribs, cantilever) or free-formed surfaces (e g undercuts, cavities) is decisive Afterwards the stored rules allow the assignment of adequate technological strategies for each class of complexity

An order-depending decision making for really applicable manufacturing sequences is based on process chains selected and evaluated before (see step two of the introduced approach)

3.2 Second Step: Selection and Evaluation of Process Chains by Cost-Benefit-Analysis

Evaluation is characterized by the tasks following:

- to achieve a number of main objectives such as low costs, high quality and time shortening which can be partly contradictory,
- to process criteria that can be quantified (e.g. costs) as well as criteria not to be quantified (e.g flexibility) and
- to consider criteria whose benefit is difficult to be quantified (benefit caused by reduced time-to-market)

Having analyzed available and applied evaluation methods, there has been chosen a method, which is based on the cost-benefit-analysis /7/. The general procedure is shown in figure 4.

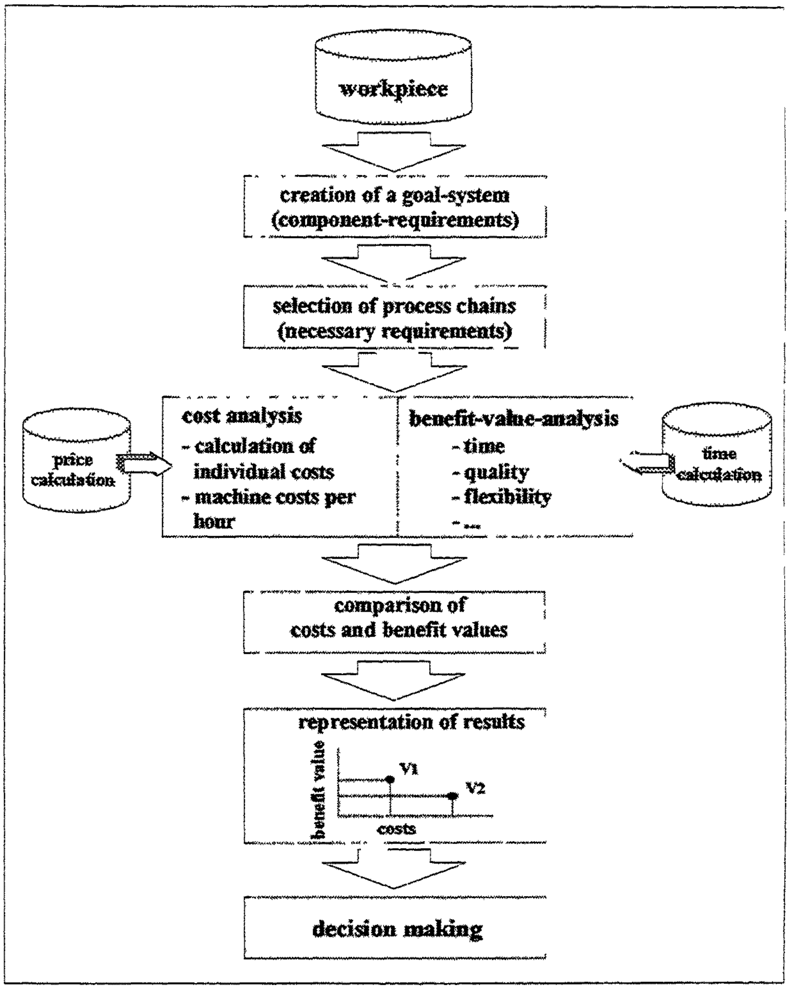


Fig. 4 : General approach to select and evaluate process chains

The consequent division of cost and benefit value analysis guarantees a high transparency of the method and the evaluation result. Many companies request a separate identification of cost indicators Furthermore, existing cost calculation systems can be included and investment (e.g. purchase of new machine tools or devices) can be taken into consideration.
All resting, particularly qualitative criteria will be evaluated by benefit-value-analysis. The final result is a comparison of generating costs and benefit values per process chain. That means, that user may select out of a sufficient range of decision variants. There is not any best solution. It is possible to validate results by benchmarking of representative parts.

Cost-benefit-analysis can be studied easily step by step. Considering the rules of probability, making mistakes can almost be excluded The procedure was applied successfully in various applications in practical use. For evaluation, an expert team is recommended. As a result, parameters (e.g the level of fulfillment) can be determined in a more objective manner

4. INTEGRATION INTO AN INTEGRATED DATA MODEL

Efficient use of the decision approach is based on an integrated data model. Thereby, data exchange inside the process chains is to be paid most attention. Losses of time as they result from extensive data transfer methods can be avoided as far as possible. Therefore, new data models are developed (process chain model) and available experience resulting from previous projects can be used /8/.

Development of the data model was mainly directed to consider entire process chain. For selection and evaluation, there are mostly relevant those data which refer to complete manufacturing sequences.

This demand is taken into consideration by "global process chain data" and "product data of finished part". An overview and classification of process chain information is given in table 2.

Data types	Explanation	Example
Global process chain data	Information/data, describing the entire process chain	-manufacturing time -manufacturing costs -reference companies -service enterprises
Global process step data	Information/data, describing a process step and which are generally valid (independent of the used process chain) for the process step	-data of technologies and manufacturing methods such as layer manufacturing processes (applicable material, maximum laser power of the device, etc.)
Specific process step data	Information/data, describing a process step and are valid for the corresponding process chain only	-specific parameters of a machine or device inside of a process chain
Product data (of finished part)	Information/data of the final product of each process chain	-achievable product properties (accuracy , surface quality,..)

Tab. 2 : Classification of process chain information

5. CONCLUSIONS

In a further step the developed concepts shall be applied for selected reference parts from companies of the tool and die making industry. A research study of typical parts is planned to carried out basing on representative parts. Additionally, the models and methods are going to be introduced into a software-prototyp for decision support and offer calculation. This prototype is planned to be linked with available CAD and/or CAPP systems /8/.

6. REFERENCES

1. H. K. Tönshoff, K. R. Hennig: "Fertigungstechnologien bewerten und auswählen" VDI-Zeitschrift, Vol. 137, issue number 6, pp.30-33, 1995

2. F. Klocke, M. Weck, H. Schell, E. Rüenauver: "Bewertung alternativer Fertigungsfolgen" Zeitschrift für wirtschaftliche Fertigung ZwF, Vol. 91, issue number 7-8, pp.359-362, 1996

3. W. Eversheim, F. Klocke, T. Albrecht, S. Nöken, H. Wirtz: "Rapid Prototyping-Unternehmensspezifische Technologiekonzepte", VDI-Zeitschrift Special Werkzeug- und Formenbau, November 1995, pp.20-23

4. B. E. Hirsch, J. Bauer: "Computerbased Rapid Prototyping System Selection and Support" Proceedings of the International Conference on Rapid Product Development Stuttgart, Germany, 10./11.6.1996

5. W. Steger, T. Conrad: "Rapid Prototyping: Operative und strategische Bewertung von generativen und konventionellen Fertigungsverfahren", VDI-Zeitschrift Special Werkzeug- und Formenbau, November 1995, pp.12-18

6. H. Dürr, U. Kunzmann: "Bewertung der Umweltverträglichkeit von Prozeßketten" Die Maschine dima, issue number 10, pp.47-48, 1996

7. P. Rinza, H. Schmitz: Nutzwert-Kosten-Analyse - Eine Entscheidungshilfe VDI-Verlag GmbH Düsseldorf, Germany, 1992

8. H. J. Warnecke, H. Dürr, A. Wauer, H. Muthsam: "Ein Schritt zu höherer Flexibilität bei der Arbeitsplanerstellung", Werkstattstechnik wt, Vol. 84 , pp.409-413, 1994

FEATURE BASED OPTIMIZATION OF PART ORIENTATION FOR RAPID PROTOTYPING

W. M. Micky Ng S.T.Tan
Department of Mechanical Engineering
The University of Hong Kong
Hong Kong

ABSTRACT

In rapid prototyping(RP) such as selective laser sintering(SLS) and stereolithography apparatus(SLA), part orientation is an essential factor that affects the part accuracy and build time. Many different aspects need to be considered so as to obtain the optimal orientation. Traditionally, this decision-making process depends heavily on the knowledge of the process planner. Automatic methods have been reported. However, they require users to give other additional information. For complex parts containing many features, it is difficult to answer the subjective questions. In addition, feature technology as applied in CAD/CAM systems seems not to have been incorporated with RP.

The authors proposed a feature based optimization algorithm that focuses on the two most important factors that affect the quality of surface finish and the efficiency of production: part accuracy and build time. Objective functions in terms of feature number and weight are formulated as the primary optimization model. The secondary objective is to minimize the build time by minimizing the number of layers. The developed system also gives some guidelines to help the user select an acceptable build direction if necessary.

KEYWORDS

Rapid prototyping, orientation, feature, accuracy, build time

1. INTRODUCTION

Rapid prototyping (RP) is an emerging manufacturing process used to fabricate a physical object directly from a computer representation of the part. The prototype mainly functions as a tool for functional or aesthetic assessment. This can enhance the design process by providing rapid and effective feedback to the designer and hence reduce product development cycle [1]. Moreover, the fabricated models can be used as pattern for advanced manufacturing technologies such as soft tooling and spray metal tooling [2]. Many such processes exist, e.g. selective laser sintering (SLS), stereolithography apparatus(SLA) and laminated object manufacturing(LOM). In the case of SLS, which is used for illustration purpose in this paper, it involves using laser to fuse powder together into laminated parts.

Despite the different principles applied in different processes, they all build parts of any shape layer by layer, and in any orientation. The choice of orientation is a critical factor that affects many important aspects, such as the quality of surface finish, the build time, the amount of support structures required [3]. There is a trade-off between these different aspects. In current practice, the chosen orientation is based on the experience of the RP builder. Few automatic methods for the determination of the optimal orientation have been reported. Some of them require a user specifying voluminous additional information for the build process.

On the other hand, modeling of mechanical parts with conventional CAD systems involves repetition of standard design features such as holes, slots and pockets. Besides achieving the requirements on the surface finish of the entire part, the surface finish of some specific geometric features may be more critical. In this case, the part orientation should be determined based on the importance of features.

This paper focuses on the definition of part orientation of the SLS process by considering part accuracy and production time. The methodology for the formulation of the feature based optimization is described and the optimization procedure is presented. Examples to illustrate the algorithm are also included.

2. RELATED WORK

Hur and Lee [4] determined the optimal part orientation using objective functions which consider the importance of three factors: accuracy, build time and support structures. Two modes of optimization are provided. In single optimization, users are required to rank the factors. In multiple optimization, the weight value of each factor is specified.

Frank and Fadel [5] proposed an expert system which recommends the best build direction based on user input and decision matrix. The system allows the user to select two most significant features on the part. The preferred build direction is suggested based on the rules derived according to the arrangement of the two geometric features. This system works with only two features and the result depends on the user to identify which feature is the most significant.

Bablani and Bagchi [6] developd an algorithm to calculate the various sources of surface errors in a part in different orientations by rotating a given part about a user-selected axis / axes at different intervals of rotation. The optimal orientation is determined based on the user's selection of the minimum stair stepping effect or the minimum number of slices.

Cheng, Fuh, Nee, Wong, Loh and Miyazawa [7] proposed a multiple objective function as a formulation to derive the optimal orientation. Factors to be considered include part accuracy, building time and stability of the part. The primary objective function for part accuracy is defined as the sum of the product of the surface number and its weight for all the surfaces of the part. The minimization of the building time formulates the secondary objective. The work raises two essential characteristics in the optimization process. First, the process is automatic. Second, this method suggests an optimization algorithm based on the surface type.

3. ESSENTIAL FACTORS IN LAYERED MANUFACTURING

Many factors contribute to the definition of the orientation of build. Some of these factors include surface finish, build time, support structures, shrinkage, curling, distortion, roundness/flatness, part tolerance, resin flow, material cost and trapped volume. All these contribute to various degrees to the quality of the final product and have to be considered when developing a strategy for preferred orientation. Only three most essential factors are discussed here.

Part accuracy: It is usually a primary criterion in the part orientation. Without good accuracy, it is often meaningless to achieve speed. Because each object is decomposed into layers, for any curved surface or inclined plane, the effect of laminar build is noticed as stair steps. It is important to orient the part in a way to minimize the stair stepping effect (Figure 1).

Build time: The layer adding process which deposits the next layer of powder is the most-time consuming operation in SLS. The roller that moves to add the next layer of powder should not be too fast to prevent a rough surface on the powder bed. A cooling-down time is added to each layer at the end of the scanning process. In addition, some process software perform concurrent slicing of the part geometry files while fabricating the object. The system has to wait until the scan data of the next layer is prepared. The waiting time is significant when building a complex part. By minimizing the number of layers, we can determine the optimal orientation for the build time.

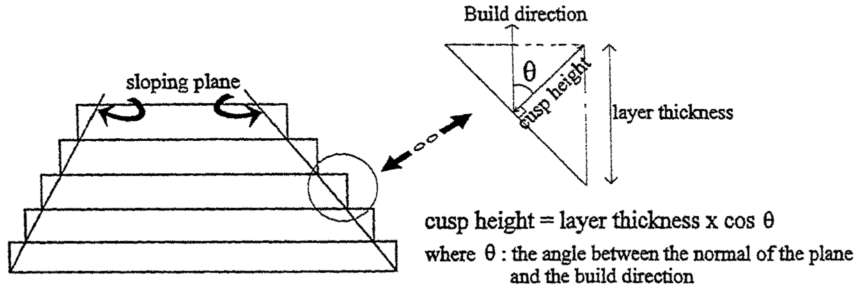


Figure 1 Stair stepping effect in layered manufacturing techniques

Support structures: Overhang features need support structures to prevent them from sagging. Support structures are also required for base support. One advantage of the SLS system is that the unsintered powder serves as a natural support for the next layer of powder and the object under fabrication. Thus our algorithm does not take this factor into account.

4. FEATURE BASED OPTIMIZATION

In the formulation of the feature based optimization model, all form features with their feature axes are identified within the part. The surface information of each feature is also extracted. The feature axes of several geometric entities are shown in Figure 2. A feature number and a feature weight which reflect the importance and the effect of a feature are assigned to each of them.

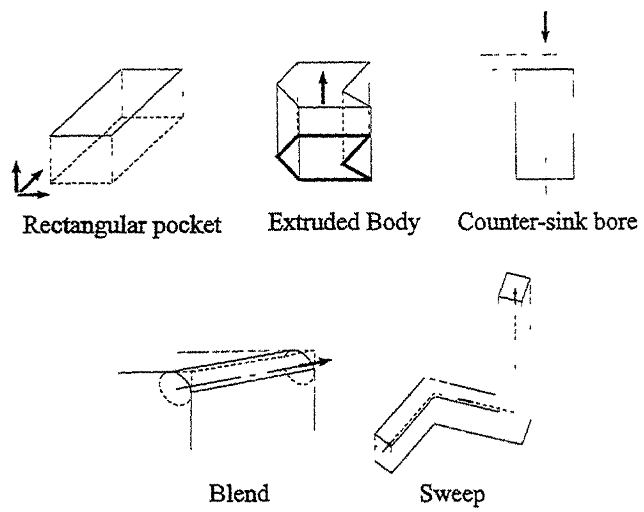


Figure 2 Features with defined feature axes

The feature number N_j is defined as:

$$N_j = \frac{A_j}{A} \qquad \text{where } A_j = \text{the surface area of the } j\text{th feature;} \qquad (1)$$
$$A = \text{the total surface area of the part.}$$

The larger the feature number is, the more important the feature is in a part.

The feature weight is the average weight of the individual surfaces corresponding to the same feature:

$$\varepsilon_j = \sum_{k=1}^{l} \frac{a_k}{A_j}\xi_k \qquad \text{where} \quad A_j = \text{the surface area of the } j\text{th feature;} \tag{2}$$

$1 \le j \le n$

a_k = the area of the kth surface of the jth feature;

ξ_k = the weight of the kth surface (from Table 1);

l = the number of surfaces in the jth feature;

n = the number of features.

Note the weight of a surface ξ is defined as the accuracy that can be achieved during fabrication. It depends on the type of surface and its orientation with respect to the slicing direction. Only two types of surfaces are considered: planar and cylindrical. Each surface contains a principal axis which defines the orientation. For a given plane, the outward normal is the principal axis. For a cylindrical surface, the axis of the cylinder is the principal axis. For curved or freeform surfaces, e.g. the cone, Bezier surface, etc., they are approximated by a number of planar facets and the weight is evaluated by averaging the weights of individual facets. For a simple implementation, only the staircase effect is considered. The relationship between the weight, surface type and principal axis is shown in Table 1, where θ is the angle between the principal axis and the build direction. A weight of 1 is set for the surface that does not contribute any errors. For instance, the vertical/horizontal plane is a good feature for high accuracy, so its weight is 1. From Figure 1, the cusp height of a plane is directly related to $cos\theta$ and so weights of $-cos\theta$ and $cos\theta$ are assigned to upward and downward sloping planes respectively.

Surface type	θ	Weight, ξ
1 Horizontal plane	0	1
2 Vertical plane	90	1
3 Upward sloping plane	(0, 90)	$-cos\ \theta$
4 Downward sloping plane	(90, 180)	$cos\ \theta$
5 Vertical cylinder	0 / 180	1
6 Horizontal cylinder	90	0
7 Sloping cylinder	(0, 90), (90, 180)	$-sin\ \theta$

Table 1 The type of surface and weight

The primary objective function is defined as the sum of the product of the feature number and its weight:

$$Q_i = \sum_{j=1}^{n} N_j \varepsilon_{ij} \qquad 1 \le i \le m \tag{3}$$

where Q_i = the overall objective value of the ith orientation of the built part;

ε_{ij} = the feature weight of the jth feature evaluated from equation (2) for the ith orientation;

N_j = the feature number of the jth feature defined in equation (1);

m = the number of possible orientations for building the part;

n = the number of features.

The individual term $N_j\varepsilon_{ij}$ in equation (3) is called the feature objective value, P_{ij}, of the j^{th} feature for the i^{th} orientation:

$$P_{ij} = N_j \varepsilon_{ij} \tag{4}$$

The direction of the recognized feature axes are taken as the possible orientation candidates. The algorithm first checks whether the accuracy is achieved in a particular orientation. If the cusp height in some region of the part is beyond the user specified tolerance, the candidate is removed. Then only orientations with an objective value greater than a threshold value are selected as candidates for determining the optimal orientation based on the build time. An accuracy factor F_{acc} is specified by the user to indicate the relative importance between the part accuracy and build time.

The overall threshold value Q_t is defined as:

$$Q_t = Q_{min} + (Q_{max} - Q_{min}) * F_{acc} \qquad 0 \leq F_{acc} \leq 1.0 \tag{5}$$

where Q_{max} = the maximum objective value among different orientations;

Q_{min} = the minimum objective value among different orientations.

If accuracy is deemed more important, a greater value should be set for F_{acc}.

Similarly, we define feature threshold value of the j^th^ feature as:

$$P_{t_j} = P_{min_j} + (P_{max_j} - P_{min_j}) * F_{acc} \qquad 0 \leq F_{acc} \leq 1.0 \tag{6}$$

where P_{max_j} = the maximum feature objective value among different orientations;

P_{min_j} = the minimum feature objective value among different orientations.

For a particular orientation i, a feature is said to be accurate if the feature objective value P_{ij} is greater than the feature threshold value P_{t_j}. A feature accuracy value Val_{acc} which is a measure of the accuracy of the j^{th} feature is evaluated from equation (7):

$$Val_{acc_{ij}} = (P_{ij} - P_{min_j}) / (P_{max_j} - P_{min_j}) \tag{7}$$

Val_{acc} can be used to determine whether a feature is accurate. If $Val_{acc} > F_{acc}$, the feature is accurate.

From the overall objective values computed from equation (3), we obtain the overall accuracy value in orientation i:

$$Q_{acc_i} = (Q_i - Q_{min}) / (Q_{max} - Q_{min}) \tag{8}$$

The larger the overall accuracy value is, the better the surface finish one can achieve. The information such as the overall/feature accuracy values obtained from equations (7) and (8) is important in the optimization process, especially when it can help a user determine the best orientation.

The secondary objective is to minimize the build time. The total build time is evaluated in terms of the total number of layers. Generally, a reduction in the number of slices will generally lead to shorter build time. Thus, the optimal part orientation is determined by minimizing the number of layers. For a constant layer thickness system, minimizing the number of layers is equivalent to finding the minimum build height. Constant slicing thickness is assumed in this work.

5. OPTIMIZATION PROCEDURES

The procedures of the algorithm are as follows:

1. Note all the features in the CAD model. The directions of their feature axes are listed as candidates for the optimal orientation.
2. Recognize the surface type of all surfaces and their corresponding feature.
3. Group the features according to the user, e.g. features of the same kind or features on a particular surface are grouped together.
4. For each orientation, check whether the accuracy of all the surfaces can be achieved in this orientation. If not, remove this orientation from the list of candidates.
5. If accuracy cannot be achieved in all orientations:

a. Report the features with accuracy not achievable in each orientation. The user is then asked to specify the critical feature(s).
b. Find the orientation(s) with primary objective value of the critical feature(s) greater than the threshold value. Obtain all possible orientations with information such as overall/feature objective values and threshold values, build time for user's guidelines.
c. Ask the user to select the best orientation. Go to step 8.

6. If accuracy can be achieved in some orientations:
   a. If the user wants the process to be a general optimization, choose automatic process. Compute the objective value of the whole part in all orientations. List the candidates with objective values greater than the threshold one. Go to step 7.
   b. If the user wants to specify some critical feature(s), choose manual process. Go to step 5b.
7. Obtain the optimal orientation with minimum number of layer among the possible candidates.
8. Report the optimization results.

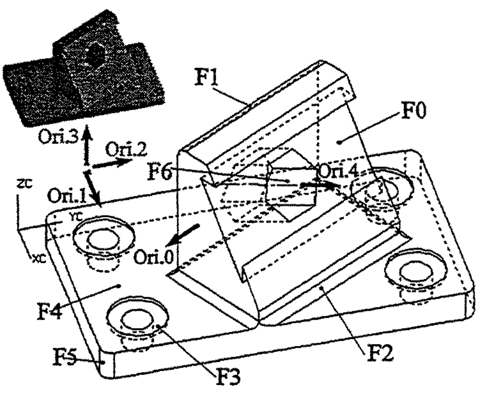
6. IMPLEMENTATION

The algorithm is developed and performed within the environment of Unigraphics(UG) system. The built-in user functions give direct access to the CAD database. Once the model is created in the UG system, the optimization runs automatically and user inputs are inquired when necessary.

7. CASE STUDY

Two examples are given here to illustrate the algorithm. Several process constraints are defined first: Cusp height = 0.2mm; Constant layer thickness = 0.1mm; Accuracy factor = 0.5.

For the part shown in Figure 3, there are 7 features with 5 possible orientation candidates. The feature number and weight for all orientations are tabulated in Table 2. From the objective values computed in Table 3, orientation 0, 1, 2 and 3 are selected for the secondary objective. Orientation 3 is the optimal orientation with the minimum build height. Some information about the optimization results is given (Table 4).



Features(F):
0: Extruded
1: Blend
2: Blend
3: Counter-sink bore x 4
4: Block (base)
5: Blend
6: Extruded

Feature group:
0: F0
1: F1
2: F2
3: F3
4: F4
5: F5
6: F6

Orientations(Ori.):
0: Feature axis of F0
1, 2, 3: Feature axes of F4
4: Feature axis of F6

Figure 3 Example1

A more complicated part with 9 features and 6 possible orientation candidates is shown in Figure 4. The features are divided into 6 groups. When a manual process is chosen, the algorithm helps the user select an acceptable build direction by providing information for user's reference. The system first asks the user to select the critical feature(s) (features in group 2 are selected). As depicted in Table 5, orientation 1 and 3 are suggested because their group 2 feature objective values are greater than the corresponding feature threshold value. A table indicating the status of each feature group is given (Table 6). Suppose the user also considers the sweep feature in group 5 to be important and neglects the amount of build time, orientation 3 is chosen as the best build direction. However, for a shorter build time, orientation 1 becomes the best orientation. Orientation 1 is the optimal orientation if the process is automatic. The optimization results are shown in Table 7.

		Orientation 0	Orientation 1	Orientation 2	Orientation 3	Orientation 4
Feature group, j	Feature No., N_j	Feature weight, ε_{0j}	Feature weight, ε_{1j}	Feature weight, ε_{2j}	Feature weight, ε_{3j}	Feature weight, ε_{4j}
0	0.30	1.00	-0.55	-0.41	0.48	0.25
1	0.01	1.00	-0.87	-0.50	0 00	0.00
2	0.03	0.43	-0.73	-0 64	0.00	-0.19
3	0.07	0.38	0.38	0.38	1.00	-0.73
4	0.53	0.64	1.00	1.00	1.00	-0.53
5	0.01	0.00	0.00	0.00	1.00	-0.87
6	0.05	-0.24	-0.43	-0.58	-0.58	1.00

Table 2 Feature number and weight of example 1

	Orientation 0	Orientation 1	Orientation 2	Orientation 3	Orientation 4
Feature group, j	Feature Obj. Val., P_{0j}	Feature Obj. Val., P_{1j}	Feature Obj. Val., P_{2j}	Feature Obj. Val., P_{3j}	Feature Obj. Val., P_{4j}
0	0.30	-0.17	-0.12	0.14	0.08
1	0.01	0.01	-0.01	0.00	0.00
2	0.01	-0.02	-0.02	0.00	0.00
3	0.03	0.03	0.03	0.07	-0.05
4	0.34	0.53	0.53	0.53	-0.28
5	0.00	0.00	0.00	0.01	-0.01
6	-0.01	-0.02	-0.03	-0.03	0.05
Overall objective value, Q_i = (Threshold value, Q_t = 0.25)	0.69	0.34	0.38	0.73	-0.22
Build height (mm) =	106.81	55.20	94.00	48.00	(84.20)

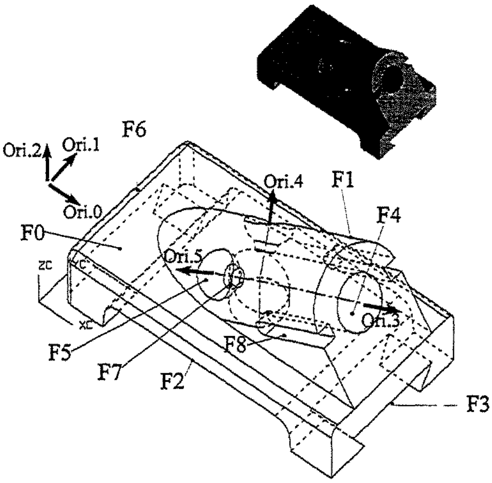
Table 3 Feature objective value, overall objective value and build height of example 1

Feature group, j	Orientation 3	Feat. Obj. Val., P_{3j}	P_{max_j}	P_{min_j}	Feat. Thres. Val., P_{t_j}	Feature accuracy, Val_{acc3j}
0	accurate	0.14	0.30	-0.17	0.06	0.66
1	less accurate	0.00	0.01	-0.01	0.00	0.46
2	accurate	0.00	0.01	-0.02	0.00	0.84
3	accurate	0.07	0.07	-0.05	0.01	1.00
4	accurate	0.53	0.53	-0.28	0.12	1.00
5	accurate	0.01	0.01	-0.01	0.00	1.00
6	less accurate	-0.03	0.05	-0.03	0.01	0.00
Build Height (mm) =	48.00					
Overall objective value, Q_3 =	0.73					
Overall accuracy value, Q_{acc3} =	1.00					

Table 4 Optimization results of example 1

8. CONCLUSION

Part orientation is important in rapid prototyping because it affects part accuracy and build time. The determination of part orientation should be automatic with user interaction. On the other hand, feature-based CAD/CAM systems provides a higher level of abstraction for design engineers and manufactures to share. Focusing on the two aspects, we developed a feature based approach to determine the optimal part orientation. Objective functions are formulated as the optimization model. Sometimes, an optimal orientation is not possible with the specified process constraints. The evaluated objective values provide a knowledge base for reference by a user before making a decision in choosing the best orientation for slicing in an RP process.



Features(F):
0: Block
1: Extruded
2: Rectangular pocket
3: Extruded
4: Sweep
5: Sphere x 2
6: Blend
7. Cone x 2
8. Rectangular slot x 2

Feature group:
0: F0
1: F2, F3
2: F5, F7, F8
3: F1
4: F6
5: F4

Orientations(Ori.):
0, 1, 2: Feature axis of F0
3: Feature axes of F8
4, 5: Feature axis of F7

Figure 4 Example 2

	Orientation 0	Orientation 1	Orientation 2	Orientation 3	Orientation 4	Orientation 5	
Feature group, j	Feat. Obj. Val., P_{0j}	Feat. Obj. Val., P_{1j}	Feat. Obj. Val., P_{2j}	Feat. Obj. Val., P_{3j}	Feat. Obj. Val., P_{4j}	Feat. Obj. Val., P_{5j}	Feat. Thres. Val., P_{tj}
0	0.29	0.29	0.29	-0.04	-0.17	-0.17	0.06
1	0.19	0.20	0.16	-0.11	-0.15	-0.15	0.02
2	-0.01	0.02	-0.02	0.02	-0.04	-0.04	0.00
3	-0.06	0.09	-0.10	0.21	0.02	0.02	0.06
4	0.04	0.01	0.01	-0.03	-0.05	-0.05	-0.01
5	-0.05	-0.08	-0.05	0.00	-0.08	-0.08	-0.04
Overall objective value, Q_i = (Thres. Val., Q_t = 0.03)	0.39	0.53	0.29	0.06	-0.47	-0.47	
Build Height (mm) =	100.01	55.20	59.28	97.80	(86.09)	(86.09)	

Table 5 Feature objective value, overall objective value and build height of example 2

Feature group, j	Orientation 1	Orientation 3
0	accurate	less accurate
1	accurate	less accurate
2	accurate	accurate
3	accurate	accurate
4	accurate	less accurate
5	less accurate	accurate
Build Height (mm) =	55.20	97.80

Table 6 Two orientations suggested by the algorithm in manual process

Feature Group, j	Orientation 1	Feat. Obj. Val., P_{1j}	Feature accuracy, $Val_{acc\,1j}$
0	accurate	0.29	1.00
1	accurate	0.20	1.00
2	accurate	0.02	1.00
3	accurate	0.09	0.54
4	accurate	0.01	0.67
5	less accurate	-0.08	0.00
		Build Height (mm) =	55.20
		Overall objective value, Q_1 =	0.53
		Overall accuracy value, Q_{acc1} =	1.00

Table 7 Optimization result of example 2

9. ACKNOWLEDGMENT

This work is supported by the Department of Mechanical Engineering, The University of Hong Kong. The support is gratefully acknowledged.

10. REFERENCES

1. S. Ashley, "Rapid prototyping systems", Mechanical Engineering, (34-43), April 1991.
2. E. Weiss, E. Levent Gursoz, F. B. Prinz, S. Fussell and E. P. Patrick, "A rapid tool manufacturing system based on stereolithography and thermal spraying", Manufacturing Review, ASME, (40-48), vol. 3, no. 1, March 1990.
3. Paul F. Jacobs, Rapid Prototyping & Manufacturing: Fundamentals of StereoLithography, Society of Manufacturing Engineers, 1st Ed., 1992.
4. J. Hur and K. Lee, "Determination of optimal part orientation for stereolithographic rapid prototyping", Technical Report, Department of Mechanical Design and Production Engineering, Seoul National University, Seoul, July 1994.
5. D. Frank and G. Fadel, "Expert system based selection of preferred direction of build for rapid prototyping process", Proceedings of the Fifth International Conference on Rapid Prototyping, (191-200), Dayton , 1994.
6. M. Bablani and A. Bagchi, "Quantification of errors in rapid prototyping processes, and determination of preferred orientation of parts", Transactions of the North American Manufacturing Research Institution of SME, vol. 23, (319-24), 1995.
7. W. Cheng, F.Y.H. Fuh, A.Y.C. Nee, Y.S. Wong, H.T. Loh and T. Miyazawa, "Multi-objective optimization of part building orientation in stereolithography", Rapid Prototyping Journal, vol. 1, no. 4, (12-23), 1995.

RAPID PROTOTYPING USING ROBOT WELDING - PROCESS DESCRIPTION

Fernando Ribeiro
Dep. Electrónica Industrial - Universidade do Minho,
4800 Guimarães, PORTUGAL

John Norrish
Faculty of Engineering, University of Wollongong
Northfields Avenue, Wollongong NSW 2522, AUSTRALIA

ABSTRACT

Rapid Prototyping is a relatively recent technique to produce component prototypes for industry in a much shorter period of time, since the *time to market* a product is essential to its success. A new Rapid Prototyping process which uses metal as the raw material had been under development at Cranfield University in the last few years. The process uses a Gas Metal Arc fusion welding robot which deposits successive layers of metal in such way that it forms a 3D solid component. Firstly, a CAD system is used to draw the solid model, then some information relative to the types of layers and dimensions is incorporated in the model and the solid model is then automatically sliced. Reports on the welding time and conditions for the component's production are automatically generated as well as the robot program.

The concept of this Rapid Prototyping process is deeply explained in this paper. Every step of the process is described by a full chart. The Hardware and Software used in this system are also described. Since a computer model is used a calibration of the system is required and therefore the most important aspects of Robot and Cell Calibration are also discussed.

KEYWORDS

Rapid Prototyping, Metal, CAD, Robotics, Welding, Computer Simulation

1. INTRODUCTION

This process was first described by Ribeiro in [1] and with some more detail some time later by Ribeiro [2]. The time to market a product is essential in nowadays so competitive industry and therefore to reduce prototyping time, it has often been stated that a need to automate the production of 'one off' components for development and evaluation is necessary. Casting technique is a time consuming process and expensive especially if used to make one component only. Most of the rapid prototyping processes evolved in response to this requirement and these use resin based materials which are not always suitable for testing purposis. Metal based prototypes are often required and additional processing is necessary to convert resin based materials to a useful form. In this process the component is formed by melting and depositing the metal using the GMA welding process. A CAD drawing system is used to create the initial solid shape and a welding robot is used to manipulate the welding torch.

The first step is to draw the component in a CAD system and then a slicing 'add-on' of the CAD program is implemented to generate the desired layers which will form the robot program. The slicing system was described by Ribeiro [3]. Additional data is required to indicate bead geometry and the material used. Since a CAD model is used a robot and workcell calibration is necessary to indicate exactly where the component is going to be built in relation to the robot. The welding conditions like voltage, current, wire feed speed and robot speed are automatically generated by the program in order to achieve the required bead geometry and stable operating parameters. Welding studies had to be carried out previously to derive the best conditions and these were carried out by Norrish [4] and the parametric equations were generated by Ogunbiyi and Norrish [5]. The robot program to build the component is automatically generated and can be computer simulated with the use of a robot simulation program to check for collisions or other problems such as access to difficult areas of the component. Should the robot program need to be changed it can be done manualy and after that it is then compiled and downloaded to the robot.

2. GENERAL DESCRIPTION

Two different CAD/CAM manufacturing processes can be used to make a component: the subtractive and the additive. In the first one, material is removed from a massive block of material. In the second one, material is added to create the desired shape. The shape of the component to be built and the material to be used influences the choice for one of these processes.

Rapid Prototyping processes are additive processes because they consist of depositing layers of material, for example, by solidifying a photopolymer or thermoplastic material or by adding sheets of paper and joining them with an adhesive. Most traditional Rapid Prototyping processes work in a similar way which consists of a movable device depositing or solidifying material on a table or container. This new Rapid Prototyping process is comparable to the traditional ones as far as basic principle is concerned and it is hereby described: a robot is used in conjunction with a turn table, the deposition device is a welding torch and the raw material will be the fused metal.

Many reasons do not allow a human to carry out this task: the precision with which the robot has to be moved, repetition of movements, the hazardous aspects of welding, boredom and tiredness. This indicates that manual welding techniques would not be feasible for this type of process. In addition the main objective of Rapid Prototyping is to automate the whole process as completely as possible. Therefore, robots are the most desirable machines to perform this operation. They do not get tired or bored, they do not hurt themselves and they are ideal for performing repetitive and accurate tasks. In the proposed system, a robot is used to deposit layers of weld metal on a turn table.

3. PROCESS CONCEPT

Figure 1 shows each step of the proposed process. Each colour represents a different task or a different software package used. Each box represents an action to be taken, information about a certain entity or even a user of the information. The arrows show the direction of information flow.

The first step is to create a 3D model using a CAD program ('Graphical Information' box). Some components cannot be made in one go depending on its complexity. In that case the designer has to draw it in independent solids although close together. Each of these solids will be built separately one after the other. It is important to distinguish component and part - One or more *Parts* make a full *Component.*

For each part information has to be 'input' into the system ('Part Features' box). Information like part's name, width, build up sequential order and torch orientation are some of the fields to be filled in, to be used to generate the appropriate welding parameters ('Welding Parameters' box). After this INPUT task follows the slicing of each part. This is automatically done ('Slicing' box) and will generate polylines, within the CAD model which represent the welding trajectory for building up the desired part.

There is a text file containing the editable instructions for controlling the table and welding gun ('Welding/Table Instructions' box) and this is used as input to generate all the appropriate outputs given by the 'Automatic Generator' box. There are four automatically generated outputs. The first one is a DXF file format ('Slices as DXF' box) which contains the welding trajectory (this file can be imported to any robot simulator to check for collisions and accessibility). The second output is the robot program in ARLA language (although it could be any other robot language) as text file format for each part ('Robot Program' box). The other two outputs consist of two reports: one is for the welding technician and contains welding instructions for each part (welding parameters, timings, etc.) represented by the 'Individual Report' box. The other is for the production manager and contains information about the full component describing the information for each part (building time, amount of material needed, etc.) and is represented by the 'Global Report' box.

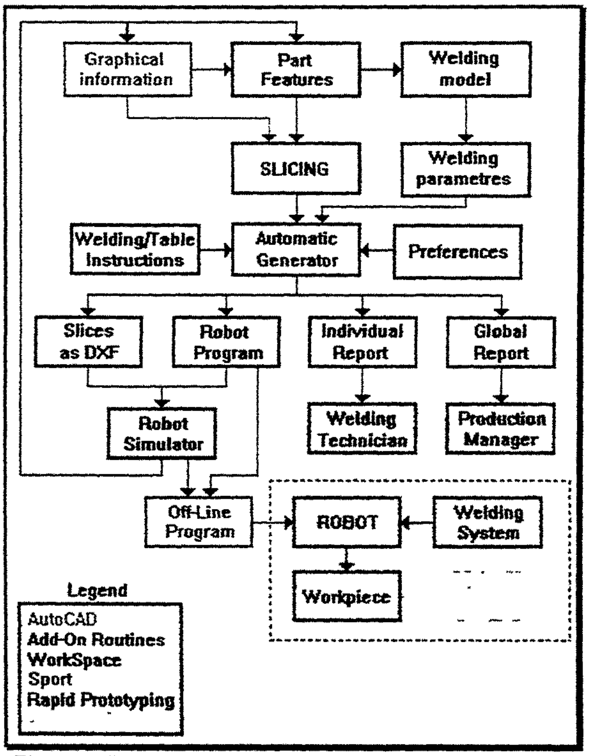


Figure 1 : 'Rapid Prototyping using Fusion Welding' Concept

All these tasks described so far were carried out within the software created by the author except the first one (drawing the model) for which AutoCAD itself was used. Should the user decide to use a robot simulator to check the robot program ('Robot Simulator' box), he has to read both 'Slices as DXF' box and the 'Robot Program' box.

After creating the robot program, this is downloaded to the robot via RS-232-C (with a serial link cable). A robot manufacturer software is necessary to perform this task because the binary file format is unique to each robot manufacturer. This task is represented by the 'Off-Line Program' box. The robot program should now have been transferred to the robot (represented by the 'Robot' box) and is ready to run. The welding apparatus (welding power source, welding torch and consumables) is represented by the 'Welding System' box. If the desired component has more than one part, the same number of robot programs have to be run, one after the other. All the robot programs can be downloaded beforehand.

The 'Welding Control' box is not part of this work but is seen as a very important improvement to this system. The idea is to read the component's shape by some means of vision or any other sensor in such a way that could give some feedback to the welding system (to correct welding parameters) and if possible to the robot as well (to correct its movement or speed).

4. HARDWARE

This workcell contains a welding robot connected to a computer via a serial cable RS-232-C, and the computer is also connected to a printer. The power source and welding consumables are connected to the welding torch which is mounted on the robot arm. The robot builds the component on a table. Safety measures should be seriously taken into account (like fencing, UV filters, protecting switches, etc). This table should ideally be controlled by the robot (dotted arrow). Figure 2 represents a graphical description of the hardware used (the dotted arrows represent ideal situations and not the real work cell).

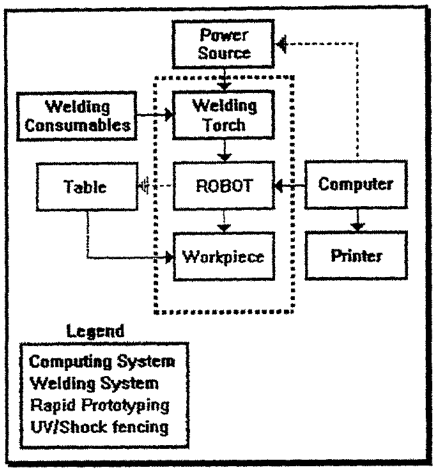


Figure 2 : Rapid Prototyping work cell Hardware

The individual system components used, were:

An ASEA IRb 2000 robot from ABB with 6 degrees of freedom with an S3 controller. It has only 64 Kbytes of memory and this represents a major limitation in the amount of programs which can be stored in it. For this process this is a major limitation because very large programs can be generated depending on the complexity of the components drawn. A turn table could be used either a stand alone or a robot linked one but for most of the samples made no turn table was used.

The welding Power Source was a Migatronic BDH 320 although some tests were later on carried out with a Migatronic BDH 550. The Welding Torch was a BINZEL PUSHPULL torch.

The only computer used to run all the software (CAD, off-line programming and robot downloading software) was an Intel 80486 microprocessor PC based with a 66 MHz clock and 16 Mbytes of RAM memory. The hard-disk capacity was 250 Mbytes. This computer proved to be enough.

The work cell had a fence around it with ultra violet filtered (UV) glasses to protect the operators' eyes. The robot and table alone had another fence with another UV filter. This second fence was also supposed to protect for physical unexpected robot movements. Should this fence opened, a circuit would go off stopping the robot and welding apparatus immediately. The consumables were all the necessary ones to a Gas Metal Arc Welding process like the wire, gas and contact tips.

5. SOFTWARE

The software used consisted of the AutoCAD ™ package, the 'add-on' developed in this work, a robot simulation software and the compiler/download software for the used robot. Figure 3 shows a graphical description of the software used.

The main reason for using AutoCAD was for its world wide use, easy interface, being an open system and accept many file formats. It is also in continuous development (although it is now in release 13 this work was carried out using release 12). Three Dimensions solids are used as well as surfaces although these were only combined in one kernel in release 13. AutoCAD has AutoLisp as its own programming language which makes it possible to customise programs according to specific application needs and this was the language used to develop the slicing routines.

WorkSpace 3.0 was used as robot simulation software. This was used for collision checking and timings control although it is not essential to this process.

Two different packages were used to compile and download the robot programs: OLP 3.0 (Off-Line Programming) is a DOS (Disk Operating System) program supplied by ABB and SPORT 1.0 (Windows based program developed by LUND University in Sweden) which performs more or less the same functions as OLP 3.0. These two programs have a slightly different dialect for the ARLA robot programs generated. This needed some considerations like generating two different versions of robot programs according to the off-line program used.

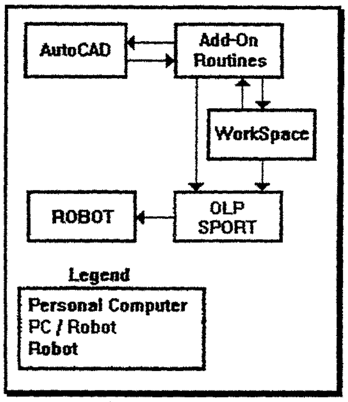


Figure 3 : Rapid Prototyping work cell Software

6. ROBOT AND WORK CELL CALIBRATION

Some work cell calibration had to be carried out to ensure correct location for build up and repeatable results and the methods employed are described in this section.

After have drawn the component in the CAD package, it is important to indicate the robot where to build it within the work cell. Therefore, the table centre co-ordinates (where the component will be built) have to be accurately located using the robot. The robot is moved to the centre of the table and the wire coming out of the torch (considering the stand-off) must touch the base plate where the component is to be started. The co-ordinates measured by the robot are read from the controller and input into the set up menu option of the software. The zero co-ordinates of the CAD model will correspond to these table centre co-ordinates. If the turn table also tilts, it must be tilted 90 degrees and the centre co-ordinates must be taken again. This is required to identify the centre position of the table when tilted.

The flatness of the table was also checked. Once this is set up, it should not need to be checked again until the table is moved although in production a periodic check should be carried out just to make sure that the table has not moved. The table surface must also be parallel to the XY plane of the robot.

If the table cannot be moved and its surface is not parallel to the robot co-ordinate system, the solid model in the CAD program can be rotated or tilted to match the location and orientation of the table. Thus, when the robot program is generated it takes into account the position and orientation of the table.

The length of the Welding Torch has to be accurately measured after mounting it onto the robot. These displacement values are represented as a (X, Y, Z) vector and set up the correct TCP in the robot memory. Once this has been set, it does not need to be reset until the torch is changed or moved, although a periodic check should be carried out just to make sure that the torch has not moved.

Each part can have a different welding orientation which can vary according to the gas flow requirements and angle of the filler wire which may influence the deposition itself. Therefore, the welding torch angle was first decided by placing the torch in the desired orientation and reading it from the robot controller. This orientation was then input into the slicing routines. A full component can have several different welding orientations but a part can only have one unique orientation. Most of the times, the orientation is kept constant for all the parts of a component.

To calibrate the robot itself an example of robot calibration to be used is Tracker system described by McMaster and Ribeiro [6].

7. CONCLUSIONS

This process was deeply tested and used, and several components were built. Some examples were described by Ribeiro in [7] and [8]. All these components were made completely automatic, from the 3D drawing up to the final component.

The main advantages of the slicing program used was that the slices were automatically created, the ARLA robot program was generated completely automatically and it was not essential to use a robot simulation package to test it, although simulation can be used to save on line time.

Should a different robot be used, this means that a different robot language has to be used. With only a few lines of code changed the system will work fine because the 'moving' instructions and the welding start/stop instructions are parametrised in an editable text file. In half an hour the system is ready to work with a different robot.

If the table is changed, this has to be calibrated for the first time, only to tell describe to the robot where is the table within the work cell. The system is not dependent on the welding consumables and therefore any wire/contact tip/gas. The torch can be any one although its size needs to be known and input into the slicing routines.

For all these reasons, this process proved to be a very flexible rapid prototyping system and has been very successful for all the components made so far.

8. ACKNOWLEDGEMENTS

Mr. Ribeiro would like to thank John Norrish for his support in this project and his expertise in the welding field and also Dr. McMaster for his academic help. My very special thanks go to John Savill for his strong support in welding all of the components and for solving in an efficient way the practical problems which arose as well as encouraging me to new ideas. A special thank you for my wife is also deserved for her understanding during the difficult periods of this project.

9. REFERENCES

[1] Ribeiro, A. F. M., Norrish, J. and McMaster R., "*Practical case of Rapid Prototyping using Gas Metal Arc Welding*", 5th International Conference on Computer Technology in Welding, Paris, France, 15-16 June 1994.

[2] Ribeiro, A. F. and Norrish, J., "Rapid Prototyping Process using Metal Directly", Solid Freeform Fabrication 1996 Symposium - Austin, Texas, USA, 12-14 August 1996.

[3] Ribeiro, A. F. and Norrish, J., "Rapid Prototyping using Robot Welding - Slicing System Development", 3rd France-Japan Congress and 1st Europe-Asia Congress, Besançon, France, 1-3 October 1996.

[4] Norrish, J., Advanced Welding Processes, Institute of Physics Publishing, 1992.

[5] Norrish, J., and Ogunbiyi B., "*An Adaptive Quality Control concept for robotic GMA Welding*", 5th International Conference on Computer Technology in Welding, Paris, France, 15-16 June 1994.

[6] McMaster, R. and Ribeiro, A. F., "Cell Calibration and Robot Tracking", IEE Colloquium 'Next Steps for Industrial Robotics', Professional Group C15, 3/1-3/7, London, 17th May 1994.

[7] Ribeiro, A. F. and Norrish, J.,"Metal Based Rapid Prototyping for More Complex Shapes", 6th Biennial International Conference on 'Computer Technology in Welding' - Lanaken, Belgium, 9-12 June 1996.

[8] Ribeiro, A. F. and Norrish, J.,"Case Study of Rapid Prototyping using Robot Welding - 'Square to Round' shape", 27th International Symposium on Industrial Robotics, Milan, Italy, 6-8 October 1996.

AN EXTENDED SLICING METHOD FOR RAPID PROTOTYPING

Kwan H. Lee and Il-sang Yoo
Kwangju Institute of Science and Technology
572 Sangam-dong, Kwangsan-ku, Kwangju, 506-303, Korea

ABSTRACT

Provision of an appropriate CAD data is essential for making accurate RP parts. The CAD data, however, have inherent inaccuracies due to the use of STL file format which approximates the actual shape of the model by tessellation. In addition, the constant thickness often used for the slice file for RP doest not guarantee the making of special features and at times produces a part that differs significantly from the original one. A few approaches have been developed to correct these problems. An enhanced method that contains the idea of both direct and adaptive slicing is proposed in this paper. This method minimizes the stairstepping error by using the mid-planes between the calculated contours as the basis for building layers. In addition it uses a two-sided slicing procedure to improve the z-axis accuracy of part features.

KEYWORDS

Rapid prototyping, Sectioning, Direct slicing, Adaptive slicing, Part accuracy

1. INTRODUCTION

1.1 RP Process and and the Problems in Data Preparation

The rapid prototyping process generally consists of three stages: data preparation, model building, and post processing. A part is first modeled by a 3D CAD software and converted to the STL file format. The STL file represents the part by approximating the part surface with many of small planar triangular patches. The STL file transferred from the CAD system is then converted to a slice file which consists of a series of layers at the RP workstation. When the slice file is ready with the appropriate parameters set for the machine, the part can be built by drawing each successive layer using various different techniques such as curing the photosensitive polymer or sintering the thermoplastic powder by a laser beam. The part built by RP processes usually needs post treatment. Sanding and polishing or coating is often used to obtain smoother surfaces.This paper deals with the accuracy problem in preparation of CAD data.

The data preparation stage consists of facetting and slicing. Facetting (also called tessellation) refers to the making of STL file format from a 3D solid model. The STL file approximates the surface of a solid model using many triangular facets. Each triangular facet is represented by the x, y, and z coordinates of the three corners of the triangle and its surface normal. The accuracy of a curved surface depends on the number of triangular facets used to represent the surface. Generally, smaller triangles allow finer resolution but increase the file size and slow down the processing time. The STL file format has been proven practical for many purposes but it has serious drawbacks due to the nature of the file that only contains unsorted, unconnected, and unrelated triangles. The file loses geometrical and topological information and cannot retain exact surfaces or features of the original CAD model.

The tessellated part model needs to be sliced by intersecting it using a horizontal plane. The thickness of each slice depends on the capability of the machine to make a layer, and the finite thickness of a layer inherently causes stairstep effect. In the case of SLS machine from DTM, the minimum allowable thickness is about 0.1mm. When a part is built without considering part features, large stairstep effect occurs, which usually results in errors such as the elimination peak features.

1.2 Prior Work

Several researchers have dealt with the data preparation problems discussed in the previous section. The prior research in this area can be classified into three areas: adaptive slicing, direct slicing, direct and adaptive slicing. Adaptive slicing uses variable layer thickness that considers the geometry of the part. It reduces the stair stepping error and also the model building time compared to the uniform slicing method. Dolenc and Mäkelä [1] presented an adaptive slicing algorithm that can handle flat areas, peaks, and stairstep effect while using the STL file. The layer thickness is computed by considering only the user-specified tolerance (or cusp height) and normal vector. Direct slicing, on the other hand, uses the exact contour of the CAD model instead of using the STL file. Since it does not involve the tessellation process, it also reduces data preparation time. Guduri et al. [2] proposed a direct slicing method that can provide more exact laser beam paths by slicing the CSG representation of a part. Vuyyuru et al. [6] directly sliced a solid model built using the SDRC's I-DEAS and segmented the NURBS-based contour curve. They could not prove the improvement in terms of dimensional accuracy but the NURBS based parts showed a better surface appearance. Adaptive slicing mainly attempts to reduce the z-axis errors while direct slicing tries to reduce the errors on the x,y plane. A few researchers investigated the use of both direct and adaptive slicing methods. Suh and Wozny [5] directly sliced the CAD model and determined the layer thickness considering the minimum vertical curvature at the sampled points and the given cusp height. They also suggested subregioning of the model in order to build peak features accurately. Jamieson and Hacker [3] determined the layer thickness by comparing the edges of the successive contours and slicing the layer by half when the error is bigger than the given tolerance. Kulkarni and Dutta [4] developed a slicing formula that considers the stair stepping effect and the containment problem for analytical models. The containment problem occurs since the RP machine builds each layer by curing a certain amount of thickness from a specific cross section of the model. The upper hemisphere of a sphere, for example, is built with the stairsteps created inside the original CAD surface and the lower hemisphere shows the stairsteps outside the surface. A part can be built with the negative tolerance in one area and the positive tolerance in others. This results in a skewed part substantially different from the original CAD model, which causes difficulty in post treatment. They proposed a slicing method that can generate either negative tolerance or positive tolerance consistently for a part.

2. A TWO-SIDED DIRECT AND ADAPTIVE SLICING METHOD

Existing direct and adaptive slicing methods often result in layers that are too thin to be built by the RP machine. This we call remnant layers. The z-axis accuracy of a model decreases as these remnant layers accumulate errors. In the case of the SLS machine by DTM, one remnant layer can be as large as 0.0762mm. In order to achieve precise dimensional accuracy in z-axis, these remnant layers have to be processed correctly. Existing methods mostly overlook this problem. Dolenc and Mäkelä [1] suggested the use of a float slice that can be adjusted to include the last thin layer. They discarded the last slice if the resulting layer is not thick enough and shifted all the slices or a subset of them in the given interval. The float slice can violate the cusp height between the subregions where the enforcement of tolerance is more critical than the other areas since they usually involve drastic slope or curvature changes. Discarding the last thin layer can cause inaccurate z-heights of peak features as well as the cusp height violation due to the shift of layers. The two-sided slicing method provides a better solution for this problem.

In addition, most of the existing methods do not deal with the containment problem. Kulkarni and Dutta proposed a method that can provide either positive tolerance or negative tolerance considering the post treatment. The post treatment of RP parts, however, is mostly to improve the appearance of a part and it would take considerable effort and time to finish a part with a complex geometry within a given tolerance. Furthermore, some small and/or internal part features allow limited access for finishing operation. In this paper, we are proposing a method that gives more accurate geometry using a mid-layer. The procedure used for the two-sided slicing method is described below.

2.1 Determination of Subregions

The stairstepping effect has resulted in a poor surface finish of RP parts. But it also affects immensely the manufacturing of peak features. These peak features can be incorrectly built or even disappeared if the slicing is performed with no consideration on them. Therefore, these peak features need special care in order to increase the accuracy of RP parts. In this research, subregions are created based on these peak features to avoid the slicing of these features. The peak features in the model such as horizontal face, horizontal edge and vertex are identified and they are sorted based on the z-height and subregions are defined accordingly.

2.2 Determination of the Thickness of the Next Layer

After the subregions are identified, the thickness of each layer in a subregion is determined based on Suh and Wozny's method [5]. The height of the next slice, Z_{j+1}, is determined from the contour of the current slice Z_j. Let C_j be the contour of the slice with the height Z_j. The first point P_i is selected arbitrarily and the next sampling point P_{i+1} is determined by approximating the contour by a circle with the curvature κ at the point P_i. A set of sampling points along the contour is calculated by repeating this same procedure. For each sampling point, the layer thickness is calculated based on the vertical curvature at each point. Among these values, the minimum layer thickness is selected for the next layer. The detailed equations for this can be found from Suh and Wozny's method [5].

2.3 Determining the Height of Mid-Layers

For each slicing plane calculated for the next layer, the height of the mid-layer is determined by lowering the slicing plane until it reduces the given cusp height by half. The calculation of the height of the mid-layer is as follows:

(1) When $\rho = \infty$

When the curvature $\rho = \infty$, the height of the mid-layer, ℓ_{intra}, becomes the half of the height of the next layer ℓ as shown in Figure 1.

$$\ell_{intra} = \frac{\delta}{2\cos\theta} = \frac{\delta}{2N_z} \quad \text{if } N_z \neq 0 \qquad (1)$$

$$\delta_p = \delta_n = \frac{\delta}{2}$$

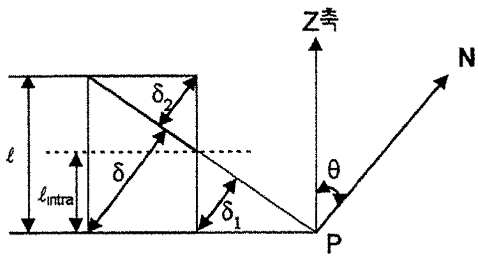


Figure 1. Calculation of the mid-layer for $\rho = \infty$

(2) When $\rho \neq \infty$

The upper hemisphere:

$$\ell_{intra} = -\rho\sin\theta + \sqrt{\rho^2\sin^2\theta - 2\delta'\rho - \delta'^2} \qquad (2)$$

$$\delta' = \delta_p = \delta_n = \frac{\delta}{2} - \frac{\delta^2}{4\rho}$$

The lower hemisphere:

$$\ell_{intra} = \begin{cases} \rho\sin\theta - \sqrt{\rho^2\sin^2\theta - 2\delta\rho - \delta^2} & \text{if } \rho^2\sin^2\theta - 2\delta\rho - \delta^2 \geq 0 \\ \rho\sin\theta & \text{otherwise} \end{cases} \qquad (3)$$

$$\delta' = \delta_p = \delta_n = \frac{\delta}{2} + \frac{\delta^2}{4\rho}$$

Figure 2 shows how the mid-layer is calculated when the layer is located at the upper hemisphere. A similar procedure applies to obtain the height of the mid-layer at the upper hemisphere.

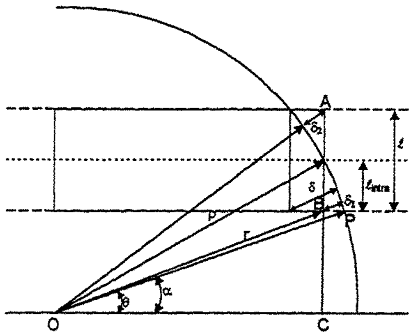


Figure 2. Calculation of mid-layer at the upper hemisphere

2.4 Handling of Remnant Layers

The remnant layers, when they appear between subregions or at the end of a part, can cause serious z-axis inaccuracy. The two-sided slicing method provides a better solution for this problem. When the algorithm begins the next slicing plane is calculated from both the bottom and the top of a subregion respectively. Since the next slicing planes are generated from both directions, the algorithm needs to carefully handle the cases when the downward slicing plane meets with the upward one. The algorithm proceeds with considering the following cases:

Case 1: When the upward and the downward slicing planes do not cross each other - it continues to calculate for the next slicing planes.

Case 2: When the next slicing planes cross each other while being located between the current slicing planes - it takes the shorter layer from the upward or the downward plane and continue.

Case 3: When the next slicing planes coincide each other - the slicing process terminates with no remnant layers.

Case 4: When a next slicing plane violates the current slicing plane from the opposite direction - when a next slicing plane is located between the current slicing planes, it is taken as the next slicing plane, but the one violates the current slicing plane is discarded.

Case 5: When both the next slicing planes violate the current slicing plane from opposite directions - it is treated differently based on the thickness of the remnant layer, r.

1) when $\ell_{min} <= r <= \ell_{max}$

Since r is smaller than the next slices calculated from the current slicing plane, the cusp height of this remnant layer should be smaller than the user-specified cusp height δ. Since it is a manufacturable thickness, it is taken as the last slice and the slicing algorithm terminates.

2) when $r < \ell_{min}$

Since r cannot be built with its own thickness, it needs to be handled with the neighboring slices as follows:

(a) $\ell_{u,previous}$ or $\ell_{d,previous} \geq 2\ell_{min} - r$

Merge with the thicker slice from the neighboring upward or downward slice and divide the thcikness so that the thickness of the remnant slice, r, becomes ℓ_{min}

(b) $\ell_{min} \leq (\ell_{u,previous}$ and $\ell_{d,previous}) < 2\ell_{min} - r$

(b-1) when $\ell_{u,previous} + \ell_{d,previous} + r \geq 3\ell_{min}$

Adjust the neighboring slices so that the remnant slice, r, becomes ℓ_{min}.

(b-2) when $\ell_{u,previous} + \ell_{d,previous} + r < 3\ell_{min}$

The sum of the neighboring slices with the remnant is smaller than 3 ℓ_{min}, hence r cannot be increased to ℓ_{min}. The remnant slice is eliminated by dividing the sum by the two neighboring slices.

3. CONCLUSION

The exisiting slicing methods showed a containment problem as well as the improper handling of remnant layers. These problems need to be corrected in order to make more accurate RP parts especially to improve the accuracy in z-axis. The two-sided adaptive and direct slicing method proposed in this paper reduces the cusp height error by half and also treats the remnant layers properly. This will improve the geometrical accuracy of part features that are primarily sensitive in z-axis.

REFERENCES

1. Dolenc, A., and Mäkelä, I., "Slicing Procedures for Layered Manufacturing Techniques," *Computer-Aided Design.*, Vol. 26, No. 2, pp. 119-126, February 1994.
2. Guduri, S., Crawford R.H., and Beaman J.J., "A Method To Generate Exact Contour Files For Solid Freeform Fabrication," *Proceedings of Solid Freeform Fabrication Symp.*, 3-5 August 1992, The University of Texas at Austin, Austin, Texas, pp. 95-101, 1992.
3. Jamieson R. and Hacker H., "Direct Slicing of CAD Models for Rapid Prototyping," *Internet Conference*, 1995.
4. Kulkarni, P., and Dutta, D., "An Accurate Slicing Procedure for Layered Manufacturing," *Computer-Aided Design.*, Vol. 28, No. 9, pp. 683-697,1996.
5. Suh, Y.S., and Wozny, M.J., "Adaptive Slicing of Solid Freeform Fabrication Processes," *Solid Freeform Fabrication Symp. Proc.*, 8-10 August 1994, The University of Texas at Austin, Austin, Texas,1992, pp. 404-411, 1994.
6. Vuyyuru, P., Kirschman C.F., Fadel, G., Bagchi, A. and Jara-Almonte C.C., "A NURBS-Based Approach for Rapid Product Realization,"*Fifth Int. Conference of Rapid Prototyping*, June 1994.

Rapid prototyping on a large scale

I. Gibson and G. Mensing

ABSTRACT

Rapid prototyping moves more and more into mainstream engineering, for instance in the form of rapid tooling. Many tools are not of the small size that can be made by current rapid prototyping systems. This means there is a market for large scale rapid prototyping. Rapid prototyping has proven that cutting down on lead time is very profitable, which also applies to larger prototypes.
To build larger prototypes the current rapid prototyping systems are not big enough or fast enough. This is because they use layered based manufacturing where the layer thickness can not be increased beyond a few tenth of a millimeter. Milling is an option that allows the use of thicker layers. Milling can do contours very accurately and with present day milling machines and tools this can be a competitive alternative for rapid prototyping, especially because of the wide variety of materials that can be milled.

KEYWORDS

large scale, rapid prototyping, milling, adaptive slicing.

1. INTRODUCTION

Rapid prototyping (RP) is the process of automatically converting computer models into physical solid models. Rapid prototyping is a CAD/CAM process that makes it possible to fabricate models independent of complexity of form. Models containing many features with complex, multi-directional contours will take as long to make using rapid prototyping as a simple prismatic block of equivalent volume and external dimensions. Processing times are such that simple parts are made slowly (compared to conventional machine-shop technology) whilst complex parts are made quickly. The key to the process is the deconstruction of the original CAD model into layers. Each layer is a prismatic approximation of the volume around the cross-section of the original model. Rapid prototyping is successful because the thickness of these layers is very small. In fact, much research and development has focused on making these layers thinner. The majority of systems are designed to produce components 250mmx250mmx250mm or smaller. Larger systems are available but even the largest machine is incapable of making parts greater than $1m^3$. There are many applications that would benefit from making parts even bigger, utilising the basic, layer-based concept of rapid prototyping to make complex forms quickly. This is particularly true as rapid prototyping becomes more widely used for tool making purposes, where many tools can be of a much larger dimension.

2. CURRENT SYSTEMS

Some of the major systems available are 3D-system's StereoLithography Apparatus (SLA), DTM's Selective Laser Sintering process(SLS), Helisys' Laminated Object Manufacturing (LOM) process and Stratasys' Fused Deposition Modelling (FDM). These and other RP-companies are continually doing research to improve their systems. Research has concentrated on accuracy of build, material properties and speed of build. Although the speed of build is improving, there is a limit to how fast a part can be built. The only logical conclusion for further improvements in speed would mean increasing the layer thickness. For straight parts thicker layers can be used, but for slopes and profiles the step-effect of thicker layers would be undesirable.

In an attempt to illustrate what would happen if these systems were larger, table 1 shows the time needed to build one 100x100x0.1 mm layer. This is then extrapolated to 1 m^3 using 0.1 mm layers. Note: These are not accurate times, just estimates.

Table 1. Calculated times for current systems.

	Maximum build envelope LxWxH mm^3	**Time needed for a layer of 100x100x0.1 mm^3 (approx.)**	**Extrapolated time for 1x1x1m^3**
DTM S.S. 2500	381x330x424.	laser sintering + adding powder = 10 + 20 = 30 sec.	(laser sintering + adding powder) number of layers = (1000 + 60)10,000 = 122 days
3D Systems SLA-500	500x500x584	Drawing time + re-coating time = 30 + 60 = 90 sec.	Laser-speed 5m/sec. Max beam ∅ = 0.25 mm. This is 800 sec/layer. 800x10000 = 92 days
Helisys LOM-2030H	813x559x508	Paper 200x200 mm. Cutting speed = 500 mm/sec. Cutting + paper feed + pressing = 42 sec. (Including cross hatching)	Paper 1100x1100 mm. (Cutting + paper feed + pressing) x number of layers) = 16 days. (Including cross hatching).
Stratasys FDM -1650	254x254x254	25.4^3 mm^3/hr. means for the layer : 220 sec.	2500 days. With Genisys 1250 days. Using zig-zag 125 days (estimate).

The data in table 1 shows that RP is not quite so rapid if used on a large scale. The data are extrapolations from times/layer or times/volume given by the companies which developed the systems. With Stratasys, when using the Genisys machine and a zig-zag fill instead of a solid material (i.e.with air gaps) the build time is improved. For solid material the rule of thumb is 1 cubic inch per hour. The Helisys system is much quicker than the others due to vector scanning rather than raster scanning. It is difficult to estimate the complete Helisys time, because the outer cross-hatching of waste material is a function of the part geometry and the remaining space. If the cross-hatching were restricted to the vicinity of the part the time required would be more predictable.

3. CURRENT APPROACHES TO LARGE SCALE PROTOTYPING

Some systems are already available for large scale prototyping. In the Netherlands, Delft Spline Systems [1] is using a program called DeskProto to calculate milling paths from 3D CAD models. First the parameters (accuracy, tools, etc.) are given to the program. Next DeskProto calculates the milling path and then a robot mills the object. Because a robot is not a rigid machining system the accuracy of the final part can not be expected to be high. Furthermore the materials used are soft (foam., plastics, etc.).

In Australia, at Queensland University [2] a large scale, layer based rapid prototyping is used. By using a water jet cutter the features are cut, similar to the Helisys system. The layers are glued together manually. A problem here is that because thick layers are cut, there is significant 'stepping'. This is reduced by cutting at an angle, although these are still straight cuts. The material used is foam.

A similar system has been developed at the University of Utah [3]. Thick layers of thermoplastic foam (1/4" to 2" thick) or paper are cut with a knife and glued together manually. This is commercially available as the Shapemaker.

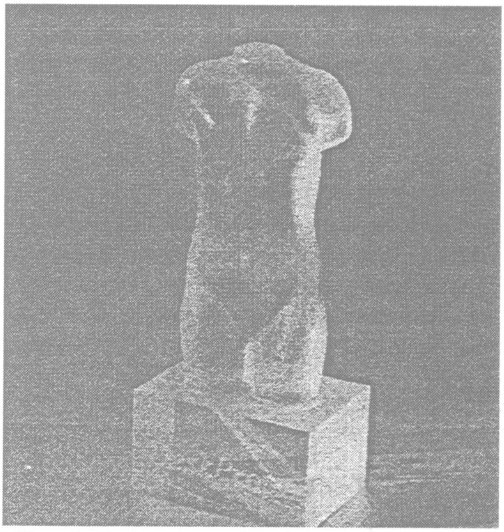


Photo 1. The human torso here is 90 mm. high. It was milled in about 10 hours. This time is similar to what a layer based RP system would use. Delft Spline Systems

Also in America, Formus [4] is using a very different approach. They use a two step shell process, rather then a solid direct process. This Topographic Shell Fabrication (TSF) process prints the outer surface of a shape using molten wax in a bed of sand. When done, the residual sand is drained, and this gives a "TSF shell". The shell gets scraped, surface waxed, and polished before fiber-glass lay-up is done (expanded foam, plaster, or concrete can be used as well, depending on the application). Most applications appear to be in marine, aerospace, some automotive, concrete moulds, film-sets, fiberglass.

Ford Research [5] put together a project team from design, machining, manufacturing, engine/process, materials systems reliability, and testing departments to produce what was termed the "100-Day Engine" process. CAD data was used to create layers for machining which were then joined by vacuum brazing. Advantages included development artefacts that could be sliced apart and modified for in-process modifications. Prototype engine heads have been made that have achieved 360 hours in durability runs. However, this process was developed for a very specific project with little attention to extending the range of applications.

There is scope therefore for the development of a RP system for large scale prototypes using a hybrid of layer based techniques and conventional machining technology. This would in effect be a more integrated and flexible version of the Ford project.

4. MILLING

Milling is a conventional technique that can do contours and slopes. This means with milling it is possible to use thicker layers without 'stepping' effects. Milling machines today exist with multiple axes, high power (up to 150 hp), high feed rates (10 m/min, depending on the material), advanced software techniques ('LookAhead') and new cutting tools (CBN = Cubic boron nitride), This makes milling a competitive alternative to RP.

The one thing that makes rapid prototyping so flexible is the layer based manufacturing. This principle has made it possible to make complex parts just as easy as simple parts.
Multiple axes milling machines and robots have many more degrees of freedom compared to the virtually two dimensional manufacturing of rapid prototyping machines. If there are internal features then a process of component-based construction needs to be used (e.g. the use of layers). Accuracy is also an important parameter for the type of milling machine used. If the accuracy is worse than 0.1 mm then a robot can be used. Otherwise a multiple axes milling machine should be used due to its rigidity.

5. THE CAD MODEL

Milling can be used as part of the hybrid technology and it can be used on thicker layers. What has been suggested is a hybrid system, combining the flexibility of layered manufacturing with the processing speed of machining. In order to handle internal features and complex geometry, the original CAD model must be sliced, but with more intelligent software than is currently available. The slicing algorithm must include knowledge of the following constraints:

- machine/tools available
  - degrees of freedom
  - accuracy
  - speed
  - milling tools
- slices available
  - slices (variable thicknesses) available in stock
- features of the model (holes, curves, surfaces, etc.).

This suggests a generic approach that can be adapted to suit available machining technology. Before slicing can take place a choice has to be made on which machine is going to be used for the processing. The machine chosen will have a specified number of degrees of freedom, capable of processing a limited geometry. This geometric limitation will determine the complexity of feature that must be processed and therefore where the part must be sliced in order to reduce this complexity to within the machine capability. Other factors may also limit the choice of machine, for example the desired accuracy of machining, the milling speed and the number of tool changes.
The proposed algorithm for the CAD slicing must know the slice thicknesses available in stock. One possible solution for what is essentially an adaptive slicer [6] is to start with very thin slices and then increase the layer thickness in the algorithm to calculate the maximum thickness possible, depending on the slices in stock. The adaptive slicer must also calculate the times needed for every slice being milled and this way come up with the minimum time needed for the whole milling process.
There follows a description of the proposed process for layer-based machining.

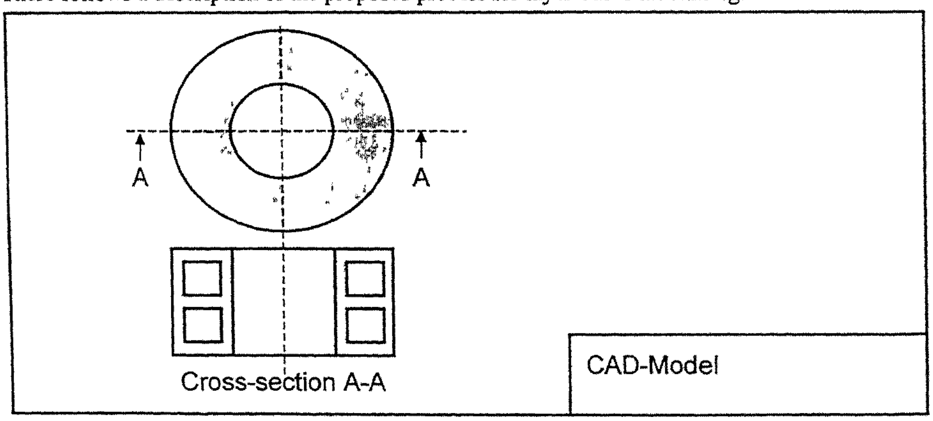


Figure 1. The CAD model.

The simple CAD model of Figure 1 will be sliced by the software depending on the constraints. The independent layers are then sent to the milling machine.

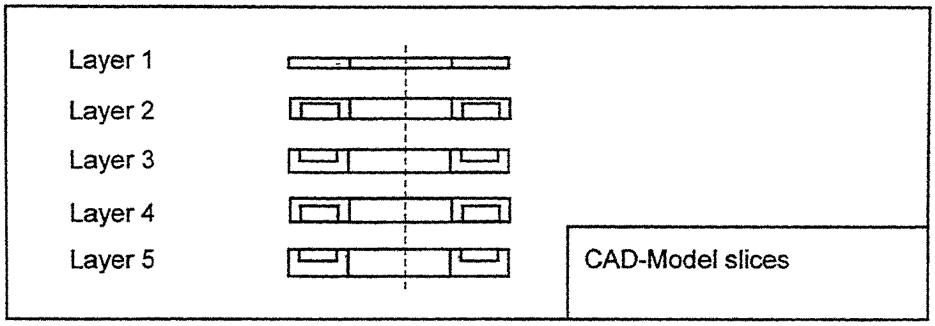


Figure 2. The sliced model.

The milling machine mills the interior of the CAD model in the slices of the material.
Only the inside of the model is milled because this way it is easier to fix the sheets of material on the milling machine.

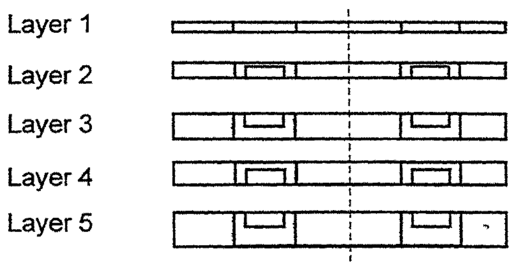


Figure 3. The internal features are milled.

If all the slices are milled then they are stacked and 'glued' together so that the internal features of the object already exist.

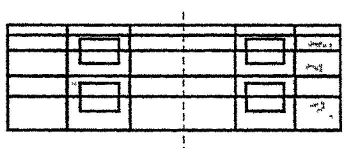


Figure 4. The sheets are stacked and 'glued' together.

Finally the outside features are milled. Depending on the object the 'block' has to be turned around so that the other (bottom) side(s) can be milled

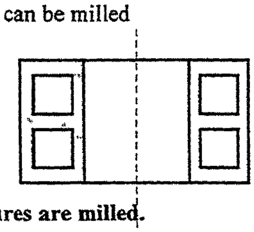


Figure 5. The outside features are milled.

6. CONCLUSION

This paper shows that for large scale modelling, conventional rapid prototyping machines cease to become rapid. Whilst some of the processes may be able to process individual layers quickly, the mere fact that these layers are very thin makes large scale modelling difficult (or impossible) to achieve in an acceptable time. The suggestion is that there is a transition in viability from layer additive techniques for small complex models to layer subtractive techniques for larger complex models. This is supported by the Helisys, LOM process showing high speed for large model processing - LOM being a hybrid additive/subtractive process.

Some large scale rapid prototyping systems are available or under development. Those systems known about do exhibit certain limitations in that they are only surface based, or are inaccurate when compared with the original CAD model, or are only semi-automated and dedicated to very specific applications. We therefore propose an automated layer based machining approach which could be generic to all milling processes. Its features will include:

- The ability to do large scale. Milling machines exist up to the size of bowling alleys.
- The ability to handle internal cavities. Because layers are being used it is possible to do internal features.
- A wide variety of materials from prototyping foams to metal.
- Automated software which is based on adaptive slicing with machine and layer thickness constraints. The software determines the slice thickness depending on the machine, tools, accuracy, degrees of freedom, stock, etc.
- Automated process of internal and external machining. The whole process of creating the slices, milling the inside, gluing the layers together and milling the outside could be automated.

REFERENCES

[1] Delft Spline Systems: http://www.spline.nl/

[2] University of Queensland: http://student.uq.edu.au/~s315413/research.htm

[3] University of Utah: http://stress.mech.utah.edu/home/novac/mpl.html

[4] Formus: http://www.agathon.com/formus.htm

[5] Press release: Ford Research, Ford Motor Company The American Road, Rm 904, Dearborn, Michigan 48121, USA. Conference: Integration of FEA and Structural Testing with Rapid Prototyping (IFSTRP) October 30-31, 1995, The Hyatt Regency Dearborn Hotel, Dearborn, Michigan

[6] Voulelaud, T. and Bagchi, A., "Adaptive lamina generation for shape dependent process control and or object decomposition", Patent Number 5,432,704, 1995.

RESEARCH OF FEATURE CONVERSION AND REPRESENTATION OF PRODUCT PROCESS INFORMATION MODEL

Jian Gao, Detao Zheng, Ping Zhang, Guowang Zhu and Jian Sun
Department of Mechanical and Electronic Engineering
GuangDong University of Technology, GuangZhou, China
Xiankui Wang
Qinghua University

ABSTRACT

Feature models are domain dependent. The design feature model is not satisfied the requirement of process planning. In order to get manufacturing feature automatically, a feature conversion module is needed. It is organized into three_step process functions: filtration, selection and aggregation. Through these functions, the attributes useful to form manufacturing features can be selected and aggregated by some matching rules. and the manufacturing features can be converted finally These manufacturing features will be the basic unit of process information model. After an analysis of function view and information view about the process information model, a binary tree data structure is proposed to represent it. The process information model can be directly referenced by applications of process design, fixture design, etc.

KEYWORDS

Feature Conversion, Process Information Model, Binary Tree

1. INTRODUCTION

Feature can be used for different applications to capture the engineering signification of a part model. In process planning for manufacturing, feature has been used to identify areas in a product that can be manufactured with a specific sequence of machining operations. In design stage, design feature allows the designer to model objects with elements that are on a higher level, and closer to his way of thinking, than the lower level geometric elements used in solid modeling. Each application has its own view to describe the object. An example is given in figure 1. It is a workpiece with a number of ribs. In design view, three protrusions can be considered as ribs features. The workpiece can, however, be manufactured not by attaching ribs to the workpiece, but instead by removing the spaces between the ribs by milling operations. So in the process planning view, the slots between the ribs will be considered as features. From the example, it can be seen that one geometric element can be considered as different features in different views, further that different combinations of the same geometric elements can be considered as different features in different views.

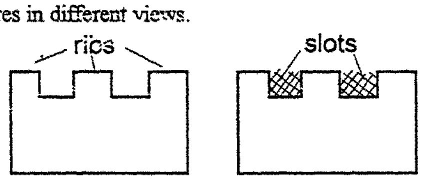


Fig.1: Two different features.

Product information modeling using feature representation can not only support product data needed by different applications, but also provide with high_level description terms according with human's thought[1]. As mentioned above , for different applications, different features are required. Each application has its own set of features, such as design feature sets and manufacturing feature sets. Therefore, there is a need for automatic feature conversion. This feature conversion is defined as the process of converting features from one application into another application for the same component. It is of particular importance for the integration of CAD/CAPP system.

In this article, the design feature model of the component is constructed by the Pro/Engineer feature modeling system. Based on this model, and in combination with a feature conversion module, application_specific feature models can be driven. The feature conversion module is illustrated in section 2, an application_specific feature model --process information model is discussed and described in section 3, an application of the model is introduced in section 4, and the conclusion is in section 5.

2. FEATURE CONVERSION MODULE

In the environment of Pro/Engineer feature modeling system, a feature conversion frame is proposed in the article , which is illustrated in figure 2. It can be used to convert design features provided by Pro/Engineer system into manufacturing features needed by the application of process planning. This feature conversion module is organized into three_step process functions: filtration, selection and aggregation, in which filtration function is used to filter information about face sets of features out of the design feature model built in Pro/Engineer system; selection function is used to select faces and their attributes which need to be machined, and these attributes are obtained by the functions provided by the Pro/Develop module; aggregation function is used to aggregate faces and

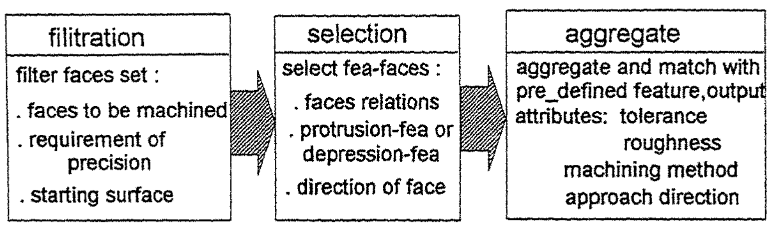


Fig. 2: feature conversion with three_step functions.

matches those faces with pre_defined manufacturing features which are represented in process information model.

Through feature conversion module, design features can be converted into manufacturing features, these manufacturing features will be the basic units of process information model. Although STEP standard used in product data exchange has been defined, the standard of process information model still haven't been completely defined. The requirements of process modeling and the structure of process information model will be next discussed in details.

3. PROCESS INFORMATION MODEL

3.1 Requirement of Process Modeling

Process information model will be the kernel model of process planning system, so the function view and information view about the system are analyzed firstly.

3.1.1 Function View

Function view is used to analyze function and relation of CAPP system, and identify the necessary information and methods to realize these functions. Figure 3 shows the function of process planning which is represented by IDEF0(a kind of information modeling language).

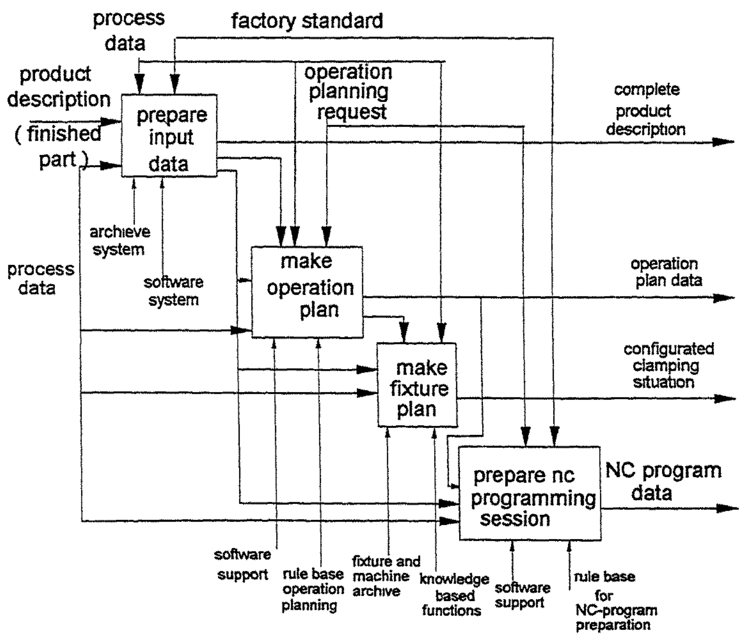


Fig. 3: Function view described by IDEF0.

3.1.2 Information View

The aim of analyzing data information is realizing information modeling between functions. The flow of information about process planning can be divided into three parts: product data, production activity data and equipment data. Product data comes from CAD system. Production activity data is involved in all data about process design. Equipment data includes the data of machine、 fixture、 tool and operation, which forms the possibility of product manufacture.

According to above analyzing, process information modeling method should satisfy three basic requirements:

· represent real information related to process design;

· convenient for keeping product process information in the archives;

· provide the decision module of process design so as to support the representation of design change and improve the flexibility of product design.

3.2 Process Representation Model

3.2.1 Composition of Model

On the basis of above analyzing about process planning system, a process information model is proposed, which includes three kind of information: component level information, feature level information and feature relations. The detailed contents of the process information model are illustrated in figure4.

	component level info	feature level info	feature relations
process inform-ation Model	identification feature bill material heat_treatment batch stock	feature attributes: dimension form tolerance coordination identification feature type: basic feature process feature combined feature duplicated feature process plan fragment	geometric relations position dimension position tolerance datum feature

Fig. 4: The contents of process information model.

In the process model, component level information is used to describe the management information about finished component. Manufacturing feature is the kernel unit of the model which can be divided into four types:

1) basic feature: it is used to represent the basic geometry of the component, such as block feature ;

2) process feature : it is used to represent the stanarded process structure which can be clearly described by several simple parameters, a typical example is A_type centric hole with thread, and as a process feature, the parameters of standard code and diameter are enough;

3) process duplicated feature: it is used to represent the regular combination of the same process features, such as eight holes distributed equally on a circle are served as a duplicated feature with the attributes of diameter and numbers of hole.

4) function combined feature: it is used to describe more complicate component. It contains some process features, and has more regular machining method.

In process information model, features are dependent on each other. There are several relations between features: father_child relation, position relation, dimension and tolerance relation, datum, etc. These relations are more useful to the application of process planning, which will be considered in the problem of interface between tool and workpiece and used to determine machining sequence , etc.

3.2.2 Binary Tree of Model

In order to represent and operate conveniently process information model , a binary tree data structure is used to describe this model . An example is shown in figure 5(a). On the figure, the component has a basic feature(block feature), which is described by six orientation faces. On its top_face, there are three process features (hole1, hole2, slot) which are also described by their own faces. These features should be child features of the top face , this father_child relation can be clearly described in binary tree structure. The binary tree of this component can be represented in figure5(b).

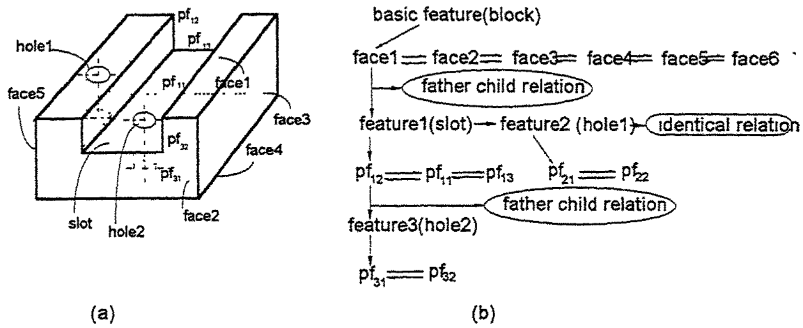


Fig. 5: An component example and its structure.

4. APPLICATION OF PROCESS INFORMATION MODEL

Once the process information model is constructed, some applications such as process design, fixture design can be implemented on it. Based on this model, a structure of process design module can be established. It adopts a kind of muti_layers and muti_decision way to produce process plan of a component. The detailed structure is illustrated in figure 6.

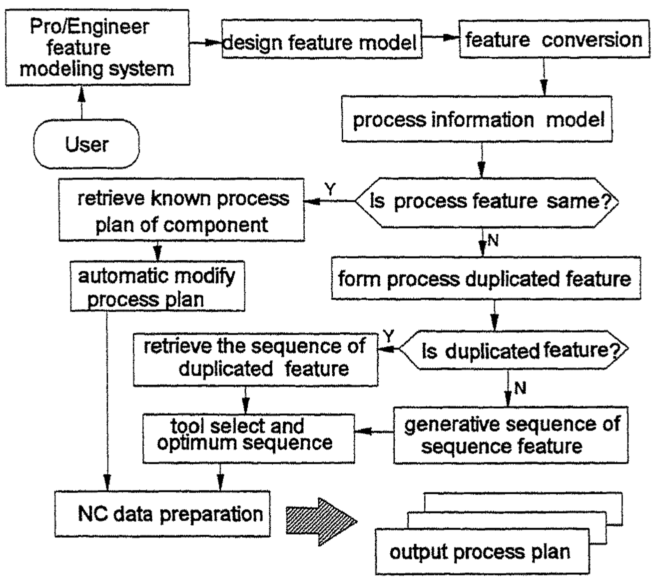


Fig.6: An application of process model

5. CONCLUSION

It is known that the Pro/Engineer feature modeling system doesn't have the function of automatic CAPP. Its manufacture module can not realize process automation. In order to realize automatic process planning system and CAD/CAPP integration , a feature conversion module and a process information model must not be absent. The system is implemented in C language on Sun Sparc20 workstation.

6. FEFERENCES

1. Shah,J.J. , Hsiao,D. and Leonard,J., A Systematic for Design Manumacturing Feature Mapping, Geometric Modeling for Product Realization,pp.205-221, IFIP,1993.

2. Detao Zheng, Jian Gao, Study on Mechanism of Feature Mapping and its Application in Domains between Product Design and Manufacturing ,Chinese Journal of Mechanical Engineering, pp.63-67, Vol.32, No.4, Aug. 1996.

3. Bronsvoort,W.F., Tansen, F.W., Feature Modeling and Conversion — Key Concepts to Concurrent Engineering , Computers in Industry, pp.61-86, Vol.21, 1993.

Research on Customer-Oriented CAPP System Modeling Technique of Prismatic Parts Based on Feature

GAO Juhong and Prof. XU Yanshen
Department of Mechanical Engineering
Tianjin University
China

ABSTRACT

The Modular Flexible Manufacture Cell (MFMC) system is developed to meet with the various demands of customers. This paper describes the design scheme of an intelligent system from product model to manufacture cellar. Agilely providing design scheme is the necessary factor for the construction period. This paper proposes a subsystem Customer-Oriented CAPP of MFMC system. On the CAD/CAPP/CAM integrated environment, the Object-Oriented technique is applied to build the unified product definition model. The model based on the abstracted feature attribution and classification of the prismatic parts machining information is developed in this subsystem. The process planning unit model is formed by Work Element (WE), has unified layers on design and manufacturing. The adjacency list that the digraph of weighted has been established. The topological sort and manufacture resource's knowledge library are used to generate the WE sequence. The principle for drawing up the process planning and the balance for the process meters both are the contract factor for the formation of foundation model of manufacture units.

KEYWORDS

Feature Technology, Product Definition Model, Work Element, Process Planning

1. INTRODUCTION

The customer-oriented CAPP system modeling technique has been one of the key means of research & development on the manufacturing technology and equipment of sedan engine key parts and components. By means of advanced design and manufacture technology and in the lights of the demand of standardization, seriation and generalization, research and develop component modules of machine tool for prismatic parts, make up various functions modular machine centers and special purpose NC machine tools, so that according to technology specification, production batches and variety of manufacturing customer's part provide rapidly and reasonably design schedule, lay a foundation for manufacture.

The MFMC (Modular Flexible Manufacture Cell) system is developed to meet with various demands of end-users. The construction period from product model to manufacture cellar is the necessary factor for agilely providing design schedule.

To solve these problems, the formation of the integration of process planning is considered as constraint condition on meeting the customer's demands and aim based on assignment of the equipment in the workstation, which addressed in component module's base of machine tool. Instead of the process planning module according to existing static shop floor situation, the method is provided that it was too difficult to concern the impact on the overall performance of the production aspects by a process planner. CLPP[1], RTCAPP[2]. Although these models can not solve the problem well, they are good enough to spark light for trends of future research.

The main concern of this paper is to show a structured approach of integration to developing a process planning unit model associated with the three level function models, including part level, process planning level, and operation level at first, and then a method of the process planning is discussed in detail.

2. THE PROCESS PLANNING UNIT MODEL FOR PRISMATIC PARTS

The process planning unit model is considered as both feature model and process planning model. Feature, in its general sense, has a wide variety of definitions. The feature has different perception in domain-design and domain-manufacturing. The feature is difficult to unite about function-oriented standard descriptions of geometric and manufacture-oriented feature information useful to design. It thus only provides a one way communication but fails to aid the information feedback from manufacturing to design. After then, the issue has been frequently addressed in literature[3][4]. They point out emphatically that feature must be treated as characteristics of design entities with intrinsic connection to manufacturing activities, but not only design oriented. These methods based on available resources and small batch manufacturers of mechanical parts.

For customer-oriented and large batch manufacturers of sedan engine key parts and components, in our opinion, feature attribute must based on user's need and be restricted by the decision of process planning. On the basis of the feature attribute and the similarity of the process, the integrated process of the process planning model for prismatic parts is established as illustrated in Fig. 1.

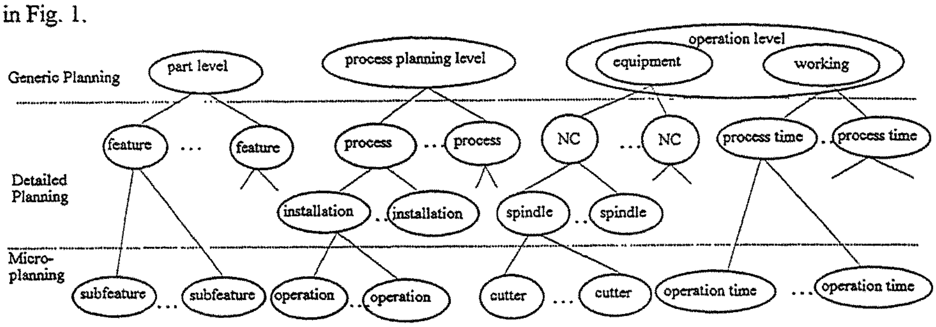


Fig. 1: An hierarchical process planning model of design/manufacturing integration.

This model embodys the communication here between part model and process planning model, that is, feature analysis and process planning information relationship is depicted graphically in Fig. 2.

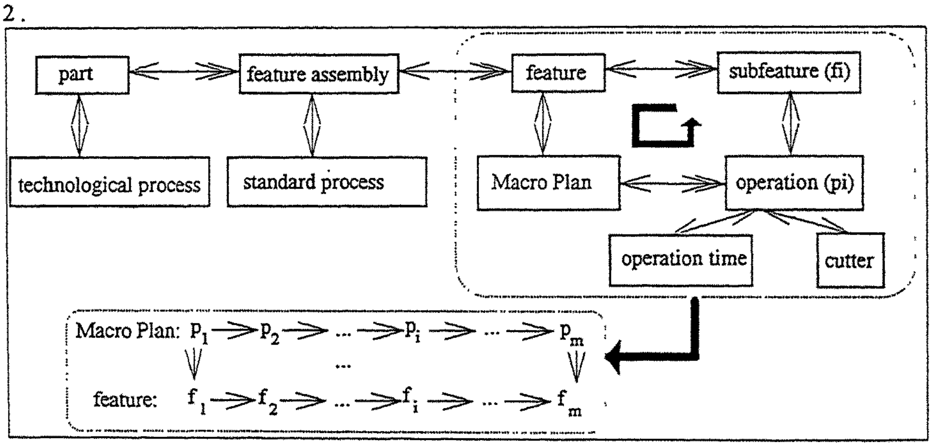


Fig. 2: Part & process planning relationship model.

On the basis of Macro Plan of components and with a corresponding feature, we have built a relationship one by one, which integrated operation (pi) with subfeature (fi), link with a certain cutter and operation time that machining subfeature on the part feature. For instance, if there is a hole in the part side, and finished completely, it needs drilling and reaming that machining operation twice.

In the process which drilling corresponds to the surface part, called "subfeature (f1)," and reaming corresponds to the hole of the surface part, called "subfeature (f2)". The hole as a feature (F) composed of f1 and f2. The problem can then be formulated as follow:

$$F=\sum_{i=1}^{m} f_i$$, where m represents the operation number of Macro Plan.

To achieve the above process, it can be defined as " Work Element—WE "(As shown in Fig. 3 .)

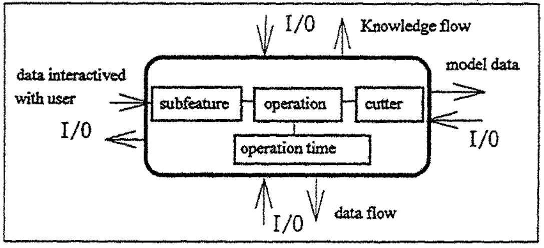


Fig. 3: Work Element—WE.

The WE represents the base-level concept or terminology, integrated feature in the component and process planning and all procedures that can be machinable in NC lathe. The WE see through the appearance to the essence for manifold process planning in the forms and the compose, and depict the intrinsic connection to manufacturing activities. For the approaches of integration, we can disperse the two complicated "object," which the component and manufacturing activities.

Moreover, The WE communicate with flow of external information in order to ensure reliability of the formation of the fundamental unit.

According to this form of fundamental unit, more and more WE fabricated in flexible manufacturing system/cells. So the limited and orderly part features assemble constitute a sequence of nodes based on the WE form. Under the same condition, these nodes correspond to a sequence of machining operation in manufacturing sense. When a process planner uses the WE to model a component both process planning and selection the equipment in all procedures could be generated along the way.

3. THE INTEGRATED APPROACH OF USER-ORIENTED CAPP SUBSYSTEM

On the basis of the above the WE model, a customer-oriented CAPP subsystem used for prismatic part is developed via C++ language. The subsystem has been put in operation of the MFMC system. The subsystem can complete feature extraction and the process planning based on the determination of equipment condition of every procedure in manufacturing, such as the type of spindle and working table and cutter tool base or cutter. The structure of the subsystem is show in Fig. 4. It consists of three models, a feature extraction model, the WE assembly model and a sequence procedure's model.

3.1 Product Definition Model For Prismatic Parts

3.1.1 Feature Extraction

To the process of feature extraction, there are two steps as follows:

(1) Geometric Entity Extraction.

(2) Feature Extraction.

First, prismatic components with exterior features can be represented in one view by two kinds of drawing entity, namely plane (including groove/slot) and hole. Thus, based on need of customer's machining operation, extracting corresponding geometric entity. The entities that have been extracted , not only including geometry data but also machining specification such as surface roughness, manufacturing tolerance that are necessary to process planning, and stored in the data structure are re-ordered according to certain machinable surface of the part.

Next, from the extracted geometric entities, feature that belongs to a redefined set of primitive feature and compound feature must be constructed. Four criteria for specifying the feature taxonomy of prismatic components are presented as follows:

(1)Feature taxonomy is understandable and usable by process planner.

(2)Extracting feature, it is required to generate the process planning sequence efficiently.

(3)Different feature must directly response to corresponding machining cutter, including standard cutting tool and compound cutting tool.

(4)In CAD/CAPP/CAM integrated environment, the general mode of production of the NC lathe machining must be considered, including single or several spindles and the mode of working table and cutter tool base allocated to NC lathe machining.

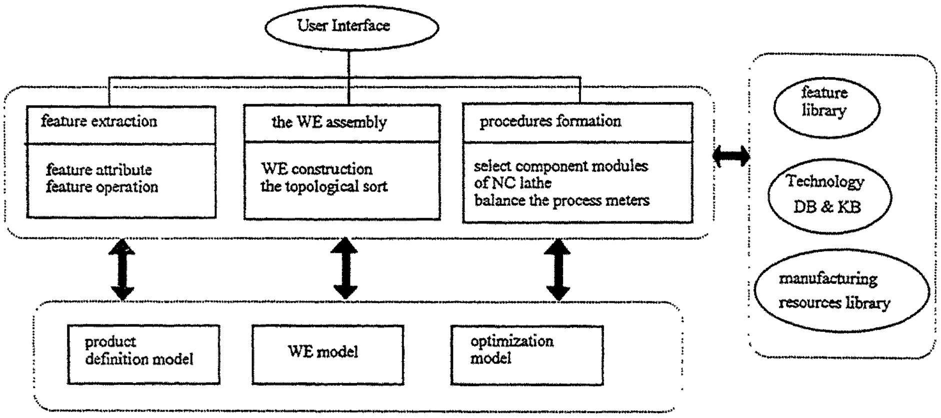


Fig. 4: Subsystem structure.

3.1.2 Product Definition Model For Prismatic Parts

By means of the analysis of feature for prismatic parts, we have built a generic product definition model using Object-Oriented system analysis method. A part is diffident as a hexahedron of function blocks according to given spatial or layout description about the blocks. A function block is composed of features based on user's demands. The case of prismatic parts is shown in Fig. 5 and Fig. 6.

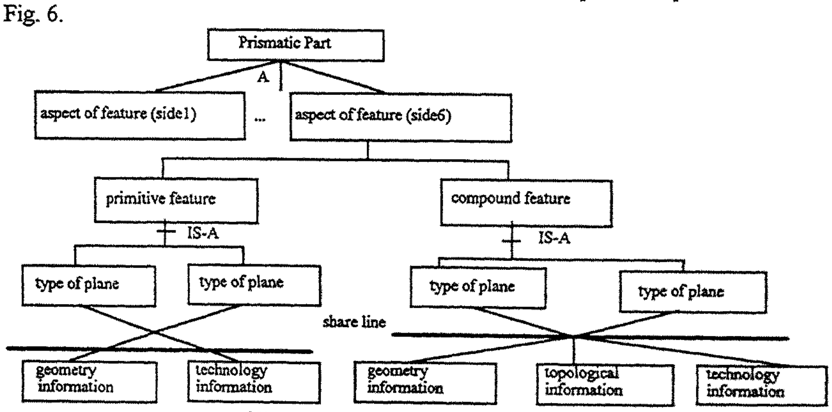


Fig. 5: Product definition model for prismatic part.

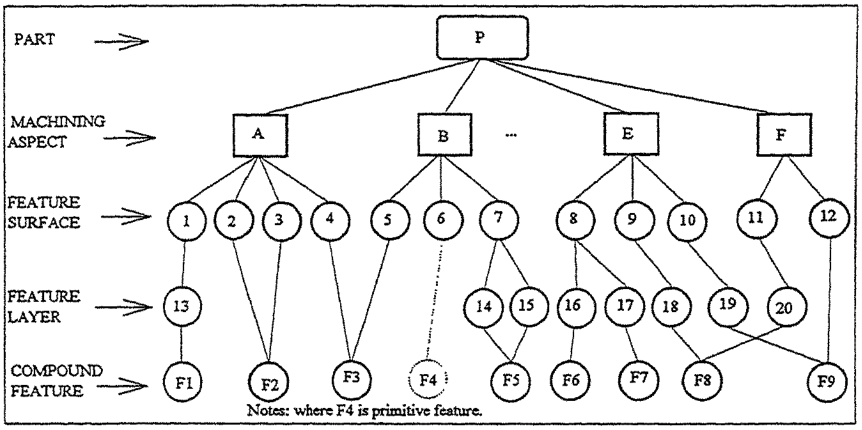


Fig. 6: Topological relation graph.

3.2 The WE Assembly

For the WE assembly, there are two steps as follows:

(1)The adjacency list that the digraph of weighted has been established.

(2)The Topological sort.

The first step is according to manufacturing tolerance, we have built a direct graph of weighted, which weighted represents manufacturing tolerance, each coloration node represents a feature and which located a surface of the part side. With an example explaining, the structure of the digraph is shown in Fig. 7.

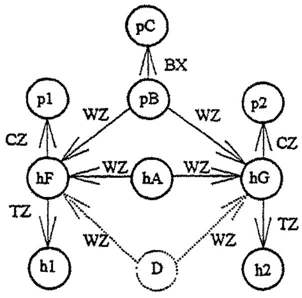


Fig. 7: The weighted digraph.

Notes: where BX represents //, CZ represents ⊥, WZ represents ⌖, and TZ represents ◎. A, B, D, F, G represents a datum (⚲) and is also a feature, and e.g. the direct of arrow point explained that h1 is based on hF and has ◎. The prefixes (such as p and h) represent type of plane or hole. The p1, p2, h1, h2 represent separately others general plane and hole. The datum "D" represents a datum line of no machining (showing with dotted line).

Then each feature as a node constitutes the adjacency list set defined in Table 1.

head node

vexdata	indegree	firarc

list node

adjvex	info	nextarc

Notes: where info reprensents weight such as WZ, BX etc

Table 1: The adjacency list.

The second step is According to the principle of the topological sort and the decision of process planning, and based on relationship of machining operation of every feature. Finally, we gained a sequence of node, that is a sequence of feature

pB → pC → hA → hF → h1 → p1 → hG → h2 → p2,

correspond to a sequence of WE, that is $\{WE_{11}, WE_{21}, WE_{mn}\}$, where n represents the number of features in the part.

3.3 A Working Procedure Formation Model—The WE Combinations

In the implementation of process planning function system, according to characteristic and arrangement in this system, the four matrixes of which operation sequence (P), subfeature sequence (F), cutter sequence (D) and operation time sequence (T) are built. (As follow:)

Operation sequence matrix: $P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ & \cdots & \cdots & \\ p_{m1} & p_{m2} & \cdots & p_{mn} \end{bmatrix}$, where F , T and D are the same as P.

Owing to development customer-oriented process planning modeling, production batches affect directly the formation of procedures. The process meter controls the working time of every working procedure. To meet the customer's demands, manufacture resource's knowledge libraries are used to generate the WE sequence, that is technology process sequence. The decision mechanism of the process planning, which corresponding optimization model, and the balance for the process meters both are the contract factor for the formation of foundation model of manufacture units. Therefore, we can achieve equipment condition of every procedure in manufacturing, such as the type of spindle and the mode of working table and cutter tool base allocated to NC lathe machining For the MFMC system, the determination of those machine component modules, that is, the type of spindle and working table and cutter tool base or cutter, realizes the manufacturing flexibility

4. CONCLUSION

In attempting to address the integration of CAD, CAPP, and CAM, this paper provides a process planning unit model, that is the WE, to the integration of part feature model and process plan model and control of the mode of equipment operation for a hierarchical control architecture for prismatic parts. The information integration based on the WE is presented, in which the machining operation for prismatic parts is the core link. The WE built up at process stage is directly available to downstream machine component module selection and manufacturing activities The integrated approach of the customer-oriented CAPP subsystem is proved to be feasible for agilely providing the MFMC system design scheme in practical use.

5. REFERENCES

1. Hong-Chao Zhang, IPPM-A Prototype to Integrate Process Planning and Job Shop Scheduling Functions, Annals of CIRP, Vol. 42/1/1993.
2. Khoshnevis. B., Integrated Process Planning, Proceedings of Manufacturing International 90, ASME, Atlanta, Georgia, pp.243-248, March 25-28, 1990.
3. Kewei Lai, Rui Li, Towards Knowledge Integration of CAD/CAPP/CAM, IFIP WG 5 2 Working conference on Intelligent Computer Aided Design, Collumbus, Ohio, USA, Sep 30-Oct 3, 1991
4. Huilin Wu, The Study on CAD/CAPP/NCP System for Bracket Parts, master's thesis, Tianjin University, March, 1993.

THE STRUCTURE OF FEATURE LAYERS AND ITS APPLICATION IN CAD/CAM INTEGRATION

Guo Qun, Zhong Yifang, Wan Li, Zhou Ji
School of Mechanical Science and Engineering
Huazhong University of Science and Technology
P.R.China

ABSTRACT

The essence of CAD/CAM Integration is information integration. Feature acts as an information carrier because it possesses characteristics of rich semantics and information representation in multiple levels. However, Feature definition is neither perfect nor uniform, and it lacks adaptability as well as variability in its conveyance of information. Because of these feature deficiencies, the CAD/CAM Integration isn't fully realized. So author proposed a new method which describes feature by layer structures. This method divides feature into three layers: feature information layer, feature constraint layer, geometric element layer. It can provide information of the different layers for application fields. On the basis of creating the feature model, the corresponding part geometric information model is also created. This geometric information model not only describes feature geometric shape and topologic structures but also includes production technological information . On the basis of the above research, the author thinks that the difficulties of CAD/CAM Integration may be solved.

KEYWORDS

Feature Technology, CAD/CAM Integration, Feature Model, Feature Definition, Feature Layer Structure

1. INTRODUCTION

Integration technology of CAD/CAM is one of the key techniques that realize Computer Integration Manufacturing (CIM). Before developing CAD/CAM integration technology, we usually use some commercial drawing software system such as AutoCAD. In integrated CAD/CAM technology, people tend to extract features from solid models of current drawing systems, namely feature recognition technology. Because of the independence and limitations, it is very difficult to put feature recognition technology into practice. In order to solve the integration of high-level design and technological information, many leading analysts propose feature-based design.

The basic idea of the feature-based design is: some features are defined in advance and stored in the feature library. The geometric forms of these features have been defined, but their sizes are regarded as parameters to be considered in design. In the course of design, these feature parameters are instantiated. This process is called parametrical design. Because these features are closely related to engineering application fields and its basic information is included in the feature model, they are widely applied in designing, process planning and NC programming.

The feature-based design can combine the geometric form and technological information of the feature very well. As information conveyor, it solves most of the problems of the information transmission in the course of CAD/CAM integration and overcomes the limitations of the conventional CAD. It is an efficient way to realize CAD/CAM integration in developing CAD. Because feature definition isn't unified and feature technology isn't fully developed, every application system isn't the same in practice need and the feature types are finite in designing, the feature space of engineering application space is, however, infinite, the features are actually variable. So there are some urgent problems to be solved in applying feature design. These problems are as follows:

(1) No uniformity of the feature definition and imperfection of feature technology result in the differences in feature connotation, feature quality, feature structure and feature expression in every engineering application field. In the course of feature application, the different problems encountered in

engineering application fields result in the discrepancies in feature comprehension and demands. So it is very difficult to transform feature information among assorted engineering fields in the course of integration, which generates the faults and errors making integration rather difficult.

(2) The lack of adaptability and variability in the course of feature information conveyance has two reasons. The first reason is that the feature which is regarded as information entity has stable geometric form and technological information attached to geometric form. The second reason is that features in the feature model are the independent geometric and information entity. And they are lack of connection with product geometric model. These cause the conversion difficulties in the application fields. So there is no way to solve the problems on using finite feature types to satisfy infinite feature space.

In this paper, on the basis of the above feature analysis, the method of feature layers (MFL) is presented to describe the features. It divides the features into three layers: feature information layers, feature constraint layer, geometric element layer. In the event that non-uniformity of the feature definition and imperfections in feature technology are encountered, this descriptive method can provide information pertaining to the layers of each independent engineering application field. In solving the problems on adaptability and variability, we present the structure of feature layers to create feature-based model On the basis of this feature model, the creation of the concomitant geometric information model can be undertaken. This geometric information model not only includes feature geometric form and topological information in a conventional sense, but also includes a geometric information model as well as product technological information. There will be transforming process from feature model to geometric information model. In this program, the geometric and non-geometric information in feature model will be transformed to corresponding geometric information model effectively. There will be a great adaptability to convey the information to other application fields . Because the feature form can be regenerated in this model, there will indeed be variability

At last, by using this feature analysis method of feature and geometric information, we create both the feature and geometric information models, and point out their relationship. The simultaneous creation of these two models can help to overcome their concurrent deficiencies and provide a means of assisting one another. On the basis of solving the above problems feature technology will be better applied in CAD/CAM integration system.

2. FEATURE DESCRIPTION

Feature technology is sure to rise with technological improvement. On the one hand, solid modeling is based on geometric representation and operation. The actual geometric operations without application meaning contradict with the engineers' design concepts and methods. On the other hand, CIMS has been developed greatly in the past ten years. Besides satisfying its own information perfection, CAD system are required to provide non-geometric information such as material and tolerance that can reflect the designer's intentions for other systems such as Computer Aided Process Planning (CAPP) and Computer Aided Manufacturing (CAM).

The feature introduction provides high-level interactive language that coincides with engineers' intentions. It takes the place of low-level interactive design methods based on geometry and topology. This allows engineers to focus on dealing with high level design problems, thus enhancing the efficiency of design production. The design quality is assured. On the other hand, because feature is a high-level design concept, it includes a lot of engineer's intentions. The intentions have important significance in design maintenance and downstream analysis and synthetic process. It can also improve the automatic degree of CAD systems and solve discontinuity between CAD and CAM[1].

A clear and precise definition of feature technology is non-existent because it is a relatively new field. The different applications form different feature definitions. From the view of machining, feature is defined as part form related to machining operations, tools and technological characteristics. From the view of form modeling, feature is composed of a group of interrelated geometric and topological elements. From the view of engineers, features are the basic elements used as design analysis and evaluation. Although feature definition is different in every field, a common recognition has basically been formed at present. Feature is a group of solid entities that have special attributes. That means that the feature contains a set of objective elements. These elements are: feature geometric form and special machining elements, (namely feature technological information). Feature description not only reflects feature geometric attributes, but also reflects its non-geometric attributes. In order to describe feature, a lot of descriptive methods are presented. These descriptive methods can be divided into three types: rule-based feature representation scheme, graph-based feature representation scheme and syntactic-pattern-based feature representation scheme[2]. Among the three representation schemes, there are common deficiencies as follows:

(1) The three representation schemes emphasize the geometric representation of feature geometric form. The feature can be perfectly expressed out in computer. It describes feature only from the view of modeling. The description of feature connotation isn't perfect and complete.

(2) The features haven't been considered from the overall situation in the three feature representation schemes. The requirements of downstream engineering aren't considered in the three feature representation schemes, which may limit its application.

In order to overcome the above deficiencies, MFL used to describe feature structure can better satisfy this kind of requirement. The main idea of MFL is: feature geometric form as feature kernel consists of feature form skeleton. Other information according to their attributes and application background are formed to information entities. These entities surround the kernel. All information entities and kernel form feature information layer. The number of feature information entities isn't limit. On the basis of the special connotation and comprehension of every application field, every engineering application field can make its information form entity. Then we can put the entity into feature information layer. Every engineering application field can look up its information requirement through hierarchical relationship. Figure 1 show the constitution of feature information layer.

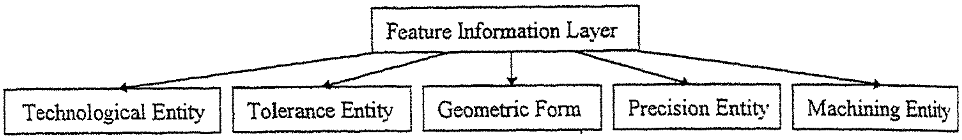


Fig.1: The Constitution of the Feature Information Layer

The layer below the feature information layer is the feature constraint layer. It includes two parts: the constraints among geometric elements (CGE) and the constraints of the technological attributes (CTA).CGE is divided into two types: form constraint and position constraint. CTA includes form tolerance constraint, position tolerance constraint, roughness constraint Figure 2 shows the structure of constraint layer.

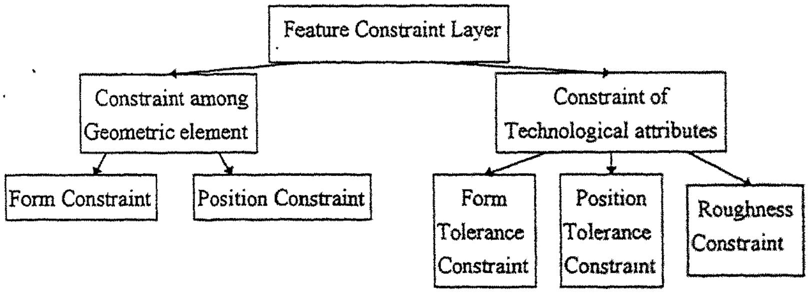


Fig.2: The Structure of Constraint layer

The layer below the feature constraint layer is the geometric element layer. It describes geometric element types and the number of features, the form constitution, topological relationship and the relationship between relevant elements of features. This structure is described in constructive solid geometry (CSG) and boundary representation (B-rep). We used winged-edge (WE) representation to describe feature geometric elements and topological relationships. In this structure, the technological attributes are added to this WE in the course of description of feature geometric elements so that it can provide both geometric attributes and technological attributes for downstream CAPP and CAM. In this WE representation, the elements are described according to the relevant scale of features with which they belong . The topological relationship of relevant elements are recorded so that the WE representation may provide more detailed information for feature recognition. This WE representation is called enhanced winged-edge representation (EWE). Figure 3 shows this structure.

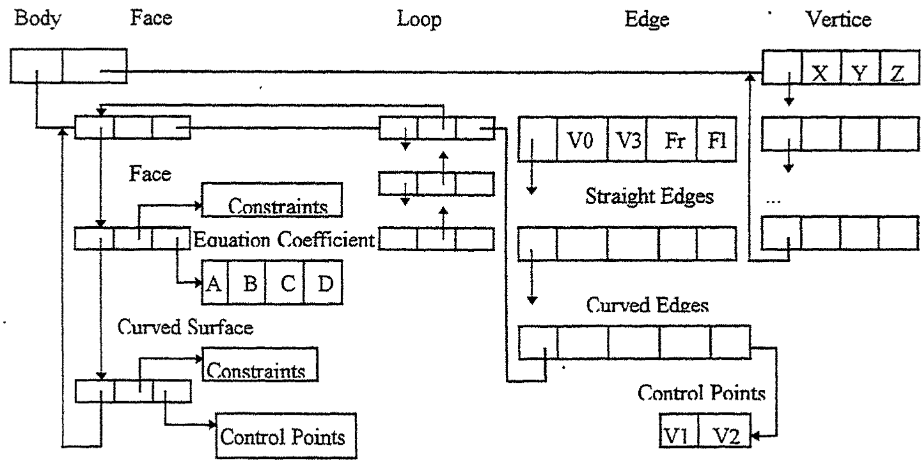


Fig.3: The structure of the enhanced winged-edge representation

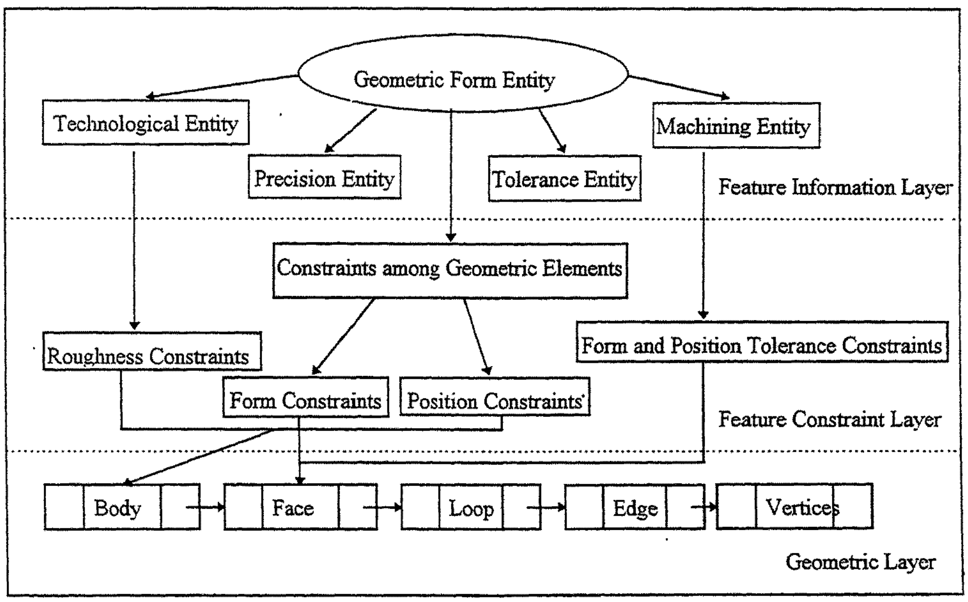


Fig. 4: The structure of MFL

The differences of this EWE compared to the conventional WE are as follows:

(1) In EWE, we regard part precision and tolerance as a kind of constraint similar to element

topological relationships. we call it technological attribute constraint and add it into EWE.

(2) In the conventional WE, The product is designed as a whole. In this paper, we try to describe a product as a whole, divide it into features. then record the geometric information in the scale of feature.

(3) In this paper, we emphasize the element relationship which includes geometric elément topological relationships and constraint relationships between features. It will provide an objective basis of feature recognition for feature adaptability and variability.

The relationship among three layers of MFL is depicted in figure 4.

3. THE APPLICATION OF MFL

MFL summarizes the description feature idea from different views. Its three-layer structure can provide information of the three different layer for engineering fields. Every layer is closely related to another layer. Feature information layer can better take advantage of feature conception. Feature constraint layer can better express the relationships of feature composition and constraint. Feature geometric layer precisely describes geometric elements by using WE.

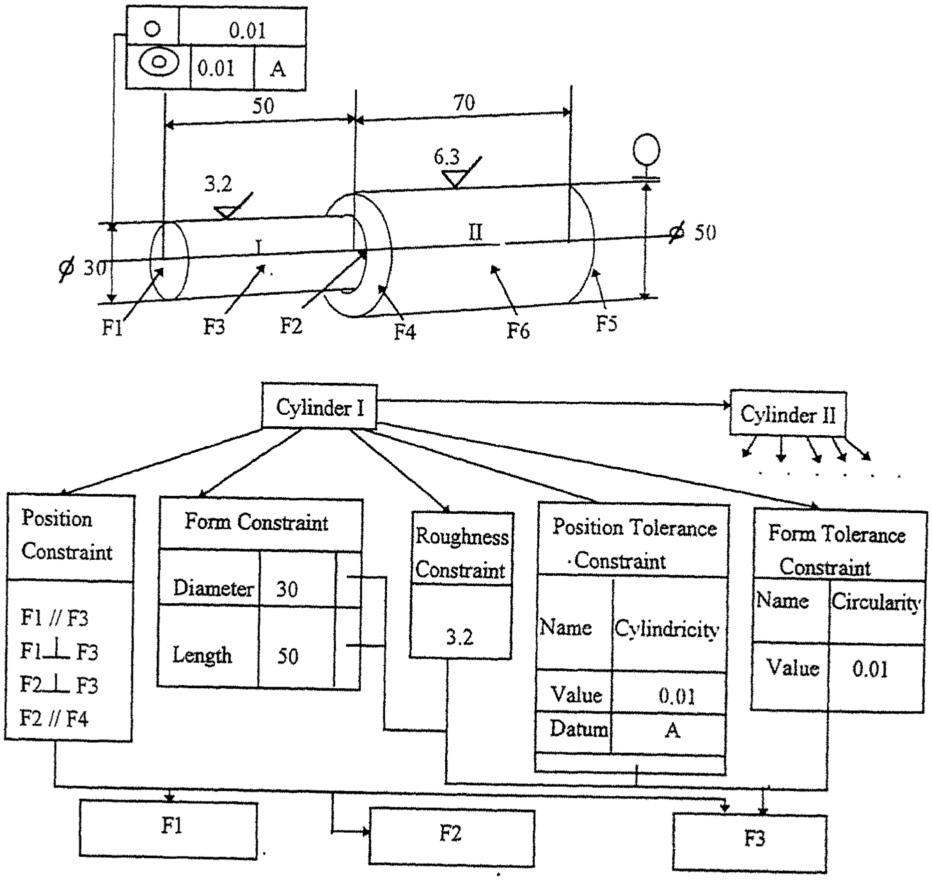


Fig. 5: A part example of applying MFL

In applying MFL, we first design the parts by feature modeling. Then feature is decomposed according to feature attributes, the topological relationships, technological attributes. The constraint relationships among elements are classed and summed up to form constraint rules.

The form features constitute the part geometric shape. Every form feature has its special geometric elements. The interaction between the geometric and topological relationships of feature geometric elements and feature geometric and technological constraint relationships form the feature having engineering significance.

The relationships among features are depicted through topological and constraint relationships of feature special geometric elements. The feature types which are needed in engineering fields can be regrouped, extracted and recognized according to these topological and constraint relationships. So features will have great variability and adaptability. In figure 5, a simple part is expressed by using MFL.

4. CONCLUSION

Through the above research of feature representation method, these information integration problems of CAD/CAM can be solved and the following goals can be gotten:

(1) In case of feature difference and imperfection, the feature demand of engineering fields is fully considered in MFL. MFL can provide different layer information and uniform expression platform of integration for engineering fields.

(2) The same design feature can satisfy different fields in feature comprehension and demands and make feature have variability and adaptability.

Because MFL can depict feature clearly and contain a great deal of information, the relationship between features and between elements is certain. It can provide a basement for mapping from design space to other application field spaces and make features be regrouped according to feature layers. relationships between features and elements in different application demands. This objective solves the contradiction of the numbers of design space features being infinite while feature types are finite.

5. REFERENCES

1. Sun Jiaguang and Yang Changgui, Computer Graphics, pp443-448, Tsinghua University Publishing House, Beijing, 1994.
2. M C Wu and C R Liu, Analysis on Machined Feature Recognition Techniques Based on B-rep, Computer-Aided Design, Vol.28, No.8, pp.603-616, 1996.
3. C.J. Su, T.L. Sun and C.N Wu, An Integrated Form-feature-based Design System for Manufacturing, Journal of Intelligent Manufacturing, Vol.6, pp.277-290,1995.
4. Jerry Y H Fuh,C.H.Chang and Michel A melkanoff, The Development of an Integrated and Intelligent CAD/CAPP/CAFP Environment Using Logic-based Reasoning, Computer-Aided Design, Vol.28, No.3, pp.217-232,1996.
5. Chongsu Kim and Peter J O'Grady, A Representation Formalism for Feature-Based Design, Computer-Aided Design, Vol. 28, No.6/7, pp 451-480,1996.

3D OBJECTS FEATURE MATCH BASED ON RELATIONAL GRAPH

Xinhua Wang Chuanhue Huo Yuguang Li
School of Electronic Engineering and Precise Mechanics
Wuhan Technical University of Surveying and Mapping (WTUSM)
430070,Wuhan, P.R.CHINA
Deren Li Peizhi Huang
School of Photogrammetric and Remote Sensing, WTUSM

ABSTRACT

In this paper, authors present an algorithm for feature-based stereo image match. The algorithm is implemented for a portable digital stereo camera which is for the purpose of 3D industrial measurement, and also can be mounted on a mobile robot. In the procedure of feature match, the feature description and the relational description of features are firstly built. Then a match measure is defined to determine the base match. Finally, the feature matches for all line segments in the images pair are realized through matching conveying.

KEYWORDS:

Stereo Vision. Feature-based Match. Relational Graph, Base Match, Match Conveying

1. INTRODUCTION

Visual inspecting, recognition and manipulating for industrial parts are the important unit of CIMS. The key step to realize the stereo vision is stereo match. So the theory and methods of image match have been one of the hot research topics in the past decades. The most common means of achieving stereo vision includes the methods of image density-based match and feature-based match. In some special situation, for instance, the industrial purpose, the latter has some advantages because it can not only realize stereo match, but also involve the relational description of features.

In the past decades, many methods of feature-based stereo match have been developed. Kreigman etc's [1]passive feature-based stereo system have previously been used to help guide a robot in performing prespecified tasks. The stereo match algorithm may return a set of potential matches when the set of directed graphs constructed in the matching process leads to more than one high-quality path. A resolution of such ambiguity is suggested via the use of motion correspondence, if such information is available. Huynh etc.[2]developed the line labeling and region segmentation algorithm in stereo image pairs. The algorithm to match feature with salient structure between left and right images reduce the degree of matching ambiguity. Horand and Skordas[3] presented a stereo matching algorithm through feature grouping and maximal cliques. The purpose is also to reduce the degree of matching ambiguity. But the spatial position of feature must be previously determined. Moreover, Li and Zhuo[4] suggested a 3D reconstruction algorithm based on line photogrammetry. Its advantages is that 3D reconstruction can be achieved directly from original image. Wang[5] developed a stereo vision system (SPPS) for industrial inspecting and measurement by using Li's method.

This paper reported a feature-based match technique that has been developed as a stereo vision camera for industrial 3D shape measurement. The idea is to use the relational description of features and the geometric constraint to realize the feature match and results in reducing the match ambiguity. The procedure of feature extracting and relation description building is introduced in the section 2. The focus of section 3 is on the feature match algorithm based on the relational graph. Finally, a practical prototype and some conclusion remarks are put forward in the section 4.

2.EXTRACTING FEATURE AND BUILDING RELATIONAL GRAPH

2.1 Extracting Feature

Chen [6] implemented a feature detector based on zero crossing operator. The process can be divided into two steps. First step involves the convolution of image with a pair of 3D masks of Gaussion function. This is followed by using zero crossing operator to detect the image edges. The feature detector has a powerful positioning ability because it is based on the precise definition of edge.

Edges are detected from each image pairs using the aforementioned feature operator. The line segment extraction process is divided into linking and merging stages, which follows the technique of Yu's[7]. Thus we can get two line segment sets for left and right images, expressed as $L = (L_1,\cdots,L_n)$ and $R = (R_1, \cdots, R_m)$.

2.2 Building Relational Graph

The 2D projection of industrial parts expresses as the construction features, which are line segments. In order to realize the stereo match and 3D reconstruction, we must build the feature attribute description and the relational description, which supply the important information for subsequent stereo match process.

2.2.1 Feature Attribute Description

Each line segment L_i in left image (or R_j in right image) has associated with a line equation:

$$\rho_i = x\cos(\theta_i) + y\sin(\theta_i) \qquad (1)$$

and the endpoints

$$(x_{si}, y_{si}) \quad \text{and} \quad (x_{ei}, y_{ei})$$

Where: (x, y) are the coordinates of points in the image plane. (ρ_i, θ_i) are the coefficients of line equation L_i, (x_{si}, y_{si}) and (x_{ei}, y_{ei}) are the coordinates of starting and ending points of the line segment L_i.

2.2.2 Building Relational Graph.

Indeed, the procedure of building relational graph is to determine the feature relational description. Building relational graph process works as follows:

- A starting edge L_s is chosen at random from L.
- Traverse along the clockwise direction (Fig.1) from the endpoint of L_s to find whether any line segment L_c is linked or merged with L_s in the traversed direction of L_s. This step mainly involves checking the following hypothesis:

1. L_c must have a endpoint falling within a neighborhood centered at the end of L_s.
2. The points of L_c, when substituted into the line equation of L_s, must yield values that have the same sign and be greater than a tolerance.

If the two hypothesis is fulfilled, then L_c is accepted as a child node of graph tree.

- If there are more than one line segments which are linked with L_s, we name it merging link. Then the searching process always firstly traverses in the right-hand direction, and more than one child node are recorded. The process carries on from L_c onward for building the descendant node at levels further down the searching tree, until one of the following condition is met.

1. Line segment at the root node (L_s) becomes a child node of the searching tree. To avoid looping, L_s is not inserted as a leaf node into the searching tree.
2. The child line segment intersects is very close to the boundary of the image, and no more descendant line segments in the vicinity can be found to fulfill.

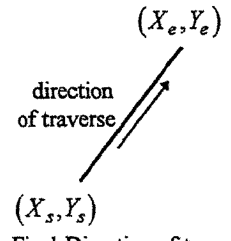
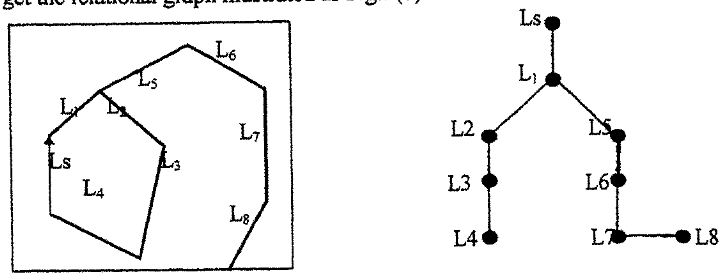


Fig.1 Direction of traverse

Fig. 2(a) is an example of building relational graph. Apparently, L_1 is a child node of L_s. Both L_2 and L_5 satisfy the hypothesis for L_1. They are therefore included in the searching tree as two child node of L_1. The same process in finding child node is applied in a recursive manner. Finally, we can get the relational graph illustrated as Fig.2(b).



(a) (b)

Fig.2 An overview of building relational graph. (a) shows the features in image; (b) shows the link relation of features.

Finally, we can get the feature description and the relational description which have a data structure shown as following:

Id	Par ...	No-node	Id_1	Id_2	...	Cop-Id

Where: Id is the identification of line segment.

Par is the parameters, include $\rho_i, \theta_i, x_e, y_e, x_s, y_s$.

No-node is the child node number of line segment Id.

Id_i is the child node identification of line segment Id.

Cop-Id is the identification of node in right image which is correspondence to the Id.

Thus we can get two unit sets $(L_1, L_2, \ldots. L_n)$ and $(R_1, R_2, \ldots. R_m)$ in the left image and right image, respectively.

3. FEATURE MATCH

The process is first to determine a base match, and then to realize the all feature matches through the match conveying based on the relational graph of features.

3.1 Base Match

Lets see a simple example illustrated as Fig.3. If the attitude and orientation parameters of camera are known, To the any point a lied on the line L in the left image, its correspondence in the right image must met the epipolar line equation (2):

$$y' = (A/B)x' + (C/B)f \qquad (2)$$

Where: $A = V_a W_s - W_a V_s$

$B = U_a W_s - W_a U_s$

$C = V_a U_s - U_a V_s$

$$\begin{bmatrix} U_a \\ V_a \\ W_a \end{bmatrix} = M_{21}^T \begin{bmatrix} x_a \\ y_a \\ -f \end{bmatrix} \qquad \begin{bmatrix} U_s \\ V_s \\ Ws \end{bmatrix} = M_2^T \begin{bmatrix} X_s - X_s' \\ Y_s - Y_s' \\ Z_s - Z_s' \end{bmatrix}$$

(x_a, y_a) is the coordinates of point a in the left image. (X_s, Y_s, Z_s) and (X_s', Y_s', Z_s') are the coordinates of projection centers S_1 and S_2. M_2 is the rotation matrix of right image, and M_{21} is the rotation matrix of right image related to left image.

Apparently, assume that line R in the right image is the correspondence of L in the left image, thus the correspondence of point a in the right image must be the intersect of epipolar line K and the structure line R. Especially, the ambiguity problem occurs when more than one structure lines exit in the right image(Fig.4). In this condition, the procedure of searching base match work as following:

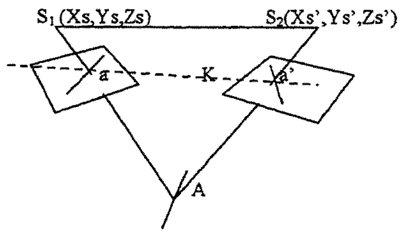


Fig.3 An overview of the projecting relation

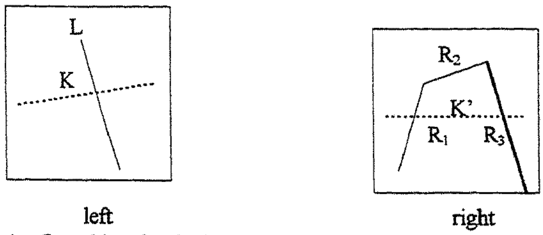


Fig.4 Searching for the base match from multrifuli correspondence

1. Epipolar constraint: If the line R in right (or left) image has some points which met the epipolar equation determined by the points of line L in left (or right) image, R can be accepted as one possible correspondence of L.

2. Parallax constraint: The parallax of L and R should be within a range that can be determined by the priori knowledge. If no priori knowledge, R is accepted as one possible correspondence of L. Generally, assume the unit subset R_j is the possible correspondence set of the unit subset L_i, The possibility of match is measured by the following match measure:

$$I(L_i) = \frac{N[L_i, R_j]}{M[L_i]} \qquad I(R_j) = \frac{N[L_i, R_j]}{M[R_j]} \qquad (3)$$

Where: $M(L_i)$ and $M(R_j)$ are the pixel number of L_i and R_j. N is the matched point number.

Thus R_j which has the maxim I is selected as the maxim possible match (base match). If all I are less than a thresholding value T (0<T<1), the base match is discarded, and another unit subsets L_{i+1} and R_{j+1} are taken to check its match possibility measure. Notice that we always select the unit which has single leaf node and is perpendicular to the epipolar line direction as the first base match.

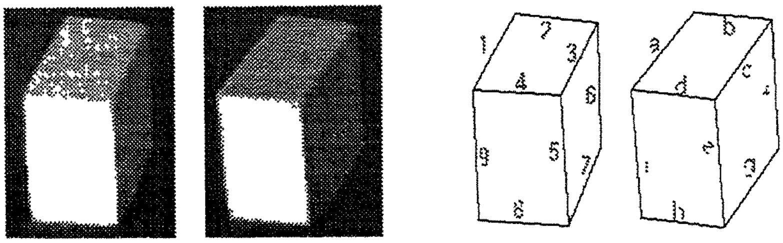
3.2 Match Conveying

After obtaining the base match, the relational graph can be used to realize the match conveying. Simply, if each units of the base match pair (L_i, R_j) has single leaf node L_{i+1} (for L_i) and R_{j+1} (for R_j), then the L_{i+1} and R_{j+1} are expected as the possible match pair. Subsequently, select L_{i+1} and R_{j+1} as the new base match. The same conveying process is applied.

Especially, if the base match pair (L_i, R_j) has more than one leafs, we will search a new base match from all leaf nodes. Assume that L_i has e leafs $(L_1, ..., L_e)$ and R_j has f leafs$(R_1, ..., R_f)$, executing formula (3) to each possible pairs (L_k, R_t) $(k = 1 \cdots e; t = 1 \cdots f)$ will get a group of possible match pairs which have the maximal I as the new base match.

4. PRACTICAL PROTOTYPE AND CONCLUSION REMARKS

Above match algorithm has been involved in our digital stereo vision camera which is developed for the 3D industrial measurement in Wuhan Iron and Steel Company. The image taken and the line segment detected are given in the Fig.5 (a)--(d) for the left and right image. Tab.1 illustrates the matching possibility measures and the final match results. It should be noted that features 2,4,8 are not selected as the base match features because they are almost parellel to the epipolar line. It is obvious that the final match results is correspondence to Fig.5 (c) and (d). Indeed, we always select the unit which has single leaf node and is perpendicular to the epipolar line direction as the first base match.



(a) left (b) right (c) left (d) right

Fig.5 Original images and the Features extracted from image pair

Based on above theory and analysis as well as our practical experiences, we can try to get the following conclusion remarks:

- As the number of taken for matching in each image is reduce significantly, the computation time for stereo matching process reduces and false matches are largely eliminated.
- In the match procedure, no priori constraints are necessary because the match process is based

Tab.3 Match results

	Match possibility measures									Final match results
Feature	a	b	c	d	e	f	g	h	i	
1	0.98	0.07	0.88	0.11	0.26	0.81	0.0	0.0	0.0	1-a
2										2-b
3	0.87	0.07	0.97	0.14	0.0	0.65	0.0	0.0	0.08	3-c
4										3-d
5	0.10	0.0	0.08	0.79	0.99	0.39	0.20	0.05	0.95	5-e
6	0.57	0.08	0.65	0.08	0.39	0.98	0.06	0.0	0.47	6-f
7	0.0	0.0	0.0	0.0	0.64	0.20	0.97	0.07	0.68	7-g
8										8-h
9	0.0	0.0	0.0	0.07	0.85	0.44	0.54	0.07	0.0	9-i

on the geometric constraint and the structure constraint which are inherent feature in the vision system, not on the similarity of images. So it greatly improve the trade off problems between the match precision and the matching difficulty.

- The match process based on the relational description of features can be conveniently joined with the process for more complex objects recognition[9]. This work is developing.

5. ACKNOWLEDGMENTS:

This paper was supported by the Chinese National Natural Science Foundation[Certificate

No:49601017] and the Chinese National Natural Science Foundation[Certificate No: 4960j1050]

6. REFERENCES

1. Kreigman D J. Etc.," Stereo vision navigation in building for Mobil robots", IEEE Trans. Robotics & Automation, Vol 5, No.6, pp.792-803, 1989.
2. D. O. Huynh Etc., " Line labeling and region segmentation in stereo image pairs", Image and Vision Computing, Vol.12, No.4, pp.213-224, 1994.
3. Horaud etc., "Stereo correspondence through feature grouping and maximal cliques", IEEE Trans PAMI, Vol.11, No.11, pp.1168-1180, 1989.
4. L. Deren Etc. "A feasibility study on measurement and reconstruction of object primitives using line photogrammetry", ACTA GEODAETICA et CARTOGRAPHICA SINICA, Vol.23, No.4, pp.267-247
5. Xinhua etc.,"CAD-based line feature objects measurement and reconstruction", SPIE Proc. of Int Conf. on the CAD/CG, pp.701-706 1995.
6. J. S. Chen etc., "Fast convolution with Laplacian of Gaussion masks", IEEE Trans. PAMI, Vol.9, No.4, pp.584-590, 1987.
7 Y. Lei etc. " The processing of broken, fault edge segments and structure points in edge detection", Pattern Recognition and Artificial Intelligence, Vol.8, No.2, pp.165-170,1995.
8 H Peizhi & Xinhua W. " new method for detecting dominant points", SPIE Proc. of Int. Conf. on the AD/CG, pp.294-298, 1995.
9 Xinhua W and L. Deren . "Relation data structure and consistent labeling algorithm for 3D industrial object", Journal of WTUSM, Vol.21, No.3, pp.242-247, 1996.

The Feature Modelling For a Sculptured Object

C. K. AU, M. M. F. Yuen
Department of Mechanical Engineering
Hong Kong University of Science and Technology

ABSTRACT

This paper describes the application of semantic features to the modelling of sculptured objects. A semantic feature is defined as a list containing two sets of information, namely semantics and geometry. A feature model is an aggregation of the semantics features. The geometry of a feature model consists of infra-feature geometric and inter-feature geometric information. The infra-feature geometric information consists of the list of characteristic topological entities forming the feature, the geometric and topological relationships governing this group of entities, and their basic geometry definitions specified in parametric form. The inter-feature geometry describes the topological and geometric relationships among features within the same object. On the other hand, the semantics of the feature model is the aggregation of each of the individual feature semantics. The scope of the semantic information covers a wide domain of knowledge including design, manufacturing, assembly and inspection. The current discussion will only be limited to the context of design information that is driven by the functionality of the parts. It dictates the position and continuity requirement relating to the features. This set of information is substituted into the constraints prescribed in the feature geometry.

KEYWORDS

Sculptured object, semantic features, feature model , continuity

1. INTRODUCTION

Feature technology is applied to various application domains such as group technology coding[1], tool path generation[2], machinability determination for the generalized mechanical parts[3], automatic process planning[4,5], automatic inspection[6], assembly modelling[7] and finite element analysis[8]. The majority of the works for features modelling concentrates on regularized object which have simple analytical geometry with the C^0 and C^1 continuity requirement between surfaces.

Although features for sculptured object modelling are included in the taxonomy[9], little attention has been received. Until recently, Jones, et al[10,11,12] discussed the feature-based approach for sculptured products. The system is developed based on a commercial CAD CAM software with conventional modelling operations.

2. PROBLEM DEFINITION

Feature model of a regular shape object is obtained by aggregating a group of pre-defined features by a set of Boolean operations. However, such an approach is not applicable to the sculptured object modelling because the geometric orientated features are not suitable to deal with the shape irregularity and high variation of the objects.

A semantic feature is a reasoning unit of an object feature model. It is an aggregation of both geometry and semantic information. This paper discusses the application of semantic features to the modelling of a sculptured object. The scope of the semantic information covers a wide domain of knowledge including design, manufacturing, assembly and inspection etc. The current discussion will only be limited to the context of design information that is driven by the functionality of the part. It dictates the position and continuity requirement of geometric entities.

3. FEATURE MODELLING

A feature model **C** is an aggregation of a set of features describing the characteristics of an object. It is defined as

$$\mathbf{C} = \sum_{ij} \varphi_{ij}(\mathbf{F}_i, \mathbf{F}_j) \tag{1}$$

where Σ is the aggregation symbol,
$\mathbf{F}_i$ is a semantic feature
φ_{ij} is an inter-feature association containing both topological and geometric information between feature $\mathbf{F}_i$ and $\mathbf{F}_j$

A semantic feature is a set of abstract data with attributes, it is generalized as

$$\mathbf{F}_i = \begin{bmatrix} \mathbf{r}_i \\ \mathbf{G}_i \end{bmatrix} \tag{2}$$

where $\mathbf{G}_i$ is the feature geometry
$\mathbf{r}_i$ is the semantics of the feature.

The feature geometry $\mathbf{G}_i$ can be expressed as

$$\mathbf{G}_i = \sum_{lk} \pi_{lk}(\overline{\mathbf{q}}^{i,l}, \overline{\mathbf{q}}^{i,k}, a_{lk}) \tag{3}$$

where $\overline{\mathbf{q}}^{i,l}$ is the l-th geometric entity of the feature $\mathbf{F}_i$,
π is an infra-feature association function between the geometric entities $\overline{\mathbf{q}}^{i,l}$ and $\overline{\mathbf{q}}^{i,k}$. This association describes the relationship between these two geometric entities. It is a graph with the entities $\overline{\mathbf{q}}^{i,l}$ as the nodes and association as the attributed edges with value a_{lk}.

This semantics $\mathbf{r}_i$ are categorized into:

(i) infra-feature semantics which consist of the feature identification, functionality and topological relationships between the geometry,
(ii) inter-feature semantics which is the topological relationship with its neighbour features.

Hence, a feature model consists of infra-feature and inter-feature geometry. The infra-feature geometry consists of

(i) the list of characteristic geometric entities forming the feature and their basic geometry definitions,
(ii) the geometric relationships governing this group of entities.

as illustrated in equation (3). The inter-feature geometry describes the geometric relationships among features within the object.

Figure 1 shows the relationship between an object, features and geometric entities by using object oriented modelling technique.

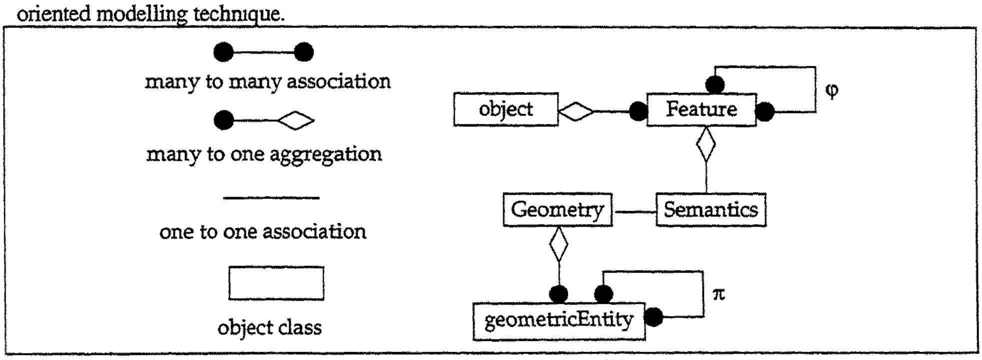


Figure 1: Object oriented model of a feature model

The semantic links of a feature, infra-feature and inter-feature association are represented by a set of attributes which can be expressed in predicates format:

AttributeTitle(Entity, AttributeValue)

For instance, the predicates generated from equation (1), (2) and (3) are

$\theta_i(G_i, r_i)$

$\pi_{lk}(\overline{q}^{i,l}, \overline{q}^{i,k}, a_{lk})$

$\varphi_{ij}(\mathbf{F}_i, \mathbf{F}_j)$

Each of the predicates is associated with a set of constraints relating the geometric characteristics of the geometry and the attribute values. Therefore

(i) $\theta_i(G_i, r_i) \Rightarrow \quad \Phi_{i,l}(\overline{q}^{i,l}, r_{i,l}) = 0 \qquad \forall l=1,\ldots,M_i,\ r_{i,l} \subset r_i$ (5)

(ii) $\pi_{lk}(\overline{q}^{i,l}, \overline{q}^{i,k}, a_{lk}) \Rightarrow \Omega_{i,lk}(\overline{q}^{i,l}, \overline{q}^{i,k}) = 0 \qquad \forall l=1,\ldots,M_i,\ k=1,\ldots,M_i$ and $l \neq k$ (6)

(iii) $\varphi_{ij}(\mathbf{F}_i, \mathbf{F}_j) \Rightarrow \psi_{ij,lk}(\overline{q}^{i,l}, \overline{q}^{j,k}) = 0 \qquad \forall l=1,\ldots,M_i,\ k=1,\ldots,M_j$ (7)

where $\Phi_{i,l}$, $\Omega_{i,lk}$ and $\psi_{ij,lk}$ are the constraint equations,

$\overline{q}^i$ are the geometric entities within feature $\mathbf{F}_i$,

r_i is the infra-feature attribute value of the predicate and

M_i is the number of geometric entities within feature $\mathbf{F}_i$

The geometry of an object model is constrained by a system of equations according to the semantics. A unique shape can be generated for a fully constrained system. Otherwise, a solution set will be obtained if the system is under-constrained.

4. FEATURE MODEL OF A SCULPTURED OBJECT

A sculptured object is characterized by its continuity. The discussion will be limited to order one continuity of the model geometry. A sculptured object is given as

Σ_{ij}Interaction($\mathbf{F}_i$, $\mathbf{F}_j$) (8.1)

where Interaction is the inter-feature association between F_i and F_j which gives a sub-association denoting the continuity between two surfaces of the two features

Continuity($\overline{q}^{i,l}$, $\overline{q}^{j,k}$, order) $\forall l,k$ (8.2)

A semantic feature for a sculptured object is defined as:

FeatureName(Feature, $\mathbf{F}_i$) (9.1)

GeometricEntity($\mathbf{F}_i$, $\mathbf{G}_i$) (9.2)

CharacteristicPointSet($\overline{q}^{i,h}$, $\mathbf{r}_{i,h}$) (9.3)

Continuity($\overline{q}^{i,h}$, $\overline{q}^{i,h+1}$, order) (9.4)

Interaction($\mathbf{F}_i$, $\mathbf{F}_j$) (9.5)

where $\mathbf{F}_i$ is the name of the feature

$\mathbf{F}_j$ ($\forall$j) are the neighbouring features of $\mathbf{F}_i$,

G_i contains a list of surfaces $\overline{q}^{i,h}$ ($\forall$ h),

Continuity is the infra-feature association,

FeatureName, GeometricEntity and CharacteristicPointSet are the semantic links within the feature,

order is the continuity order between two surfaces $\overline{q}^{i,h}$ and $\overline{q}^{i,h+1}$, and

$\mathbf{r}_{i,h}$ is an array of points characterizes the feature surface $\overline{q}^{i,h}$.

Figure 3 shows a feature model of a sculptured object with two features. There are two levels of abstract data: semantic and geometric. The topology of the model is defined by the semantic links and associations. A constraint graph of the geometric entities with the surfaces as nodes, semantic link and associations as links can be derived according to the feature graph.

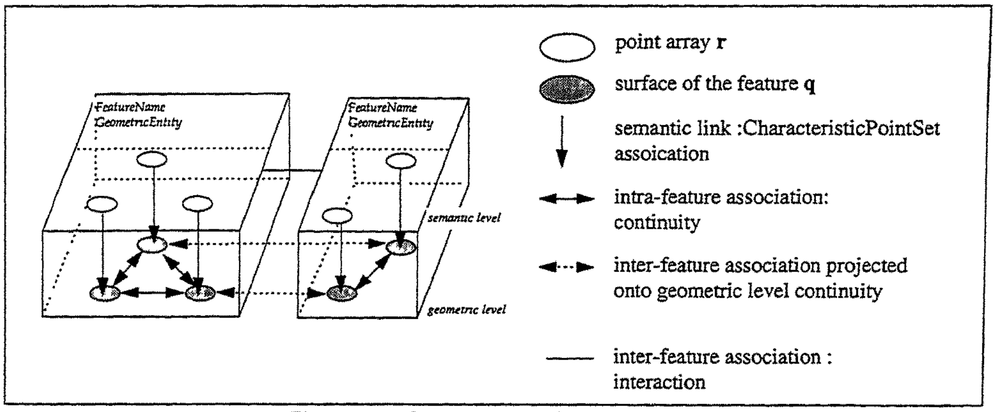


Figure 3 : A feature model with two features

Two types of geometric constraints are generated by the predicates Continuity and CharacteristicPointSet to constrain the surfaces:

(i) continuity($\overline{q}^{h}, \overline{q}^{h+1}$,1) - continuity of order 1 between two surfaces $\overline{q}^{h}$ and $\overline{q}^{h+1}$. The implied constraint is

$$\frac{\partial \overline{q}^{h}}{\partial u} + \alpha \frac{\partial \overline{q}^{h+1}}{\partial v} + \beta \frac{\partial \overline{q}^{h+1}}{\partial u} = 0 \qquad (11)$$

where α and β are polynomials.

(ii) CharacteristicPointSet($\overline{q}^{1,h}$, **r**) - an array of characteristic points on the surface $\overline{q}^{1,h}$. The implied constraint is

$$\overline{q}^{1,h} = \sigma(\mathbf{r}) \qquad (12)$$

where σ is a function relating the surface $\overline{q}^{1,h}$ and point set **r**.

A system of equations can be set up with equation (11) and (12) to constrain the surfaces in the geometric level of the feature model.

5. EXAMPLE

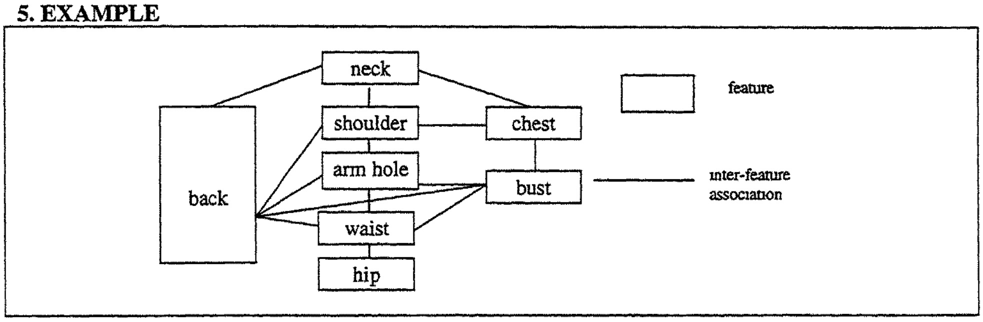


Figure 5: Feature graph of a (half) mannequin

A mannequin is used as an example to illustrate the formalism representation of feature modelling for a sculptured object. A mannequin is a reference model for garment design which has a similar form of a person. Figure 5 shows the feature graph of a mannequin feature model. Each of the features consists of a list of B-spline surfaces with continuity and characteristicPointSet as infra-feature association and semantics. The feature interaction is also defined by the continuity between the surfaces of various features. A shaded image is shown in figure 6 after solving the constraint graph of the feature model.

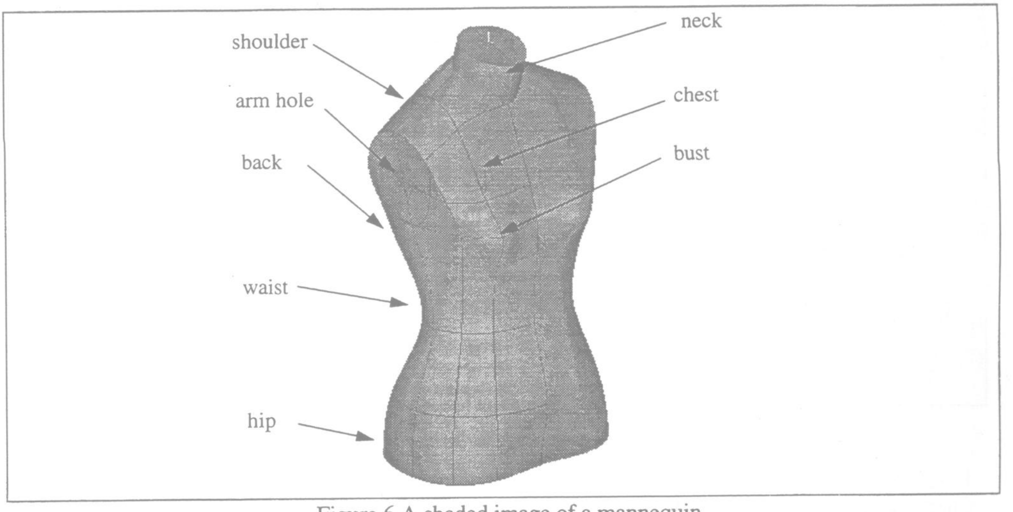


Figure 6 A shaded image of a mannequin

6. DISCUSSION

The traditional features are geometry oriented, features are defined according to their geometry and then infra-feature semantics are attached to the geometry. Since the regular shape objects consist of lower (2 or 3) order geometry such as the analytical surfaces, a set of pre-defined features with these shapes is usually enough to describe the objects. This approach is particularly suitable for the feature model of a regular shape object in the machining domain because of the simple one-to-one mapping between the geometry and machining operations. However, such approach is not applicable to sculptured objects because of their free-form shape. In addition, the feature graph for a sculptured object is product structure dependent. There is a fixed feature graph for a specific object even their geometric entities are not yet known. A feature will lose its natural if it is not at the specific node of the graph that it should be. Therefore, a semantic oriented feature is needed to describe a sculptured object. Unlike the geometric oriented feature, both of the infra-feature and inter-feature semantics are included in the feature which defines the feature geometry.

The other application of the semantic feature is to model the deformable objects as cloth modelling. The shape of the cloth is due to the combination of the repulsion, stretch, bending, trellising and gravitational effects on the particle of the cloth. These properties are included in the semantics of a piece of cloth feature to define its shape at a particular instance.

The geometry of a model is constrained by two types of constraints:
(i) geometric constraints such as position of a data point, continuity between two surfaces etc.,
(ii) metric constraints such as length of a curve, area of a surface etc.

These two types of constraints have different roles in the geometric model. The geometric constraints define the shape of the model and the metric constraints size of the model. The geometric constraints are established during model design while the size variation is achieved through the metric constraints. So far, the discussion concentrates on the object shape definition. Size variation can be achieved by introducing a set of parameters in the semantic level relating to the surface which will be discussed in the follow-up work.

7. CONCLUSION

A formalism representation of feature modelling for sculptured objects is discussed. Most of the feature technology researches are concentrated on a set of generic features such as holes, slots, pocket, etc. to model a class of generalized mechanical objects with regular shape. This set of generic features can then be associated with the semantics of a particular domain such as machining operations. These features are geometry oriented.

Unlike regular shape objects, sculptured objects set is a huge set and has unlimited form variations. Using geometry oriented features to model sculptured object limits the modelling coverage. Instead of using a set of pre-defined features, a semantic oriented feature is proposed to model the sculptured objects.

8. REFERENCES

1. Srikantappa, A. B. and Crawford, R. H., 'Intermediate geometric and interfeature relationships for automatic group technology part coding', ASME Computers in Engineering Conference. I pp.245-251, 1992.
2. Yuen, M. F., Tan, S. T., Sze, W. S. and Wong, W. Y., 'An octree approach to rough machining', Proceedings of the Institute of Mechanical Engineering 201(B3) pp.157-163, 1987.
3. Su, C. J. and Mukerjee, A. 'Automated machinability checking for CAD/CAM', IEEE Trans. on Robotics and Automation 7, pp.691-699, 1991
4. Wu, M. C. and Liu, C. R., 'Flexible process planning for finish machining based on process requirements modelling', International Journal of Computer Integrated Manufacturing, 4, pp.121-132, 1991
5. Park, J. Y. and Khoshnevis, B., 'A real time computer-aided process planning system as a support tool for economic product design', Journal of Manufacturing System, 12, pp.181-193, 1993
6. Park, H. D. and Mitchell, O. R., 'CAD based planning and execution of inspection', Proceedings of Computer Vision Pattern Recognition Conference, Ann Arbor, MI, pp.858-863, 1988
7. Lin, A. C. and Chang, T. C., 'An integrated approach to automated assembly planning for three dimensional mechanical parts'. International Journal of Production Research, pp.1201-1227, 1993
8. Henderson, M. R. and Razdan, A., 'Features based neighbourhood isolation techniques for automated finite element meshing', Geometric Modelling for Production Enginnering, M.J. Wozny, J. U. Turner and K. Preiss, pp.301-319, North Holland, Amsterdam, 1990
9. Gindy, N. N. Z., 'A hierarchical structure for form features', International Journal of production Research 27, pp.2089-2103, 1989
10. Jones, R., Mitchell, S. R. and Newman, S., 'Feature-based systems for design and manufacture of sculptured products', International Journal of Production research, 31(6), pp.1441-1452, 1993
11. Mitchell, S. R., Jones, R., and Newman, S., 'A structured approach to the design of shoe lasts', Journal of Engineering Design, 6(2), pp.149-166, 1995
12. Mitchell, S. R., Jones, R., and Hinde, C., 'An initial data model, using object-oriented paradigm, for sculptured-feature-based design', Research in Engineering Design, 7, pp.19-47, 1995

A method for recognizing feature interactions

C. Zhang, K.W. Chan, and Y.H. Chen
Dept. of Mechanical Engineering
The University of Hong Kong

ABSTRACT :

Handling feature interaction is an unsolved issue in feature recognition approach. This paper presents a method for recognizing the presence of feature interactions. Firstly, based on the convex hull concept, a so-called Reference face is defined. Secondly, by adding the Reference face into the AAG(Attributes Adjacency Graph), a modified AAG is obtained. Two general feature types, namely depression and protrusion features, are identified by the Reference face The basic features such as Slots, Pockets and Bosses are represented by the modified AAG Any features that remain unrecognized by the modified AAG are regarded as interacting features. The types of reference faces and feature face are also classified. Based on the kind of face classification, the interacting features are finally recognized via a process of virtual face extension and volume addition.

KEYWORDS:
Feature Recognition, Convex Hull, Attributed Adjacency Graph, Feature Interaction

1. INTRODUCTION

The concept of feature for supporting design and manufacturing applications has received much attention. The use of features is seen by many researchers as the key to a genuine integration of CAD and CAM. Generally speaking, the definition of feature is application dependent In our work, we consider form features which can be defined as a part's geometry associated with process planning entities such as Slots and Pockets.

In practice, the definition of feature is highly dependent on the feature representation methods (1). According to Shah (2), the methods of modeling features can be classified into three groups:

1) Interactive feature definition in which a CAD model is created first and features are then defined manually by means of picking entities on an image of the model rendered on the display screen.
2) Automatic feature recognition in which a CAD model is processed by a computer program for discovering and extracting features automatically.
3) Design by features in which a CAD model is defined directly in terms of features from the beginning of the design process.

The automatic feature recognition approach is attractive because the feature recognition algorithm can be constructed to suit different applications. Most existing feature recognizers deal with basic features which are pre-defined, they run into difficulties when the basic features interact each other. Due to the interactions, some of the faces that belong to a feature may be absent, partially missing, or split into several sections. Hence, one of the major difficulties found in most feature recognition methods is in dealing with complex feature interactions that do not conform with the predefined feature patterns

This paper firstly presents a review of typical research work on feature recognition and feature interaction. Our approach of representing and recognizing feature and feature interaction is then presented and discussed with some illustrations. The final section concludes our work and gives recommendations foı future work.

2. REVIEW OF FEATURE RECOGNITION WORK

Several methods of feature representation and recognition have been used such as graph-based, syntactic pattern, cavity volume and rule-based. Some typical examples are briefly described below:

In Floriani work (3) a method for extracting certain classes of form features from a boundary model was described. Feature recognition is based on the identification of the so called cutnode and biconnected components in a face-oriented graph of the boundary model. For example, for an object containing a Pocket, the face that has an inner edge loop due to the opening of the Pocket depression is a cut node in the face-oriented graph. After removing the cut node, the graph would be divided into two biconnected components.

Choi et al (4) used an approach known as syntactic pattern recognition which involves searching for string patterns of elements that correspond to some parts in the B-rep model. Chuang et al (5) also presented a syntactic pattern method to automatically recognize compound features using Web grammar parsing on a graph.

Ferreira et al (6) presented a method based on the convex hull concept. The method involves the determination of the convex hull for the component's faces and the edges within the convex hull that form a feature. The feature recognizer merges two or more features together, if they have some common geometric characteristics. But the method is not sufficient in dealing with feature interaction.

Donaldson (7) adopted a rule-based approach for feature recognition. He characterized a richer set of shape properties on which search of features was based. A feature analysis system was incorporated in a 2.5D CAD system for generating a feature description to support NC part programming.

A common weakness of the above reported techniques, however, is that they have not sufficiently dealt with the issue of feature interaction. Published work that have attempted to tackle the issue of feature interaction are briefly reviewed below.

Joshi et al (8) described the use of an Attributed Adjacency Graph (AAG) for representing feature pattern. The AAG is basically a face-edge graph in which the edge convexity or concavity had been added as an edge attribute. Joshi used some heuristic rules to direct the recognition of some feature interaction.

Prebhakar et al (9) developed a feature recognition algorithm using neural-network technology. A suitable net architecture which is similar to the multi-layer perceptron model was used. The algorithm can solve some feature interactions based on certain defined conditions. Feature interactions that involve the destruction of one or more feature faces can not be handled by the algorithm.

Suh et al (10) presented an approach to handle the feature interaction problem by classifying the feature space into three groups: (1) features remaining without interaction, (2) features to be removed, and (3) features remaining after being partially removed. The last group is further classified into three subgroups: (i) only the face set being removed, (ii) only the boundary edge path being removed, and (iii) the face set including the boundary edge path being removed. It further classifies the last group into only the face set being removed, only the boundary edge path being removed, and the face set including the boundary edge path being removed. The three interaction cases in the last group require modification procedures that perform three functions: (a) decide whether the remaining part of an existing feature is valid for a feature definition, (b) update it as a new feature, and (c) define the feature's relationship.

Gu et al(11) made use of connectionist modeling methods, based on neural network modeling for pattern recognition to deal with feature recognition. The attributed adjacency graph(AAG) extracted from a (B-Rep) solid model is converted to attributed adjacency matrices(AAM) that can be used as input data for the neural network model to train and recognize feature patterns. With this technique, the system can self-reconstruct its recognition abilities for new features by learning without a priori knowledge and recognize and decompose intersection features.

Kumar et al (12) presents a novel approach to recognizing interacting features. Based on the AAG representation and two defined types of feature intersections, they developed a recognition algorithm which involves the splitting of some faces, creation of virtual faces, recognition and manipulation of the AAG in a matrix form and finally, the recognition of feature.

Although the above reported methods can solve some feature interactions, however, recognition of features which involve the destruction of the basic feature patterns still remains an unresolved issue. We present a different approach to solving the feature interaction recognition in the sequel.

3. FEATURE REPRESENTATION

Prior to the description of our feature representation scheme, several terms need to be clarified first. These terms are reference face, RAAG, cavity volume and Extensible face.

3.1 REFERENCE FACE

A 2D convex hull (6) can be defined as the minimum convex polygon that encloses a face. A convex polygon is one in which a straight line from one point to another lies wholly within the face. Thus, the convex hull of the face shown in Fig. 1(a). is the boundary ABCD of the face itself because any lines in Fig. 1(a) which connect two points are within the face. The convex hull of the face shown in Fig. 1(b). is given by the polygon EFHI, because in Fig. 1(b), the line connecting the points E and F is outside the face. In the latter case, the face is said to be a non-convex face and the edges FG and GH that do not belong to the boundary of the convex hull are termed as non-convex hull edges. A face that has one or more inner edge loops is also considered as a non-convex face.

In our work, it is assumed that the non-convex edges of a non-convex hull face are connected to some features faces. In other words, the presence of non-convex face indicates the presence of features. The non-convex face is therefore called the reference face. Also for simplicity reason, reference face is abbreviated as rface in the text. Furthermore, two rfaces are called relevant if they are adjacent to one or more common feature faces.

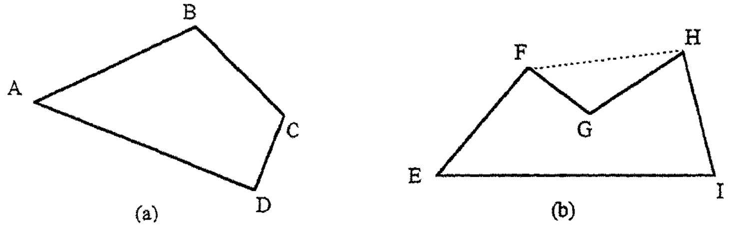


Fig. 1: The concept of convex and non-convex face.

3.2 RAAG

The AAG as originally proposed by Joshi (8) is essentially a face-edge graph representation of an object's boundary. The Attributed Adjacency Graph(AAG) can be defined as a graph G=(N, A, T) where N is the set of nodes, A is the set of arcs, and T is the set of attributes to arcs in A, such that

- For every face f in F, there exists a unique node n in N.
- For every edge e in E, there exists a unique arc a in A, connecting the nodes n_i and n_j, corresponding to face f_i and f_j which share the common edge e.
- Every arc a in A is assigned an attribute t, where t=0 if the faces sharing the edge form a concave angle and t=1 if the faces sharing the edge form a convex angle.

We consider that the AAG is weak in discerning feature interaction conditions due to its limited information. In order to facilitate our feature representation and enhance the recognition, we modify the original AAG by including the following attributes in the edge arcs and face nodes of the graph:

Attributes attached to the edge arcs

1. t = 0 edge arc is a concave edge
2. t = 1 edge arc is a convex edge
3. t = 2 edge arc is a virtual arc linking two relevant rfaces which are anti-parallel to each other.
4. t = 3 edge arc is a virtual arc linking two relevant rfaces which are perpendicular to each other.

Attributes attached to the face nodes

1. a = 0 face is an rface and does not contain an inner edge loop
2. a = 1 face is an rface and contains an inner edge loop

The modified AAG is named RAAG as an indication of the inclusion of the rface information in the graph. We use the RAAG for feature representation and recognition as shown in Fig. 3.

3.3 CAVITY VOLUME

One of the most widespread used methods of classifying form features has been based on the geometry and topological arrangement of the boundary faces that constitute a form feature. For instance, a rectangular Pocket can be considered to consist of a set of wall faces and a single bottom face. Topologically, the wall faces are adjacent to the bottom face and geometrically, the wall faces are perpendicular to the wall faces. The specific set of faces that constitute the unique form of a feature are actually present on the object boundary and they are hereby termed as feature faces. If the boundary of the feature faces are extended, new virtual faces can be created. A virtual volume can also be formed by considering the volume bounded by the feature faces and the extended virtual faces as shown in Fig. 3.

Such a virtual volume is termed a cavity volume which is equivalent to the concept of volume decomposition approach reported in (13), (14).

3.4 CLASSIFICATION OF BASIC FEATURES

We make use of the concept described above to define seven basic features as shown in Fig. 2. The corresponding RAAG representations of the seven basic features are illustrated in Fig. 3.

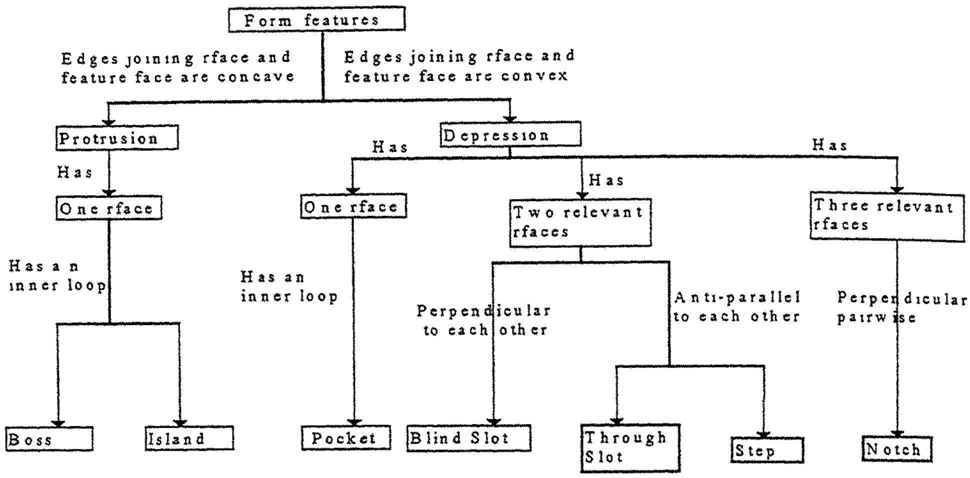


Fig. 2: The basic feature taxonomy.

For example, the Pocket shown in Fig. 3 has a rface rf1 which contains of an inner edge loop. The feature faces f1, f2, f3 and f4 are connected with rf1 via the inner edge loop. The edge convexity of the edges are all convex. The Through Slot shown in Fig. 3 has two relevant rfaces rf1 and rf2. The dotted line joining rf1 and rf2 has an attribute value of 2. It indicates that rf1 and rf2 are anti-parallel to each other since their surface normals are opposite in direction. Rfaces rf1 and rf2 are connected with feature faces f1, f2 and f3 with convex edge. Rfaces rf1 and rf2 are also said to be relevant because they are adjacent to the common feature faces f1, f2, and f3. The Notch shown in Fig. 3 has three rfaces rf1, rf2 and rf3. The dotted lines joining the rfaces have an attribute value of 3 which indicate that three rfaces are mutually perpendicular to each other. Rfaces rf1, rf2 and rf3 are also said to be relevant because they are adjacent to the common feature faces f1, f2 and f3.

3.5 CLASSIFICATION OF RFACES

In our work, the designed objects are assumed to be 2.5D (15) and contain basic form features as defined and depicted in Fig. 2. However, as feature interaction causes some feature faces being partially or wholly removed, or split into two or more faces, the relationship of features is complicated. The relationship of features can be reflected by three types of rface as defined below.

1) Simple rface.

A rface that is connects one or more simple features that are spatially disjoint.

2) Compound rface.

A rface that is connected with two or more simple features that are adjacent to each other.

3) Combined feature condition.

A rface that is connected with a combination of simple features and compound features.

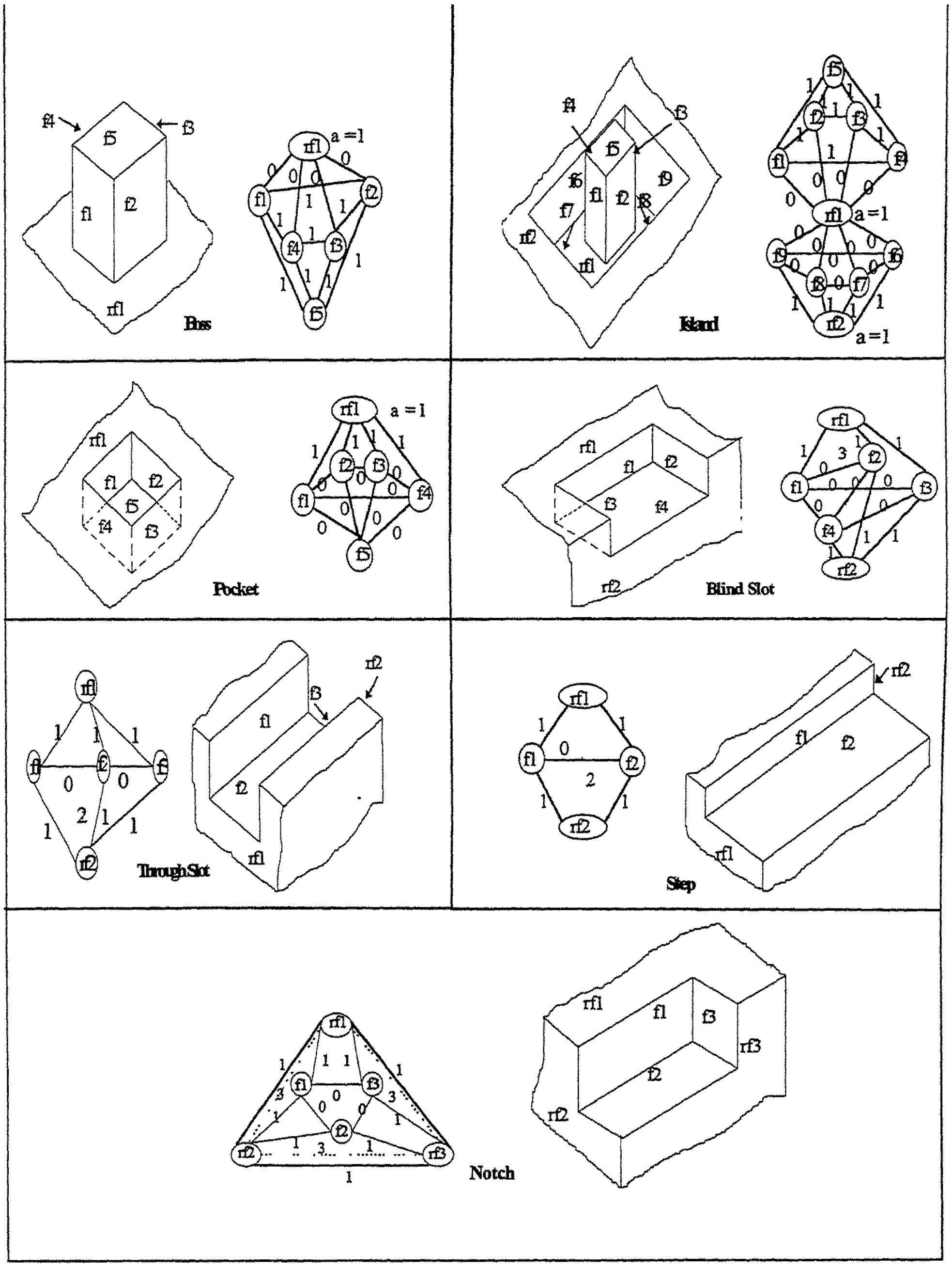


Fig. 3: The RAAG patterns of the 7 basic features

A rface can also be considered to be passive or active. A passive rface is one that does not constitute a feature face, while an active rface is one that constitutes a face of at least one feature. It should be noted that both passive rface and active rfaces may be simple rfaces, compound rfaces or combined rfaces.

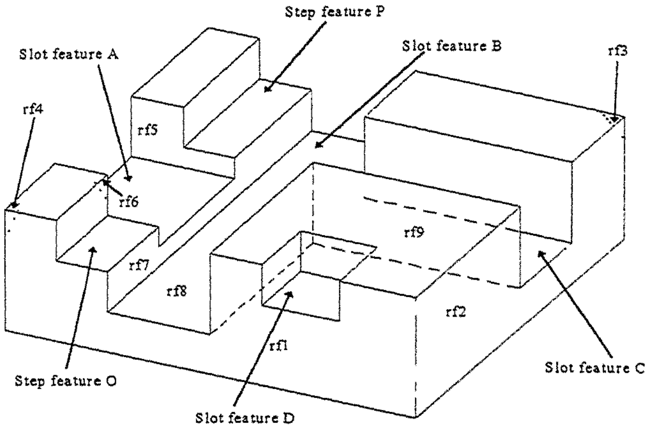


Fig. 4: An example to illustrate the feature interaction conditions.

As shown in Fig. 4. rf1 and rf3 are passive compound rfaces because they do not constitute a feature face and they are connected with Step features P and O and Slot feature B. rf2 and rf4 are passive simple rfaces because they do not constitute a feature face and they are connected with Slot features A and C respectively. rf5, rf6 and rf7 are active simple rfaces because they constitute a feature face and they are connected with Step features P and O and Slot feature A respectively. rf8 is an active rface shared by Slot features B and C. Slot feature A has two rfaces, namely rf5 and rf6. rf1 is a passive combined rface because it does not constitute a feature face and it is connected with Step feature O, Slot features B and D. Step feature O and Slot feature B are compound feature while Slot feature D is a standalone basic Slot feature.

Also we use 2 additional clues to identify the feature components in a compound feature by examining the convexity of the common edges relevant to the rfaces of the compound feature:

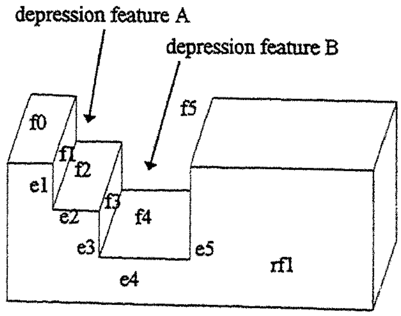


Fig. 5: An example to illustrate the two additional clues.

1) In the case of depression features, a convex edge indicates the end of one feature and the start of another feature.

2) In the case of protrusion features, a concave edge indicates the end of one feature and the start of another feature.

To explain this by an example shown in Fig. 5, the non-convex hull edges of rface rf1 are e1, e2, e3, e4 and e5 which are all convex to rf1, so they indicate the presence of depression features A and B. As the edge between f0 and f1 is convex, by applying the first clue, f1 is the start face of depression feature A. Similarly, since the edge between f2 and f3 is convex, f2 is therefore the end face of depression feature A and f3 is the start face of depression feature B.

3.6 EXTENSIBLE FACE

When two features interact each other, the feature faces which belong to different features interact each other. As a result of interaction, some faces in the feature are partially or wholly removed. The resulting faces of one feature are convex to the resulting faces of the other feature separately. We defined this kind of feature faces as Extensible faces, namely Eface. Only Eface can be extended for recovering the missing face. An active rface may be an Eface if it is convex to other feature faces. An Eface must have a corresponding Eface which belongs to another feature convex to it. We defined the edge between the two corresponding Efaces as an Eedge. The Eface can only be extended along the Eedge.

For example, in Fig. 7, faces f5 and f4 are Efaces because they are faces of two features and they are convex to each other. E1 is an Eedge.

4. FEATURE RECOGNITION ALGORITHM

Our feature recognition algorithm is devised according to the feature representation and classification of rfaces defined above. The basic mechanism of the algorithm is to make use of the classification of rfaces to locate the presence of basic features as well as feature interactions. The presence of the basic feature which are intermingled in a feature interaction is exposed by means of Eface extension along the Eedge and merging operation for forming the missing feature faces. Recognition of features is based on the pattern matching method using the basic RAAG patterns defined in Fig. 3. Cavity volumes of the recognized features are then constructed to fill the volume of the original object, thereby simplifying the original object geometry during the feature recognition process.

In detail, the algorithm starts with the generation of the RAAG of the modeled object and putting all the determined rfaces into a set. The rface set is then subdivided into two subsets, namely an active rface set and a passive rface set. The active rface set is evaluated first followed by the evaluation of the passive rface set. In each step, the Efaces and Eedge set should also be determined in order to decide which face should be extended to build the virtual face.

Step 1) In the active rface set if there is one or more rfaces shared by two or more features, then the relevant rfaces of each rface candidate in the active rface set are considered for determining which feature faces can be extended. Virtual faces are built by extending the Efaces along the Eedge in order to divide the rface into two or more faces which may or may not be rfaces. The extended faces may need to be merged together to form a new face. If any simple feature can be recognized from the extended face configuration by means of matching the extended face RAAG with the basic RAAG patterns shown in Fig. 3., then the corresponding cavity volume of the recognized basic feature is determined. The original volume of the object is then unioned with the cavity volume to form a new object. The RAAG of the new object is built. This step is repeated for the new object until there is no more rface shared by two or more features in the active rface set. The procedure then goes to step 2).

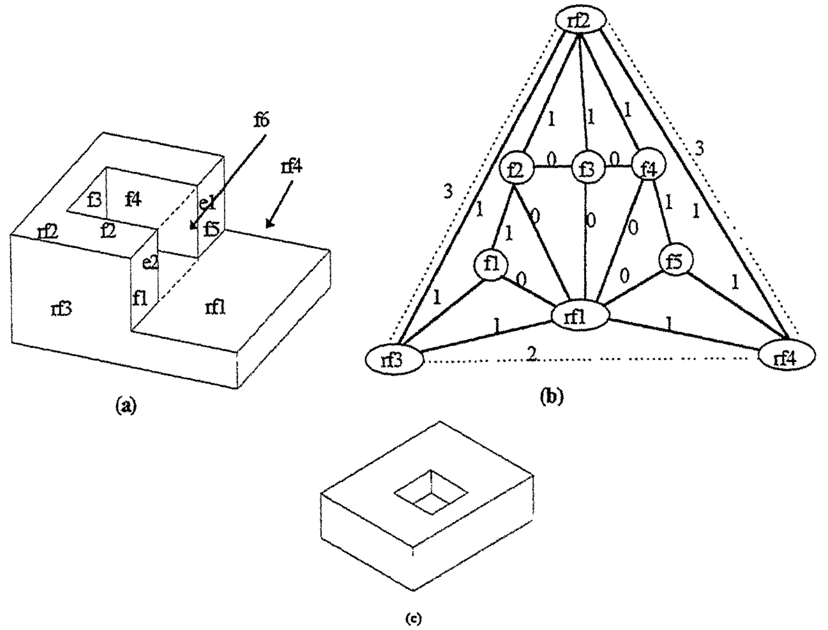


Fig. 6: An example for illustrating Step 1) of the algorithm.

For instance, for the object shown in Fig. 6(a)., the active rface set contains only rf1 which is shared by two features, feature1(f1 & rf1) and feature2(f5 & rf1). The relevant rfaces of rf1 are rf2, rf3 and rf4. rf3 is anti-parallel to rf4. From the RAAG of the object shown in Fig. 6(b)., it can be seen that the subgraph (f1, rf1 & rf3) and the subgraph (f5, rf1 & rf4) indicate the possibility of two Step features. The Eface set contains f1, f2, f4 and f5. The Eedge set includes e1 and e2. A virtual face f6 is built by extending Efaces f1 and f5 along the Eedges e1 and e2 separately. This virtual face divides rf1 into two faces. As a result of building this virtual face, a Step feature becomes apparent. Hence, the corresponding cavity volume of the Step feature is added to the original object giving a new object shown in Fig. 6(c).

Step 2) In the active rface set if there is one or more disjoint simple rfaces and each of them is connected with one or more basic features, then the connected basic features of each rface candidate are recognized by means of feature pattern matching with the basic RAAG patterns shown in Fig. 3. The corresponding cavity volumes of the recognized disjoint features are determined and are then unioned with the original volume of the object to form a new object. The RAAG of the new object is then built and this step is repeated until all the rface candidates are treated. The procedure then goes to step 3).

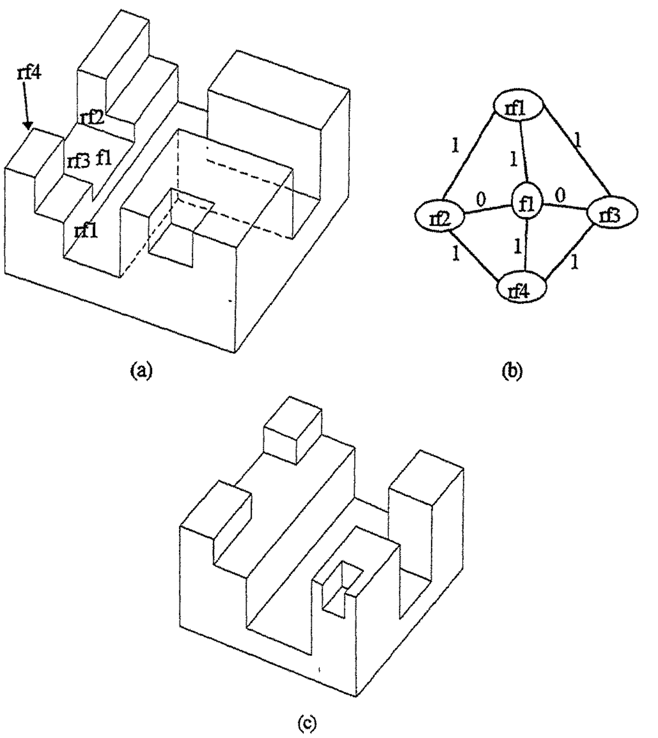


Fig. 7: An example for illustrating Step 2) of the algorithm.

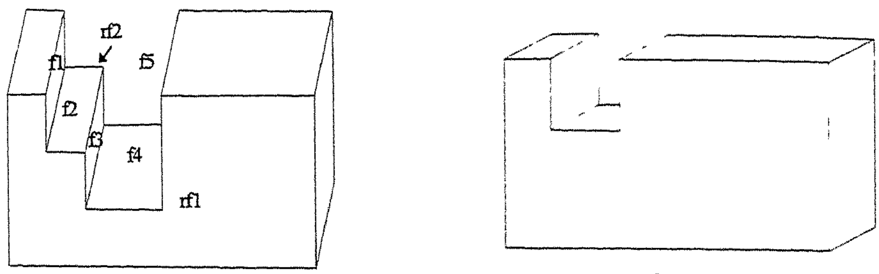
As shown in Fig. 7(a)., rfl is a simple active rface and rf4 is its relevant rface because rfl and rf4 are anti-parallel to each other. From the RAAG subgraph shown in Fig. 7(b)., it can be seen that faces (fl, rfl, rf2, rf3 & rf4) constitutes a Through Slot feature which can be recognized by matching with the basic RAAG patterns shown in Fig. 3. The cavity volume of the recognized Through Slot feature is generated and is added with the original object to give a new object as shown in Fig. 7(c).

Step 3) In the active rface set if there is no compound rface, procedure will go to step 4), otherwise, any basic features contained in the the compound features of each rface candidate are recognized by means of applying the 2 additional clues described earlier and the graph matching with the basic RAAG patterns shown in Fig. 3. The corresponding cavity volumes of the recognized features are added to the original object to yield a new object. This step is repeated until there is no more compound rface left in the active rface set. Procedure then goes to step 4).

Step 4) In the active rface set if there is no combined rface, procedure will go to step 5), otherwise, the procedures as described in step 2) and 3) are activated to deal with the basic features and the compound features. This step repeated until there is no more combined rface left in the active rface set. Procedure then switches to step 5).

Step 5) When arriving at this step, all the active rface candidates should have been handled by the above procedures. The remaining procedures therefore aim to tackle the passive rface set. In the passive rface set if there is no simple rface, procedure will go to step 6), otherwise, the basic features of each passive rface can be easily recognized by matching with the basic RAAG patterns as shown in Fig. 3. New object is built by adding the corresponding cavity volumes of the recognized features to the original object. This step is repeated until there is no more simple rface left in the passive rface set. Procedure then goes to step 6).

Step 6) In the passive rface set if there is no compound rface candidate, procedure will go to step 7), otherwise, any basic features contained in the the compound features of each rface candidate are recognized into the corresponding basic features by means of applying the 2 additional clues described earlier and the graph matching with the basic RAAG patterns shown in Fig. 3. Similarly, new object is built by combining the corresponding cavity volume of the recognized features with the original volume of the object. Again, this step is repeated until there is no compound rface left in the passive rface set. Procedure then goes to step 7).



(a) (b)

Fig. 8: An example to illustrate Step 6) of the algorithm.

For example, in Fig. 8(a)., rf1 and rf2 are compound rfaces. Faces f3, f4 and f5 are recognized as a Slot feature and the corresponding cavity volume is added to the original object forming the new object shown in Fig. 8(b). which contains only one Slot feature.

Step 7) Finally, in the passive rface set if there is no combined rface, the algorithm will terminate, otherwise, the procedures as described in step 5) and 6) are activated to deal with the basic features and compound features. This step is repeated until there is no more combined rface candidate left in the passive rface set. The feature recognition algorithm finishes here.

5. PLAN OF IMPLEMENTATION

An experimental prototype system will be developed to study the feasibility of our feature recognition algorithm for extracting features from a CAD database. The system is integrated with AutoCAD by using the Application Programming Interface(16) of Advanced Modeling Extension of AutoCAD as development tools. AutoCAD is used as a design platform for modeling the desired object and for providing the boundary information to the feature recognition module.

6. DISCUSSION AND CONCLUSION

In this paper, we described a method which uses a modified AAG to facilitate the representation and recognition of depression features specifically from a feature interaction condition. The modified AAG contains more attributes which are determined based on the convex hull faces. Consequently, the modified AAG provides more clues in the process of feature detection and recognition. The recognition of features is enabled by re-constructing the necessary boundary faces that might have been lost due to feature interaction. The extraction of features is facilitated by adding the cavity volumes of the recognized features back to the original volume of the object. The geometry of the original object and the corresponding modified AAG representation are therefore progressively simplified during the feature recognition process.

There is still a problem to know the likelihood of which basic feature should be chosen to match the component of the unrecognized interacting feature. When this can not be processed automatically by the algorithm, the decision has to be performed manually for the current implementation of the algorithm. Our plan of dealing with this face extension problem is by making use of the learning ability of a Neural Network model with the RAAG matrix input. Neural Network has been proven more successful in classification. As the Neural Network ability mainly depends on the training set, we can develop the Neural Network to deal with our problem with a relatively complete training set. At the same time, the architect of Neural Network is also very important.

REFERENCES

1. Case K. and Gao J., "Feature technology: an overview", Int. J. Computer Integrated Manufacturing, Vol.6, Nos.1&2, pp.2-12, 1993.
2. Shah J.J., "Assessment of features technology", Computer-Aided Design, Vol.23, No.5, June, pp.331-343, 1991.
3. De Floriani L, "A graph-based approach to object feature recognition", Proc. 3rd ACM Symp. Computational Geometry, Waterloo, Canada, pp.100-109, 1987.
4. Choi B., Barash M. and Anderson D., "Automatic recognition of machined surfaces from a 3D solid model", Computer-Aided Design, Vol.16, No.2, pp.81-86, 1984
5. Chuang S.H. and Henderson M.R., "Compound feature recognition by Web grammar parsing", Journal of Research in Engineering Design, Vol.2, Iss.3, pp.147-158, 1991.
6. Ferreira J.C.E. and Hinduja S., "Convex hull-based feature-recognition method for 2.5D components", Computer-Aided Design, Vol.22, No.1, pp.41-49, 1990.
7. Donaldson I.A. and Corney J.R., "Rule-based feature recognition for 2.5D machined components", Int. J. Computer Integrated Manufacturing, Vol.6, Nos.1&2, pp.51-64, 1993.
8. Joshi S. and Chang T.C., "Graph-based heuristics for recognition of machined features from a 3D solid model", Computer-Aided Design, Vol.20, No.2, pp.58-66, 1988.
9. Prebhakar S. and Henderson M.R., "Automatic form-feature recognition using neural-network-based techniques on boundary representations of solid models", Computer-Aided Design, Vol.24, No.7, pp.381-393, 1992.
10. Suh H. and Ahluwalia R.S., "Feature modification in incremental feature generation", Computer-Aided Design, Vol.27, No.8, pp.627-635, 1995.
11. Z. Gu, Y.F. Zhang and A.Y.C. Nee, "Geneeric form feature recognition and operation selection using connectionist modelling", Journal of Intelligent Manufacturing, pp263-273, 1995
12. Senthil Kumar A., Salim F. K. and Nee A.Y.C., "Automatic recognition of design and machining features from prismatic parts", The Internationa Journal of Advanced Manufacturing Technology, Vol 11, pp136-145, 1996.
13. General Dynamics Corporation, "Volume decomposition algorithm - Final report", CAM-I Report R-82-ANC-01, 1985.
14. Woo T. C., "Feature extraction by volume decomposition", Proceeding of the conference on CAD/CAM, USA, pp39-45, 1982.
15. Chan, K.W.A., "A knowledge-based approach for the extraction of machining features from solid models", Ph.D. Thesis, Loughborough University of Technology, May 1993.
16. "Advanced Modeling Extension Reference Manual", Autodesk, Inc, AutoCAD Release 11, 1991.

STANDARDIZATION OF BASIC DESIGN SEQUENCES OF HEAT-EXCHANGER FOR POWER PLANT

Hideyuki Yamaguchi† Zhihua Zhang†† Isao Nagasawa†††
Masanobu Umeda†††

†Department of System Technology Developing. Choryo Design Ltd. Co., Japan,
email:yamaguti@cad.mse.kyutech.ac.jp
††Graduate School of Faculty of Computer Sci. and System Eng . Kyushu Institute of Technology, Japan
†††Faculty of Computer Sci. and System Eng.. Kyushu Institute of Technology, Japan

ABSTRACT

Main machine of power plant design business is generally single piece of order from customs. Design is just performed when the order is accepted. To make the design work more efficient, something has been done in division of design work, but as the side effect, it become more difficult in technique inheritance for recruit designers. (1). We have developed a method which make it possible both in getting good efficiency in design and the comprehensibility for technique inheritance by systematically sort out the design knowledge from experienced designers with the aim of computer-based support. (2). We give a definition on design space by limiting the operating scope of design solutions and described the device on design calculation and the design space reduction to help reduce the quantity of calculation.

By the systematic method of design sequences, we give an experiment on the description on design sequences for function design of heat-transfer component of shell and tube type heat-exchanger, and got good results from view of comprehension, quality of design solution, efficiency of design calculation.

KEYWORDS

Design Space, Generate-and-Test, Design Sequence, Design Calculation

1. INTRODUCTION

In design of main power plant machine, the design aim is to get high reliability for power providing and low cost of product for holding high competition. Because some of them are custom-made machine, it is also key item to improve the productivity of design work itself. In order to satisfy these requirements, machine makers have done something in standardization of design approach as well as the division of design work. But the standardization work on design approach in the past is just concerned with the design sequence for application, and the division of design work has been done too over, therefore, it become more difficult for designers to get design knowledge in wide scope. Because of this, the important design technique may be lost gradually. More, in recently, designers who can do new developing design work become few, it become more and more difficult to cope with the problem like the variety of customers needs and internationalization of design business.

On the other hand, the intelligent CAD[1] is hoped to be the tool which can improve the complicated design work, and promote the quality of design as well as productivity of design business.

In this paper, we describe the systematic method for design sequences which are got by analyzing the basic design work in the design of shell and tube type heat exchanger for power plant.

2. DESIGN SPACE

In design work, designers treat the design objects by abstracting the design object according to the nature of design work. Here, the design object is represented by a set of attributes[2], and there are subordinate relation among the design attributes.

Now, let us consider a design problem of a cylinder type water tank with the outer diameter do, thickness of board t, height h, from these attributes we can get the inner volume of the tank v. The design condition is to ensure the storage of water volume is satisfied with $2000 \leq v \leq 3000$, and the size of the tank are described as the expressions of $do \in \{10, 20, 30\}, t \in \{1, 2, 3\}, h \in \{10, 20, 30\}$.

The design variables and the subordinate relation among of them are shown as Fig. 1.

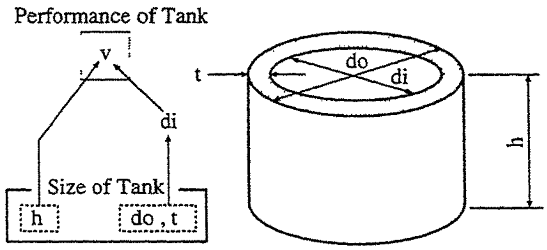


Fig. 1 Relationships between attributies of functional design of tank

In this occasion, the design solution can be represented by four design variables of do, t, h, v, and there exist relation expressions like $di = do-2*t, v = \pi/4*do^2*h$ among the design variables.

Based on the design variables, the design space can be defined by a four-group set as $\langle P, D, F, C\rangle$. Where, P is the set of design variables, D is the set of domain of the variables(in the following, we call as domain constraints), F is represent the subordinate relations, and C is used to describe the conditions which the design variable should satisfy. By this, we can represent the design space of previous problem as follows.

$$
\begin{aligned}
DS &= \langle P, D, F, C\rangle \\
P &= \{do, di, t, h, v\} \\
D &= \{do \in \{10, 20, 30\}, t \in \{1, 2, 3\}. h \in \{10, 20, 30\}, v \in Real\} \\
F &= \{di = do - 2 * t, v = \pi/4 * di^2 * h\} \\
C &= \{2000 \leq v \leq 3000\}
\end{aligned}
$$

Generally, design problems in the basic design, can be formalized using the concepts of design space.

3. GENERATE-AND-TEST

Here, Let us consider about the solving method for design problems.

In design calculation, many problem can not be easily solved by analysis method, at this occasion, designers usually solve it by assumption of design attribute values firstly, and then to test them iteratively if the required design conditions are satisfied or not in trial and error way. This design calculating activity including the trial and error behavior can generally be represented as Generate-and-Test method, that is, a method which give assumption of solution at first and then to test the constraints[3].

The sequences of design calculation can be described by the dependency relation (which is called as data flow in following). Compare with the procedural representing method, the representation of data flow owns higher abstract ability, The representation with different form by control flow may have the same form by data flow representation, and this approach is the most suitable method for representing design calculation.

As an example, the sequences of design calculation about the problem described in previous section can be shown as in Fig. 2.

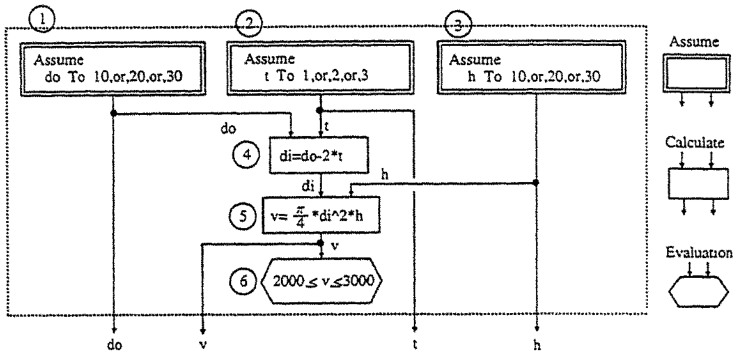


Fig. 2 Calculation sequences using generate-and-test method

where, ①,②,③are the assumptive value of design variables *do. t. h* respectively. ④ do the calculation of *di* from values of *do* and *t*. ⑤ is for the calculation of the inner volume *v* from *h* and *di*. After this, in ⑥, do the test to verify the design constraints are satisfied or not. The design solutions can be got after the sequences are performed.

To execute the calculating sequences, we can acquired two solutions as $\{20, 18, 1, 10, 2010.62\}$ and $\{20, 16, 2, 10, 2544.69\}$ from all the candidate design solutions with the numbers of $3\times3\times3 = 27$.

Like the demonstration in previous, the generate-and-test method can give a representation on design calculation in terseness way, but the method is not good for complex problem with large search space because of the problem of combination explosion.

4. TO REDUCE THE QUANTITY OF DESIGN CALCULATION

Many of machinery design problems can be processed using generate-and-test method, but in basic design of power plant machine discussed in this paper, there are big quantity in design calculation. To solve this problem, it is necessary to give a device on design calculation by designers.

For this, various of design calculation techniques have been provided by considering the complexity of calculation, the severity of design constraints and the different numbers of design variables and their domain. For example, the inverse operation method for the reformed problem using generate-and-test method; in accumulative calculation problems, the method of giving constraints on each element instead of giving constraints on the grand total and cutting the search branch at different stage; and the method of cutting branch of search before giving loose constraints on the problem, etc.

In this section, we give an introduction about the inverse operation method.

Inverse Operation If we take a look at the results of the problem described in section 3., we can discover that the value of the design attributes *do* and *h* have only one solution. The quantity of the calculation can be reduced if we can know this previously. This method is to reverse certain design attribute for a problem, solve the reformed problem and get a small scope results of it, then use the generate-and-test method to re-solve the given problem.

The calculating sequences using inverse operation method is shown in Fig. 3 for the example described in previous section which is used to indicate the generate-and-test method. In Fig. 3,*ProbA* is the reformed calculation part and *ProbB* is the original design calculation problem.

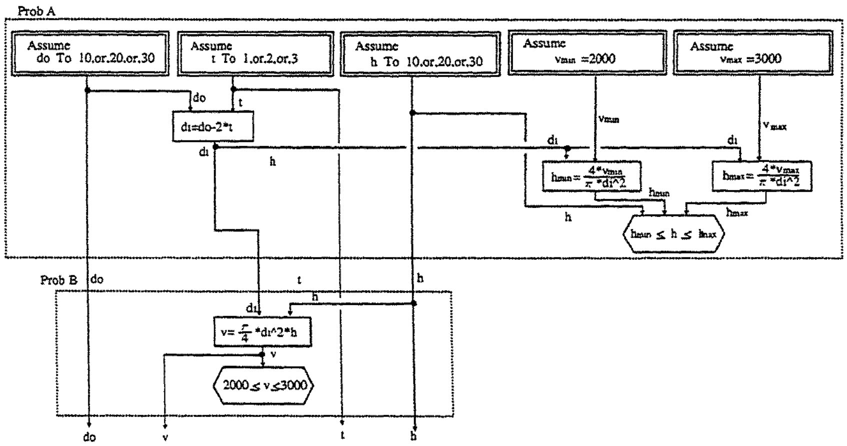


Fig. 3 Design sequences using inverse operation method

To execute the calculation of this problem, we can know that the combination numbers of attribute values is 18-groups in *ProbA*, and for *ProbB* it is 2-groups, we can understand the quantity of calculation has been reduced by comparing with the method in 3. which has 27-groups.

As the method of reducing calculating quality, besides the search method, there are some other methods which make the design be easy by processing the design space. In next section, we described some of the method in design space processing.

5. REDUCTION OF DESIGN SPACE

The good solutions with a balance in manufacturibility, performance, strength and simplicity of design are high quality design solutions. In design, if only designers can get the representative solutions with a good balance, the design can be considered as a good design. This idea is well used in the reduction of design space so as to reduce the quantity and complexity of design calculation.

The methods about the reduction of design space can be mainly divided into two kinds. One of the method is to reduce the design space of object into smaller design spaces and solve them in step by refinement way(which is called design space division method in following), and the other one is to use a approximate design space to solve the problem, by giving a detailed analysis on the solutions, the correctness of the approximation can be verified(which is called as design space reduction on approximation method in the following).

5.1 Design Space Division Method

The design space division is makeup numbers of design space as $P' \subset P$. The general method in space division is like dividing the design of container into the body design and cover board design problem in the structure, or the other division aspects based on different kinds of design requirements like functional design and strength design, etc. In addition, there are other methods, for example, in the problem of inner volume design for container, firstly to solve the area of cross section problem and make the design requirement to be satisfied, then give adjustment on the height of the container until the overall requirements are satisfied.

5.2 Design Space Reduction on Approximation Method

Design space reduction on approximation method is by adding new design variables to design space which have a approximate value of the design attribute, as the representative variable, it take the place of a group of variables such as all the arrangement states, and by doing like this, a design space can be made with reduced number of design variables. Generally, after getting the design solutions, in order to confirm the correctness of the approximate value of this variable, more detailed analysis will be performed. The approximate value of this variable can be acquired by the idea like the following description. If we can think the error is small to design solutions when the non-linear problem like the distribution of temperature and the distribution of load are considered as a linear problem, we can use the average value as the representative value to deal with the problem.

In design, we use both of the design space reduction method described in this section and the calculation methods which are described in previous section as well, and have greatly reduced the calculating quantity.

6. DESCRIPTION EXPERIMENT

A description experiment has been done to the function design(case A) of main components of container, the function and strength design(case B,C) of heat transfer part by adapting the concept of design space, problem solving methods for design calculation and the reduction of design space method to them. The numbers of design variables and the numbers of settled design problems about the problem done in the experiment are given in the Table 1.

Table 1 Experimental Results

experimental object	numbers of design variable	numbers of design problem
case A	11	2
case B	977	32
case C	45	12

(note). the container component consists of 2 parts,
and the heat transfer part consists of 357 parts.

We got the following result through the experiment.

The number of design variables in function design of heat transfer part is very big, this is because there are many attributes about the space for flow of fluid and the different arrangement states for the different position of heat transfer pipe. For this, many occasions are determine the representative sections and reduce the design space by approximate method.

7. CONCLUSION

In this paper, a method to formalized the design problem by design space has been presented, and some methods for designers to determine the design sequences have been proposed. Based on this, the comprehension about the design sequences can be promoted. But we can not go on with the model and method for getting a optimal design sequence. In the future, it is necessary to adapt the representation method about design sequence to many other design cases in order to verify the capability, also it is necessary to do the study on getting the optimal method. More, it need to do some research on designers education and the development of intelligent CAD system.

REFERENCES

[1] Hiroyuki Yoshikawa, Tetuo Tomiyama: Intelligent CAD(1)-Idea and Paradigm-, pp2-52, Asakura Shoten Ltd., Tokyo, 1989.

[2] Isao Nagasawa, Masatoshi Ito: Attribute Oriented Modeling, Simulation, Vol.10, No.2, pp110-118, 1991.

[3] Yoshiaki Tegoshi, Isao Nagasawa, Junji Maeda, Minoru Makino: An Information Processing Technique For A Searching Problem of An Architectural Design, Collected Papers of the Planning Department of the Association of Japan Architecture Society, No.405, pp157-165, 1989.

IMPLEMENTATION OF DESIGN FOR ASSEMBLY : ASSEMBLY-CENTRIC MODELING AND DIGITAL MOCK-UP SYSTEM

Yoong Ho Jung
School of Mechanical Engineering
Pusan National University
Republic of Korea

ABSTRACT

Most design changes are due to interference and fit-up as parts are assembled if a large product is developed by many concurrent engineers. These design errors are from the following reasons : (1) In design stage, designers cannot fully consider assembly sequence which is carried out in down stream process. (2) Designers release their drawings for manufacturing without fully checking interferences among parts including others'. (3) Initial design changes makes the secondary design changes if they are not completely resolved.

This paper first proposes the methodology of assembly-centric modeling with CAD systems in order for the assembly sequence to be considered in design stage. In the later part, the paper presents the implementation of a digital mock-up system in order to eliminate the second and the third reason of the design changes. Even though some of CAD systems partially present their digital mock-up functions, but they don't release details about implementation. In this paper, based on the requirements from designers, the digital mock-up functions are implemented with CATIA as CAD system, and ORACLE as database management system.

KEYWORDS

assembly, virtual, mock-up, database, concurrent, CATIA

1. INTRODUCTION

In the development of such a large complex product as an airplane or an automobile, a deep analysis reveals that tremendous potential for saving money in large machine development exists, as shown in Fig. 1, in reducing design changes which are from following reasons. First, designers who are in design stage cannot fully consider the assembly sequence of down stream process without making physical mock-up or real assembly. Second, designers release their drawings without fully checking the cross-functional interferences among parts and assemblies. The typical development process for large machine starts to divide design tasks by discipline, then assigns design engineers to each discipline. Structures are designed by one group of engineers, pneumatic systems by another, electrical subsystems by an others. Hundreds of disciplines come together in the product. But, it is impossible for an engineer in one discipline group to design with fully considering parts designed by others, which results in many troubles and engineering changes. Third, when initial design changes are not completely resolved, they makes the secondary design errors in neighboring parts.

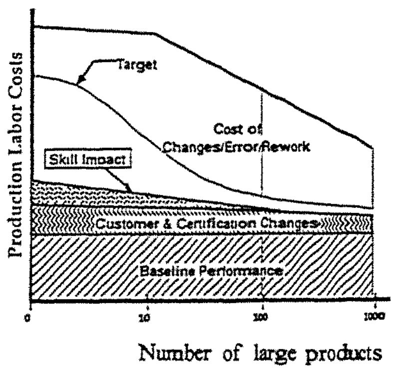


Fig. 1 Cost of design errors

Recently, many researches and methodologies are devoted to resolve these problems. Some of them are on the automation of design process with the aid of rule based system[1,2]. And others are on the methodology of design for assembly[3,4,5]. But their focus is not on eliminating the design errors, but on improving existing design with some design criteria. On the other hand, the concept of concurrent engineering is proposed. That is to say, designers and manufacturing engineers have to work together from design stage. In order for concurrent engineers to design with this concept, various virtual mock-up functions are developed by commercial CAD vendors[6] and dedicated programs[7,8]. But they don't release the details about implementation.

This paper first proposes the methodology of assembly-centric modeling with CAD systems, in order to reduce the first reason of design changes. Unlike part-centric modeling, the method is to model a part to have geometric data and positioning data seperately. Then the part positioning data is to be the relative orientation w.r.t. its assembly coordinates. In later part of this paper, the implementation of a digital mock-up system is presented to reduce the second and the third reasons of the design changes. The system uses relational database tables to manage all of the attribute data of a part with its geometric and positioning data, which are proposed to be separated in the methodology.

2. ASSEMBLY-CENTRIC MODELING METHODOLOGY

There are two approaches to develop a product : bottom-up and top-down approach. In bottom -up approach, individual components are first modeled, and then brought together. Designers with this approach are not expected to have a good grasp of cross-functional relation and assembly sequence of parts, even when developing a large product which is more than ten assembly levels. But in top-down approach, which is typically used in aerospace industry, the hierarchy of parts is defined first, then design is proceeded from the assembly level to detail part level. As a result, designers can consider the interrelation and assembly sequence of parts at the design stage.

The proposed methodology adopts the latter approach for developing large assemblies. That is to say, from Bill of Material, it first defines the assembly tree structure as shown in Fig. 2 (a), in which a leaf node represents a part, others represent assemblies, and a branch represents father-child relationship which means the position of part w.r.t. its assembly. Then with the given tree structure, a designer starts the process creating coordinates X_A of assembly A at the same position of main assembly coordinates X_M. Next, he moves X_A to the position where the assembly A is to be located. Similarly, he creates the coordinates X_P of part P at the same position of assembly coordinates X_A, then models the geometry of part P w.r.t. its own coordinates X_A. After modeling part P, he moves the coordinates X_P to the position where part P is to be located for assembling.

In order to realize this methodology, each part and assembly is modeled as following scheme. First, position data of part are isolated from geometric data of a part in order for the position data to represent the assembling process. The geometry is modeled w.r.t. its own reference which is used in

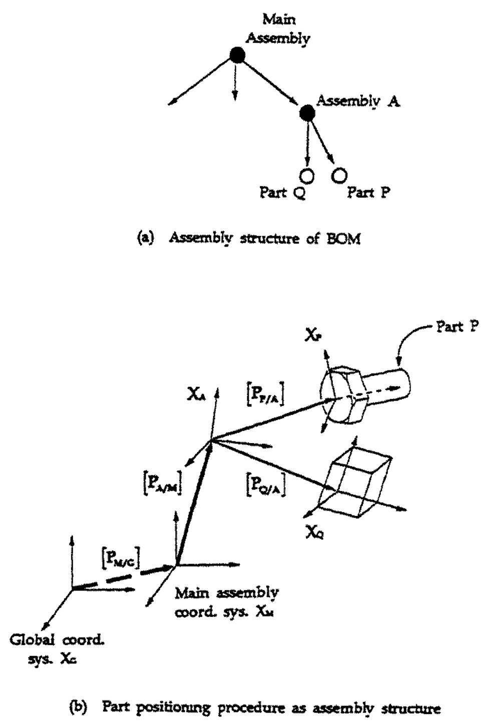


Fig. 2 Assembly-centric modeling procedure

down stream process of NC manufacturing, inspection, and assembling. As a result, the geometry can also be used in other assembly without any data transformation. Second, assembly is similarly modeled as a part except 'NULL' geometric data, so that it has its own position data w.r.t. its father assembly. The main assembly has its position data w.r.t. absolute coordinates. Third, the position data of a part and a assembly from its father is represented as matrix which has rotational and translational information. When all of these matrices of parts and assemblies are stored in database, the absolute position of a part can be calculated by multiplying the position matrices from main assembly to itself. The database to store the geometric data and position matrix are explained in next section.

The proposed modeling methodology has following advantages. First, designers can prevent potential design errors by considering the assembly sequence in design stage, since the positioning process of part and assembly is similar to real assembling process. Second, designers can easily resolve an interference among neighboring parts. Since all of the parts in the assembly are automatically moved together by moving assembly position, designers can adjust part and assembly position with ease and downward direction from assembly to part, which results in minimize the rework to resolve the interference problem.

3. DIGITAL MOCK-UP SYSTEM

If a designer can simulate assembling parts with solid models, he can find design errors before manufacturing parts. A digital mock-up system is a tool for eliminating the second and third reasons of design changes as mentioned before, since it creates complete prototype of assembly in computer screen, which is used for interference and clearance checking, simulating assembling, visualizing, and calculating mass properties of assembly, and so forth. There are two approaches to implement a digital Mock-up system - to use a solid modeler or to implement dedicated program. In this study, we adopted the former approach because of following reasons. First, model created with a solid modeler can be directly used for a digital mock-up, since the model is as accurate as the solid modeler allows. Second, even though designers have to create digital mock-ups as frequently as possible, they do not want to change their design environment to another in order to use digital mock-up functions. Third, since all of the solid modeling functions of visualization, evaluation, and modification can be used in the digital mock-up system, it is not necessary to develop such functions.

We developed a Digital Mock-up system in CATIA[9] environment with ORACLE [10], the relational database management system (R-DBMS), to manage geometric and attribute data of parts as shown in Fig. 3. It is easy to add and delete an item in relational database tables, and CATIA provides their program of CDM [11] (CATIA Data Management) to interface with ORACLE.

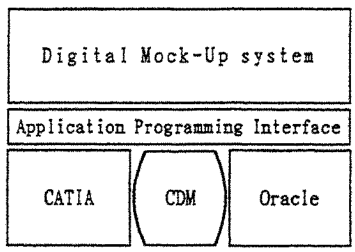


Fig. 3 Development environment

3.1 Requirement of specification

Before developing a digital mock-up system in order to support the proposed assembly modeling methodology and to be able to check cross-functional interferences, we gathered the requirements for the system from field engineers as follows.

First, the system has to create digital mock-ups in various ways. Since designer would like to check the interferences among parts not only designed by himself but also designed by other concurrent engineers, the system has to be able to assemble parts independent of owners. For this purpose, we proposed three different ways of creating digital mock-ups as shown in Fig. 4. That is, when a designer specifies a region in main assembly, the system creates a digital mock-up with parts in that region. This function enables designer to check cross-discipline interference of parts. Next, when a designer specifies function name of parts, the system creates a digital mock-up with only parts of that function. This enables designers to check interference among parts of the same function. Finally, the system creates a digital mock-up with parts whose names or identification numbers are selected by the user in a part list. This function enables designers to assemble any parts which are independent with discipline and location.

Second, the system has to write and read digital mock-ups. During the design process, designers have to create digital mock-ups as frequently as possible. But if he has to remember and input the same searching criteria for the mock-up which was constructed before, it makes designers avoid the checking process. So, the system has to be able to write and read a digital mock-up. But, the system has to write and read the criteria for searching parts, instead of the digital mock-up itself, since other designer may add new parts or change the old ones before the owner read the mock-up.

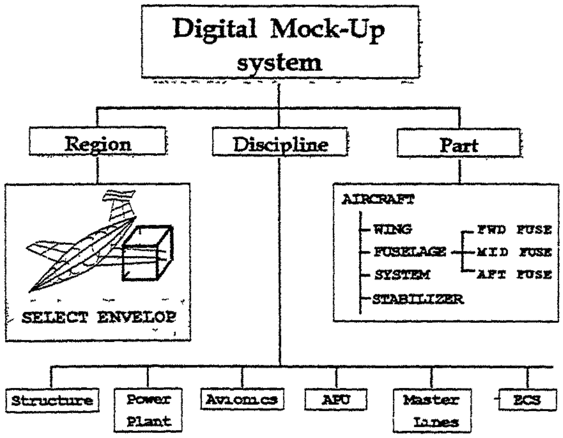


Fig. 4 Creating a digital mock-up scheme

3.2 Database table design

Developing large complex product creates tremendous part data - part numbers, part names, geometric models, relational data between assembly and parts. and other attributes - which are to be systematically controlled. In order to develop a digital mock-up system with the requirements, we designed relational database tables to manage those data as follows : (1) 'PART TABLE' to manage the data of Bill of Material, (2) 'MODEL TABLE' to manage geometric data, (3) 'RELATION TABLE' to manage the relational data between assembly and its parts, (4) 'MOCK-UP TABLE' to read and write the criteria for a constructed digital mock-ups, and (5) 'LOCK TABLE' for accessibility of data. 'PART TABLE', 'MODEL TABLE', and 'RELATION TABLE' have the same entity of ' COID' which represents the same part, since the part data in those tables have to be consistently managed. The detail of some important tables and the relation among the tables are explained below.

1) PART TABLE

'PART TABLE' plays an important role of the basis for database management. It stores, as shown in Table 1, all of part list data : part number, part name, revision number, status of approval, part type, material, and function name. The 'status' of part represents which stage in approval process : in designing, checking, or approved. When a designer wants to create a digital mock-up with parts of specific function, the system searches this table for parts with the same function name (which is defined as 'DISCIPLINE' in Table 1). When a designer selects some of part numbers in a part list for the mock-up, the system also searches this table for parts of selected part numbers. The geometric data of searched parts are stored in 'MODEL TABLE' which is explained in next subsection.

2) MODEL TABLE

A part can have several models : analysis model, exact model, faceted model, and others. This table stores and manages, as shown in Table 2, model data : model name, model type, geometric data, density of the part, envelope data, and mass properties. Geometric data is stored in "Long Field' which is provided in ORACLE to contain variable size data. The 'MIN_POINT' and 'MAX_POINT' in Table 2 represent minimum and maximum coordinates of box enclosing the model. When a designer wants to create a digital mock-up with parts which are in a specified region, the system compares the coordinates of the input region and enclosing box coordinates of each model in this table. If the enclosing box of a model is included in or overlapped with the region, the model is gathered to be a digital mock-up.

Attribute Name	Data Type	Description
COID	Char*08	Part I.D
COMPID	Char*08	Component I.D of the model in 'Part' Table
PART_NUMBER	Char*16	Part number
PART_NAME	Char*16	Part name
REVISION	Char*02	Revision number of the part
STATUS	Char*04	Current state of design process (WRK / CHK / APP)
PART_TYPE	Char*04	Part type (MASM / ASSM / COMP)
DISCIPLINE	Char*08	Discipline name
MATERIAL	Char*08	Part material

Table 1

Attribute Name	Data Type	Description
COID	Char*08	Part I.D
COMPID	Char*08	Component I.D of model in Model Table'
MODEL_NAME	Char*16	Model name
MODEL_TYPE	Char*08	Model Type (SOLID/DRAFT /SURFACE/MESH/NC/..)
MODEL_GEOM	Long Data	Geometric data of the model
MODEL_DENS	Float	Density of the model
MIN_POINT	Point	Envelope coordinate of model
MAX_POINT	Point	
VOLUME	Float	Volume of model
WEIGHT	Float	Weight of model
CG_POINT	Point	Center of gravity coordinates
IXX	Float	Moment of inertia
IYY	Float	
IZZ	Float	

Table 2

3) RELATION TABLE

'RELATION TABLE' defines the parent-child relationship between parts and their assembly, and relative orientation of a part w.r.t. its parent coordinates. An example shows a assembly tree structure in Fig. 5 (a), and some parts : P0, P1, P2, and P3. The relationship of a part consists with as many rows in the table as the number of children. In Fig. 5 (b), the part P1 relationship is defined in second and third row, since it has two children. That is, to define the P0-P1-P2 relationship, the identification number of P1 is stored in first column, and parent P0 in forth column, and one of children P2 in third column of the second row. And the twelve positional matrix elements of P1 relative to P0 are stored in the same row. The P0-P1-P3 relationship is defined similarly, except P3 instead of P2. After defining the part relationship in the table, we can calculate the position of a part by multiplying positional matrices from the main assembly to the part.

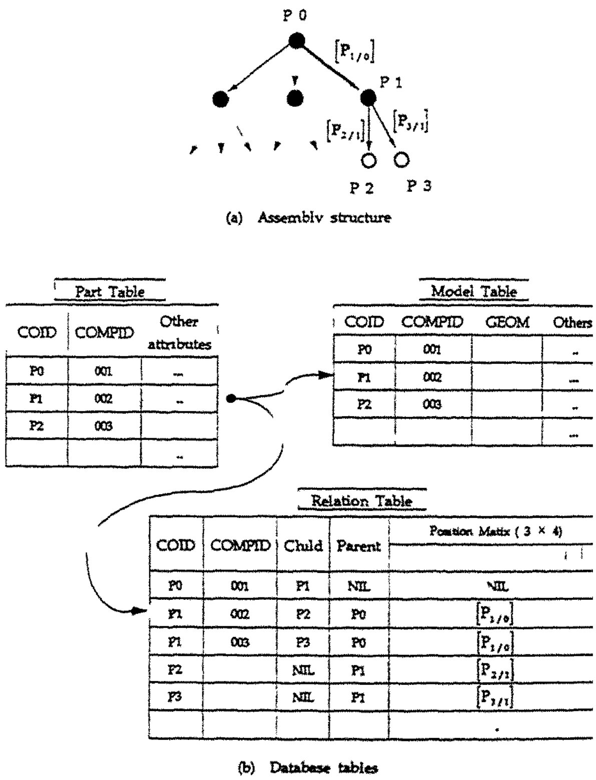
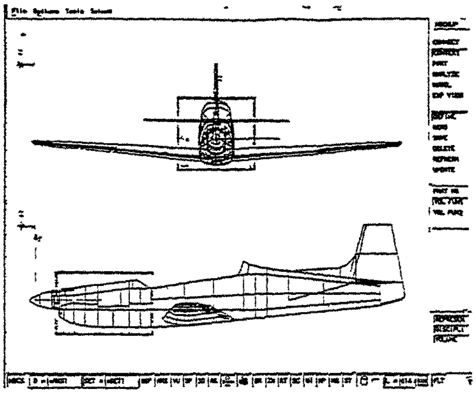


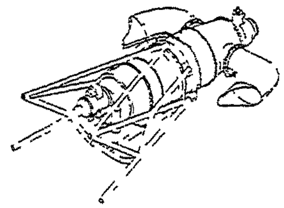
Fig. 5 Relation of database tables

4. RESULTS

To verify digital mock-up functions with the implemented system, about 150 parts of light aircraft are modeled with CATIA. When parts are stored in database, such attributes as part number, part name, discipline name, and parent-child relation are input by designer. When a designer designates a region in front and side view of the aircraft as shown in Fig. 6 (a), Fig. 6 (b) shows the digital mock -up with parts which are located in the designated region. When a designer selects the 'landing' item in the discipline menu as in Fig. 7 (a), Fig.7 (b) shows the digital mock-up of parts with the 'landing' function name. Also, when a designer selects some parts in a part list, he can create a digital mock-up with the selected parts. When we checked the interference of Fig. 7 (b) assembly, we found an interference among parts which are designed by field engineers as shown in Fig. 8



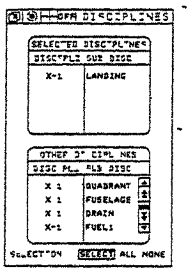
(a) Envelop defined in front and side view of the main assembly



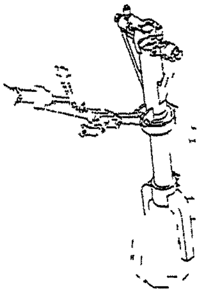
(b) Resulted Digital Mock Up with the envelop information

Fig. 6 Digital mock-up with envelope criterion

It took about 28 sec. to create the digital mock-up as in Fig. 6 (b), and 35 sec. for Fig. 7 (b) on UNIX workstation (RAM : 128 MB, CPU : Power PC 604). It took longer to obtain the result of Fig. 7 (b) than Fig, 6 (b), since the models in Fig. 7 (b) are more complicate to display in screen.



(a) Selecting the discipline of 'LANDING'



(b) Resulted Digital Mock-Up with the discipline

Fig. 7 Digital mock-up with discipline criterion

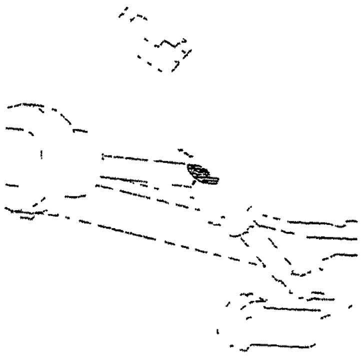


Fig. 8 Interference detected with the digital mock-up

5. CONCLUDING REMARKS

When such a large complex product as an automobile or an aircraft is developed, the most of cost loss is due to engineering changes which are from interference and fit-up problem. In this paper, the implementation of a digital mock-up system with assembly-centric modeling methodology is presented in order to reduce the engineering errors.

The advantages of the proposed modeling methodology are as follows. First, designers can consider the assembling sequence at the design stage, because the part positioning process is similar to real assembling. Second, designer can easily change the assembly position to resolve interference problem. That is, if he changes coordinates of assembly position, all of the part and sub-assembly position included are automatically changed, which makes the designer to adjust part position to resolve the interference problem in the way of top-down approach. Finally, the geometric data which are proposed to be isolated from positioning data can be used in other assembly without any manipulation of model data, because the geometric data is independent of the part position where it is located.

Even though some of commercial CAD system and dedicated programs provide digital mock-up functions, the detail of implementation is not released. In this paper, a digital mock-up system in CATIA environment is implemented based on the designers' requirements. We used ORACLE, the relational database management system for managing model and attribute data. But the more parts in database, the slower the system response is, since relational database management system basically searches all of the items in a database table. In future work, the efficiency can be improved by using object-oriented database management system.

6. REFERENCES

1. Hisup Park, Conru,A.B., Cutkosky,M.R., and Soo-Hong Lee, "An Agent-Based Approach to Concurrent Cable Harness Design," Artificial Intelligence for Engineering Design, Analysis and Manufacturing (AI EDAM), No. 8, pp. 45-61, April, 1994
2. Cutkosky,M.R., Engelmore,R.S., Fikes,R.E., Genesereth,M.R., Gruber,T.R., Mark,W.S., Tenenbaum,J. M., "PACT: an experiment in integrating concurrent engineering systems," Computer, Vol. 26, No. 1, pp. 28-37, Jan. 1993
3. Boothroyd, G., "Product design for manufacture and assembly," Computer-Aided Design, Vol. 26, No. 3, pp. 505-520, Jul. 1994
4. Sturges Jr, R. H. and Kilani, M. I., "Towards an integrated design for an assembly evaluation and reasoning system," Computer-Aided Design, Vol. 24, No. 2, pp. 67-79, Feb. 1992
5. Hsu, W., George Lee, C. S. and Su, S. F., "Feedback approach to design for assembly by evaluation of assembly plan," Computer-Aided Design, Vol. 25, No. 7, pp. 395-410, Jul. 1993
6. 'Pro/Fly-Through,' Parametric Technology Corporation, 1996.
7. 'dV/Reality,' Division Limited. 1996.
8. 'Interactive PreAssembly,' Immersive Design, 1996.
9. CATIA Solution, IBM, 1995.
10. Programmer's Guide to the ORACLE Call Interfaces, ORACLE, 1994.
11. CATIA data management library - Reference manual, Dassault Systemes. 1994.

INTEGRATED CAM DESIGN AND MANUFACTURING SYSTEM

L. Tang
Department of Production Technology
Massey University
New Zealand

ABSTRACT

This paper presents the development of a cam design and manufacturing system. The advantage of this cam system is that it is a window-based, modularised and integrated cam design and manufacturing system. Three modules are implemented in the package, which are cam design, cam system animation and manufacturing module. NC programs generated by this package can be directly used on $2\frac{1}{2}$D or 3D milling machine tools with no automatic tool compensation function. Testing has been made on this package, which gaves an encouraging result both on cam design and manufacturing aspects.

KEYWORDS

Cam Profile, Spline, Offset Curve

1. INTRODUCTION

There are many approaches in cam design and manufacturing. However, as cam speed, accuracy and dynamic requirements become more and more critical, especially for high speed, high precision and complex cams, classical cam design and manufacturing techniques tend to be inadequate. Cam designers and manufacturers are keen for new techniques which can reduce cam design and manufacturing cycle time, improve cam precision and provide better dynamic characteristics.

Cam design is a process which could be very involved and mathematically complex, as the output motion requirements become critical. To achieve a satisfactory cam system performance, factors such as maximum pressure angle, minimum radius of curvature, maximum allowable cam size, forces and torques occurred have to be considered. In recent years, researches on obtaining accurate cam profile have achieved significant progress using computer techniques. Different types of cam design software packages are available on the market which can deal with cam design and output the geometry of cam profile. However, from the manufacturing point of view, the surface or the profile of a cam is formed by the cuter path. In cam manufacturing process, the final cam profile is formed through a number of machining process which may involve rough, semi-finishing or finishing process depending on the requirements. More often, cutters with different size are used to produce a cam. In most of the cases, cutter radii are not the same size as the radius of the roller follower, therefore, cutter radius offset has to be considered. This is particularly important when output NC data to a machine tool, which has no automatic tool compensation function, to produce a cam through several cuts (rough cut, finishing cut and possibly semi-finishing cut) with different size cutters.

The cam system presented in this paper is developed using Microsoft C/C++ Version 7.0 and Windows MAKER Professional Version 5.0. IBM PC-486 was chosen as the development platform. This cam system integrated cam design and cam manufacturing under one platform. Commonly used cam laws are covered in this system and B-spline is used to obtain the cam profile. The output NC programs are in the standard format which is compatible with most of machine tools. In the case, the NC format is not compatible with the machine tool, the cam system can be easily customised to suit the particular machine tool.

2. GENERATE CAM PROFILE

For a given array of vertices $\mathbf{V}_0,\ldots,\mathbf{V}_n$, (n ≥ 3), a composite cubic B-spline curve can be defined by a matrix form of a parametric equation:

$$\mathbf{R}_i(t) = \mathbf{VMT} \tag{1}$$

$$\mathbf{R}_i(t) = \begin{pmatrix} x_i(t) \\ y_i(t) \\ z_i(t) \end{pmatrix}, \qquad \mathbf{M} = \begin{pmatrix} 1 & -3 & 3 & -1 \\ 4 & 0 & -6 & 3 \\ 1 & 3 & 3 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathbf{T} = \begin{pmatrix} t^0 \\ t^1 \\ t^2 \\ t^3 \end{pmatrix}$$

$$\mathbf{V} = (\mathbf{V}_{i-1,}\mathbf{V}_i, \mathbf{V}_{i+1}, \mathbf{V}_{i+2})$$

where $0 \le t \le 1$ and i= 1, ..., n-2

The vertices act as control points which define the shape of the spline. For a composite cubic B-spline, the i th segment $\mathbf{R}_i(t)$ is determined by four consecutive vertices. To apply this technique to construct a smooth closed curve, there are different approaches. This paper presents a technique which can produce a smooth closed curve and also confines the matrix form in a simple format.

To construct a B-spline curve, a set of control vertices need to be known and the vertices are normally physically off the curve. In cam design, the discrete points of the locus of the roller follower are known through the motion requirements. Cam profile is an offset curve of the locus. How to obtain the interpolation points between these discrete points using B-spline technique that follows the desired shape with the vertices are unknown is the case of cam design. Since the profile is described by a B-spline, it means that the B-spline must pass through the discrete points (the locus of the roller follower). In other words, the discrete points are the knots on the B-spline. In this case, the problem becomes knowing the knots on the B-spline to find the control vertices.

Suppose the knots are $\mathbf{P}_0,\mathbf{P}_1,\ldots\mathbf{P}_n$, the B-spline described by Eqn.(1) will satisfy:

$$\mathbf{R}_i(0) = \mathbf{P}_i, \qquad \mathbf{R}_i(1) = \mathbf{R}_{i+1}(0)$$

where i=0, ..., n-1

Applying these conditions to each of the knots on the B-spline, a system of simultaneous continuity equations is constructed and has the form:

$$\begin{bmatrix} & & & - & & & & \\ 1/6 & 2/3 & 1/6 & & & & & \\ & 1/6 & 2/3 & 1/6 & & & & \\ & & & & & & & \\ & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ & & & & 1/6 & 2/3 & 1/6 & \\ & & & - & & & & \end{bmatrix} \begin{bmatrix} \mathbf{V}_{-1} \\ \mathbf{V}_0 \\ \\ \\ \\ \mathbf{V}_n \\ \mathbf{V}_{n+1} \end{bmatrix} = \begin{bmatrix} - \\ \mathbf{P}_0 \\ \mathbf{P}_1 \\ \\ \\ \mathbf{P}_n \\ - \end{bmatrix} \qquad (2)$$

Obviously, to solve the equations uniquely, two more conditions are required. as cam profile is a closed curve, the end condition can be specified as:

$$\mathbf{R}_{n-1}(1) = \mathbf{R}_0(0)$$
$$\mathbf{R}_{n-1}(1) = \mathbf{R}_0(0)$$
$$\mathbf{R}''_{n-1}(1) = \mathbf{R}''_0(0)$$

Suppose the second derivative at $\mathbf{P}_0$ is known as **A**, then

$$\mathbf{R}''_0(0) = \mathbf{R}''_{n-1}(1) = \mathbf{A}$$

Therefore,

$$\mathbf{V}_0 = \mathbf{P}_0 + \mathbf{C}$$
$$\mathbf{V}_n = \mathbf{P}_0 + \mathbf{C}$$

where $\mathbf{C} = f(\mathbf{A})$

With these two additional equations, system (2) becomes a solvable system. The coefficients are diagonally dominant which can be solved as a tridiagonal system.

However, until now, **A** is still unknown. In order to find **A**, an additional B-spline is constructed by a set of points p_0, p_1, p_2, p_3

$$\mathbf{L}_i(t) = \mathbf{UMT}$$

where $\mathbf{p}_0 = \mathbf{P}_{n-1}, \mathbf{p}_1 = \mathbf{P}_0, \mathbf{p}_2 = \mathbf{P}_1, \mathbf{p}_3 = \mathbf{P}_2$

By applying $\mathbf{L}_i(0) = \mathbf{p}_0$, $\mathbf{L}''_i(0) = \mathbf{0}$ and $\mathbf{L}''_2(1) = \mathbf{0}$, a simple solvable tridiagonal system will be constructed which will give the answer of **U**. And finally **A** is obtained by

$$\mathbf{A} = \mathbf{L}''_0(1) \quad \text{or} \quad \mathbf{A} = \mathbf{L}''_1(0)$$

Once the control vertices **V** is known, interpolation between the knot points on the B-spline can be carried out. The cam profile is then generated by offset the locus of the follower.

3. CAM DESIGN AND MANUFACTURING PACKAGE

The presented cam system is still under development although it can produce sample cams. The system consists three modules. Each module can be invoked individually. However, the design procedure should be started from the design module in order to generate the data for animation and manufacturing modules. The cam design work is carried out by the design module. Once the design output is satisfied, the design result is saved in the database. Animation module can then be invoked to simulate the cam system. The manufacturing module is focused on material selection, cutting tools, machining parameters, machine tools, machining process and generating NC programs. A general structure of the system is given in Fig. 1.

Cam design module covers all the calculations for designing the cam. Cam type, motion type, the direction of cam rotation and the size of the cam are defined in this module. Besides, information of the cam system such as the size of the follower, system mass and stiffness, etc, are also input to the design module through dialogue boxes. Once the cam design calculation is completed, the display function will show the design result on the screen. Seven curves are displayed. They are displacement, velocity, acceleration, jerk, pressure angle, radius of curvature, force and torque. Based on these information, modifications or even redesign can be made before further work is carried out. Until the designer is satisfied with the design result, the data is then stored in the database to be used by other modules.

The animation module is for the dynamic check of the contact between the roller follower and the cam surface. One complete rotation of the cam is displayed on the screen for visual inspection to ensure the roller and the cam has a proper contact. Once the animation module is invoked, the system will read the cam data and the design result from the database. The corresponding calculations are then carried out. The simulation is made based on even angular increments. The simulation will show the cam rotates in the desired direction and the corresponding roller position at each angular position.

The manufacturing module of this cam package focuses on $2\frac{1}{2}$D and 3D CNC milling machines. Since these CNC milling machines with no automatic tool compensation is commonly found in industrial companies, the manufacturing module is designed in the way that can generate cutter path for a group of cutters with different size. When manufacturing module is invoked, a set of information needs to be provided. There are three machining process available, which are roughing, semi-finishing and finishing machining. According to the precision requirements, cam material, cutter and machine tools, user needs to chose the machining process first. Then input detailed information such as cutting speed, feed rate, allowance, etc.. The output of the manufacturing module is the NC program with standard format. For a CNC controller with special NC format, the system can be customised to generate the compatible NC code.

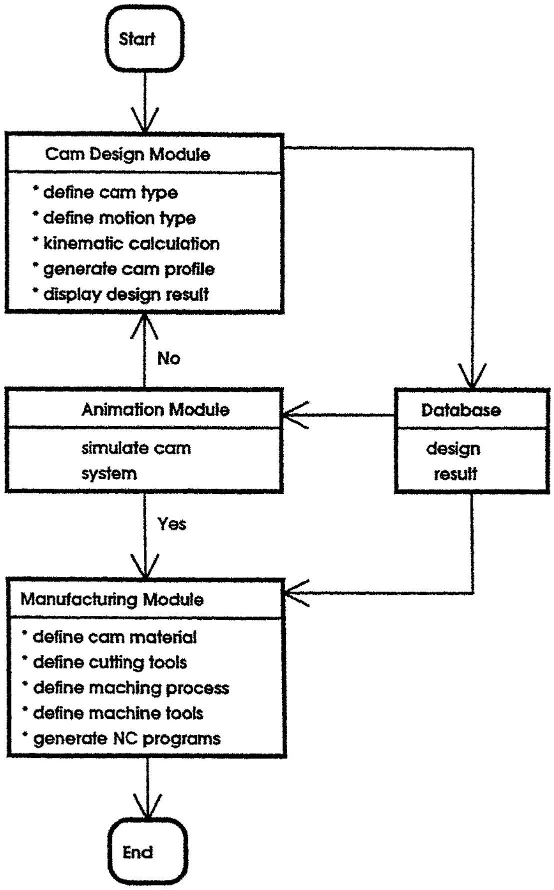


Fig. 1: Structure of the cam package.

4. DISCUSSION AND CONCLUSION

Test samples have been made using the presented system. The result is positive and encouraging. Since this cam system is an integrated cam system, it overcomes the problems caused by using one system for cam design, then process the cam data by another system to generate the adequate NC code for a particular controller. It is therefore a more practical system for cam design and manufacturing.

This cam system currently only covers disk and face cam. To enhance and extend the capabilities of this cam design and manufacturing system, further work such as including more cam types, enhancing the system compatibility with other CAD/CAM systems and CNC controllers are obviously needed. Work in this aspects is in process. The task is to produce a further improved cam design and manufacturing package.

5. REFERENCES

1. Liang Z., and Quinn, C. Jack, "Accurate Design of a Cam Profile on the CAD system", Journal of Manufacturing Systems, Vol. 10/No. 6, 1991, pp. 501-508.

2. Chen, F. Y., "Mechanics and Design of Cam Mechanisms", Pergamon Press, 1982.

3. Pham, B., "Offset approximation of uniform B-splines", Computer-Aided Design, Vol. 20/No., 8, Oct. 1988.

4. Klass, R., "An offset spline approximation for plane cubic splines", Computer-Aided Design, Vol. 15/No.5, Sep. 1983, pp297-299.

5. De Boor, C., "On Calculating with B-splines", Journal of Approximation Theory, Vol. 6, 1972, pp 50-62.

6. Farin, G., "Curves and Surfaces for Computer Aided Geometric Design", Academic Press, Inc., 1993.

7. Shikin, E. V. and Plis, A. I., "Handbook on Splines for the User", CRC Press, Inc., 1995.

TOLERANCE DESIGN AND ANALYSIS USING THE TOLERANCE REGIONS

Jyun-Ming Chen **Chi-Hui Chen**
Department of Computer Science and Engineering
Tatung Institute of Technology
Taipei, Taiwan

ABSTRACT

Tolerance design is an indispensable task for mechanical product development. Unnecessarily tight tolerancing results in low yield rates and high manufacturing costs. Yet, tolerancing should ensure the functional performance of the product to stay within satisfactory range, subject to the inevitable manufacturing deviation. If the specified tolerance values do not meet the functional objectives, a systematic approach should be devised to tighten the tolerance values in the most economical way.

In this paper, the computation of a new deterministic representation of tolerancing, the tolerance region, is proposed for analyzing the results of the dimensional tolerances. Graphical representation of the tolerancing result is the principal feature of this new method. Designers can visually verify the feasibility of the design by inspecting the tolerance regions. The computation of the tolerance region uses the Minkowski sum operators. Stacked tolerances are dealt with by the concept of "stack level", which associates a partial order for all geometric features in the configuration. A prototype has been implemented to demonstrate the capability of this new representation. The prototype system covers two dimensional parametric tolerancing with linear and circular geometry.

KEYWORDS

tolerance design and analysis, geometric modeling

1. INTRODUCTION

Due to finite manufacturing precision, tolerance management and analysis is indispensable for mechanical deign. Unnecessarily tight tolerances result in low yield rate and high manufacturing costs. Yet, tolerance values should be set to ensure the functional performance of the product to stay within the satisfactory range, subject to the inevitable manufacturing deviations. Thus, the analysis of tolerance values is important. Also, if the specified tolerance values do not meet the functional objective, a systematic approach should be devised to tighten the tolerance values in the most economical way.

In recent years, the use of advanced CAD systems in product design has become a standard practice in many production areas. However, the modeling and treatment of tolerance is not yet seriously considered in most design systems. Solid modeling based CAD systems deal with nominal geometry. Tolerances are often regarded as textual attributes, to be displayed on the drawings. In engineering practices, tolerance values are often assigned based on previous design experiences. The effect of such a tolerancing scheme is rarely evaluated in the design process. This negligence often causes quality problems during production.

This paper summarizes the tolerance research[2] at the CAD laboratory of Tatung Institute of Technology. The goal of this study is to analyze the effect of tolerance values of dimensional tolerancing of practical designs. The tolerances are analyzed using the new concepts of *tolerance regions*. A salient feature of this new approach is the graphical depiction of the "allowable regions." Sensitivy analysis and parametric adjustment of functional attributes can also be achieved by considering the interaction of the tolerance regions.

2. RELATED WORKS

In the literature, the approaches for analyzing tolerancing are often categorized into two kinds: statistical and deterministic approaches[1,6]. The former (see Ref. 7 for a recent review) takes into account the probabilistic nature of the associated manufacturing process. The latter analyzes the tolerance by seeking a conservative bound of the tolerance scheme. The works by Requicha[8], Turner[3,4,10], and Jayaraman[5] are examples.

Variational geometry is another important approach for tolerance analysis. The functional objective in question and the underlying geometric constraints derived from the configurations constitute a nonlinear programming (NLP) problem. Sapossnek[9] uses constraint dynamics to implement a variational tolerance analysis system with interactive user interface. The technique of vertex analysis [11] is frequently used in commercial systems. Taking advantage of the (relatively) small tolerance values, the aforementioned NLP problem can be simplified to an LP problem, whose extreme value of the objective can be found by visiting the vertices of the feasible space. The computation effort of vertex analysis grows exponentially with the system size for a naïve implementation.

3. THE COMPUTATION OF TOLERANCE REGIONS

The relevant terminologies in computing the (dimensional) tolerance regions are defined below.

- Geometric Feature (GF): any feature that has significance in the tolerance analysis.
- Characteristic Component (CC): the components that define the "characteristics" of a GF Non-ideal shape deviation is not considered in dimensional tolerancing. For instance, the CCs of a line segment are its two end points. The characteristic components of a rounding arc are its two component lines. The 0D and 1D CCs are also called characteristic points (CP) and characteristic lines (CL) respectively.
- Characteristic Point Region (CPR): the possible location of a CP according to the tolerance specification.
- Dimensional Region (DR): In a fully dimensioned two dimensional drawing, each CP has exactly two conditions specifying its two degrees of freedom. A *condition* can be a geometric constraint associated with the GF (e.g., a segment being vertical), or a dimension associated with the CP. Each dimension specifies a particular location (DR) where the CP can happen. The CPR is derived from the information of DR and constraints.
- Geometric Feature Region (GFR): the possible location of a GF.
- Minkowski sum: Given two point sets A and B, the Minkowski sum [12] of A+B is defined mathematically by the set: $S = \{a+b; a \in A, b \in B\}$. The term "sweep" is often used interchangeably with the Minkowski sum operator.

3.1. Bottom-up Tolerance Region Evaluation Algorithm

The algorithm for computing the tolerance regions requires a set of fully and explicit dimensions. The issue of implicit dimensioning schemes is discussed in Sec. 6.2. The geometric elements addressed in the prototype is confined to line segments and circular arcs. The algorithm is outlined below.

Step 1: Preprocess the dimensions: assign the "stack level" of each dimension on the drawing.

Step 2: If the configuration is not fully dimensioned, prompt the user on the incompletely dimensioned areas. Stop the program.

Step 3: In ascending order of the stack level, compute the dimension region of each dimension. The computation of DR should consider the tolerance region of its defining feature, as will be explained in Sec. 3.3.

Step 4: When the two DRs of a CP are both computed, compute the CPR as the intersection of the two DRs. When one DR of a CP and a constraint of GF is available, compute CPR by considering the interactions.

Step 5: When all CPs of a GF have been computed, compute the GFR of the GF. If a GFR is associated with some zero-toleranced constraint (e.g., verticality), the GFR should be labeled as such.

Step 6: Repeat steps 3-5 until all GFRs have been resolved.

3.2. Stack Level

The stack level defines the partial order of all dimension and features on an engineering drawing. Datum dimensions are of stack level zero. The stack level of each dimension is the stack level of the referred feature plus one. For a completely determined configuration, the stack level of each CP is the higher stack level of the two defining dimensions. The computation sequence is determined by the stack level. The higher stack level entity can only be determined after the lower-stack-level entities have been computed. Also, via stack level assignment, one is able to detect whether each CP is exactly specified by two conditions, the necessary and sufficient condition of a completely determined configuration.

3.3. Dimension Region Computation

Dimension schemes can be classified as baseline and non-baseline types. The former positions a geometric feature against a datum, a zero-toleranced geometric feature. The latter positions a geometry against another geometric feature which is likely to deviate from the nominal position. The dimension region computation differs in these two dimension schemes.

Fig. 1(a) shows the baseline dimensions of distance, length and angular dimension . The DRs of these baseline dimensions are defined mathematically as the infinite point set prescribed by a slab (from D-t to D+t), a concentric ring, and a wedge, respectively. If the datum is a toleranced feature, then the resulting regions are obtained by sweeping the baseline DRs against the GFR of the datum, as shown in Fig. 1(b). Fig. 2 illustrate the above procedure with a simple case of a rectangle.

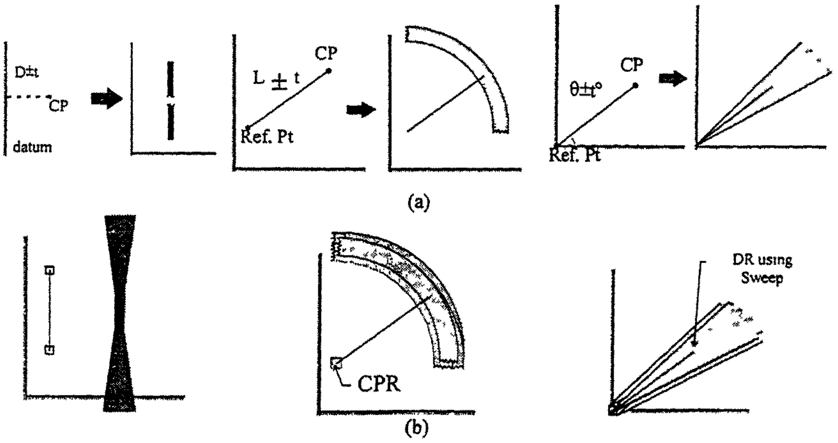


Fig. 1 DR Computation of Distance, Length, and Angular dimensions. (a) baseline; (b) nonbaseline.

3.4. Terminating Conditions

If some CPs are defined in a "coupled" fashion, repeating steps 3 through 5 will not be able to resolve all GFRs. In such cases, the solution of simultaneous equations is required to resolve a range of a particular dimension, hence decoupling the computation. The solution procedure proceeds as usual afterwards.

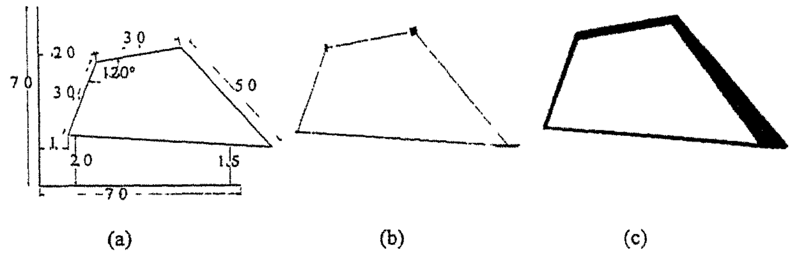


Fig. 2 Computing CPR & DR of a Rectangle: (a) configuration; (b) CPR; (c) GFR.

4. TOLERANCE DESIGN USING TOLERANCE REGIONS

As mentioned earlier, the salient feature of this approach is to present the possible outcome of the tolerance scheme in a graphical form. The designer can then use the result to evaluate the functional concern of the product design, such as clearance assessment of fitting, collision detection in assembly, and so on. With the help of a solid modeling environment, the functional concerns can be quantitatively assessed by evaluating the tolerance regions. For instance, for collision detection, the area of the two overlapped regions can be identified as a functional "signal".

If the signal in question fails to stay within satisfactory range (the definition of the range depends on the application), parametric values of the tolerances should be adjusted. To identify the contributing parameters of the signal is straightforward. In the bottom-up evaluation procedure, the effect of each tolerance parameter is recorded in the DR. This information is later propapaged to CPRs and GFRs. Attribute handling during the sweeping and Boolean operations of the computation is the key for making this data management possible.

Sensitivity analysis is often required for making efficient parametric adjustment. One needs to identify the dominating factors contributing to the functional objective. Mathematically, this corresponds to finding the partial derivatives of the signal with respect to the contributing parameters. In the tolerance region approach, sensitivity can be obtained by observing the objective variation in accordance to the parametric variation. The computation is quite time consuming, although performed in automatic fashion. Multivariate analysis (such as the Taguchi method) may be employed to cut down the computation time for more complicated configurations.

5. EXAMPLES

Fig. 3(a) illustrates a well-known benchmark for stacked tolerance analysis. The goal is to examine the height of the circle subject to variations of all eight dimensions. Fig. 3(b) and (c) show the CPR and GFR of the configuration. Table 1 lists the result of the sensitivity analysis.

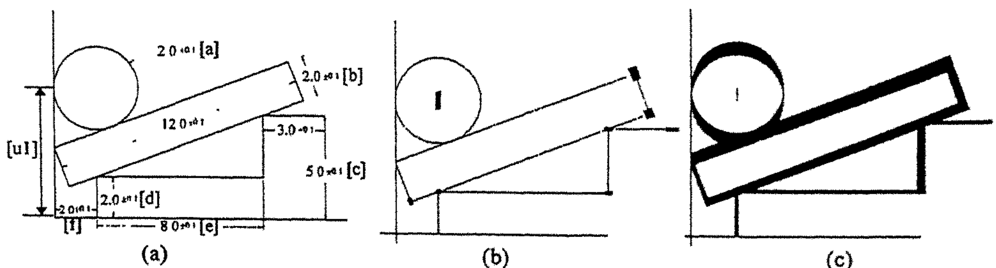


Fig. 3 Example of Stacked Tolerance Analysis: (a) configuration; (b) CPRs; (c) GFRs.

Dimension	Sensitivity to U_1
f	7.3221e-2
e	1.4347e-2
d	1.6078e-1
c	3.4608e-2
b	5.0048e-2
a	1.4400e-1

Table1. Sensitivity Analysis of Fig 3.

6. DISCUSSIONS

6.1. Tolerance Region Analysis vs. Vertex Analysis

Both tolerance region and vertex analysis concern only the effect of parametric variation, withou taking into account the shape imperfection. As mentioned in Sec. 2, vertex analysis often deals with the problem of discrete parameter analysis. Such results can be read off from the tolerance regions. Tolernace region approach presents a more information complete representation, which can be useful for othe purposes.

6.2. Implicit Constraints

In engineer practices, designers tend to use minimum set of dimensions to illustrate the configuration. For instance, the outer contour of Fig. 4 contains five circular arcs and two line segments and contains 30 degrees of freedom, but only 7 dimensions are marked. The rest of the degrees of freedom are resolved from engineering assumptions (parallel lines, tangency, continuity, etc.). Usually these "design intents" are recorded in the design interface as geometric elements are incrementally constructed. The resolution of nominal geometry is straightforward. But the conversion of tolerance values is non-trivial. Furthermore, the determination of the stack level of the underlying geometric features depends on the semantics of the design intent. Thus, the tolerance analysis of such kind of parts is a lot more difficult and reward further research.

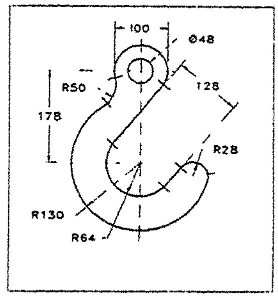


Fig. 4 Example of Implicit Dimension Scheme

6.3. Implementation Issues

The computation of tolerance regions involves many geometric computations. The current prototype utilizes a non-manifold geometric modeling kernel, Shapes (Xox Corp., U.S.A.). The computation of DR and GFR involves sweep and Boolean operations. To speed up the computation, geometric simplification procedures are due. As tolerance values are quite small compared to the nominal dimension, nonlinear (e.g., circular) components of the region boundary are simplified as linear segments. This simplifies the further sweeping operation.

7. CONCLUSIN AND FUTURE WORKS

This paper presents a new tolerance analysis method, the tolerance region. The tolerance regions present a graphical result to the designers for visual inspection. Systematic methods of parametric design in tolerance adjustment is also presented. The computation of this method involves many geometric computations.

Extending this methodology to other geometric domain is the next focus. We intend to expand on the domain of geometric tolerancing and three dimension tolerancing. Also intended is the functional analysis of tolerance design. Functional analysis is an important objective for any geometric related design. In this paper, we pointed out collision detection and dimensional verification as the functional objective. In fact, tolerance design also plays an important role for life cycle engineering concerns. For instance, heat related product should take into account the thermo-deformation. How such deformation should be modeled is important and still an open area. It is hoped that this region-based method can be applied to that type of analysis.

8. REFERENCES

1. Chase, K. W. and A. R. Parkinson (1991). "Survey of Research in the Application of Tolerance Analysis to the Design of Mechanical Assemblies." *Research in Engineering Design* 3(1): 23-37.

2. Chen, J. H. (1995). "Tolerance Design and Analysis Using the Tolerance Regions", M.S. Thesis, *Computer Science and Engineering*, Tatung Institute of Technology.

3. Guilford, J. and J. Turner (1993). "Representational Primitives for Geometric Tolerancing." *CAD* 25(9): 577-586.

4. Gupta, S. and J. U. Turner (1993). "Variational Solid Modeling for Tolerance Analysis." *IEEE CG&A* May: 64-74.

5. Jayaraman, R. (1989). "Geometric Tolerancing: Virtual Boundary Requirements." *IBM J. RES. DEVELOP* 33(2): 90-104.

6. Juster, N. P. (1992). "Modelling and Representation of Dimensions and Tolerances: a Survey." *CAD* 24(1): 3-17.

7. Nigam, S. D. and J. U. Turner (1995). "Review of Statistical Approaches to Tolerance Analysis." *Computer-Aided Design* 27(1): 6-15.

8. Requicha, A. A. G. and S. C. Chan (1986). "Representation of Geometric Features, Tolerances and Attributes in Solid Modelers Based on Constructive Geometry." *IEEE Journal of Robotics and Automation* RA-2(3): 156-166.

9. Sapossnek, M. (1993). "An Interactive Approach to Geometric Tolerance Design and Analysis." Ph. D. Thesis, *Electrical and Computer Engineering*, Carnegie Mellon University.

10. Sodhi, R. and J. U. Turner (1994). "Relative Positioning of Variational Part Models for Design Analysis." *Computer Aided Design* 26(5): 366-78.

11. Soin, R. S. a. R. S. (1988). *Tolerance Design of Electronic Circuits*, Addison-Wesley.

12. Srinivasan, V. (1993). "The Role of Sweeps in Tolerancing Semantics." *Manufacturing Review; Manuf. Rev. (USA)* 6(4): 275-281.

AN APPROACH TO AUTOMATE THE BENDING SEQUENCE OF PROGRESSIVE DIES

C.C. Chan
CAD/CAM Centre
The Hong Kong University of Science & Technology
Clear Water Bay, Hong Kong

C.H. Lai
Intelligent CAD/CAM Technology Limited
616, Tower B, Regent Centre, Kwai Chung, Hong Kong

ABSTRACT

This paper describes an approach to automate the design of progressive dies strip layout. The design of progressive dies involves designing sheet metal part, developing flat pattern, nesting of flat pattern, designing strip layout, designing staging plan, selecting punches, creating detail drawings, machining die components, assembling die, and testing. This approach will focus on determining the bending sequence of progressive die by resolving the unfold sequence of the part during flat pattern development. A solid modeller is used to build the part. With reference to the part geometry and topology, the system performs analysis based on knowledge and rules, and obtains the unfold sequence. The result can be used to aid the design of staging plan. Problems such as intermediate bends and interference checking are also addressed. The result can also be displayed as a series of solid models to simulate the bending operation, when the strip moves from one station to the other.

KEYWORDS

Progressive dies, CAD/CAM, solid modelling

1. INTRODUCTION

The design of progressive dies for stamping is a highly complex planning process. In the manual design process, tedious and enormous geometrical calculations are required, and generally relies on the skill and experience of the die designer.[1] The designer starts with the desired sheet metal part, unfolds it and then designs the strip layout. Once a feasible strip layout is obtained, staging design which performs piercing, blanking and forming can be determined accordingly. The next step involves the decomposition of the design into standard and non-standard components. Detail part drawings are then generated for each component. Finally, the non-standard components are machined and other die components are selected to form the die assembly. The whole die assembly is mounted on a suitable press for trial run. These processes have to be repeated several times until satisfactory production is achieved.[2]

2. CAD/CAM TECHNIQUES FOR PROGRESSIVE DIE DESIGN

Over the years, CAD/CAM technology was introduced to assist designers and engineers to automate modelling, drafting and manufacturing activities. A CAD/CAM approach to progressive die design is shown in Fig. 1. The starting point of die design is the 3D model of a product to be manufactured. Many commercial 3D CAD systems are available for product designers to create 3D model of the product. In many cases, the CAD system being used will also allow the user to unfold the product into a flat pattern (stage 1). The flattened 3D model is then passed to some nesting algorithm for better material utilisation.[3] Efficient algorithms for nesting have been developed and some of these are bundled in commercial CAD/CAM system (stage 2). The output, which is typically a 2D drawing of the strip layout is then used to aid the designer for staging design. Staging design is a complex process which involves a lot of decision making and user interaction (stage 3). Currently, this process is performed interactively by experienced designers using generic CAD packages. Since output from the nesting algorithm is a 2D flat pattern, the resulting staging plan is also a 2D drawing. Based on this plan, standard punches and non-standard punches are chosen to perform various operations such as piercing and bending. This highly iterative task is normally accomplished by skilled designers. Once the punch selection task is completed, the die layout can be determined. All other die components can then be derived from the layout to form a complete die-set. Then, part drawings and NC toolpath are generated for component manufacturing (stage 4 & 5). Specific commercial CAD/CAM packages designed for the sheet metal industry are available in the market to automate these processes. However, human expertise is needed to arrive at the final design, particularly the staging design.

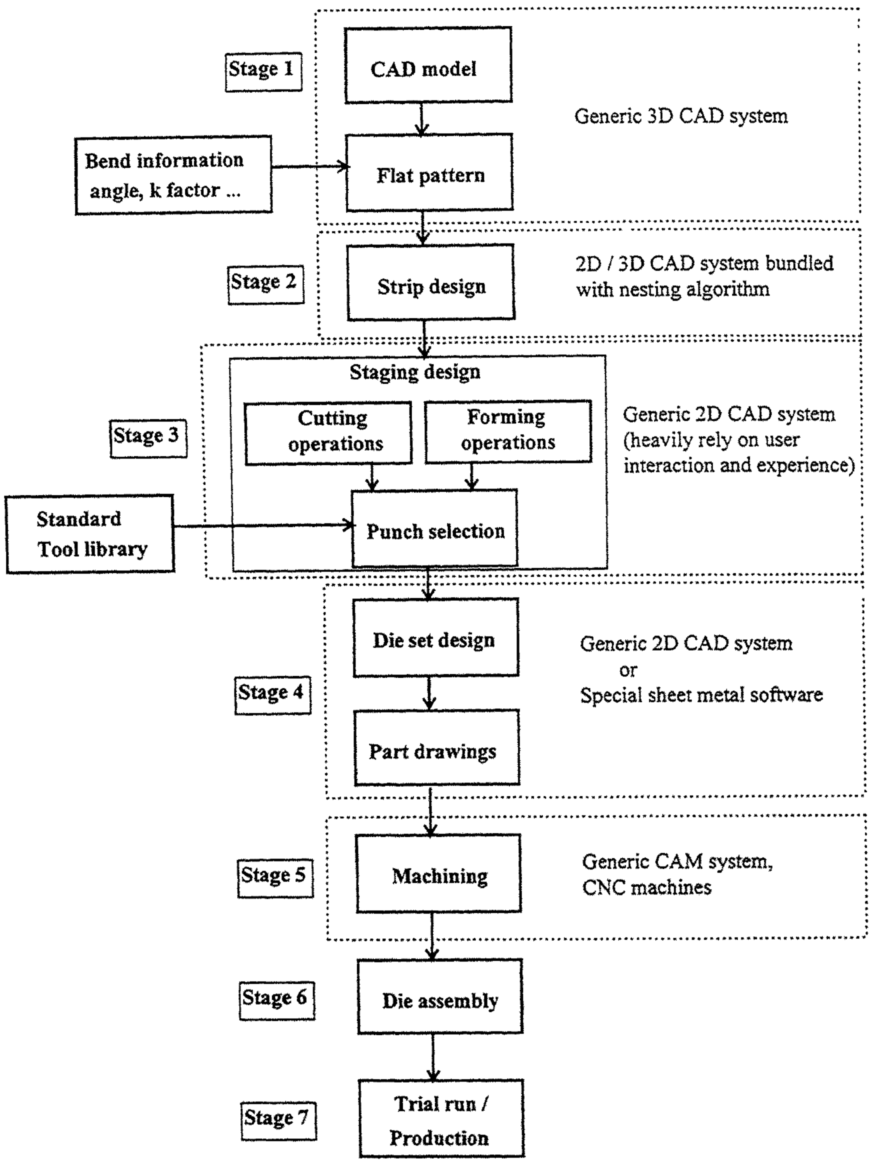


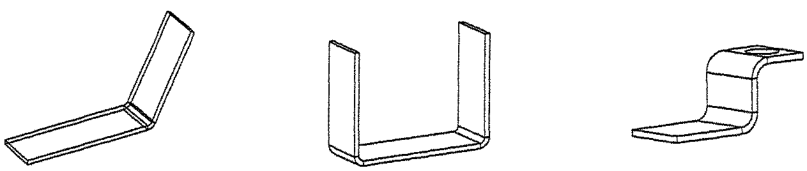
Fig. 1 CAD/CAM techniques in progressive die design

3. STAGING DESIGN

The staging process is the planning of the die in which the die operations are sequenced in a progressive manner. Cutting and/or forming operations will be performed at each stage when the strip pass through the die. Finally, the required part is produced at the last stage. Nee and Foong [4] present an approach to develop an automatic staging and punch shape design algorithm based largely on rules, shape decomposition and pattern-matching technique. In their system, internal and external profile are decomposed into elemental shapes and being checked against a standard punch library. This system automates the selection of punches. Cheok *et al* [5] enhanced the system by integrating the geometrical information from CAD drawing with knowledge-based development tool to automate the punch modelling process. Staging of the punching operations is addressed by Cheok *et al* [6] based on heuristics rules. In their system, topological information was extracted from the part and then feed into a knowledge-based system (KBS) to generate the staging plan of the die. This system emphasise the decomposition of cutting features and punch selection which perform the cutting operation. Staging plan is then derived based on the selected punch and heuristics rules. Not much was mentioned on the staging design of punches which performs the bending operation. On the other hand, selection of punches should not only depend on the geometry of the flat pattern but also the sequence of punching operations such as bending, lancing and piercing. In the next section, we shall discuss the bending operation and explain why it is so important in designing the staging plan.

4. BENDING OPERATION

Most of the sheet metal part produced from a progressive die will contain one or more bends. These bends are formed by the bending operation. When a 2D strip is subjected to a bending operation, it becomes a 3D model. Bends can be classified into three major classes: 1) L-shape, 2) U-shape, and 3) Z-shapes as shown in Fig. 2. The U-shape and Z-shape are actually combinations of different L-shapes. Similar to cutting, forming operations have to be sequenced in a proper manner. In fact, the forming sequence determines the cutting sequence and hence determines the shape of the punches. A bend can be considered as either internal or external. For internal bends, both cutting and bending operations will be performed by the same punch and is known as lancing. For external bends, open edges should be cut prior to the bending operation. Fig. 3 illustrate these two bends and their corresponding cutting operations.



(a) (b) (c)

Fig. 2 Different kind of bends a) L-shape b) U-shape c) Z-shape

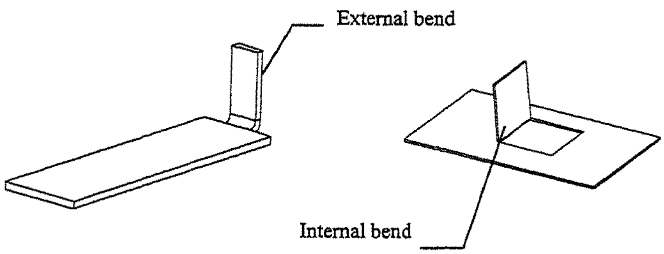


Fig. 3 External and internal bends

In general, we can stage all cutting operations before the bending operations. However, there are cases where the piercing/blanking operations should be performed after the bending operation. As shown in Fig. 4, if piercing is done first, the distance between the two holes cannot be guaranteed. To achieve better accuracy, hole #2 should be punched after the bend is formed. Therefore, both forming and cutting operations should be considered during the staging design phase. However, mixing these two operations together complicated the staging design. For example, folded flanges may interfere with other part of the die while moving from one stage to the other.

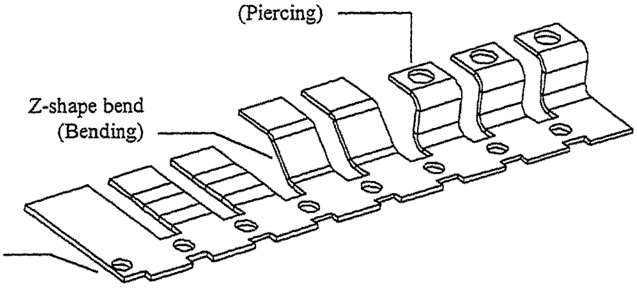


Fig. 4 Bending before piercing for better accuracy between hole #1 and hole #2

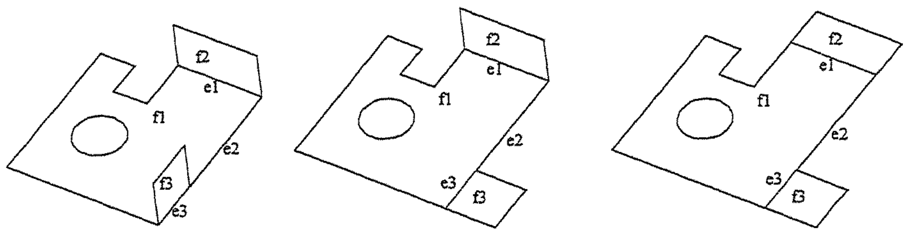
5. THE BENDING SEQUENCE

With a 2D flat pattern or strip layout, it is difficult to imagine how the pattern will fold to form the final part. Partially folded strip move along the die may interfere with die plates or other die components. Generic 2D CAD packages do not provide tools for interference check. Although 3D CAD packages can be used to detect interference, it take longer time to create 3D solid models. In either case, the designer has to iterate several times before he can come to the final design.

The folding or bending sequence can be determined by reversing the unfolding sequence of the sheet metal part if we unfold the part in a proper manner. In fact, there are rules which we can follow to govern the fold/unfold process. Many CAD/CAM system can unfold a 3D sheet metal part, however, only the final flat pattern is presented, the intermediate states or the sequence are not shown. In this paper, we propose an approach to determine the bending sequence of the progressive die by determining the unfolding sequence. The bending sequence is then used to generate a series of 3D models to simulate the bending operations performed on the strip when it moves along the progressive die. These models provide sufficient information on the geometric shape that is required to cut the strip prior to these bending operations. Cutting operations can then be inserted into the bending sequence to form the staging layout. Since all the models, including punches, die plates, strip and other die components are created as 3D solids, interference detection algorithm can be applied at any time to avoid collision between various die components. This approach provides a new method to automate the staging design process with least human intervention.

5.1 B-rep model as the link

Traditionally, sheet metal parts are represented as either 3D wireframe or surface models and almost all the down stream die design processes are completed in 2D. With the advancement of solid modelling techniques, most of the sheet metal part can be created using solid modeller. In this approach, a boundary representation (B-rep) solid modeller is used to represent the part. A B-rep solid modeller stores not only the geometrical information of the solid, but also the topological information. Topology refers to the spatial relationships between various geometric entities in the model. Consider a sheet metal part and its intermediate unfold states as shown in Fig. 5. The geometry at each state are different, however, they shared the same topology: f1 is connected to f2 along e1 while f3 is connected to f1 along e3. In other words, the topology remains the same during the unfold process. This properties will be used to construct a tree diagram for determining the unfold sequence as discuss in the next section.



(a) (b) (c)

Fig. 5 Different shapes share the same topology

5.2 Determine the unfolding sequence

Based on the geometrical and topological information, the following 7 steps are used to determine the unfolding sequence.

1. Read in the 3D solid model
2. Extract all the bending information such as angle, radius, and k-factor
3. Extract the inner and outer shell
4. Determine the datum face manually or automatically. In both cases, the following selection and prioritisation rules apply:
   a) Area - higher for larger face
   b) Number of bends within the same face - higher for higher number
   c) Centre of bending - higher for shorter distance from centroid
   d) Visibility from punch direction - higher for higher visibility
5. Search for faces connected to the datum face

   A tree structure of all the faces is constructed with the datum face as the root node. Faces connected to the datum face are represented as child nodes. This search goes on until all the faces have been assigned to the structure. A parent node may have more than one child while there is one and only one parent for each child. (Fig. 6)

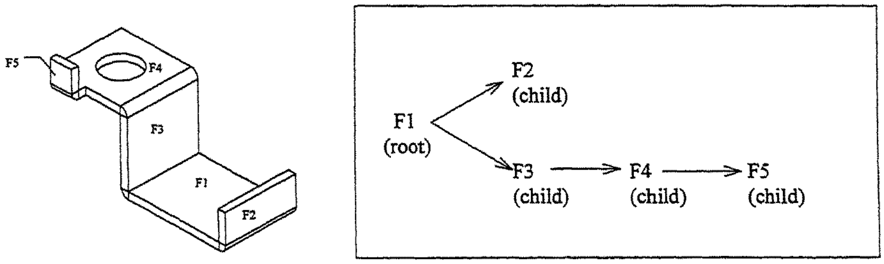


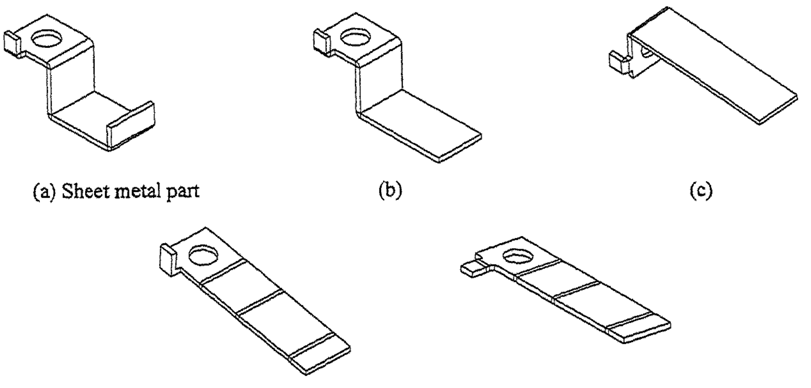
Fig. 6 Tree structure of the sheet metal part

6. Determine the unfolding sequence

   The farthest node represent the last bend, the next farthest node represent the second last bend. This search goes on until it reach the root node. If there are several nodes at the same level of the tree, special rules are used to determine the sequence. Some of the rules are:

a) which has two or more bends combined together,
b) which contains a longer sub-structure,
c) which is closer to the centroid, ... and so

7. This unfolding sequence is then used to create a series of 3D solid models. Each of these models represents an intermediate state of the strip between the final part and the flat pattern. (Fig. 7)



(a) Sheet metal part (b) (c)

(d) (e) Flat pattern

Fig. 7 A sheet metal part and its corresponding unfolding sequence

6. PROOF-OF-CONCEPT SYSTEM

A proof-of-concept system which run on a Win95 platform is developed to determine the unfolding sequence of a sheet metal part. The system is built on ACIS, using Microsoft Visual C++ 4.0. ACIS is a B-rep based geometric modelling toolkit designed for use as a geometry engine. AutoCAD Mechanical Desktop is used to create the solid model of the sheet metal part. The design is then saved in the ACIS native SAT format.[7] With reference to the geometrical and topological information stored with the solid model, the system performs analysis based on knowledge and rules to obtaining the unfolding sequence. The resulting solid models of each stage is then written to files in SAT format. These solid models can then be retrieved in AutoCAD for down stream die set design.

7. REFERENCES

1. Paquin, J.R., Crowley RE. Die Design Fundamentals, Industrial Press Inc., 1986.
2. Choi, S.H., Wong, K.W., "A CAD/CAM package for sheet metal blanking dies", International Conference on Manufacturing Automation, 1992, pp.674-679.
3. Nee A.Y.C., "Computer-aided layout of metal stamping blanks", Proc. of the Inst. of Mech. Engrs., Vol.198B, No.10, 1984, pp.187-194
4. Nee, A.Y.C., Foong, K.Y., "Some considerations in the design and automatic staging of progressive dies", Journal of Materials Processing Technology, Vol. 29, 1992, pp. 147-158.
5. Cheok B.T., Foong K.Y., Nee A.Y.C., Teng C.H., "A knowledge-based approach to automate the punch modelling process for progressive die design", J. Inst. Eng. Singapore, Vol. 33, No. 4, 1993, pp. 65-72.
6. Cheok B.T., Foong K.Y., Nee A.Y.C., "An intelligent planning aid for the design of progressive dies", Proc. of the Inst. of Mech. Engrs., Vol.210B, No.1, 1996, pp.25-35.
7. ACIS save format manual, Spatial Technology Inc., Version 2.0, 1996.

SIMULATION APPROACH IN PLANNING AND DESIGN OF MANUFACTURING CELLS - A CASE STUDY

F. T. S. Chan[1] and R. W. L. Ip[2]

[1] Department of Industrial and Manufacturing Systems Engineering, The University of Hong Kong, Hong Kong
[2] Department of Manufacturing Engineering, City University of Hong Kong, Hong Kong

ABSTRACT

This paper reports a case study performed at a company manufacturing washing machines. The purpose of this study is to investigate possible solutions for the reduction of costs in producing those components for the washing machines. This has been achieved through the optimisation of various facets of manufacture. It has been observed that the current system has no set methodology in either the layout or the scheduling system. The redesigned or new system is a cellular manufacturing system, based on part type, with push (MRP) scheduling, multi-skilled, multi-tasked and multi-celled operators.

An extensive comparison between the existing and the proposed new systems was performed with the assistance of computer simulation. The tangible comparison factors, such as number of finished products produced in a specific time, number of required operators for a specific production rate, etc. have showed that the new system is superior to the current system. Furthermore, the new system has also provided many positive intangible features. It has been recommended that the implementation of the new system should occur as the comparative data suggests. The study showed that the implementation of cellular manufacturing systems are viable solutions to problems concerning manufacturing scheduling, layout and efficiency.

KEYWORDS

Simulation, Scheduling Systems, Cellular Manufacture

1. INTRODUCTION

The washing machine manufacturer studied, produces washing machines for Australia and overseas. The factory is separated into a variety of production shops, which include the machine shop, assembly line and die casting shop. The machine shop has a variety of machines, such as grinders, induction hardeners and lathes; some of them are semi-automated. This machine shop is used to produce a variety of parts for washing machines, some of which are, pulleys, gears and shafts. The shop is relatively unchanged since its establishment, now the company is looking for improvement in the whole factory. Improvements have been sought and simulation has been used to identify problems in the current manufacturing system and verify the new system developed. The steps(1) in the redesign of a factory floor are shown in figure 1.

A range of manufacturing strategies, both push and pull systems, have been investigated. The Kanban production scheduling system(2) has been analysed to determine why the implementation of the single Kanban system failed in the workcentre. Other manufacturing strategies(3) are also studied, such as MRP-II, Cellular Manufacture(CM) theory and Optimal Production Theory(OPT), as the redesigned system utilised the optimisation characteristics of these manufacturing strategies.

The purpose of this study is to investigate possible solutions for the reduction of costs in producing components for the washing machines. To fulfil this objective an emphasis was placed on improving the accuracy and efficiency of the workcentre activities, increase human resource utilisation, reduce the number of operators, remove bottlenecks and optimise part flows.

2. OVERVIEW OF CM/MRP/OPT

2.1 Cell Manufacturing(CM) Theory

The basic methodology of CM theory is for machines to be grouped so that Work-In-Progress (WIP) and human resources are optimised. The grouping of machines can be performed generally in two methods.

1. Machine Requirement Cellular Manufacture (MRCM). This method of grouping machines is dependent on their requirements for production, e.g. loading, unloading, setup, machining, etc. This method is specifically suited to simulation, where quick evaluation of alternatives can be obtained and thereby making optimisation easy and tangible.

2. Part Type Cellular Manufacture (PTCM). Machines are grouped dependent on the manufactured parts features, e.g. Size, workflow, operation requirements, physical geometry, etc. The grouping can be performed with a variety of criteria and it is up to the designer to determine the importance of the evaluation criteria for the basis of the cells. Commonly there will be at least one machine which will perform a common process on a variety of parts. This theory suggests that this machine should be isolated and create its own cell, as the machine will be a possible bottleneck.

2.1.1 Multi-celling

An extension to the CM theory is the optimisation of human resources. Multi-celling is the optimisation of labour by assigning more than one cell per operator, opposed to more than one operator per cell. It is based on the fact that some of the cells may have longer lead times than others. The operators in the shorter lead time cells, when finished the production requirement(s) will assist in another cell which has longer lead time.

2.2 Material Requirements Planning (MRP)

MRP-II, closed loop MRP, leaves the scheduling decisions within each workcentre or workshop, reliant on the workshop leader (team leader) or supervisor. This scheduling system primarily provides the supervisor with information on two facets of manufacture.

1. Quantity required of each type of part in a specified time period.
2. Schedules the parts or raw material to arrive at the workcentre with enough time so that part production will be able to meet the demand. Then it is the supervisor's responsibility to produce the parts.

MRP is suited to production in which the part mix does not vary significantly. It is also suited to workshops which perform intermediary processes to final products or assembly, especially when the demand for the intermediary parts far exceeds the capacity of the intermediary workcentre. With such a high demand, a pull system would not be beneficial due to the size requirement of buffers.

2.3 Optimal Production Theory(OPT)

OPT was developed by Goldratt and Cox(4) in a novel named *"The Goal"*. It is a system that schedules production off-line and takes into account of utilisation and resource dependencies. The benefits used from this theory are the optimisation of part flow and the concept of determining and the removal of bottlenecks.

- A bottleneck is any resource whose capacity is equal or less than the demand placed upon it. The effect of a bottleneck may be reduced by increasing the capacity of the machine or increase part type workflow through the machine.

- To optimise WIP and increase workflow through bottlenecks, the theory suggests to continually reduce batch size until maximum production is obtained or a batch size of one is reached. If machines have high setup times, this theory would not work, thereby limiting the methodology to a production floor with limited flexibility and machines that require little or no setup.

3. CASE STUDY

3.1 Current Manufacturing System

The current manufacturing system can be best described as MRP based quantity scheduling, with random and arbitrary batch sizes, operations performed on a First Come First Serve (FCFS) basis in a job shop layout. The random and arbitrary batch sizes have occurred from the attempted implementation of a single Kanban card scheduling system. Unfortunately, the principles of the Kanban system have generally been ignored and now the purpose of the Kanbans (cards) is only for the identification of parts in a bin. The failure of this system arose from the pressure placed on the machine shop by the assembly line as, the initial implementation caused scheduling problems which eventually causing part shortages. The natural evolutional solution was to push more parts into the system and thus create more WIP. Also due to the lack of evaluation and addressing the problems, consequently the Kanban system demise. A schematic diagram from the simulation software is shown in figure 2 and a legend for the diagram is depicted in figure 3.

The manufacturer produces three basic sizes of washing machines, *large, medium* and *small,* where the size is based on the tub/bowl size inside the washing machine. The large and medium machines are further broken down into two other categories which based on bowl type, i.e. *stainless steel* and *plastic.*

To create the model initially, the current MRP data and the production rate were ascertained. The actual production rate of the various washing machines is presented in table 1. The allocation of human resources was the responsibility of the team leader of the workcentre, keeping with the MRP system methodology. Generally the allocation was one operator per machine or machine group. The number of operators was determined by the loading (demand for parts) of the machine shop which in turn was determined by the MRP data.

It has been observed from the simulation output that many operators had excessive amount of idle time and this was validated by observations from the machine shop floor. The high amounts of idle time can be accredited to the fact that the majority of the machines have automated processing capabilities and the only operation for the operator was to load and unload the machines.

The validation of the model was performed with two features of manufacture.

- Finished Goods. The comparison between the amount of parts produced in the time allocated, with allowance of ±5% variation. The variation allowance was due to RM allocation assumption and the stochastic feature in the system.

- Machine Utilisation. Charts of the machine utilisation were given to the workshop leader to verify for the particular production rate. This was performed so that some real comparative input could be incorporated into the model.

3.2 New Manufacturing System

The modified system is a cellular manufacturing system with MRP scheduling and multi-skilled, multi-tasked and multi-celled operators. The modified layout is a CM based on part type, hence the methodology of PTCM has been used. PTCM is utilised as the basis for the design methodology as proposed the most favourable results from the alternative designs created. PTCM is feasible because of the specific nature of the parts and manufacturing requirements.

- Parts which are for the small washing machine were easily isolated and could have specific dedicated machines.
- Parts which are for the large and medium washing machines were easily grouped due to their similar geometry, processes plans and machine requirements.
- Machines could be dedicated to particular part groups.
- The immediate advantages obtained from the CM theory could be achieved, specifically with the physiological effect on the employees.
- The neatness of the layout and easy positional determination of particular component groups.

Figure 4 shows the new design of the layout. The cells with dashed lines separating are the cells linked by multi-celled operators. The number in the diagram corresponding to the allocated operator number in the simulation.

The redesign is based on the same parts produced and the same processing times as in the current system. However, management suggested that the new system should be capable to achieve a higher production rate. The decided production rate for the medium and large washing machines is 851 machines per day(table 2). The small washing machine is decided to have the same production rate as the current system, 14 washing machines per day.

To optimise the labour allocation, the operators have to be multi-skilled. The determination of the number of operators is an iterative process of simulating a variety of alternatives. Initially each cell was allocated a single operator per cell, eg. Agitator Shaft Cell had one operator servicing the whole cell of seven machines. This scenario was simulated and it was discovered that some operators had experienced a very high idle time, but others could not service all the machines in their cells. One alternative design has been considered such that operators with high idle time were linked with another cell when they became idle. Several different alternatives of multi-celled operators were allocated. The alternative that gave the most favourable production figures was the multi-celling system.

The production system utilised is the MRP, FCFS system. This system was used for a variety of reasons as the manufacturing scenario suited MRP scheduling better than any other scheduling theory.

- The demand rate for the parts from the assembly line far exceeded the capacity/supply capabilities of the machine shop.
- The part mix is limited. Many parts produced are similar in nature and quantity.
- The machine shop is an intermediary workcentre to the assembly line.
- Operator allocation has been determined by PTCM methodology.
- Kanban system has already failed.
- FCFS process scheduling at machines, due to the inability of MRP to accurately schedule and control flow of WIP.

When creating the production system, it was found that part workflows had to be optimised to maintain cohesion with the PTCM theory. This was established by dedicating machines to processes and this action removed some machines from the manufacturing system, as they created excess capacity and would inhibit production efficiency.

3.3 Comparison Of Models/Results

The following parameters were entered into the simulation for comparison.

- 851 large and medium, plus 14 small washing machines made per day.
- The simulation was run for the equivalent length of one month, 20 days. The results were then divided by 20 to obtain an average production rate per day.
- One day consists of three shifts, 1440 minutes with 120 minutes breaks in total.

3.4 Number Of Parts Produced

Figures 5 to 7 show the comparison of the quantity of parts produced for each size of washing machine. It can be seen that the current system was substantially inferior than the new system. The delays in production for the current system can be accredited to production being performed in batches and the high quantity in each of the containers. The new system did not exhibit these problems due to the optimisation of partflow and the dedication of machines.

3.5 Number Of Operators

Reducing the number of human resources is the most substantial cost saver. The new system reduced the number of operators by 4, for the simulated production rate of 851 plus mini.

- Current System = 34 operators
- New System = 30 operators

Under the new system the operators have been allocated differently as compared to the current system, giving the new system higher shift penalty rates. It was approximately determined that the new system would save $1,460 per week, which amounts to over $70,000 per year.

3.6 Analytic Hierarchy Process(AHP) Model

The AHP technique was developed in the 1970's by Thomas L. Saaty at the Wharton School of the University of Pennsylvania. AHP is a multi-criteria decision method that allows the consideration of financial and non-financial factors of selected options or decisions. The features considered for the comparison of the current and new systems are in table 3.

Many of the features considered in this model are difficult to measure and the figures used are subjective, the solution of this AHP model gives an indication that the new system has some superior qualities compared to the current system.

3.7 Intangible Aspects Of The New Design

- Promotes "Poke Yoke".
- Increases employee, responsibility, morale, management of production, multi-skilling and friendship.
- Highlights setup problems.
- Easier to determine processes and machines which cause quality problems.
- Unreceptive attitudes from the employees of the machine shop due to the perception of; *"if the current system works why change it?"*

- New system can cause too much competitiveness between operators in cells. Therefore continual changing of operators between cells must occur. However if the operators are rotated between cells too frequently, this will remove some of the positive employee aspects.

4. RECOMMENDATIONS

- The gearbox assembly area should be moved into the machine shop. This would allow increased control over the production volumes of the gear box and more importantly increased control over the parts for the gearbox that the machine shop manufactures.
- Ideally the process based workcentres should be removed and changed to part type production areas, eg. remove the machine, die cast shops and gearbox assembly area and create a gearbox workcentre.
- Remove the small washing machine from production, due to its extremely low requirement and high cost of production. This will increase capacity of the system, especially the grinders. The extra capacity will allow for a comprehensive implementation of the CM theory which will decrease the lead time of parts and possibly further reduce the number of operators.
- Ideally replace the gear box with a suitable simple alternative as the direct drive system. The gear box is the most expensive component of a washing machine that the machine provides. The gear box utilised 2/3 of the machinery in the machine shop and is approximately 1/2 of the total production of parts for the large and medium washing machines.
- The Kanban production methodology would probably not work to the theory's potential in the machine shop due to the production scenario that the machine shop exhibits. This is because Kanban systems depend on the pre selling of the item or close demand and supply rates between dependant workcentres. Washing machines are an *'off the shelf'* product and the only way the Kanban system would work if retailers could order smaller quantities and at an increased frequency. Therefore the current running of the MRP system provides an adequate scheduling system especially with the PTCM methodology developed.
- A preventative maintenance system should be implemented. The manufacturing system developed is a high productivity system heavily dependant on preceding machines. If a breakdown occurs, this could have implications through the whole production floor for the washing machines.
- The new system should be implemented, due to savings and advantages it has primarily shown. If the small washing machine is removed from production, there should be no hesitation to redesign a more idealistic CM system with the increased capacity and implement this system immediately.

5. CONCLUSIONS

The requirement for manufacturing companies to change their traditional production systems and methods is becoming more evident in the competitive nature of the market. With more imports being purchased and tariffs being reduced, the requirement for a cheaper and high quality product is required to obtain enough market share for a company to survive. With the implementation of the new system, there is a reduction of operators. This means a decrease in unit production cost and therefore increase the profit to the company. If the washing machines costs are reduced, this will increase the market share as cost is the biggest driver for consumer purchases.

This study utilised simulation to its full potential. Simulation provided a model which closely represented the current manufacturing system and allowed the quick, inexpensive, evaluation of alternative systems. It also provided a means of comparison so that the best alternative redesigned system could be optimised and compared to the current system.

The implementation of the new system should proceed immediately. The redesign also established many positive intangible features towards manufacture. The study showed that the implementation of Cellular manufacturing systems, Optimal Production Theory and Multi-celling of human resources are viable solutions to the problems that machine shop had occurred.

6. REFERENCES

1. Baudin M, *"Manufacturing Systems Analysis - With Application to Production Scheduling"*, (1990) Prentice-Hall Inc.
2. Singh, N., Kwok, H.S. and Meloche, D., The development of a Kanban system: A case study, Int. J. of Operational and Production Management, Vol. 10, No. 7, 1990, pp. 28-36.
3. Flapper, S.D.P., Miltenbury, G. J. and Wijngaard, J., Embedding JIT into MRP, Int. J. od Production Research, Vol. 29, No.2, 1991, pp. 329-341.
4. Goldratt E M and Cox J, *"The Goal"*, Second Edition (1993), Gower Publishing Company.

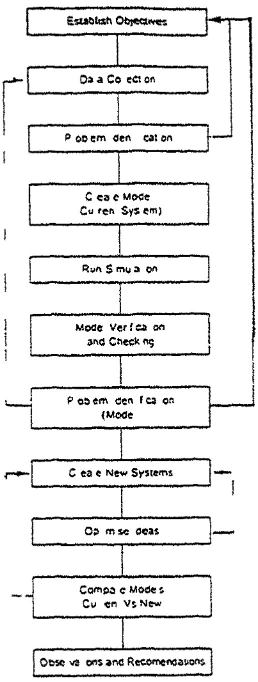


Figure 1 - Flow Diagram for the Redes[illegible]

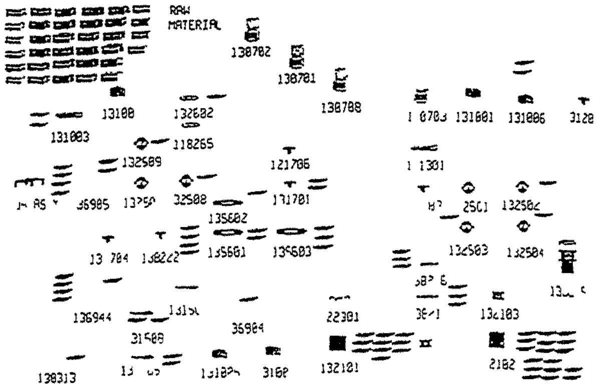


Figure 2 - Current Workshop Layout

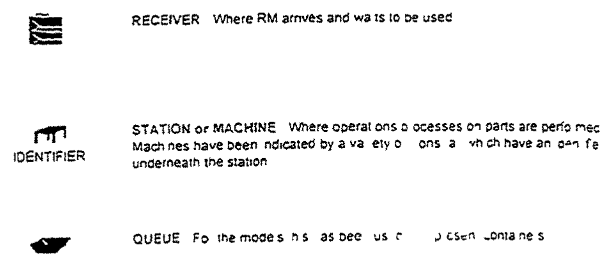


Figure 3 - Legend for Simprocess Output

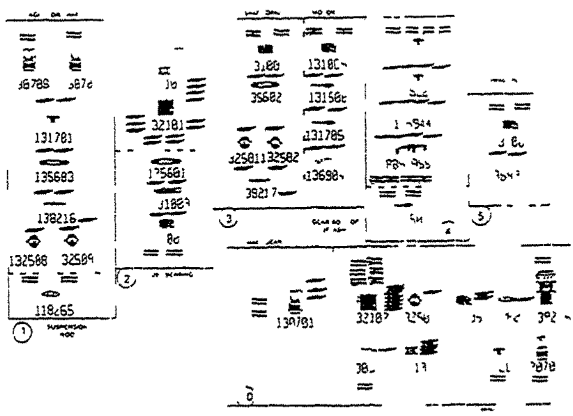


Figure 4 - Redesign New Workshop Layout

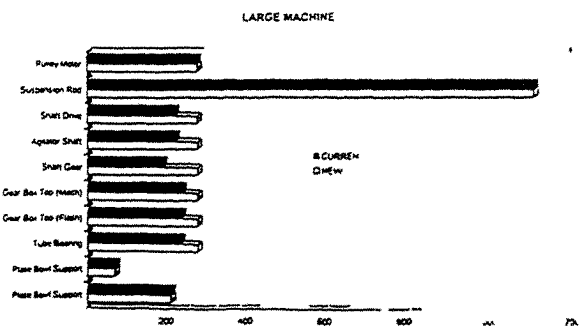


Figure 5- Large Washing Machine Part Comparison

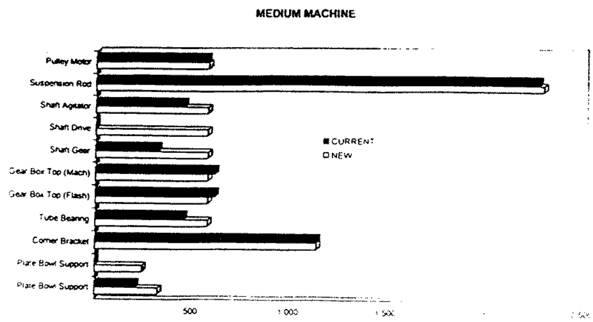


Figure 6 - Medium Washing Machine Part Comparison

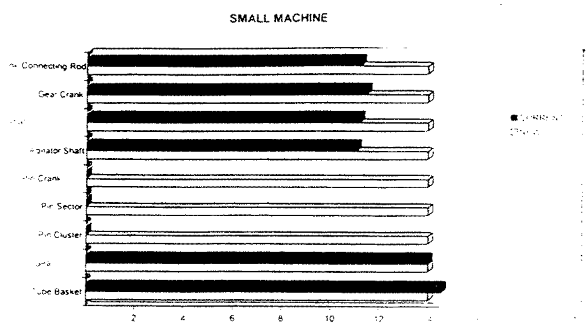


Figure 7 - Small Washing Machine Part Comparison

WASHING MACHINE TYPE	*AMOUNT (per day)*
Large	225
Stainless	170
Plastic	55
Medium	459
Stainless	262
Plastic	197
Small	14

Table 1 - Production Rate of Washing Machines to Establish Model

WASHING MACHINE TYPE	*AMOUNT (per day)*
Large	280
Stainless	211
Plastic	69
Medium	571
Stainless	326
Plastic	245
Small	14

Table 2 - 851 plus small Production Rate Breakdown

CATEGORY	CRITERION	WEIGHTING	RATING Current	RATING New
Costs Weight = 0.767	Labour Efficiency	0.431	0.455	0.545
	Shift Penalties	0.079	0.001	0.999
	Non Value Added Work	0.105	0.545	0.455
	Operating Expenses	0.086	0.479	0.521
	WIP	0.299	0.400	0.600
Intangible Features Weight = 0.08	Outcome	0.376	0.455	0.545
	Training	0.210	0.455	0.545
	Multi Skilling	0.293	0.385	0.615
	Material Handling	0.088	0.143	0.857
	Space Utilised	0.033	0.100	0.900
Intangible Costs Weight = 0.153	House Keeping	0.030	0.417	0.583
	Performance	0.207	0.345	0.655
	Supervision	0.120	0.167	0.833
	Balance of Labour	0.253	0.091	0.909
	Competition	0.176	0.900	0.100
	Flexibility	0.214	0.400	0.600
Overall Rating			0.406	0.594

Table 3 - AHP Analysis Data

THE APPLICATION OF GENETIC ALGORITHM IN THE PLANNING OF FMS SYSTEM

K.W. Keung[+], W.H. Ip[+], Y. Li* and D.W. Wang*
[+]Department of Manufacturing Engineering
The Hong Kong Polytechnic University, Hong Kong
*Department of Automation Control
Northeastern University, China

ABSTRACT

Flexible Manufacturing System (FMS) is a research area in manufacturing system design which has attracted many research interests. Recently, Genetic Algorithm (GA) is regarded as a powerful tool to solve complicate system design problems. In this paper, the author attempts to verify that GA is applicable to solve FMS problems. In this paper, GA is proposed to address the problem of production scheduling and planning (PSP) in FMS as an example. FMS often produces many types of products with various lot-size to support different market requirements. Earliness and Tardiness Production Scheduling and Planning (ETPSP) is one kind of methods which can be applied to solve FMS PSP problems. However, the PSP problem is a large scale problem where conventional approaches to ETPSP cannot effectively solve large scale problem. This paper outlines the application of GA and demostrates through a multi-product FMS production problem.

KEYWORDS

Flexible Manufacturing System (FMS), Earliness and Tardiness Production and Scheduling Problem (ETPSP), Genetic Algorithms (GA).

1. INTRODUCTION

A Flexible Manufacturing System (FMS) is a manufacturing system which is similar but somewhat different from conventional manufacturing system. It consists of a set of numerically controlled machines which are connected through an automated storage/retrieval system. Each process in FMS is considered to be computerised. Yet, the use of FMS still involve many problems as stated in [1] This paper is focused on planning and scheduling problems of FMS. In fact, FMS planning problems can be classified into long range planning problems and short range planning problems [2]. Long range planning problem refers to the planning and design of the FMS while short range planning problem refers to the master production scheduling, material planning and lotsizing, detailed capacity planning and short range scheduling of FMS.

Out of the short range planning problems, some scheduling and planning problems could be solved by Earliness/Tardiness Production Scheduling and Planning (ETPSP) method. However, existing researches on ETPSP problem have been mainly focused on scheduling problems of single-machines and parallel multi-machines with earliness and tardiness penalties, process capacity is simply regarded as a constant rather than a variable and lot-size is not considered [3,4,5,6,7]. Meanwhile, a new extensive model is developed in this paper to address the same ETPSP problem incorporating lot-size consideration and multi-process capacity balancing using a Genetic Algorithm (GA) approach.

GA was invented and developed to mimic some of the processes observed in natural selection, initially by J. Holland and his associated at the University of Michigan in 1960s [8]. It is stochastic, discrete event and non-linear process. GA have been used on machine learning, artificial intelligence, pattern recognition and operation research, etc. [9]. There are also some papers that discussed the application of GA to scheduling problem, most of them considered job-shop, flow-shop or linear balancing sequencing, etc. [10,11]. Few papers discussed their application in the ETPSP problem. In this research, the ETPSP problem is considered as an non-linear and discrete optimal problem, which can be solved by GA. GA is a novel approach of solving ETPSP problem with lot-size consideration and multi-process capacity balancing. The experimentation results demonstrate the effectiveness of using this novel approach to solve the ETPSP problem.

2. PROBLEM AREAS IN FMS

The task of running a FMS smoothly is a great effort because there exist many problems inside. A brief introduction on the problems of FMS has been described in the previous section. In fact, FMS problems mainly can be divided into four main areas. The first problem area begins with the design of FMS. This is followed by planning problem area. Then, there is scheduling and control problem area. Finally, there is also day to day operation problem area. *Figure 1* shown below summarizes the major problems in each area of FMS [1].

DESIGN PROBLEM AREA

- what components shall be manufactured
- what processes are necessary
- what type and number of machine tools are needed
- what type of different flexibilities are required
- what type of FMS shall be developed
- what is the type and capacity of material handling system
- layout of the FMS
- number of pallets and fixtures

PLANNING PROBLEM AREA

- select the part types for immediate and simultaneous manufacture
- partition the machines of each type into machine groups
- determine the production ratio at which the selected set of part types should be maintained at over time on the system
- allocate the limited number of pallets and the fixtures of each fixture type among the selected part types
- allocate the operations and associated cutting tools of the selected part types among the (grouped) machines.

SCHEDULING AND CONTROL PROBLEM AREA

- input sequencing problem: the order of releasing into the system the parts that have been selected at the planning stage must be determined
- operation sequencing problem: some operations of the part process plans may be interchanged
- workstation selection problem: when an operation can be performed on several workstations, the one to use has to be determined
- part sequencing problem· the part that must be processed first on a work station has to be selected from the input queue at this station
- material handling carrier selection problem: in case AGVs are used in FMS, the cart that must come to pick up a part has to be determined

DAY TO DAY OPERATION PROBLEM AREA

- aggregate planning: for a given day's requirement, a algorithm should be applied to determine the workload requirements such as the machine capacity, the current on hand inventory and the number of working machines
- machine loading: each operation and its associated cutting tools needs to be allocated to one or more machines subject to time and tool magazine capacity constraints
- fixture management: if there are fixtures sharing among part types, then fixture allocation problem exists
- inventory management: lot-sizing problem is one of the inventory management problems

Figure 1: Summarize of FMS problem areas

Out of the above mentioned FMS problems, the author takes aggregate planning as an example to illustrate that GA can be applied to solve this problem through ETPSP method. The following section described briefly the ETPSP model.

3. THE ETPSP MODEL

3.1 Background of ETPSP Problem

Assuming that a FMS produces N types of products (Product 1, Product 2, ..., Product N) to supply market requirements in a horizon [1, T] which contains Period 1, Period 2, ..., Period T, and each product from raw-material to final product has to pass M processes or assembling stages (Stage 1, Stage 2, ..., Stage M). According to the order and the information of the FMS, some useful symbols are described as follows,

$d_i(k)$ ($i=1,2,...,N$, $k=1,2,...,T$) : the requirement quantity of Product i in Period k,
$c_j(k)$ ($j=1,2,...,M$, $k=1,2,...,T$) : the available capacity of Process j in Period k,
w_{ij} ($i=1,2,...,N$, $j=1,2,...,M$) : the unit capacity requirement of Product i for Process j[1],
l_i ($i=1,2, ..., N$) : the initial inventory of Product i[2].

In fact, the available FMS capacity may fail to meet the customers' requirements at any time. To overcome capacity shortages and surpluses, either early production or tardy delivery is necessary.

3.2 Optimization of ETPSP Problem

The principle of ETPSP is to minimize the total penalties of earliness and tardiness in the FMS. Considering that,

s_i ($i=1, ..., N$): the lot-size of Product i,
α_i ($i=1,2,...,N$) : the unit time earliness penalty of Product i,
β_i($i=1,2,...,N$) : the unit time tardiness penalty of Product i[3].

The aim of ETPSP is to find an optimal lot-size production schedule in the horizon [1,T], then the total cost of earliness and tardiness penalties are minimized and the manufacturing process capacity constraints are confirmed. Let $p_i(k)$($i=1,2,...,N$, $k=1,2,...,T$) be the planning production quantity of Product i in Period *k*. This is discrete due to the different lot-size. Then, the ETPSP problem can be described by an objective function as formula (1),

$$\min_{p} = \sum_{i=1}^{N}\sum_{k=1}^{T}\{a_i[l_i + \sum_{t=1}^{k} p_i(t) - \sum_{t=1}^{k} d_i(t)]^+ + \beta_i[\sum_{t=1}^{k} d_i(t) - \sum_{t=1}^{k} p_i(t) - l_i]^+\} \qquad (1)$$

and its constraints are subject to formulae (2) and (3),

$$\sum_{i=1}^{n} w_{ij} p_i(k) \le c_j(k),\ j=1, 2, ..., M,\ k=1, 2, ..., T, \qquad (2)$$

$$0 \le p_i(k) \in S_i, S_i = \{r \cdot s_i, r \ge 0, i = 1,2,...,N\},\ i=1, 2, ..., N,\ k=1, 2, ..., T. \qquad (3)$$

where, $(x)^+ = \max\{0, x\}$. Formulae (1), (2) and (3) are regarded as a programming Problem (P_0) which will be proposed to apply GA to solve it.

4. OVERVIEW OF GENETIC ALGORITHMS

GA is firstly invented by John Holland [8] and his associates at the University of Michigan in the 1960s. He is generally regarded as the father of Genetic Algorithm although he did not give its name. GA is inspired by the mechanism of natural selection, a biological process in which the rule of it is the fittest will be survive. It weeds out the bad and tends to produce more of the good individuals. Not only does it produce more of the good solutions but better and better solutions. This is because it combines the best traits of parent individuals to produce superior children. This combination operator is called crossover. The term

[1] The unit capacity requirements of all types of products should be kept in constants along a horizon [1,T].
[2] l_i<0, means the initial shortage of Product i.
[3] Generally α_i>β_i, α_i and β_i can be determined by the inventory cost and tardiness compensation in practice.

"genetic algorithm" comes from the fact that individuals are represented as strings of bits analogous to chromosomes and genes. In addition to recombination by crossover, we also throw in random mutation of these bit-strings every so often. This keeps the GA from getting stuck at good but non-optimal solutions.

GA is inspired by the mechanism of natural selection where stronger individuals are likely to be the winners in a competing environment. GA uses a direct analogy of such a natural evolution. GA presumes a potential solution as an individual which can be represented by a vector. This idea is familiar to biology applications, which can be termed as the genetic structure of a chromosome. Throughout the genetic evolution, starting from a population of chromosomes, some fitter chromosomes tend to yield good quality offspring, and this means better solutions to the problem [9,12,13]. *Figure 2* illustrates the logical structure of GA.

1 Initialize a population of individuals
2. Evaluate all individuals of the population.
3 Select a sub-population for offspring production
4 Create new offspring by mating current sub-population
  4 1 recombine the chromosomes of the current sub-population by crossover operator
  4 2 agitate the mated population stochastically by mutation,
5 Delete members of the parents population to make room for the new individuals
6 Evaluate the offspring individuals created
7 Select the new individuals and insert them into the population
8 If the stopping criteria is met, stop and return the best individual, else restart at step 2

Figure 2: The Logical Structure of GA

5. SIMULATION RESULTS

Having established the required operations for the GA approach, the overall programming procedures follows that as expressed in *Figure 3*.

Consider that the FMS consists of two-type products and 10 assembling stages in a 12-period horizon. According to different types of products, the required parameters are listed as follows,

- ***Earliness and tardiness penalties:*** $\alpha_1=10$, $\alpha_2=10$; $\beta_1=15$, $\beta_2=12$;
- ***Capacity requirement*** $w_{ij}(i=1,2; j=1,2,...,10)$, *see Table 1-(a);*
- ***Order quantity*** $d_i(k)(i=1,2; k=1,2,...,12)$, *see Table 1-(b); and*
- ***Available capacity*** $c_j(k)(j=1,2,...,10; k=1,2,...,12)$, *see Table 2.*

Using the proposed GA method, the ETPSP with lot-size consideration and multi-process capacity balancing can be obtained as shown in *Table 3*.

Table 1-(a) Capacity requirement w_{ij}

i\j	1	2	3	4	5	6	7	8	9	10
1	1.0	0.6	0.8	0.3	0.7	1.5	1.2	1.1	0.9	0.4
2	0.6	0.8	1.3	2.0	0.7	2.1	0.6	0.8	0.2	0.1

Note i is the product type, j is the assembling stage

Table 1-(b) Order quantity $d_i(k)$ and total requirement

i\k	1	2	3	4	5	6	7	8	9	10	11	12	**Total Req't**
1	0	0	20	0	0	0	40	0	0	0	20	0	**80**
2	10	0	0	50	0	0	0	0	20	0	0	5	**85**

Note: i is the product type, k is the planning period

Table 2. Available capacity $c_j(k)$

j\k	1	2	3	4	5	6	7	8	9	10	11	12
1	30	30	30	30	30	30	30	30	30	30	30	30
2	18	28	18	18	18	18	18	18	18	18	18	18
3	34	44	34	34	34	34	34	34	34	34	34	34
4	24	34	24	44	19	24	54	20	24	24	24	24
5	26	36	26	46	26	26	26	26	26	26	26	26
6	60	70	60	60	60	60	30	30	60	60	60	60
7	19	29	19	39	19	59	16	99	19	19	49	19
8	36	46	36	56	36	16	36	26	36	36	36	36
9	20	30	20	40	20	20	20	10	20	20	20	0
10	18	28	18	30	18	18	18	14	18	28	18	18

Note: j is the assembling stage, k is the planning period

Table 3. Production Quantity $p_i(k)$, lot-size s_i and total production quantity

i\k	1	2	3	4	5	6	7	8	9	10	11	12	s_i	**Total Prod. Qty.**
1	10	0	10	**0**	15	5	5	5	10	15	5	0	5	**80**
2	10	*10*	*10*	**20**	0	*10*	0	10	10	0	0	10	**10**	**90**

Note. i is the product type, k is the planning period

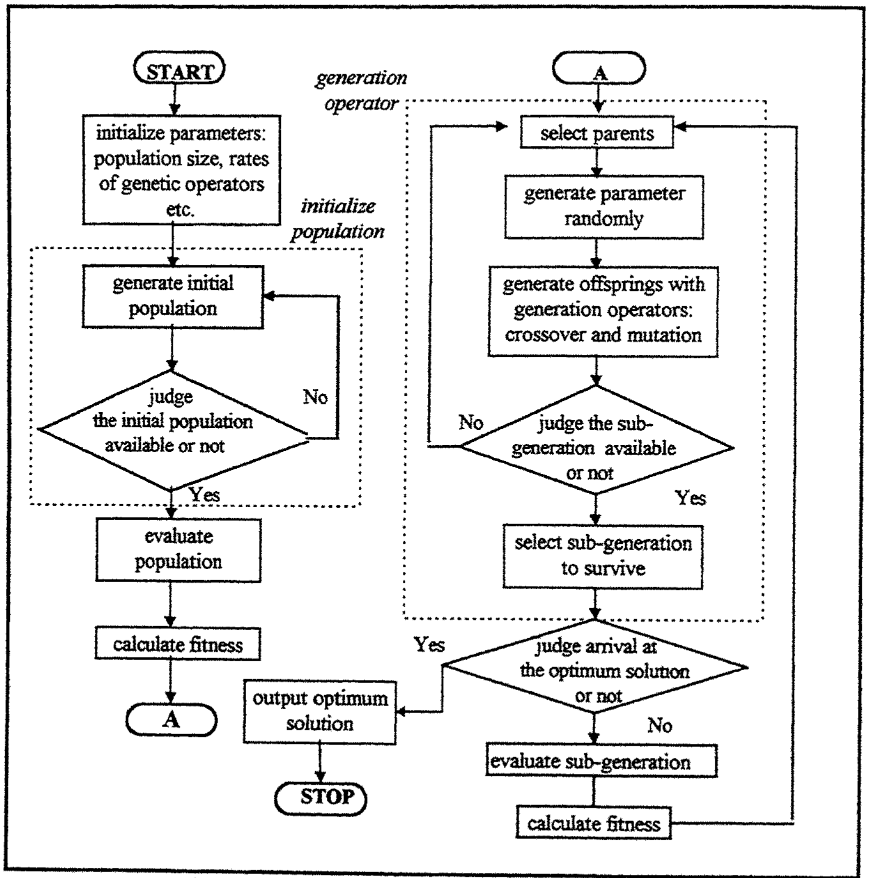


Figure 3: Programming flow chart of the proposed GA

It can be seen from *Table 1-(b)* and *Table 3* that the difference between total production quantities and the total requirements is not more than just one lot-size, i.e., 5 or 10 units. That is,

$$\sum_{k=1}^{12} p_i(k) - \sum_{k=1}^{12} d_i(k) < s_i$$

Taking an example of Process 3 in Period 4 and the production schedule in the order indicated in *Table 1-(a)*, we find that there exists a shortage process which is,

$$c_j(t)\Big|_{j=3,t=4} < \sum_{i=1}^{N=2} w_{ij} \cdot d_i(t)\Big|_{j=3,t=4}$$

If production is arranged according to the ETPSP indicated in *Table 3*, the capacity shortage can be overcame due to the condition,

$$\sum_{i=1}^{N=2} w_{ij} \cdot p_i(t)\Big|_{j=3,t=4} < c_j(t)\Big|_{j=3,t=4}$$

It is clearly demonstrated that the ETPSP is an effective means of solving the process capacity shortage problem by early or tardy production in FMS. Furthermore, *Table 1-(b)* and *Table 3* also indicate that the GA approach to ETPSP not only satisfies the customers requirement, but also offers a minimum in regard to the total early and tardy penalties.

CONCLUSIONS

A GA approach to solve the FMS, ETPSP problem with lot-size consideration and multi-process capacity balancing is proposed. The simulation results show that this scheme is effective as well as efficient. The proposed method does not require any unrealistic assumptions on the objective functions such as linearity, convexity and differentiability. It can even compute highly complicated and non-linear functions of the measuring performance. In addition, a large scale ETPSP problem with multi-type products can be solved, and a near-optimal solution is reached in a short time. It is a noted improvement on any existing techniques by modifying the manufacturing process, so that the manufacturer can respond to the changing market requirements quickly and fulfil the need of the customers.

REFERENCES

[1] K.E. Stecke, "Design, Planning, Scheduling, and Control Problems of Flexible Manufacturing Systems", Annals of Operations Research 3, pp.3-12, 1985.

[2] H. Tempelmeier, H. Kuhn, Flexible Manufacturing Systems. Decision Support for Design and Operation, Ch. 2, John Wiley & Sons, Inc., 1993.

[3] R. Baker, G.D. Scudder, "Sequencing with Earliness and Tardiness Penalties: a Review", Operation Research 38, pp. 22-36, 1990.

[4] J. De, B. Ghost, C.E. Wells, "Scheduling to Minimize Weighted Earliness and Tardiness about a Common Due-Date", Computer Operation Research 18, pp. 465-175, 1991.

[5] Hall, M.E. Posner, "Earliness-Tardiness Scheduling Problem I: Weighted Deviation of Completion Times About a Common Due Date", Operation Research 39, pp. 836-846, 1991.

[6] Hall, W Kubiak, S.P. Sethi, "Earliness-Tardiness Scheduling Problem II: Deviation of Completion Times About a Restrictive Common Due Date", Operation Research 39, pp. 847-856, 1991.

[7] De, Jhosh, C.E. Wells, "General Solution for a Class of Early/Tardy Problems", Computer Operation Research 20, pp.141-149, 1993.

[8] J. Holland, Adaptation in Natural and Artificial Systems, MIT Press, 1975.

[9] C.R. Reeves, Modern Heuristic Techniques for Combinatorial Problems, Halsted Press, New York, 1993.

[10] Y. Hashimoto, I. Nishikawa, H. Tokumaru, "Line Balancing Using a Genetic Evolution Model", Control Engineering and Practice 3(1), pp. 69-76, 1995.

[11] Mulkens, "Revisiting the Johnson Algorithm for Flow-Shop Scheduling with Genetic Algorithms, Knowledge-Based Reactive Scheduling", Elsevier Science B.V., North-Holland, pp. 69-80, 1994.

[12] Man, K.S. Tang and S. Kwong, "Genetic Algorithms: Concepts and Applications", to be appeared in IEEE Transactions on Industrial Electronics, 1996.

[13] Davis, Handbook of Genetic Algorithm, Van Nostrand Reinhold, New York, 1991.

AN FMS PROCESS PLANNING SYSTEM BASED ON COLOURED PETRI NETS

T.N.Wong[1], K.H.Ng[1], M.C.K.Kwok[2]
[1]Department of Industrial and Manufacturing Systems Engineering
University of Hong Kong
Pokfulam Road, Hong Kong
[2]Department of Manufacturing Engineering
Hong Kong Technical College (Chai Wan)
Chai Wan, Hong Kong

ABSTRACT

This paper discusses the design of a dynamic process planning system and its integration into the production planning of an FMS. Coloured Petri nets and expert systems are used to model and analyse both the process planning function and the simulation aspects of FMS. A coloured Petri net based simulation tool that simulates the dynamic operations of FMS and provides the dynamic status of an FMS is also presented. Planning and decision are carried out in different stages by using domain-specific knowledge methods. The stages include FMC selection, process planning, routing, production scheduling, etc. By integrating these various sub-systems together, it is envisaged that more accurate and efficient plans can be generated.

KEYWORDS

Dynamic CAPP; FMS; Simulation; Coloured Petri nets; Expert systems

1 INTRODUCTION

This paper describes a prototype dynamic Computer Aided Process Planning (CAPP) system for FMS. The aim of the prototype system is to generate process plans that represent more realistic situations in FMS operation. Hence the prototype system is developed in which both the FMS configuration and the dynamic status of FMS are considered by simulating these FMS operations. Petri net methodology is used for the process planning system and FMS simulation. Modularised coloured Petri nets are built to simulate the dynamic operations of different FMCs and hierarchical coloured Petri nets are constructed in the planning and decision section of the prototype system. In the prototype CAPP system, scheduling mechanism in terms of rules or heuristic knowledge, and the technological and sequencing constraints of process planning are simultaneously considered. As conflicts always exist between the scheduling and process planning aspects, knowledge based expert system is used to resolve such conflicts. These expert rules for planning and conflict resolution are transformed to parts of the structures of the Petri net based simulation tool. The status of the dynamic resources and decision of the FMS configuration constitute the knowledge for process and production planning of the FMS.

2 MODELING METHODOLOGY

2.1 PETRI NET MODELLING

Petri nets provide an elegant and mathematically rigorous modelling framework for discrete event dynamic systems. Petri net models have been used as a very promising performance modelling and analysing tool for systems which exhibit causal dependency, concurrency, conflicts and synchronisation [1,2]. The execution of a Petri net is determined by the number and distribution of tokens and by firing transitions. A transition fires by removing the tokens from the input places and then depositing the tokens in the output places. The state of a Petri net is defined by the number of tokens resided in each place. Although Petri nets are very powerful for modelling systems, it is necessary to analyse the modelled system in order to study the system behaviour. There are several important desirable properties that models of real physical systems should exhibit. Firstly, *boundedness* refers to a finite requirement of resources and the models are bounded. Secondly, *reversibility* ensures that recovery from failure states is possible. Finally, *liveness* in a Petri net implies absence of deadlocked states. These three properties can be analysed based on the finite reachability sets which are simple from a conceptual point of view [3].

Basic Petri nets provide a powerful formal method for modelling the dynamic behaviour of discrete concurrent systems. However, it is only suitable for modelling the local levels of a complex system which usually contains many activities which are similar but not identical. Basic place/transition models are capable of representing large and complex systems but the graph becomes illegible and it is difficult to model and analyse the system. The development of high level Petri nets constitutes a very significant improvement in this respect. The coloured Petri nets describe complex systems in a more compact form in which they stress the similarities between the two kinds of processes. The more compact representation has been achieved by equipping each token with an attached data value - the token colour. The data value may be of arbitrarily complex type. For a given place all the colours of its tokens must belong to the same colour set. The conciseness of the resulting model is compensated by the more complex inscriptions attached to the arcs of the nets. The guard is a predicate attached to the transition which restricts the set of possible token occurrences to be used in the transition. The guard is like an input rule which can be used to model preconditions on the execution of the transition [4,5]. Almost all complex systems consist of hierarchical structures. Petri nets can be organised using hierarchical coloured Petri nets. A large coloured Petri net can be constructed by combining numbers of smaller hierarchical decomposed nets. They can be constructed top-down, bottom-up, or by mixing these two methods [3].

2.2 PROCESS PLANNING FOR FMC

Generally, people consider process planning as an individual discrete system. In most manufacturing establishments, the process planning and the production planning functions are found to be incompatible to each other and a lot of problems arise during shop-floor for production. This is because the interactions of the process planning stage with the other production functions are not considered during system design and implementation. As a remedy, a manufacturing system has to integrate the production planning and process planning functions so that infeasible plans can be eliminated. The system can concurrently consider more downstream factors and status. Thus, such dynamic factors are necessary to keep track of and adapt to changes in the shop floor status. As a result, more efficient and accurate plans can be proposed.

Several prototype process planning models have been reported regarding the dynamic property of manufacturing systems [6-8]. Most of them consider the machine characteristics, capability and capacity in process planning. Existence of such machines are considered but actual configurations of machines and other facilities such as transport systems or buffer storage are not considered. In FMCs, machine layout, transport systems and material flow are as important as machines availability and their individual characteristics. These factors always affect process determination and sequencing. And the type of transport system used in an FMC may change the process plan. In addition, it is likely that the dynamic shop-floor status and the scheduling mechanism for an FMC are in conflict with the technological and sequencing constraints in process planning. Whenever such conflicts occur, the original process plans will become infeasible, and the mechanism or rules used to generate these plans will no longer be valid. If all the related technological, sequencing and scheduling rules are considered together, more efficient and feasible process plans will be generated.

2.3 DYNAMIC PROCESS PLANNING FOR FMC

In the prototype system, the hierarchical coloured Petri net (HCPN) methodology is used to model the dynamic process planning for FMC. Such complex systems with multiple and shared resources can be modelled by HCPN. HCPN uses the concept of distinguishing the multiple resources by colour tokens. Each token represents a class of resources and it can model larger and complex systems. Currently, Petri net models are commonly used in modelling and analysing FMCs. Simulation for the dynamic FMC operations is used in order to reflect the dynamic status of resources. Therefore, the dynamic shop-floor feedback model will become a sub-net of the whole system and different manufacturing functions modelled by Petri nets can be easily integrated. In preparing the Petri network for the specific FMC, a hybrid synthesis approach is used. The method combines top-down and bottom-up approaches while preserving the desired analysing properties with the shared resource throughout the design process [3,9].

In the prototype system, knowledge-based expert system are used to aid in decision making and conflict resolution. Applications of knowledge-based expert systems are in the following areas :

- FMC selection according to parts and their associated manufacturing features,
- process planning knowledge representations such as machining - operation, tooling, operation-selection and operation-sequence knowledge,
- scheduling mechanism in terms of rules or heuristics using in simulation tool, and
- conflicts knowledge representation.

For most of the rules, they can be directly applied as the firing function, guard functions or other areas in the coloured Petri nets such that the power of decision and conflict resolution will be improved. On the other hand, some rules may transform to CPNs. This approach builds model from the selective rules to describe suitable match of manufacturing features and machines [9]. For example, a colour Petri net is given in Fig. 1 to depict a simple rule for the selection of the drilling machine for a hole feature :

IF (hole in PLACE 1) (drilling machine in PLACE 2)
THEN (drilling process in PLACE 2)

As planning knowledge is primarily domain specific, the knowledge base of the system is designed to continuously expand to include the various planning environments of different machine configurations with different levels of complexity under FMC.

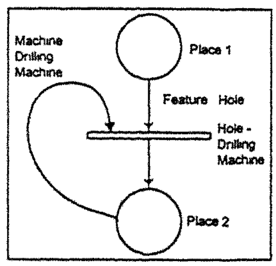


Fig. 1 : CPN for process selection rule

3. SYSTEM STRUCTURE AND IMPLEMENTATION

3.1 SYSTEM STRUCTURE

The prototype CAPP system is a generative dynamic CAPP system which designed for prismatic machined parts manufacturing in an FMC. An FMC consists of the basic elements including CNC milling machines, CNC lathes, AGVs, robots, input and output buffers, common storage, etc. Thus, the FMC can fabricate a variety of features including pocket, hole, slot, notch, step, etc. The whole system is modelled by Petri nets programmed under the C language environment and supported by CLIPS for decision making. Fig. 2 shows the system structure. There are five modules in the prototype system. The first module is the Parts Description Module in which interactive part design is facilitated. As parts are defined, the parts data are fed to the next two modules: the FMC Selection Module and the Pre-Planning Module. The FMC selection module selects the most suitable FMC and creates a set of modularised Petri nets for the corresponding configuration to simulate the dynamic shop-floor operations, and the Pre-Planning module prepares a tree of alternative manufacturing

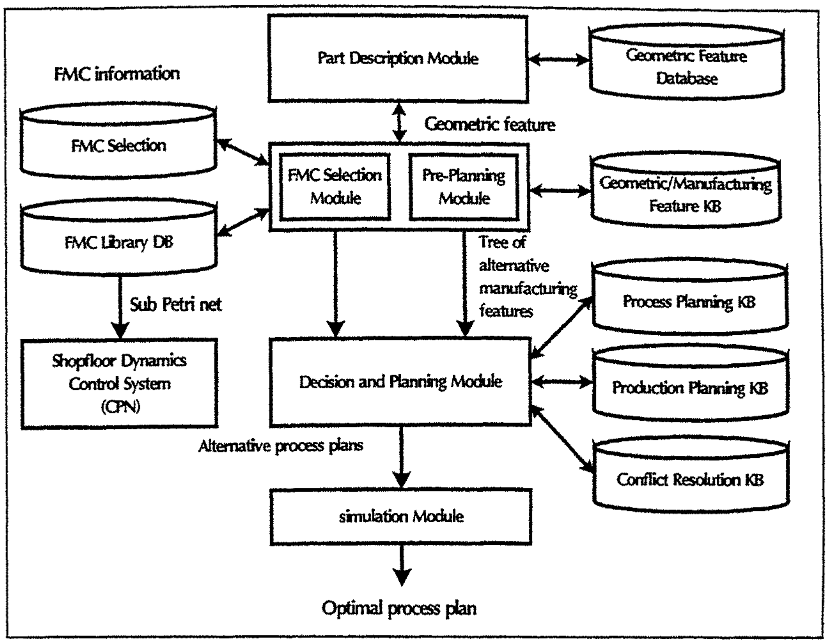


Fig. 2 : The prototype CAPP system

features for each geometric feature. Information of both modules are then merged together in the Decision & Planning Module. As a result, process and production plans are generated. In addition to feedback dynamic status, system verification and optimisation are also carried out in the Simulation Module.

3.2 SYSTEM IMPLEMENTATION

The Parts Description Module is required to completely describe the geometrical, dimensional and technological information of the parts. In the proposed system, the design by feature methodology is used to describe the parts [10]. This provides the designer with a features library which can be used with a set of operators such as add, delete and modify in order to create a feature representation. This gives a unique, pre-defined feature set for describing parts. The feature domain includes a set of holes, slot, step, pocket, and notch. Geometric feature representation database is created and the feature structure includes surface components, tolerance information, etc.

The FMC Selection Module is implemented to determine appropriate FMC to the part or part family by using group technology. The FMC selection knowledge base is built to maintain knowledge or rules in selecting FMC under CLIPS language. Generally, more than one FMC may suit for fabricating a part. Thus different rules are considering in this stage. For example, the module will select the one which has more idle machines but less utilisation. Some of the rules are as follows :

Rule 1 : if (FMC has_capability ALL-PART-FEATURES) and (FMC has_utilization LOW)
then (FMC is_selected_for PART)
Rule 2 : if (FMC-M/C has_capability PART-FEATURE) and
(FMC-M/C-TOLERANCE is_suitable_for PART-FEATURE-FINISH) and
(PART-FEATURE assigned NOT-READY)
then (FMC-M/C is_assigned_to PART-FEATURE) and
(PART-FEATURE assigned READY)

Rule 3 : if (FMC-M/C is_assigned_to PART-FEATURE) and
(FMC-M/C-TOLERANCE < 0.1) and (PART-FEATURE-DIMENSION < 10)
then (FMC-TOOL is_assigned_to FMC-M/C for PART-FEATURE)

A set of modularised Petri nets with hierarchical structure for the chosen FMC is created. Generally, these sub-nets act as a shop-floor control sub-system in order to trace the dynamic shop-floor operations and their status. These nets will merge together with the Decision and Planning Module. An FMC library database is maintained in order to generate the corresponding Petri sub-net. Different pre-defined machines or facilities and available processes can be selected to build up the FMC (Fig. 3). Fig. 4 depicts the corresponding coloured Petri sub-nets. Fig. 4(a) and (b) illustrate the two sub-nets of a machining workstation, in which physical flow of parts is dominant. The sub-net in Fig. 4(c) gives the routes of the flow. The CPNs are mainly controlled by the message flow in both machine and cell level. Apart from definition of place, transition and arc, the inscriptions such as guard functions, arc expressions and conflict resolutions must be defined.

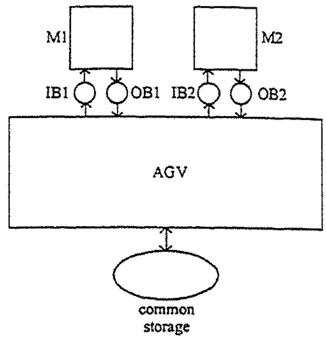
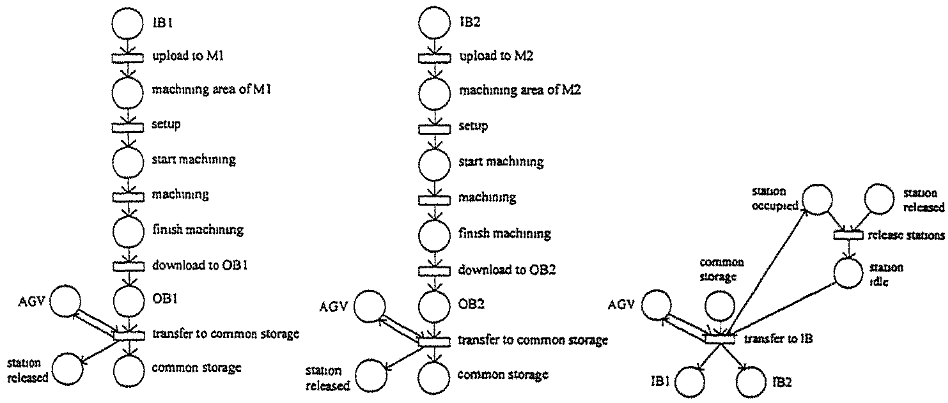


Fig. 3 : An FMC



(a) (b) (c)

Fig.4 : Sub Petri net for the FMC

The Pre-Planning Module converts each geometric feature to a tree of alternative manufacturing features. There is a geometric/manufacturing feature conversion knowledge base. The module will consider the geometric, dimensional and tolerance data and generate the tree according to different priorities. These trees will be used in the Decision and Planning Module.

The Decision and Planning Module is the main sub-system to make various logical decisions automatically in order to generate process plans for the parts manufactured in each FMC. In general, this is a coloured Petri net model which is programmed using the C language. CLIPS is also used to represent the rules which can be translated to the firing functions or guard functions in the coloured Petri net model. Three knowledge bases are implemented: one each for process planning, scheduling and conflict resolution of rules. The process planning rules include geometric constraints and technological constraints. The scheduling knowledge base includes loading, transport constraints, machine selection rules, etc. The conflict resolution portion consists of various decision rules to work around the constraints according to the company goal. In addition, the dynamic shop-floor status from the previous defined CPN models in the shop-floor dynamics control sub-system are included. As a result, a number of alternative production plans is generated.

The Simulation Module will use to simulate the dynamic environment of FMC operations. Moreover, this is to verify and optimise the resultant plans. The previously defined Petri sub-net for tracing the shop-floor status can be utilised in this module and they can be enhanced with optimisation techniques such as shortest processing time, minimisation of machine idle time, etc.[11].

4 CONCLUSION

The Petri net methodology is widely used in production planning and scheduling and found to be very useful in describing dynamic and complex systems. In modelling the prototype system, coloured Petri net models are used for both process and production planning. Coloured Petri net modelling is a powerful tool to describe the simultaneous and concurrent nature of the manufacturing subsystems of an FMC. In addition, CLIPS provides an efficient environment for integrating the expert rules into the Petri net models

Assumptions have been made in the proposed system. Firstly, each part must be manufactured in one FMC since set-up or changeover time between any two FMCs is expected to be high and unjustified. Secondly, each FMC is independent of other FMCs.

REFERENCES

1. Peterson J. L., Petri net theory and the modelling of systems, Prentice-Hall, Englewood Cliffs, N.J., 1981.
2. Viswanadham N. and Narahari Y., "Petri net models", Chapter 5, Performance Modelling of Automated Manufacturing Systems, Prentice Hall, Englewood Cliffs, N.J., 1992.
3. Dicesare F., et al., Practice of Petri nets in manufacturing, Chapman & Hall, London, 1993.
4. Jensen K., Coloured Petri Nets - Basic Concepts, Analysis Methods and Practical Use, Volume 1, Springer-Verlag, Berlin, 1992.
5. Jensen K. and Rozenberg G. (Ed.), Higher-level Petri Nets - Theory and Application, Springer-Verlag, Berlin, 1991.
6. Masood A. and Srihari K., "RDCAPP: A Real-Time Dynamic CAPP System for an FMS", Int. J. of Advanced Manufacturing Technology, Vol. 8, 358-370, 1993.
7. Hou T.H. and Wang H.P., "Integration of a CAPP system and an FMS", Computers in Industrial Engineering, Vol 20, No. 2, 231-242, 1991.
8. Zhang H.C. and Mallur S., "An integrated model of process planning and production scheduling", Int. J. Computer Integrated Manufacturing, , Vol. 7, No. 6, 356-364, 1994
9. Tzafestas S., "Petri-net and knowledge-based methodologies in manufacturing systems modelling simulation and control", Proc. 5th CIM Europe Conf., 39-50, Athens, Greece, May 17-19, 1989.
10. Dolinska M. and Besant C.B., "Dynamic Control of Flexible Manufacturing Systems", Int. J. of Advanced Manufacturing Technology, Vol. 10, 131-138, 1995.
11. Carrie A.S., Simulation of Manufacturing Systems, John Wiley & Sons, Tiptree, 1988.

DYNAMIC MIXED DISPATCHING OF FMS

A. Krishna Kishore
Research Scholar

N.Ramesh Babu
Associate Professor

Manufacturing Engineering Section
Department of Mechanical Engineering
Indian Institute of Technology
Madras, India

ABSTRACT

This paper presents an algorithm, called Dynamic Mixed Dispatching Rule (DMDR), for dispatching different parts in an FMS. This approach tries to change dispatching rules to be applied at different machines over fixed time intervals to minimise the mean tardiness of different part types to be produced in an FMS. A search algorithm is developed to obtain the rule combination quickly. Effectiveness of the present approach is explained with a case study. DMDR approach showed substantial improvement over single dispatching rule (SDR) approach and mixed dispatching rule approach (MDR).

KEY WORDS

Flexible Manufacturing System, Scheduling, Dispatching

1. INTRODUCTION

A manufacturing system is flexible if it is capable of processing a number of different workpieces simultaneously with inherent flexibilities such as process flexibility, routing flexibility, volume flexibility and machine flexibility. A Flexible Manufacturing System (FMS) is designed to combine the efficiency of a high production-line and the flexibility of a job shop to best suit the small to medium volume production of a variety of products.

Scheduling of FMS is of great importance for optimum utilisation of resources in FMS. It affects the overall performance of the system: internally affecting the utilization of expensive resources, and externally in meeting the customer demands. Scheduling of an FMS can be done in three stages namely loading, sequencing and dispatching. For dispatching different parts to resources, heuristic rules called dispatching rules are widely used to arrive at results quickly. Many dispatching rules have been employed for scheduling of an FMS. Blackstone et al. (1982) indicated that no single dispatching rule can give the best results for selected objective function.

Wu et al. (1989) proposed a multi-pass algorithm for scheduling an FMS with which dispatching rule is changed over fixed time intervals to minimise the mean tardiness and the mean flow time of different parts to be processed in the FMS. This algorithm have yielded an improvement up to 29.3% over conventional single dispatching rule (SDR) approach. Ishi et al. (1991) proposed a transient based scheduling algorithm which selects the dispatching rule dynamically over different time intervals. This algorithm has shown 16.5% improvement over SDR approach in minimising the mean tardiness of parts. Ishi et al. (1994) applied Mixed Dispatching Rule (MDR) approach for scheduling an FMS which finds combination of dispatching rules that vary from machine to machine. They indicated that MDR approach gives 4% better results over SDR in minimising the mean tardiness and the mean flow time of different parts in the system.

From the above, it is clear that dispatching the parts at different machines according to different dispatching rules can improve the performance of the system in minimising the mean tardiness of different parts to be produced in FMS. Further, the changing dispatching rules at a machine from time to time can also improve the performance of FMS in minimising the mean tardiness of different parts. This paper addresses the scheduling of FMS based on the DMDR approach with a view to examine the influence of this approach on the mean tardiness of parts processed in the FMS.

2. METHODOLOGY

DMDR approach tries to change dispatching rules to be applied at different machines over fixed time intervals to improve the performance of an FMS. The scheduling becomes complex, as the number of dispatching rules to be considered or the number of machines in an FMS increases. This complexity further increases with increased number of intervals. If schedule has to be drawn over 't' time intervals for 'n' processing stations and 'm' rules are chosen for dispatching, the number of rule combinations become $m^{(nXt)}$. Evaluating all the rule combinations is not feasible in view of its computational requirements. Hence, a search algorithm that can provide a near optimal solution quickly is implemented for scheduling an FMS with DMDR approach.

Blackstone et al. (1982) classified dispatching rules into four categories. To generate the schedule for an FMS with MDR approach, four dispatching rules are selected from different categories of dispatching rules. FCFS from democratic based rules, SPT from dispatching rules involving processing times, SLACK from rules based on the due date and NINQ from dynamic rules category are chosen for dispatching.

Different characteristics of the system considered are

- The parts to be processed in the FMS can have routing flexibility.
- Machines used in the system to process different parts are capable of performing multiple operations.
- AGVs movement is unidirectional.
- Setup time of parts is included in the processing time.
- Tool change time is considered to be negligible.
- No machine can process more than at a time.
- Machines will not breakdown while the parts are being processed.
- Arrival pattern of various part types is considered to be deterministic.

2.1 Search Algorithm

The search for best combination of dispatching rules at various machines at different instances is carried out in two stages. Primarily a search algorithm is adopted using the Bottleneck Machine Table (BMT) and the mean tardiness of various part types to arrive at MDR rule combination (Ishi et al. (1994)). BMT is prepared depending on the machine utilisations and the process sequence of a part type. BMT shows the ranking of bottleneck machines that process different part types. The machine that is utilised to the maximum extent is known as a bottleneck machine. This particular machine can become a hindrance to reduce the mean tardiness of the part type with the selected dispatching rule. To find out the different dispatching rules to be applied at different machines over different time intervals a search algorithm is developed that utilises the MDR rule combination and mean operation tardiness (MOT) of different part types processed at a machine over a particular time interval. MOT of a part type is the average positive difference of operation completion time including waiting time for that particular operation and operation due date.

2.2 Algorithm

Different steps involve: he algorithm are

1) The system is simulated w. ifferent dispatching rule combinations and a prominent MDR rule combination for different machines is obtained using the following procedure.
   a) The system is simulated using SDR method. Dispatching rule which minimises the mean tardiness of different part types is applied to all machines as an initial rule.
   b) BMT is prepared using utilisation of different machines and process sequence of different part types. The tardiness of different part types is estimated.
   c) Most tardy part type is selected for improvement. This process continues until all the part types are selected. If all the part types are selected, the search algorithm will terminate.
   d) A bottleneck machine, that processes selected tardy part type is selected using BMT. If all the machines are selected, then go to step (c) to select another tardy part type.
   e) Simulate the system with various dispatching rules at bottleneck machine. Evaluate the performance and select the most promising rule for the bottleneck machine.

Case1. : If the new rule combination shows any improvement in objective function, go to step (b).
Case2. : Otherwise, go to step (d) to select another bottleneck machine.

A set of dispatching rules to be applied at different machines can be obtained. by applying the above MDR methodology. To apply DMDR algorithm, the production period is divided in to multiple and equal time intervals. To arrive at the dispatching rules to be applied in different time intervals at each of the machines, MOT table is prepared while applying the set of dispatching rules obtained with MDR approach and the following procedure is followed:

2) Using the MOT table, the interval with highest mean operation tardiness and corresponding machine are selected to change the dispatching rule to be applied in the particular interval.
3) FMS is simulated with different dispatching rules in selected interval of selected machine and the promising rule is selected for the particular interval of the machine.
4) If all intervals with non zero mean tardiness are considered, the search algorithm is terminated. Otherwise go to step (2) to select another interval and machine.

This procedure is continued to obtain dispatching rules to be applied at different machines in different intervals to optmise the objective.

3. RESULTS AND DISCUSSION

3.1 Case Study

For the purpose of illustrating the application of the present method, the FMS with the following configuration is considered.

TABLE I : CONFIGURATION OF FMS MODEL [Ishi et al. (1994)]

LOAD/UNLOAD STATIONS (L/U) & MACHINES (M)	PART TYPES	BUFFER CAPACITIES	AVERAGE TRANSFER TIME (min.)
L/U - 2 M - 4	6	I/P Buffer - 6 O/P Buffer - 4	2

Each part type has a different processing sequence in the FMS and a different deterministic processing time at each machine. A different due date is defined for each part type so that the ratio of the due date to the total processing time is between 2.0 to 3.0.

3.2 Application of Algorithm

Step 1

By applying the single dispatching rule at all the processing stations, the mean tardiness of different parts produced in the FMS is estimated. The results are given in Table - II.

TABLE II : MEAN TARDINESS OF PARTS WITH SDR

DISPATCHING RULE	FCFS	SPT	SLACK	NINQ
MEAN TARDINESS (min.)	9.82	15.6	2.54	11.3

Among the four rules, dispatching based on SLACK rule gives a minimum mean tardiness.

b) By applying SLACK rule for dispatching parts in FMS, the utilisation of different machines for processing different part types is estimated and the results are shown in Table - III. Table -IV indicates the mean tardiness of different part types processed in the FMS by dispatching it with SLACK rule at all the stations.

TABLE III : BOTTLENECK MACHINE TABLE

Part	RANK					
Type	I	II	III	IV	V	VI
A	M4 (82.6%)	M1 (81.7%)	LU1 (65.2%)	M2 (58.3%)	LU2 (25%)	-
B	M3 (83.5%)	M4 (82.6%)	M1 (81.7%)	LU1 (65.2%)	M2 (58.3%)	LU2 (25%)
C	M3 (83.5%)	M4 (82.6%)	M1 (81.7%)	LU1 (65.2%)	M2 (58.3%)	LU2 (25%)
D	M3 (83.5%)	M1 (81.7%)	LU1 (65.2%)	M2 (58.3%)	LU2 (25%)	-
E	M4 (82.6%)	M1 (81.7%)	LU1 (65.2%)	M2 (58.3%)	LU2 (25%)	-
F	M3 (83.5%)	M4 (82.6%)	M1 (81.7%)	LU1 (65.2%)	M2 (58.3%)	LU2 (25%)

TABLE IV : MEAN TARDINESS OF DIFFERENT PART TYPES

PART TYPE	A	B	C	D	E	F
MEAN TARDINESS (min.)	2.71	2.25	0	1.69	6.5	6.25

c) Among the six different part types, E is the most tardy part type with a mean tardiness of 6.5 minutes. In order to minimise the mean tardiness of parts processed in the system, this particular part type is selected.

d) Among the different machines that process the part type E, machine M4 is utilised highly thus becoming a bottleneck. Hence, different dispatching rules are applied at machine M4 to improve the objective function.

e) Table - V shows the mean tardiness of different parts produced in the FMS applying different dispatching rules at machine M4.

TABLE V : MEAN TARDINESS OF PARTS WITH APPLICATION OF MDR AT BOTTLENECK MACHINE

RULE COMBINATION						MEAN TARDINESS (min.)
LU1	M1	M2	M3	M4	LU2	
SLACK	SLACK	SLACK	SLACK	FCFS	SLACK	2.49
SLACK	SLACK	SLACK	SLACK	SPT	SLACK	12.8
SLACK	SLACK	SLACK	SLACK	SLACK	SLACK	2.54
SLACK	SLACK	SLACK	SLACK	NINQ	SLACK	3.4

Among the different rule combinations obtained, the rule combination with FCFS rule at machine M4 gives the minimum tardiness. This particular rule is selected for dispatching parts at machine M4 to minimise the mean tardiness further. The same process is continued to obtain the dispatching rules to be applied at different bottleneck machines in order to minimise the mean tardiness of different part types to be processed in the system and dispatching rules to be applied at different machines obtained with MDR approach are shown in Table -VI.

TABLE VI : RULE COMBINATIONS AND MEAN TARDINESS WITH MDR ALGORITHM

RULE COMBINATION						MEAN TARDINESS(min.)	NUMBER OF ITERATIONS
LU1	M1	M2	M3	M4	LU2		
1	2	3	3	1	3	1.86	31

1 - FCFS, 2 - SPT, 3 - SLACK.

Step 2

Mean operation tardiness of different part types processed at different machines over a time interval is calculated and the same is shown in Table - VII. Using this table the interval with highest mean operation tardiness (MOT) and corresponding machine are selected to change the dispatching rule to be applied in the particular interval.

TABLE VII : MEAN OPERATION TARDINESS (MOT) OF DIFFERENT PART TYPES IN DIFFERENT TIME INTERVALS

MACHINE	LU1			M1			M2			M3			M4			LU2		
INTERVAL	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
MOT (min.)	0.2	0.2	0.2	9.7	0.4	2.3	0	0	7.0	5.8	6.0	25	28	46	0.6	0	0	0

Step 3

Different dispatching rules such as FCFS, SPT, SLACK and NINQ are applied at machine M4 in II interval. Mean tardiness of different part types obtained by applying different rules in II time interval at M4 is shown in Table - VIII

TABLE VIII : RULE APPLIED IN II TIME INTERVAL AT MACHINE M4 AND MEAN TARDINESS OF DIFFERENT PART TYPES

RULE	FCFS	SPT	SLACK	NINQ
MEAN TARDINESS (min.)	1.86	8.32	1.79	1.83

Application of SLACK dispatching rule in II interval of time at machine M4 yields better results in minimising the mean tardiness of different part types.

This procedure is carried out to select different machines and intervals to change dispatching rules and thereby to minimise the mean tardiness of different part types to be processed in an FMS. Different dispatching rules to be applied over different intervals of time at different processing stations are shown in Table - IX. For the purpose of comparison, the mean tardiness of parts processed in the FMS with different methods of dispatching such as SDR and MDR is also indicated in the same table. The number of iterations required to arrive at the different rule combination with each method is found and the same is also shown in Table -IX For the case study considered, DMDR can reduce the mean tardiness of parts processed in the system to the extent of 10% over the MDR and 34% over SDR approach. It may be noted that the present method gives the final rule combination for the system with four machines and two load/unload stations in 70 iterations. To arrive at the rule combination with the

MDR applied to four machines 31 iterations are needed In case of MDR, the complete enumeration procedure can give the optimal rule combination after 4^6 (4096) iterations. In case of DMDR approach, the complete enumeration procedure needs 4^{18} iterations with three intervals of time. From the above, it can be concluded that near optimal solution can be obtained with lesser number of iterations compared to complete enumeration procedure, with the present heuristic based MDR approach.

TABLE IX : RULE COMBINATIONS AND MEAN TARDINESS WITH DIFFERENT ALGORITHMS

ALGORITHM	MEAN TARDINESS (min.)	NUMBER OF ITERATIONS	RULE COMBINATION					
			LU1	M1	M2	M3	M4	LU2
SDR	2.54	4	3	3	3	3	3	3
MDR	1.86	31	1	2	3	3	1	3
DMDR	1 67	70	2,1,1	1,3,2	3,3,3	3,3,3	1,3,1	3,3,3

1 - FCFS, 2 - SPT, 3 - SLACK.

4. CONCLUSIONS

For scheduling an FMS, an algorithm that changes dispatching rules to be applied at machines over different time intervals to minimise the mean tardiness of parts to be processed in the system is presented. From the case study, it is observed that the present method reduces the mean tardiness of different part types to an extent of 34% over the SDR. Further, it is also noticed that the present method improves the objective by 10% over the results obtained with MDR approach. The search algorithm is efficient in reducing the number of iterations required to obtain the set of dispatching rules to be applied at different machines over different time intervals.

ACKNOWLEDGMENT

The authors express their sincere thanks to the Department of Science & Technology, Government of India for the financial support rendered to carryout this work under the sponsored project.

REFERENCES

1. Blackstone J.H., Philips D.T. and Hogg G.L. "A state of the art survey of dispatching rules for manufacturing job shop operations" International Journal of Production Research V20, no. 1, pp.27-45,1982.
2. Wu S.Y.D. and Wysk R.A. . "An application of discrete event simulation to on-line control and scheduling in flexible manufacturing" International Journal of Production Research V27, no. 9, pp.1603-1623, 1989.
3. Ishi N. and Talavage J.J. "A transient based real-time scheduling algorithm in FMS" The International Journal of Production Research V29, no. 12, pp.2501-2520, 1991.
4. Ishi N. and Talavage J.J. "A Mixed Dispatching Rule Approach in FMS Scheduling" The International Journal of Flexible Manufacturing Systems V6, pp.69-87, 1994.

DEVELOP SILICONE STEEL SHEET CUTTING PATTERNS WITH FILLED-IN BLANKS

Cui Yaodong
Department of Mechanical Engineering
Changzhou Industrial Technology College
Jiangsu, China
Zhu Lian,Jin Aiming,Cao Ming
Shanghai Electrical Machinery Factory
Shanghai, China

ABSTRACT

In July 1996, the authors delivered a new version (Version 2) of the software for developing cutting patterns to Shanghai Electrical Machinery Factory. The factory will begin to use it sometime in 1997 to replace the old version (Version 1) put into use in October 1994. It differs from Version 1 in that it can generate cutting patterns with filled-in blanks. In this paper, the problem of generating cutting patterns with filled-in blanks is briefly described and the mathematical formulation of the problem is presented. The solution to a sample of the factory's cutting stock data in 1995 is given. Finally, to estimate the potential saving on material by using the new version, cutting patterns are generated for all the cutting stock data of the factory in 1995 by Version 1 and Version 2 respectively. The results indicate that a large amount of material can be saved by using Version 2.

KEYWORDS

Cutting Stock Problem, Cutting Patterns with Filled-in Blanks, Computer Aided Design and Manufacture, Manufacture of Electrical Machinery

1. INTRODUCTION

Cutting stock problems have been studied extensively with the development of computer and optimization technology since late fifties [1-2]. In the literature, it is generally assumed that only blanks of the same shape will be allowed to appear in a strip in the design of guillotine cutting patterns [1-4]. For the cutting stock problem of silicone steel sheet in the electrical machinery industry, often two kinds of blanks, circular and sectorial, are arranged in a strip. In October 1994, a software developed by Cui Yaodong to generate silicone steel sheet cutting patterns was put into use in Shanghai Electrical Machinery Factory [5]. Meaningful improvement on material usage has been achieved since then, but a new problem arose as follows:

Before the application of Version 1, a large amount of residual pieces with larger sizes was generated in the cutting process because of larger blank sizes and poor hand generated cutting patterns. Some factories in the industry need smaller blanks to make their products. To reduce material cost, these factories usually purchased residual pieces from Shanghai Electrical Machinery Factory to make their blanks. Because of computer generated cutting patterns, no residual pieces with larger sizes were generated in the cutting process after October 1994. The material costs of these customers are increased because they have to purchase sheets from the market to make their blanks.

On the other hand, material can be better used in computer generated cutting patterns when the number of blank sizes (especially small blank sizes) increases. The factory and the customers will benefit from including both the blanks needed by the factory and the blanks needed by the customers in the same cutting pattern.

The authors have developed a new version of the software (Version 2) witch will be put into use sometime in 1997. It differs from Version 1 mainly in that the portion between larger blanks or between a larger blank and the edges of a strip is allowed to be filled with small blanks. The goal of

Version 2 in developing cutting patterns is to pursuit minimal material cost of blanks.

In this paper, the problem of generating cutting patterns with filled-in blanks is briefly described and the mathematical formulation of the problem is presented. The solution to a sample of the factory's cutting stock data in 1995 is given. Finally, to estimate the potential saving on material by using the new version, cutting patterns are generated for all the cutting stock data of the factory in 1995 by Version 1 and Version 2 respectively. The results indicate that a large amount of material can be saved by using Version 2.

2. THE PROBLEM

Silicone steel sheet is one of the main materials consumed by the electrical machinery industry. Often several thousand or more than ten thousand tons of sheet are needed by a middle sized factory in one year. Since reduction of only one percent in trim losses would mean a saving of a large amount of material, the solution to this cutting stock problem is of considerable realistic significance.

When the diameter size of the stator iron core is larger than one meter, sectorial blanks are used instead of circular blanks to reduce trim losses. Circular blanks are used to make rotor iron cores. So many factories in the industry need circular and sectorial blanks simultaneously. The cutting stock problem of circular and sectorial blanks of Shanghai Electrical Machinery Factory is typical in many factories in the industry.

More than ten thousand tons of silicone sheet is consumed by Shanghai Electrical Machinery Factory to make circular and sectorial blanks. Production runs are organized monthly in the factory. Blanks are of many sizes. For a given month, when the numbers of various types of electrical machinery to be produced in the month are known, the demand for every blank size can be calculated. During the cutting process, first the sheet is cut into strips by shearing machine, then the blanks are punched from the strips. When using Version 1 to develop cutting patterns, only one blank size is included in a strip for most strips. A small number of strips contain one circular blank size and one sectorial blank size, as shown in the second strip in Fig. 1. Some requirements in developing Version 2 are as follows: (1) If necessary, Version 2 should be able to generate the same cutting patterns as Version 1 does. (2) Version 2 should be able to generate cutting patterns with filled-in blanks to make a better use of the strips. (3) Some small circular blanks required by other factories in the industry should be included in the cutting patterns generated.

Some cutting patterns with filled-in blanks are shown in Fig. 1 to Fig. 3. All sectorial blanks and small circular blanks are allowed to be filled into strips relating to larger circular blank sizes in generating cutting patterns with filled-in blanks. In this case the portion between two adjacent large circular blanks and the portion near the left or right edge of the strip can be used more effectively.

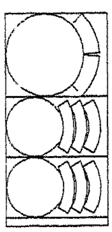


Fig. 1

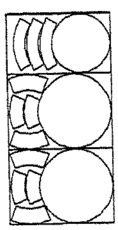


Fig. 2

In developing the software, Sectorial blanks are filled into a strip as follows. (1) The right edge portion of a strip is filled with a row of sectorial blanks of the same size, as shown in the second strip of Fig. 1. (2) The right edge portion of a strip is filled with 1 to 3 sectorial blanks of the same size. The blanks are arranged around the right side of the large circular blank, as shown in the first strip of Fig. 1. (3) Three rows of sectorial blanks are filled into the right edge portion as shown in the first two strips of Fig. 2. The size of the blanks in the middle row can differ from the size of the blanks in the other two rows. Ways in which small circular blanks are filled into a strip can be found in Fig. 3. They are as follows: (1)Fill the right edge portion with 2 circular blanks of the same size.(2)Fill the portion between two adjacent large circular blanks with two small circular blanks of the same size. (3) Fill two small circular blanks of the same size into the corners near the right edge. If possible, fill the middle with one circular blank that may differ from the other two blanks in diameter. (4) Fill more than three circular blanks of the same size into the right edge portion.

3. THE MATHEMATICAL FORMULATION OF THE PROBLEM

The mathematical model used in the software to develop cutting patterns is as follows:

$$Min\ Z = \sum_{j=1}^{m+d} W_j P S_j X_j$$

$$S.t.\quad AX = b$$
$$DX = B$$
$$X \geq 0$$

Where

W_j	=	the value factor of the sheet used for the j'th pattern. It is set to 1 when actual sheet is used and set to W when virtual sheet is used.
W	=	the ratio between the unit area price of the virtual sheet and the unit area price of the actual sheet.
P	=	the unit area price of the sheet.
X	=	$[X_1, X_2,..., X_{m+d}]^T$.
X_j,S_j	=	the number of the pieces of sheet cut according to the j'th cutting pattern, the area of the piece respectively. There are $m+d$ cutting patterns.
m, d	=	the number of internal blank sizes (needed by the factory) and the number of external circular blank sizes (needed by customers) respectively.
b	=	$[b_1, b_2,..., b_m]^T$.
b_i	=	the demand for the i'th internal blanks.
A	=	a matrix with m rows and $m+d$ columns. Its element a_{ij} indicates the number of the i'th internal blanks included in a piece of sheet cut according to the j'th pattern.
B	=	$[B_1, B_2,..., B_d]^T$.
B_i	=	the demand for the i'th external blanks.
D	=	a matrix with d row and $m+d$ columns. Its element d_{ij} indicates the number of the i'th external blanks included in a piece of sheet cut according to the j'th pattern.

In the model, it is assumed that the demands of the customers must be meet, but the factory can purchase small circular blanks (so called virtual sheet) from somewhere and sell to the customers with the same price. Z is the sum of the cost of the actual sheet consumed and the cast of virtual sheet needed. A cutting pattern relating to a virtual sheet is called a virtual cutting pattern that includes only one external blank. W is the ratio between the selling price of the external blanks and the purchasing price of the sheet. The goal function can be understand as follows: the customers pay all the cost of the blanks they needed to the factory. The total payment will be PWS, where S is the total area of external blanks. The factory will include some or all external blanks in its cutting patterns if profitable. The external blanks not included in the cutting patterns will be purchased with price W to meet the demands of the customers. This means a portion of the total payment will be given back to the customers. The cost of this portion is the same as the cost of all virtual sheet in the goal function.

When W is large enough, all external blanks must be cut from actual sheet. When W is 0, no blanks for sale will be included.

The above model is a linear programming model. The solution approach is similar to that of reference 6. The original base matrix includes $m+d$ different cutting patterns and each pattern in it holds only one blank. Actual and virtual pieces of sheet are used for cutting patterns relating to internal and external blanks respectively.

4. SOLUTION TO A SAMPLE

The actual cutting stock data of Shanghai Electrical Machinery Factory in February 1995 is taken as an example to illustrate the power of Version 2. W is set to a vary large positive value to force to include all customer demanded blanks in the cutting patterns. The cutting patterns generated are shown in Fig. 3. Details about the demands of blanks and the description of the cutting patterns are omitted.

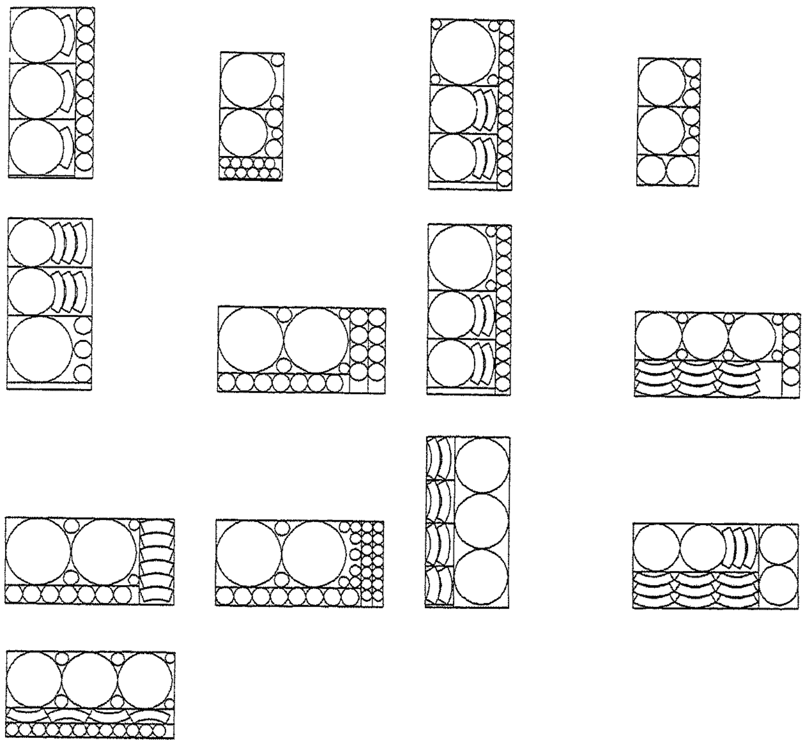


Fig. 3

5. CONCLUSIONS

When Version 2 was finished, cutting patterns were generated for the cutting stock data of each month in 1995 by Version 2 and Version 1 respectively. The computational results are as follows. (1) When only internal blanks are considered in generating cutting patterns, the saving on material can reach 35 tons per year (1.6 percent of the total amount of sheet consumed in one year) by using Version 2. (2) If the blanks of Shanghai Mingjing Electrical Machinery Factory, one of the customers, are included in generating cutting patterns, the expense on purchasing sheet can be compensated through selling small blanks. The saving on material in one year will reach 74.34 tons (3.4 percent of the total amount of sheet consumed in one year). Compared to purchasing sheet from the market directly to make blanks, the customer's saving on material will reach 42.69 tons through punching blanks from Shanghai Electrical Machinery Factory. The total saving on material of the factory and the customer will reach 117 tons!

6. REFERENCES

1. Hinsman, A.I., "The trim loss and assortment problems: A survey ", European Journal of Operational Research, vol. 5, pp. 8-18, 1980.
2. Dyckhoff, H., "A typology of cutting and packing problems", European Journal of Operational Research, vol. 44, pp. 145-159, 1990
3. Dagli, C.H, "An approach to two dimensional cutting stock problems", International Journal of Production Research, vol. 25, pp. 175-190, 1987
4. Hahn, S.G., "On the optimal cutting of defective sheets", Operations Research, vol. 16, pp. 1100-1114, 1968.
5. Cui Yaodong, " Computer aided design of silicone steel sheet cutting patterns in the electrical machinery industry ", Proceedings of the Fourth International Conference on CAD/CG, Shuzi Yang, Ji Zhou and Chenggang Li, Part 2, pp. 985-989, The International Society of Optical Engineering(USA), Wuhan, China, 1995.
6. Gilmore, P.C. and Gomory, R.E., " A linear programming approach to the cutting-stock problem", Operations Research, vol 9, pp. 849-859, 1961

AN INTEGRATED SYSTEM OF CAPP AND JOB SHOP SCHEDULING

Deng Chao Li Peigen Luo Bin
School of Mechanical Science and Engineering
Huazhong University of Science and Technology
Wu Han, Hu Bei, P.R.of China

ABSTRACT

Conventional CAPP and job shop scheduling are independent of each other and result in difficulty of integration. In this paper, an integrated system of CAPP and job shop scheduling is introduced. Integration of process planning and job shop scheduling at task stage is highlighted. The system architecture and characteristics of CAPP and scheduling under integrated environment are proposed. The integrated system with six modules is introduced in detail in this paper. A strategic method for alternative process planning using heuristic search is put forward. A scheduling algorithm based on heuristic rules is given. At last, The application of concurrent concept in the integrated system is discussed.

KEYWORDS

Production Scheduling , CAPP, Integration

1. INTRODUCTION

With the development of computer integrated manufacturing (CIM) in current manufacturing system, computer-aided process planning (CAPP) has been considered as an important part. CAPP serves as the critical link between CAD and CAM. The integration between CAPP and CAD has received remarkable progress, while the integration of process planning with production scheduling has not attained significant results. Although the problems of the integration of process planning and production scheduling have been addressed early in the middle of the 1980s and several current approaches for the integration have been proposed, some aspects of the problems have not been solved completely.

Since process planning and scheduling are carried out separately, and process plans are generated without considering the working status of job shops, some problems arise within the manufacturing environment.

So the process plan containing a linear sequence of operations is not flexible enough. Much research work has been done recently on integrating process planning with scheduling, such as non-linear process planning, alternative process planning, closed-loop process planning, dynamic scheduling, and distributed process planning (or just-in-time process planning)[2].

To achieve an optimal scheduling, it is necessary to consider process planning along with the real workshop status and capacity constrains. One should use a scheduling strategy that can benefit from the non-linear, alternative plan representation. This paper describes a prototype system to integrate process planning and job shop scheduling. The system architecture and characteristics of CAPP and scheduling under integrated environment are highlighted. Approaches for the integration of CAPP and scheduling are also discussed in detail.

2. INTEGRATED SYSTEM ARCHITECTURE

The system architecture is illustrated in Figure 1. It consists of six modules as follows.

(1) Alternative Process Planning Module — to generate alternative plans in terms of generative approach.

(2) Job-Shop Scheduling Module — to generate an optimal or satisfactory schedule which should be feasible for the shop floor control.

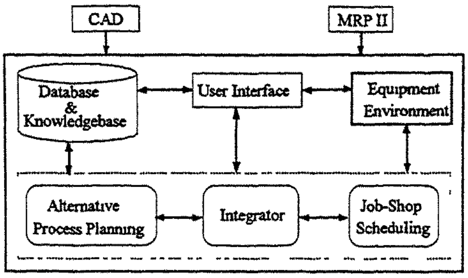


Fig.1: The Structure of Integrated System.

(3) Integrator Module — to be subjected to the constrains from both process planning and job shop scheduling and complete the dispatch of resources (time and machines).

(4) User Interface Module — to provide functions of user interaction and edit operation.

(5) Database & Knowledgebase Module — to record the information and knowledge needed in the system. The database includes the data of the alternative operations, static and dynamic facility status information, alternative machines and schedules. It also includes part feature information and task files from MRPII. The knowledgebase contains knowledge for selecting operations, machines, and tools (including the alternative operations and machine tools, dispatch rules, and criteria as well).

(6) Equipment Environment Module — to record the facility information in the shop floor, including static information and dynamic information.

The equipment environment information plays a key role in the integrated system. This information mainly contains the capability and working status of machines. It affects on process planning and job shop scheduling in the aspects of process route, machine allocation, and job shop efficiency as well. MRPII task information utilized for the system includes the part types, the due dates, and the production volume. The due date is usually considered as an important factor in common scheduling algorithms. The production volume also influences the process planning and scheduling. A process plan contains important information necessary to manufacture a specific part. It is also required by production scheduling. The scheduling information can be simulated and evaluated, and may be taken as feedback information to ask for modification of process plans. The validated scheduling information will finally drive the implementation of shop floor production.

A method for communication between process planning and job shop scheduling through the integrator has been developed. At first, according to the task from MRPII, process planning module carries on process analysis and scheduling module takes task/capability analysis. Then CAPP generates alternative process plans which will be stored in the integrator module. In the integrator, according to some heuristic rules, certain process plans of the parts are picked up for job shop scheduling. After scheduling, if the result is not satisfactory, go back to the integrator to rearrange. Otherwise, the design of process planning and scheduling can be completed.

3. CHARACTERISTICS OF CAPP AND JOB SHOP SCHEDULING UNDER THE INTEGRATED ENVIRONMENT

3.1 Process planning for alternative process plans

In the process point of view, a CAPP should generate more than one satisfactory or near-optimal plans. The main reasons of existence of alternative process plans are as follows.

1) The process route to machine a part is not unique. A part consists of several features, and there may exist several machining sequence of features. Different machining sequences result in different process plans. However, the machining sequences can not be carried out arbitrarily. There are also geometric constraints among the features. For example, if a keyway is on a cylinder, the operation to machine the cylinder should be finished before the operation to machine the keyway. These geometric constraints may effect the machining equences.

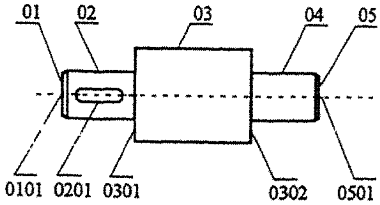


Fig.2: An Example of Shaft.

2) The method to machine a feature is not unique. There may be more than one method to machine a feature. For instance, there are several methods can be selected: planing, milling etc. to machine a flat surface. And the machining method is different, the machining equipment is different. That means a planer or a milling machine can machine the flat surface in the instance. The machining method is different, the machining process is different.

3) There exist some rules for process planning. For examples, the rough and finish operations should be split or joined. The similar operations should be split or joined. Generally, a shaft can be machined from left to right, or from right to left. It is determined by some constraint conditions, such as location and clamping direction. And the NC process principle and the common process principle are different. For example, there is a ladder-shaped shaft to machine. The cylinder features should be machined in proper order from major diameter to minor diameter on the engine machine. However, these features are profiling machined from minor diameter to major diameter on the NC unit.

According to the analysis, the machining sequence of related features may not be unique, and can be described by AND/OR graph. That means the alternative process plans can be specified. For instance, a shaft part is shown in Figure 2. The part features information is based on the shape-feature binary tree[1]. An algorithm for alternative process planning is as the following:

Step 1. Build machining sequence AND/OR graph. Alternative setups are given in form of machining sequence AND/OR graph as shown in Figure 3. Here, a frame-based approach is used to represent the AND/OR graph, as shown in Table 1. Table 1 is a feature relationship table, which is based on the analysis of features and feature relationship[1].

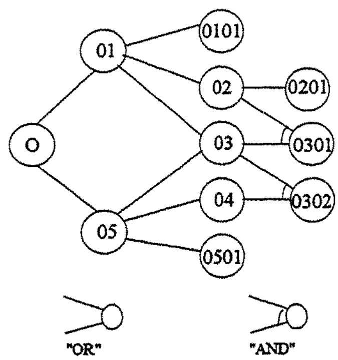


Fig.3: Machining Sequence AND/OR Graph.

Tab.1: Feature relationship.

Feature	Code	Precedence Feature
left-face	01	
center-hole	0101	01
cylinder-face	02	01
keyway	0201	02
cylinder-face	03	01 OR 05
end-face	0301	02 AND 03
end-face	0302	03 AND 04
cylinder-face	04	05
right-face	05	
center-hole	0501	05

The "Precedence Feature" column in the table means the feature in the corresponding row can be manufactured only after the precedent feature(s) in the column have already been manufactured. For instance, the 6th row means feature end-face, whose code is 0301, can be manufactured only after the feature 02 and 03 have been manufactured. While a blank means the feature can be performed instantly. The principle of the change from feature relationship table to machining sequence AND/OR graph is shown as following:

1) A feature code is described as a node. A line between two nodes is an arc.

2) A feature's precedent feature code is the parent node of the node. The two nodes are linked by an arc. If the precedent feature is blank, it is described as a "O" node.

3) The two restrictions between the precedent features are described by the input arcs, as shown in Figure 3.

All the possible machining sequences can be found from the machining sequence AND/OR graph by the heuristic search (depth-first search and breadth-first search) .

Step 2. Build process AND/OR graph. For each machining sequence from step 1, do the following:

1) Find the operation method(s) for each feature according to the list-based operation method table, as shown in Table 2. The "machining method code" column in the table means the feature in the corresponding row can be manufactured by the machining method in the column. For example, the "S2,S3,S4,S5" in row 3, column 2, means feature 02 can be manufactured by machining method S2, S3, S4, or S5. And Table3 is a machining method & its code bilingual table.

2) Change a machining sequence into a process AND/OR graph. For instance, there is a feature machining sequence from Figure 3, which is changed into a process AND/OR graph.

3) Get the initial alternative process plans by the heuristic search (depth-first search and breadth-first search) .

Step 3. Rearrange the operations for each initial alternative process plan, according to the process planning rules that have been posed formerly.

Step 4. Find the feasible alternative process plans. At first, allocate the equipment for each machining method from Table 3. Some process plans are infusible because of the equipment being without or breakdown.

Tab.2: Machining method.

feature code	machining method code
01	S1
0101	S7
02	S2,S3,S4,S5
0201	S6
03	S2
0301	S2,S3,S4,S5
0302	S2,S3,S4,S5
04	S2,S3,S4,S5
05	S1
0501	S7

Tab.3: Machining methods and equipments.

method code	machining method	equipment	
S1	rough turning	lathe	NC unit
S2	rough—finish turning		
S3	rough—semi-finish—finish turning		
S4	rough—semifinush—finish turning—grinding	grinder	
S5	rough—semifinish turning—grinding		
S6	milling	miller	
S7	drilling	driller	
S8	drilling—benchwork	driller	
S9	tapping—benchwork	driller	
···	··· ···		

Step 5. Give priority order of the alternative process plans for each part type by some principles. The priority of Pi is created.

$Pi = \{ P_{i1}, P_{i2},\dots, P_{ik},\cdots,P_{ik_i} \}$;

Where:

- i part index, $i = 1,\cdots, n$;
- k alternative process plan index of part "i", $k = 1,\cdots, k_i$;
- P_{ik} an alternative process plan of part "i".

3.2 Scheduling with alternative process plans

Scheduling a job shop manufacturing system involves concurrent use of equipments and alternative process plans. The scheduling model should accommodate simultaneously use of equipments and the availability of alternative process plans. Since alternative process plans exist, how to schedule with alternative process plans to obtain satisfactory scheduling plans, is the problem of scheduling algorithm. The flow chart of scheduling is shown in Figure 4.

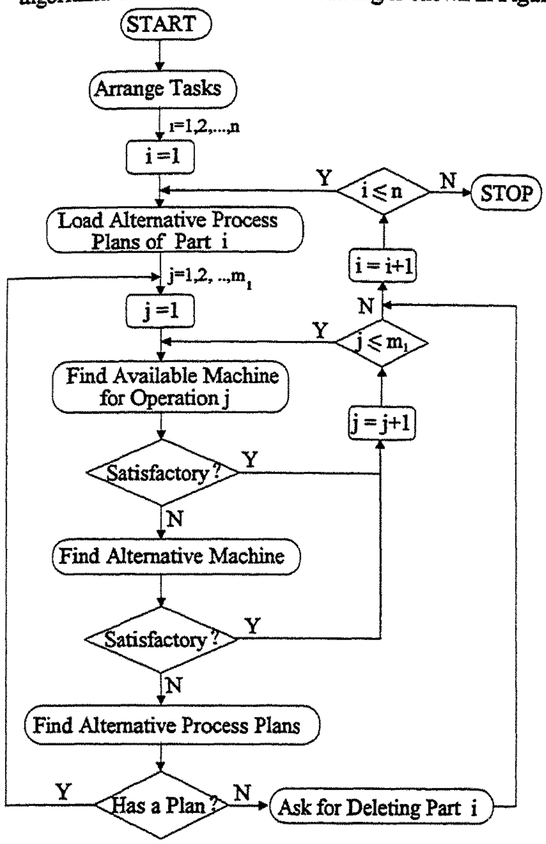


Fig.4: Flow Chart For Scheduling.

Step 1. Arrange task from MRPII in order. Give priority order of them based on some principles, such as FCFS, SPT, EDD, etc. i={1, ..., n}

Step 2. i =1.

Step 3. Load Alternative Process Plans Pi of Part i.

Step 4. Retrieve one process plan P_{ik} with a top priority from Pi.

$P_{ik}=\{ p_{i1}, p_{i2}, ..., p_{ij}, ..., p_{im_i} \}$;

where:

i part index, i = 1,···,n;

j operation index of process plan P_{ik}, j =1,···, m_i;

p_{ij} an operation of process plan P_{ik}.

Step 5. j=1.

Step 6. Find the available machine for operation j.

Step 7. Test the result with some evaluated criteria. If the result meets the criteria and get a satisfactory answer, go to step 8. If not, go to step 9.

Step 8. j=j+1. If j≤mi, go to step 6. If not, go to step 12.

Step 9. Find alternative machine for operation j. If there is an available alternative machine, go to step 8. If not, go to step 10.

Step 10. Delete the P_{ik} from Pi. If Pi is not NULL, go to step 4. Otherwise, go to step 11.

Step 11. Ask the higher authorities to delete the task of part i. Then go to step 12.

Step 12. i = i+1. If i≤n, go to step 3. If not, STOP.

There are several points should been discussed:

1) The evaluated criteria in step 7 are some scheduling rules. For example, the completed time meets the due date, or not; the machine load balanced rate is up to standard, or not.

2) If there is an unavailable machine m, which is the bottle-neck machine, find its alternative machine according to the decision rules for alternative machines In step 9, share some of the load with the alternative machines, which can lighten the burden of the bottle-neck machine. Do the steps again and again, till get the satisfactory result and go to next step.

4. CONCURRENCY IN PROCESS PLANNING AND JOG SHOP SCHEDULING

As stated above, process planning and job shop scheduling are performed separately in traditional production. It is necessary to integrate process planning and scheduling, so as to find out optimal or more feasible solution for them. There are some relationships between process planning and scheduling, which are critical to the integration.

1) Alternative process plans improve the flexibility of manufacturing environment. If the selected process plan has been considered with the due date and shop floor environment, the machining method and equipments will be selected rationally. For example, if the due date is short, the more rapid and more efficient machining way should be selected. Otherwise, the more economical one will be selected.

2) The process analysis is related to the decision in starting time of parts. If a part's technology parameter is too high to perform , or the waste rate is high, it is necessary to schedule the part starting time earlier.

3) One of the main concerns of process planning is to select rational shop floor resources for part fabricating. On the other hand, the purpose of scheduling is to determine operations with starting time, sequencing, and the resource allocation as well within the scope of objective and constraints. The two functions, the process planning and job shop scheduling, have the same goal in this aspect.

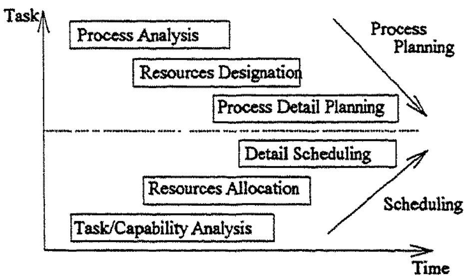


Fig.5: Concurrent Model In The System.

According to the structure of integrated system as illustrated in Figure 1, concurrence concept has been used in the integrated system. After receiving the task from MRPII, process planning module makes process analysis and generates alternative process plans, while scheduling module makes task analysis and resources capability analysis. And then coordination is carried on by the integrator to dispatch the resources for the machining operations, unlike conventional way to separately deal with process planning and scheduling. This concurrence of process planning and scheduling results in more rational resource utilization and production control. It is obvious that the procedure of the integration is actually a concurrent scheme (as illustrated in Figure 5).

5. CONCLUSION

Integration of process planning and scheduling has been recognized as playing an important role to form the integrated manufacturing. In this paper, an integrated prototype system of CAPP and job shop scheduling is introduced. The system architecture and information flow are discussed in detail. The integrating approaches of CAPP and scheduling are demonstrated to capture alternative process plans and resources based on the dynamically changing shop floor environments. Further research work will focus on process analysis and integrating method.

6. REFERENCES

1. Cai Ligang, Li Peigen, Duan Zhengcheng, et al., "An Applied Part Model Based on Form Feature Binary Tree for Integrated CAD/CAP/CAM System of Rotational Components", SPIE ICIM'95, Wuhan, China, 1995.
2. Hong-Chao Zhang, " IPPM — A Prototype to Integrate Process Planning and Job Shop Scheduling Functions", Annals of the CIRP, Vol.42, No.1(513-518), 1993.
3. Luis M. M., Joao J. S., and Carlos F. G., "Production Planning and Scheduling Using a Fuzzy Decision System", IEEE Trans. on Robotics and Automation, Vol.10, No.2(160-167), 1994.

STUDY ON NC AUTOMATICALLY PROGRAMMING FOR CARTRIDGE VALVE BLOCKS

Lin Xie
Department of Mechanical Engineering
Dalian University of Technology
Dalian, 116024, China

ABSTRACT

A CAD/CAM NC automatically programming system for hydraulic cartridge valve blocks--CVBNCP is introduced in this paper. CVBNCP is a main part of CVBCADAM, which is a CAD/CAM integrated software package dealing with the design (CVBCAD), process planning (CVBCAPP), NC programming and NC machining simulation (CVBNCP) for hydraulic cartridge valve blocks. The geometry information of CVBCAD and the process information of CVBCAPP can be translated into machining information by CVBNCP to control the manufacturing of cartridge valve blocks.

In this paper, the functions, the functional modules and the developing approach of CVBNCP are described. Examples are presented to illustrate CVBNCP and its applications.

KEYWORDS

CAD/CAM, NC Programming, Cartridge Valve Block, Machining Simulation

1. INTRODUCTION

Numerically-controlled(NC) machine tools and machine centers are being widely used to support production in manufacturing industry in the past decades. Machine centers are practically useful in the machining for hydraulic valve blocks, especially for hydraulic cartridge valve blocks--key components in hydraulic systems. However, the quality and efficiency of NC programming are not satisfactory and the errors in NC programs made by NC programmers or computers are inevitable. Therefore NC automatically programming and NC machining simulation (NCP) are of great importance in the research area of computer aided manufacturing(CAM), on which significant efforts have been made. During the last thirsty years, NCP methods have developed from the original manual process to the recently introduced CAD/CAM.

2. CAD/CAM DATA EXCHANGE

CAD/CAM data exchange is the transfer of data between dissimilar engineering computer systems, which is also a main aspect which should be considered in the development of software systems. There are two approaches to data transfer, which are described below.

2.1 Database Conversion

Direct database conversion and neutral database conversion are two methods of database conversion.

2.2 Common Database (Product Model)

This is an ideal structure or method of data exchange and integration, which requires the entire CAD/CAM system share a common database--common product model at any stage of the product lifecycle. The use of a common database linking design and manufacturing facilities forms the cornerstone of an integrated CAD/CAM system. Component geometry produced by designers and draughtsmen is used directly in process planning, jig and tool design and NC tape preparation. Other graphic and non-graphic activities can be added to this "base system" to offer an integrated manufacturing system similar to the one shown in Fig. 1.

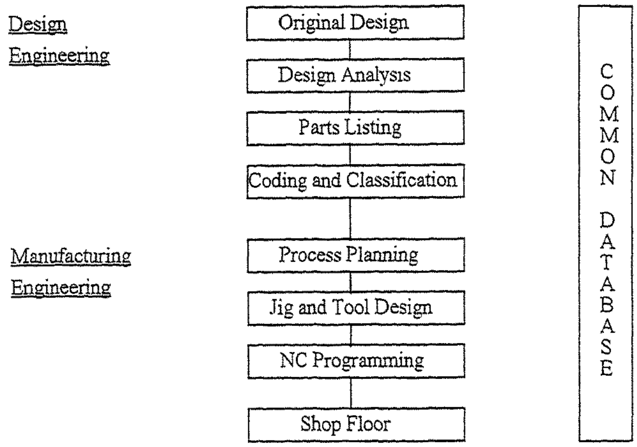


Fig. 1 : The CAD/CAM sequence using a common database

The CAD/CAM process from design to NC programming uses the same order of operations as has been used manually, but by the use of a common database many advantages are to be gained. Each stage in the sequence depends on information from other stages and the rapid retrieval and feedback of data provides a much more efficient engineering system. CVBCADAM uses the structure of common database, every module of which is integrated basing on the product model of common database.

3. STRUCTURE AND INFORMATION FLOWS OF CVBNCP

3.1 Structure of CVBNCP

CVBNCP is a CAD/CAM NC programming system and a main part of CVBCADAM. It consists of module of NC programming(NCP), module of NC machining simulation(SIMU) and module of communication between CVBNCP and machine centers(COM). The structure diagram of CVBNCP is shown in Fig. 2.

3.2 Information Flows of CVBNCP

CAD, CAPP and CAM are three important parts in production automation. There are close information relationship between them. The information to CVBNCP, which is offered from CVBCAD, CVBCAPP and others, is as follows:

1) Geometry information from CVBCAD stored in product model: Size of cartridge valve blocks, data of holes and etc.

2) Process information from CVBCAPP stored in NC process file: Name of working step, tool, speed, feed and etc.

3) Information stored in NC machining database: Standard of NC programming, standard specific to particular machine tools and control combinations.

4. MODULES OF CVBNCP

4.1 Module of NCP

NCP module is a key part of CVBNCP, which is divided into two parts: pre-processor and post-processor.

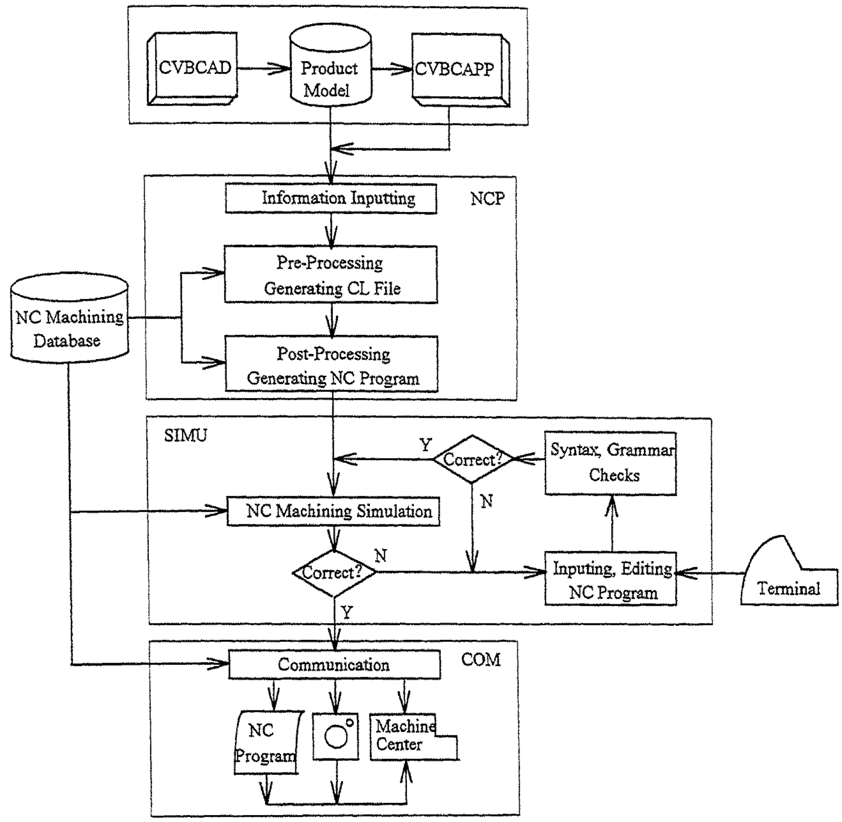


Fig. 2 : Structure diagram of CVBNCP

4.1.1 Pre-Processor

The pre-processor utilizes the geometry information from CVBCAD, NC process information from CVBCAPP, standard in NC machining database and some tool motion and machining parameters which are input interactively, constructs toolpaths and general cutter location (CL) file. Fig. 3 shows the structure diagram of the pre-processor. Table 1 is example cutter location file.

Table 1 : An example of cutter location file

FIRST	GOTO/-80,70,0
PROGRAMNO/%40571	MILLING/-80,70,0 430,70,0
PART/DG40WTKT-405A007	GOTO/430,70,150
DATE/1994.7.28	...
MACHINE/1	CUTTER/T04
PLANE/X-Y	COOLNT/ON
CUTTER/T01	FEDRAT/80
COOLNT/ON	SPINDL/800,CLW
FEDRAT/100	R02=4.0 R03=-3.0 R10=10.0
SPINDL/400,CLW	G57/Z
B=0	FROM/57.5,27.5,150
G54/Q	DRILLING/57.5,27.5,4 57.5,27.5,-3
FROM/-80,70,150	...

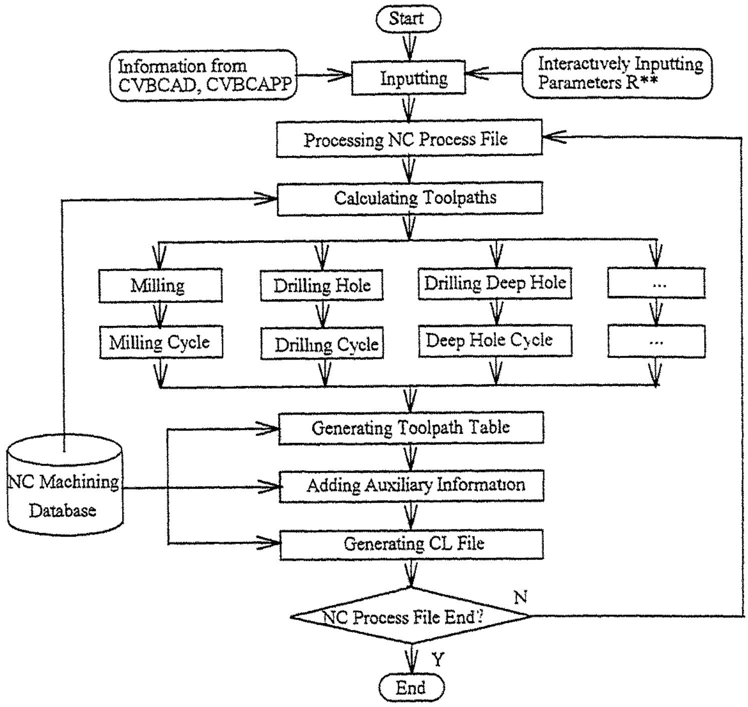


Fig. 3 : Structure diagram of the pre-processor

4.1.2 Post-Processor

The post-processor is a piece of software specific to a particular machine tool and control combination which converts the general cutter location data into the required NC tape. Fig. 4 is the structure diagram of the post-processor. Table 2 and Fig. 5 are examples of NC program and working drawing generated by the post-processor.

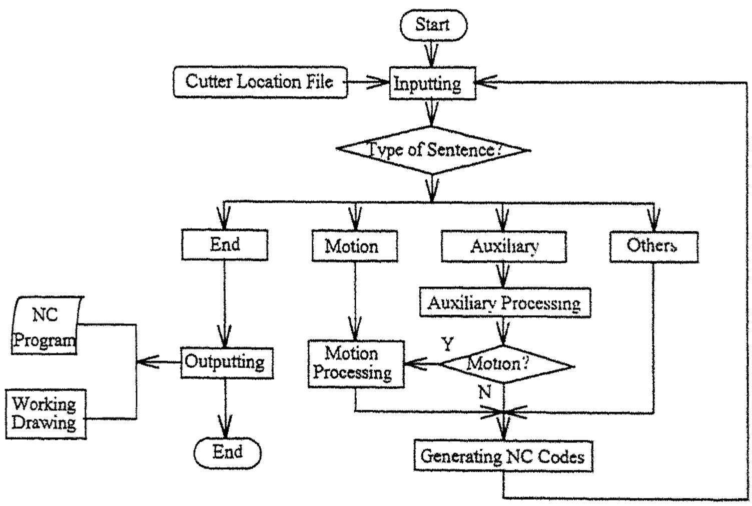


Fig. 4: Structure diagram of the post-processor

Table 2 : An example of NC program

(***)	N0030 G00 Z0
(*************First Installation*************)	N0035 G01 X430.0
(***)	N0040 G00 Z150
%40571	N0045 B90
(Part Name: DG40WTKT-405A007)	..
(Date: 1994.7.28)	(************Drilling Center Holes************)
(Definitions of Parameters R**)	N0120 B270
(R00: Stop Time of Starting Point)	N0125 L300
(R01: First Working Depth)	N0130 G00 G57 X57.5 Y27.5 F80 S800 M03 M08 T08
...	N0135 R02=4.0 R03=-3.0 R10=10.0 L81
(R16: Working Feed)	N0140 X142.5 Y27.5 L81
(R17: Return Speed)	N0145 X150.0 Y70.0 L81
N0005 L301	N0150 X100.0 Y70.0 L81
N0010 B=0	N0155 X50.0 Y70.0 L81
N0015 G17 G54 T01 (Preparing Tool)	N0160 X57.5 Y112.5 L81
(*****************Milling*****************)	N0165 X77.0 Y112.5 L81
N0020 L300 (Changing Tool)	N0170 X142.5 Y112.5 L81
N0025 G00 G54 X-80.0 Y70.0 F100 S400 M03 M08 T04	..

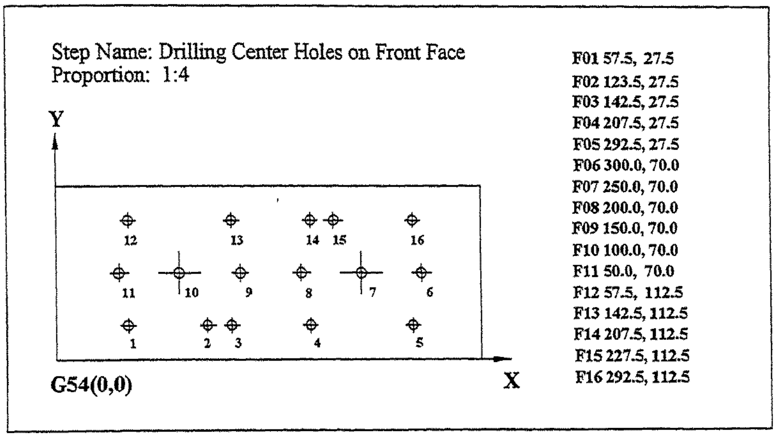


Fig. 5 : An example of working drawing

4.2 Module of SIMU

4.2.1 Importance of NC Machining Simulation

In CVBNCP, a NC program is generated by NCP module, according to the information from CVBCAD, CVBCAPP and etc. How to examine if the NC program is correct? Usually, a programmer punches the NC program on a paper tape and a machine operator feeds the tape into a NC machine center, and then a test machining is carried out on a cartridge valve block, which will take much time and waste material. If a valve block is complicated, the errors of the NC program are usually unavoidable and sometimes the test machining is dangerous and it may be done many times. Dynamic simulation of NC machining of a cartridge valve block on computers can find that if there are errors in a NC program from NCP module or programmers by checking the drawings on a computer screen and data files of the block produced by module of SIMU. Therefore SIMU is a great importance in improving the quality and productivity of NC machine centers by avoiding the damage from NC programming.

4.2.2 Functions of SIMU Module

Module of SIMU is a very useful tool for NC machining of hydraulic cartridge valve blocks and it is of the following functions:

1) According to a NC program, SIMU can vividly display the machining procedure of NC machine centers on a computer screen for cartridge valve blocks and check the errors in the program.

2) On the data area of the screen, several kinds of data information can be shown, including the name of the machined plane of the block, the coordinate values of the machined holes and the related NC codes.

3) Warning when there are errors in a NC program or the cutting force and other machining parameters are larger than demanded.

4.3 Module of COM

In CVBNCP system, a NC program which is generated and checked by module of NCP and SIMU will be transformed to machine center by module of COM to control the manufacturing of cartridge valve blocks. COM module has three functions:

1) Printing NC program for operators.

2) Saving NC program in floppy disks for operators.

3) By means of the interface of computer and the communication agreement, the NC program will be transformed directly from computer to numerically-controlled system of machine center.

5. CONCLUSION

CVBNCP is a CAD/CAM NC programming system, which has improved the NC programming quality, efficiency and standardization. The following benefits are available with CVBNCP:

1) No part geometry definition is required.

2) No manually produced source program is written.

3) No high level language needs to be learnt.

4) By means of CVBNCP, the quality and productivity of NC machine centers have been improved for avoiding the waste caused by errors in NC programs.

5) CVBNCP is also capable of teaching for NC programming and training for NC programmers.

6. ACKNOWLEDGMENTS

The author gratefully expresses her appreciation to the Nature Science Foundation of Liaoning Province for the financial assistance given to the work reported in this paper.

7. REFERENCES

1. Lin Xie, Research and Development on Feature-Based CAD/CAM System for Cartridge Valve Blocks, pp. 88-102. Ph.D. Dissertation of Dalian University of Technology, Dalian, 1995.

2. P. Gu and Kam Chan, "Product Modeling Using STEP", Computer-Aided Design, Volume 27, 3, pp. 163-179, 1995.

3. Lin Xie, Hua Dong, Nenghong Liu and Qinghua Sun, "Research on CAD/CAM for Hydraulic Manifold Blocks", Chinese Journal of Mechanical Engineering, Volume 7, 4, pp. 268-274, 1994.

4. Lin Xie, Hua Dong, Xinan Feng and Wei Xiang, "Research on Feature-based CAD/CAM for Hydraulic Valve Blocks", Proceedings of the 47th National Conference on Fluid Power, Valume II, pp. 359-367, NFPA, Milwaukee, 1996.

GENERATION OF ALTERNATIVE PROCESS PLANS BY NET MODEL

Min-Hyoung Kang, Ji-Hyung Park, and Myon-Woong Park
CAD/CAM Laboratory
Korea Institute of Science and Technology
Korea

ABSTRACT

A process planning system that generates alternative process plans offers multiple process plans for a part, thereby provides the flexibility to cope with the changes in shop floor status. In this paper, we introduce the concept of process net as a model for the generation of alternative process plans. We also show the usefulness of process net model in generating alternative process plans by implementing the developed system to construct process net, and devising an algorithm to generate alternative process plans for rotational parts.

KEYWORDS

Alternative Process Plans, Process Net

1. INTRODUCTION

Conventional process planning strategy assigns just one process plan for a product, but this method lacks flexibility and cannot cope with changes of shop floor status. Therefore, intimate combination of process planning system which provides alternative process plans and scheduling system which considers the current shop floor status.[1][2]

In order to provide alternative process plans, a variety of strategies have been proposed such as NLPP (Non-Linear Process Planning), CLPP (Closed Loop Process Planning), DPP (Distributed Process Planning).[3] An NLPP system generates every possible process plan prior to scheduling. At the moment of scheduling, the system selects a feasible process plan among the process plans made in advance. FLEXAPLAN[4] is an example of NLPP system. NLPP system requires large capacity to store the whole process plans[3]. A CLPP system generates only one process plan per product. If the process plan becomes unfeasible due to the change of shop floor status, the system modifies the process plan or generate another one.[3] A DPP system conducts process planning and scheduling at the same time in order to generate feasible process plan. IPPM (Integrated Process Planning Model) [3][5] and IPPS (Integrated Process Planning Project)[6] are examples of DPP system. CLPP and DPP systems have not been put to practical use because they require extremely high performance hardware for real time process planning and scheduling.

In this paper, we introduce the concept of process net and demonstrate the usefulness of process net by implementing a system for process net generation, generating a process net for a cylindrical part, and finally produce a process plan from the generated process net.

2. PROCESS NET

In early 1990's, Zhang and Kruth proposed a concept which is similar to process net. Zhang constructed graphs which consist of element processes and developed net search method to generate process plans from the graph.[5][6] He also proposed ALPS, a specialized language for graph search.[6] Kruth constructed Petri-net of processes for FLEXAPLAN.[7]

As is shown in Fig. 1, a process net is an AND-OR graph the nodes of which are element processes. Multiple process plans can be extracted from the process net. Fig. 2 shows process plans, in tree form, which can be extracted from the process net in Fig. 1. Those process plans can be used as input data for NLPP or NLPP-like systems.

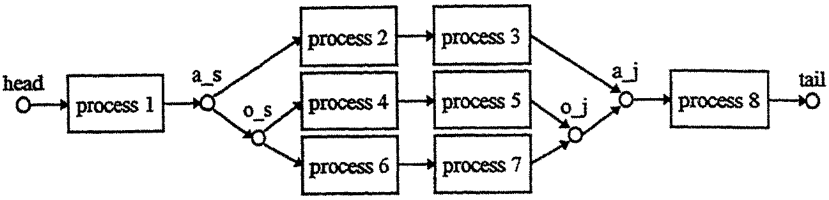


Fig. 1 : An example of a process net.

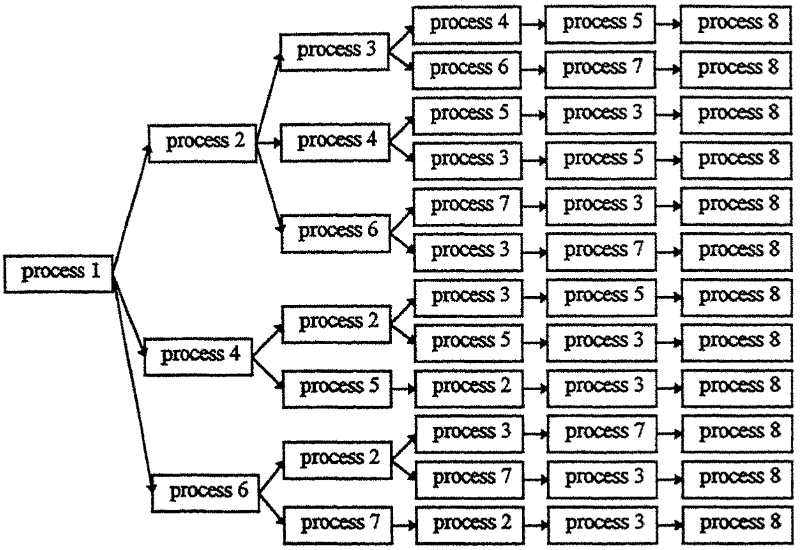


Fig. 2 : Alternative process plans generated from the process net in Fig. 1.

In this paper, we define 'feature process net' for each feature of cylindrical parts. In feature process nets, element processes for the feature are stored in net form. 'Process net' can be constructed by superposition of feature process nets. Assigning machines for the element processes in the process net, the 'machine net' is generated. We also devised an algorithm to search the process nets and the machine nets in order to produce process plans.

As shown in Fig. 1 and Fig. 2, the process net can store multiple process plans in a condensed form. Therefore, the system utilizing process net can save considerable memory capacity in comparison with conventional NLPP systems that should store the whole process plans. Moreover, just like the case of NLPP systems, the system using process net generates process plans prior to the scheduling. Hence, this method doesn't require such performance as is required in CLPP or DPP. With these merits, this method can be put to practical use readily.

3. SYSTEM CONFIGURATION

The implemented system comprises 5 modules: feature input module, process net generation module, machine net generation module, process plan generation module, and graphic output module. First 4 modules conduct, respectively, the four steps of process planning: (1) feature input, (2) process net generation, (3) machine net generation, and (4) process plan generation. Graphic output module is an accessory module which interprets the binary data of process nets and machine nets into net graph drawings. In addition to these modules, the system contains feature process net file and machine data file. In feature process net file, the feature process net of each feature is defined. Machine data file contains machine availability information in the form of available time segments of each machine.

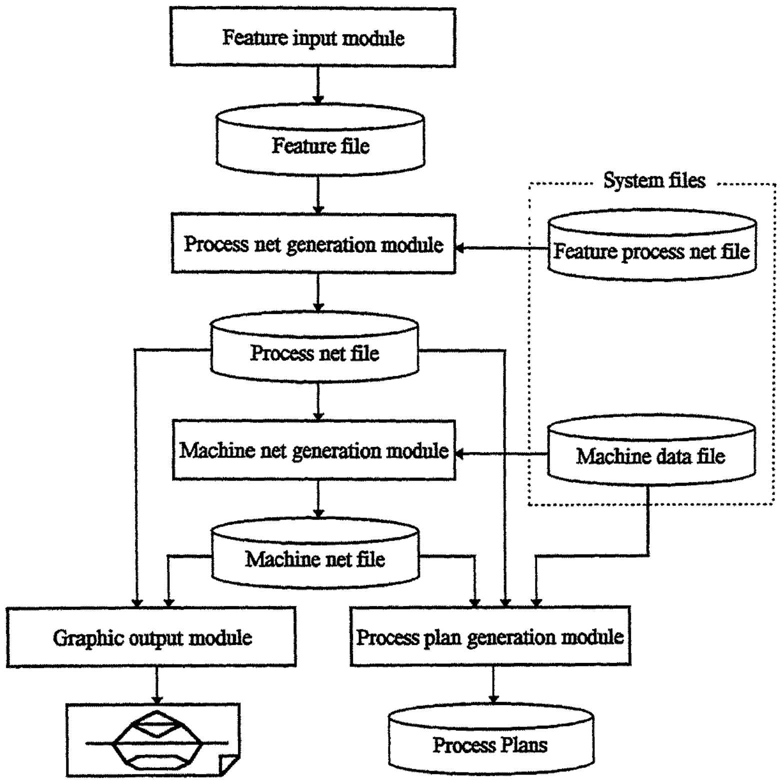


Fig. 3 : Overview of the system.

3.1. Feature Input

With the interactive feature input module, user inputs the feature information of a part: feature names(user selects among the predefined feature names e.g. OUTER_DIAMETER, INNER_THREAD, END_SURFACE, and so forth), dimensions, setup orientations, priorities, and machining times. Input data are stored as feature file.

3.2. Process Net Generation

Feature process file stores the processes of each feature in net form as is shown in Fig. 4. Nodes of the net are classified in 10 types: HEAD, TAIL, PROCESS, AND _SPLIT, AND_JOIN, OR_SPLIT, OR_JOIN, CONDITIONAL_SPLIT, CONDITIONAL_ JOIN, CONDITION.

HEAD : Start node of a feature process net.

TAIL : Terminal node of a feature process net.

PROCESS : This type of nodes represent the processes. In Fig. 4, the PROCESS nodes are expressed as rectangles.

AND_SPLIT/AND_JOIN : They are used in pairs. It connects more than one groups of processes, any group of which can be processed first, and every group of which should be processed.

OR_SPLIT/OR_JOIN : They are used in pairs. It connects more than one groups of processes, and only one group of which should be processed.

CONDITIONAL_SPLIT/CONDITIONAL_JOIN : They are used in pairs. It connects more than one groups of processes, only one group of which should be processed according to the conditions.

CONDITION : First nodes of node groups connected with CONDITIONAL_SPLIT and CONDI-TIONAL_JOIN. They indicate conditions pertaining to dimension, precision, etc.

Fig. 4 shows a trial feature process net for the inner diameter of cylindrical part. If the radius of the feature does not exceed 10, the feature should be processed by drilling, otherwise it can be processed either by turning or drilling and boring.

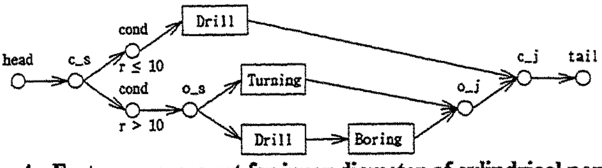


Fig. 4 : Feature process net for inner diameter of cylindrical parts.

The feature process net is retrieved from feature process net file, then all nodes except those coincide with the conditions from the feature file are deleted. For instance, in case of radius > 10, the feature process net of Fig. 4 turns into the net in Fig. 5.

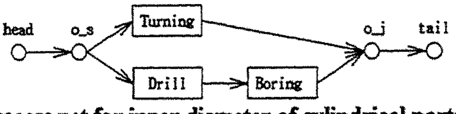


Fig. 5 : Feature process net for inner diameter of cylindrical parts in case of r > 10.

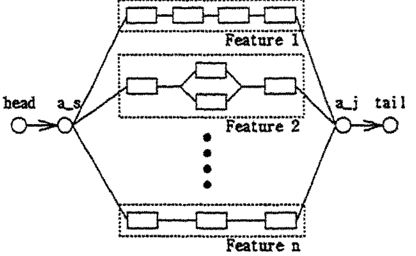


Fig. 6 : Generation of process net by combination of feature process nets.

When feature process net is generated for each feature, all the feature process nets are combined by AND_SPLIT and AND_JOIN as is shown in Fig. 6 to construct the process net of a part.

3.3. Machine Net Generation

Machine net is constructed from the process net file and the machine data file which provides the available time segments of each machine. Machine net generation module assigns available machine to each PROCESS node of process net. The out put of this step is stored as machine net file.

3.4. Process Plan Generation

Process plans are generated from the process net file, the machine net file, and the machine data file. This step is conducted at the moment of scheduling, considering the machine availability information provided by the machine data file.

First, pmatrix(precedence matrix) [P] is constructed as the following rule: in case the i-th feature should be processed after the j-th feature, $P_{ij} = 1$, otherwise $P_{ij} = 0$. When pmatrix is prepared, process plan generation module conducts net search to get the process plans.

In general, quite numerous process plans are extracted from one process net. Therefore, a net search algorithm is required to extract reasonable number of best process plans in reasonable time. The devised algorithm is a variant of depth-first-search algorithm for AND-OR graphs with a heuristic of minimal total process time. When user selects N, the total number of process plans to be generated, process plan generation module makes N process plans which have least total process times.

4. DEMONSTRATIVE EXAMPLE

To evaluate the developed system, process planning was conducted for a simple cylindrical part. This test part has five features as is indicated in Fig. 7. The pmatrix of this part is:

$$[P] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

The process net of the part in Fig. 7 is shown in Fig. 8. Fig. 9 shows the machine net of the corresponding process net in Fig. 8. The total number of process plans to be generated, N, was set to five. The five process plans which have least total process times were generated as are listed in Fig. 10.

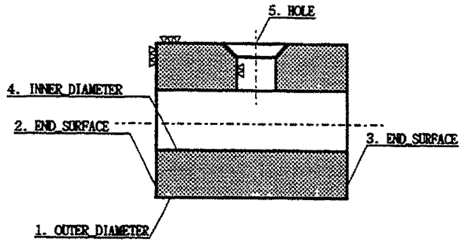


Fig. 7 : An example of a cylindrical part.

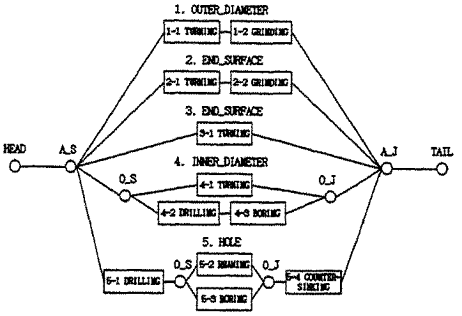


Fig. 8 : Part process net.

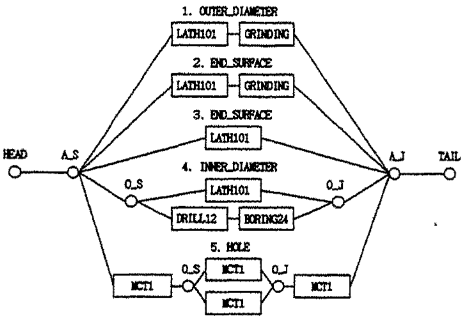


Fig. 9 : Part machine/tool net.

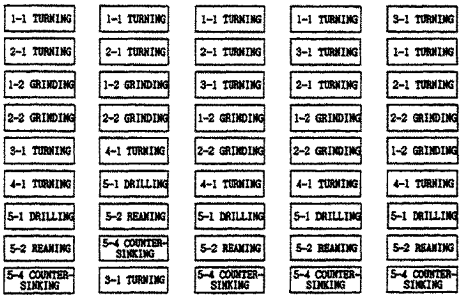


Fig. 10 : Generated process plans.

5. CONCLUSION

In this paper, a process planning system providing alternative process plans with the concept of process net has been presented. Taking advantage of process net, this approach requires less system capacity than NLPP and less hardware performance than CLPP and DPP. Also, a net search algorithm to extract best process plans in a reasonable time was developed.

Further research includes the modification of the process net to deal with prismatic parts. The structure of process net needs improvements as well, to include plastic processing, heat treatment, and so on.

6. ACKNOWLEDGEMENT

This research was conducted partly with the support of the Ministry of Science and Technology(N13610).

7. REFERENCES

1. Hou, T.-H. and Wang, H.-P., "Integration of a CAPP System and an FMS", Computers ind. Engng Vol. 20, No. 2, pp. 231-242 (1991)

2. Lee, M. S., Rho, H. M., and Kang, M. J., "An Evaluation System of Order Acceptability under Consideration of Machine Loading in Die Manufacturing", Annals of CIRP Vol. 44/1 (1995)

3. Zhang, H.-C. and Alting, L., Computerized Manufacturing Process Planning Systems, Chapman & Hall (1994)

4. Detand, J. and Leuben, K. U., "The Generation of Non-Linear Process Plans", Preprints of the 22nd CIRP International Seminar on Manufacturing Systems, Section 2 (1990)

5. Zhang, H.-C., "IPPM-A Prototype to Integrate Process Planning and Job Shop Scheduling Functions", Annals of the CIRP Vol. 42/1, pp 513-518 (1993)

6. Huang, S. H., Mei, J., and Zhang, H.-C., "IPPS: An Integrated Process Planning Project", Proceedings of Autofact, pp 123-128 (1992)

7. Kruth, J. P. and Detand, J., "A CAPP System for Nonlinear Process Plans", Annals of the CIRP Vol. 41/1, pp 489-492 (1992)

TOOLS FOR ASSEMBLY IN A VIRTUAL ENVIRONMENT

R.G. Dewar, J.M. Ritchie, I.D. Carpenter and J.E.L. Simmons
Department of Mechanical and Chemical Engineering,
Heriot-Watt University,
Riccarton, Edinburgh EH14 4AS, UK.

ABSTRACT

Three-dimensional CAD models can be re-used in a virtual reality system so that potential problems with new assemblies of components can be detected, re-design can be undertaken and planning can be started before costly prototypes are even built. Fundamental to this analysis of a new product is the ability to assemble components. This paper describes work undertaken to produce an immersive virtual environment in which manual assembly can be conducted and from which an assembly plan can be automatically generated. This system has now undergone industrial trials. The results from these trials indicate that there is a correlation between real and virtual assembly times and sequences. The fatigue of the test subjects and whether they were right or left handed appear to be significant factors affecting behaviour during virtual sessions.

KEYWORDS

Virtual Reality, Assembly Planning, Manual Assembly

1. INTRODUCTION

Virtual Reality (VR) is an exciting, developing technology with potential applications in a number of industries. In manufacturing engineering, generating the assembly plans for a new design is a lengthy and costly manual process with a direct impact on the rate of product introduction and time to market.

This paper presents an immersive virtual assembly planing system which includes two novel tools to assist users mate components in an intuitive way without the aid of synthetic haptic feedback. The results of industrial trials of the system are also presented.

A completed assembly plan includes a set of instructions describing the sequence of operations, joining methods and materials, tooling, fixturing and relevant quality assurance procedures. Attempts to automate the process through the development of conventional Computer Aided Assembly Planning (CAAP) systems have not met with widespread appeal outside the research community [1]. This lack industrial application is partly due to the computational expense of such systems [2] and partly due to the large burden on users to input data about the assembly before analysis and to sort through large numbers of feasible sequences after the analysis [3]. Furthermore, a great deal of expert knowledge is required for effective CAAP. However, this knowledge has proved difficult to collect [4].

An assembly task in the real world relies on a great deal of touch and a sense of component weight and shape (often called haptic feedback). Predictably, in the virtual world, assembly becomes almost impossible without this haptic feedback, especially when combined with relatively poor images and magnetic tracking. Although little has been written on the application of VR technology to assembly [5, 6], the system described here seeks to utilise alternative sensory cues, since haptic feedback in VR is still a technology in its infancy [7].

2. A VIRTUAL ASSEMBLY PLANNING SYSTEM

The virtual assembly planning system described here uses version 3.1 of Division Ltd.'s dVS/dVISE VR development software on a Hewlett-Packard 725/75 workstation with a PV10 graphics accelerator to enable immersive operation in a head mounted display.

An assembly planning tool needs to capture two major attributes - the assembly sequence and the methods of joining components together. The system described in this paper uses the top level algorithm given in Fig. 1. Here a user selects and moves a virtual component until it has either touched, or is reasonably close to, another virtual object whereupon the user decides if the two objects are to be joined and by what means. This assembly information, namely the sequence and method of joining is stored and an assembly plan produced.

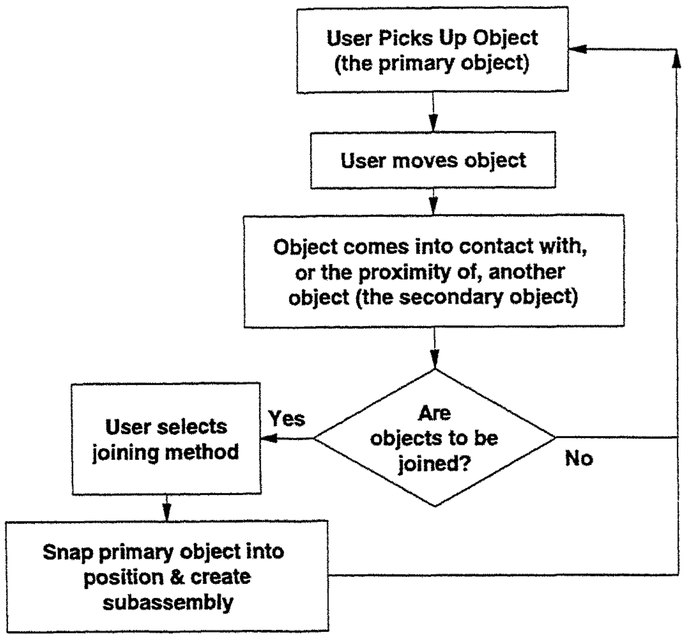


Fig. 1 : Top level assembly algorithm

In this virtual assembly system, two methods are proposed to assist the user in moving an object to its final position relative to another object so that the joining operation can take place. Both joining tools make use of a concept called *snapping*, similar to that employed by Computer-Aided Design (CAD) packages to assist in the exact placement of new vertices in a drawing. The methods here, *Collision Snapping* and *Proximity Snapping*, are based on the assumption that the fully assembled position of each object is already known [8]. This assumption is reasonable since a designer will invariably have assembled the components in a CAD package during the design process and the positional information can be re-used in the VR system.

Snapping is not the only parallel that exists between CAD and virtual joining. In many technical drawings the visual representation of fasteners (screw, rivets, etc.) is not deemed necessary. Instead the fasteners are called up in notes and/or bills of material. To see their geometry would clutter the drawing and add little to the descriptive properties of the representation. The same argument can be applied to VR. As such, fasteners are not graphically represented in the virtual environment. Instead, when a user decides that two objects are to be joined, and subsequently picks the method by which they are to be joined, this method is recorded in the assembly plan.

2.1 Collision Snapping

Collision Snapping is the simplest tool available to assist the assembly planner in joining two objects together in a virtual world. The user picks up a component, known as the *Primary Object*, and moves it towards the component to which it is to be mated, known as the *Secondary Object*. As soon as a collision is detected between the two objects the user is asked, via the toolbox shown in Fig. 2, if the two components are to be joined. Since it may not always be immediately obvious what the primary object has collided with, the secondary object is highlighted by making it flash. The collision is also accompanied by an audio cue from the system, in the form of a simple 'beep', to emphasise the event. If the two are to be joined then the method of joining can be chosen from another toolbox, which includes options such as gluing, welding, bolting, etc.

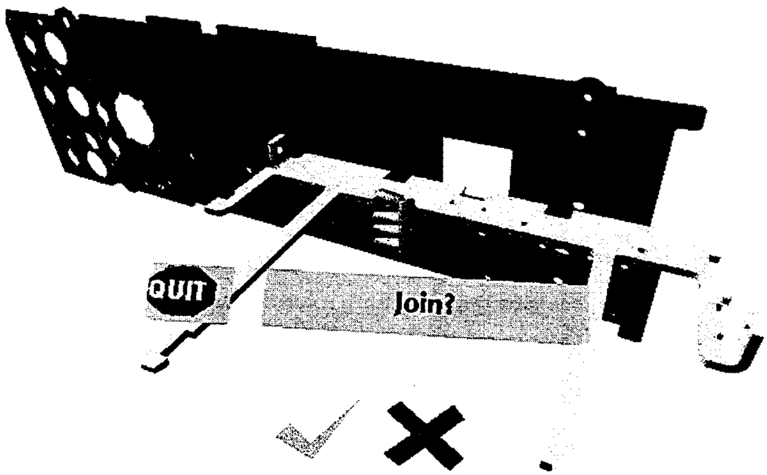


Fig. 2 : Two sheet metal components, a hand and a toolbox in a virtual environment

After the joining method is selected, the collision snapping algorithm repositions and re-orientates the primary object such that its location relative to the secondary object is the same as in the final assembled state.

2.2 Proximity Snapping

Proximity snapping offers a more realistic approach to assembly than collision snapping since it does not call up the toolbox shown in Fig. 2 unless the primary object is sufficiently close to its final position relative to one or more of its neighbours, that is until it is in the *proximity* of its final location. This maps more closely on to the real world. Furthermore, while the user is manipulating the primary object to get it close enough to its final position for joining to take place, a large number of collisions are likely to occur. Although in the real world two solid objects will not readily pass through each other, in the virtual world they will unless some action is taken when a collision is detected. Thus, the algorithm must deal with collisions as they occur to stop the boundaries of objects overlapping. When the primary object is placed in the vicinity of its final assembled position, in the same manner as for collision snapping, it is repositioned and reoriented with respect to the secondary object.

The collisions that occur as a user manoeuvres the primary object into place are dealt with as follows. To prevent object boundaries appearing to overlap, as each new frame is rendered the position and orientation of the primary object is recorded. A set of the last few positions is stored and, when a collision is detected, the primary object is moved back through these previous positions and orientations until there are no longer any overlapping component boundaries.

3. RESULTS OF PILOT TESTS

Pilot tests have recently been completed on a series of virtual assembly environments which make use of the collision snapping algorithm. These trials were intended to test the algorithm's usability and to investigate the relationship between real and virtual assembly in terms of sequences generated and times taken.

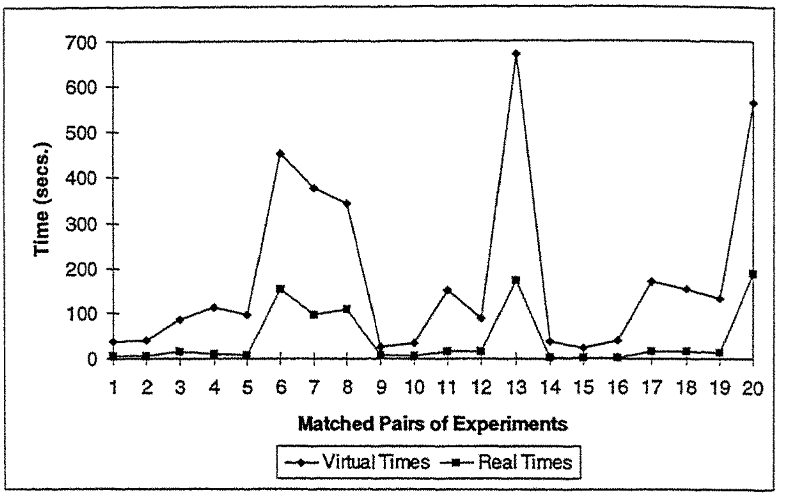


Fig. 3 : Virtual and real assembly times for pilot tests

Subjects from the user company collaborating in this research, NCR (Scotland) Ltd, were given the task of assembling simple components in both the real and virtual worlds (a repeated measures design [9]). The times taken and the sequences generated were compared. The results showed that the virtual assembly tasks took around five times as long as their real equivalents but there was a statistically significant correlation between the two. Fig. 3 clearly shows this relationship. It is predicted that the ratio of virtual to real assembly times will by halved with a streamlined user interface. It was also noted that the subjects suffered from fatigue during their virtual sessions, even though the most complex assembly only had eighteen components.

Finally, some correlation was noted between the real and virtual sequences generated by the right handed subjects. The left handed subject showed a tendency to act as right handed subjects in the real world while becoming wholly left handed in the virtual world.

4. CONCLUSIONS

Affordable VR continues to be unable to provide highly realistic images and haptic feedback. This paper recognises these limitations. Novel tools are presented to assist users of virtual environments re-use their product model database and carryout manual assembly in a meaningful, intuitive way. The ability to assembly virtual components not only allows the assembly process to be planned but also prototypes to be verified.

The approach presented here has been successfully tested in pilot trials. These trails have shown that there is a relationship between real and virtual assembly in terms of the time taken to complete a task and the sequences generated. Future full scale trials intend to investigate these relationship further so that confidence can be obtained that virtual assembly produces the same, or predictably different, results as its real world forebear.

5. ACKNOWLEDGEMENTS

The authors are grateful for the support of NCR (Scotland) Ltd., Division Ltd, and the EPSRC (Grant GR/K41823).

6. REFERENCES

1. Homen De Mello, L.S., and Lee, S., Computer-Aided Mechanical Assembly Planning, Kluwer Academic Publishers, Boston, 1995.

2. Lin A.C., and Chang T.C., "An integrated approach to automated assembly planning for three-dimensional mechanical products", International Journal of Production Research, vol.31, no.5, pp.1201-1227, 1993.

3. De Fazio, T.L., Whitney, D.E., Lui, M-C., Abell, T.E., and Baldwin, D.F., "Aids for the design or choice of assembly sequences", Proceedings IEEE Conference on Systems, Man and Cybernetics, pp.61-70, 1989.

4. Ritchie, J.M., Simmons, J.E.L., Carpenter, I.D. and Dewar, R.G, "Using Virtual Reality for Knowledge Elicitation in a Mechanical Assembly Planning Environment", Proceedings of the 12th Conference of the Irish Manufacturing Committee, De Almeida, S.M. (ed.), pp.1037-1044, University Colledge Cork, Cork, 1995.

5. Connacher, H.I., Jayaram, S. and Lyons, K., "Virtual assembly design environment", Proceedings of Computers in Engineering Conference and the Engineering Database Symposium, pp.875-885, ASME, 1995.

6. Lee, D.E. and Hahn, H.T., "Virtual assembly production analysis of composite aircraft structures", Proceedings of the Computers in Engineering Conference and the Engineering Database Symposium, pp.867-874, ASME, 1995.

7. Taylor, P, "Tactile and kinaesthetic feedback in virtual environments", Transactions of the institute of measurement and control, vol.17, no.5, pp.225-233, 1995.

8. Sato, A. and Maciejewski, A.A., "A Virtual Object Manipulation Interface for Automated Assembly Programming", Proceedings of the IEEE International Conference on Systems, Man & Cybernetics, Vol.2, pp.1826-1831, San Antonio, 1994.

9. Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. and Carey T., Human-Computer Interaction, pp.647, Addison-Wesley, Wokingham, 1994.

EXACT RECOGNITION OF COMPOUND FEATURE BY FEATURE ADJACENCY MATRIX ELIMINATION

Yong Yu, Tie-Chang Zhang
CAD/CAM Research Center
Northwestern Polytechnical University
P. R. China

ABSTRACT

Aiming at the axiom of *design for manufacture (DFM)*, this paper describes a recognition method for abstracting the compound features from a pre-generated part model and discloses the basic mechanism of multi-feature compounding, and also builds the corresponding 2D simulation model to illustrate the mutual spatial relationship among lots of features. The inner association between feature neighboring and feature compounding is deeply discussed and, based on the essential transforming rule of two neighboring features, the corresponding *feature adjacency matrix (FAM)* of multi-feature entities is generated. For the manufacturing feature converted from the pure design feature; an innovative concept of *homogenous compounding* is presented to clarify the architecture of machining domain. Then, the FAM recursive elimination algorithm is developed to determine all compound features and output a series of machining domains in the suitable machining order.

KEYWORDS

Feature Adjacency, Machining Cell Recognition, Feature Modeling

1. INTRODUCTION

Feature-based NC programming strongly bases on the strict definition of machining cell and reasonable process planning and manufacturing planning to properly select tools, sequence machining operations, and determine machining domains as well as generate tool paths and NC programs. In traditional feature recognizing[1]-[3], the following two drawbacks impedes the applying of feature technology into CAD/CAM integration: (1) Purely recognized feature does not contain the manufacturing meaning, and further process planing is needed to determine cutting area as well as cutting chain, and (2) The complexity, suited for almost all recognition methods, of feature group, especially compound feature, is very limited and requires more extensive engineering messages.

The aforementioned drawbacks are resulted mainly from that almost all kinds of traditional feature recognition methods based on BRep or CSG models have adopted the following five similar procedures: Searching low level geometry elements (vertex, edge, loop, face, etc.) from CAD database → Extracting neighboring topological models of edge with edge or face with face → Matching topologic/geometric patterns → Restoring feature models → Completing feature geometric data. Because of the unpredictable topological complexity of features compounding and tremendous number of low level discrete geometrical entities, these five steps have to confront many extreme cases or boundary programs. If only using such straggle quantitative data, it is impossible to deal with many sophisticated situations like feature overlapping or intersecting.

This paper shows the mechanism of feature compounding, builds a general method to determine sophisticated compound feature, and proposes a unified representation of manufacturing cell in the feature technology-based NC system. The feature adjacency matrix elimination algorithm is responsible for judging deeply feature compounding and variant topological types derived from it. For such mutual relationships as embracing, overlapping, direct-neighboring or indirect-neighboring between two features, particularly for poly-compounding among three or above features, this algorithm will exactly abstract the most kinds of topological type of compounding, and the obtained feature groups maintain the comprehensive manufacturing information and semantics which will aid the post process and manufacturing planning of the machining cell.

2. PRELIMINARIES

While a part is described in the feature classes, the two neighbored features maybe contact with each other as sticking, embracing (including half-embracing) or overlapping, which respectively corresponds to one Boolean operator: add, subtract and intersect. Boolean description of product model wholly orients the stage of product designing and can not be adopted in manufacturing. For example, milling of pocket bottom or open-ended face only needs to avoid overcutting the profile of compound feature groups which occurred within the limited boundary of this machining face, but does not care of their specific feature class, e.g. cylinder, cube, *L* shape or *T* shape pocket, even the other user-defined simple feature. For NC programming, only the macro type of the feature in terms of *depression* or *protrusion* will direct effect the tool path calculating and trajectory demarcating.

2.1 Relative Space Relationship of Two Arbitrary Features

Suppose F_i and F_j are two feature entities, then three spatial mutual relationships exist:

- *Isolation:* F_i and F_j do not stick or embrace, or intersect with each other, but it is allowed that F_i or F_j simultaneously sticks, embraces or intersects with other common feature respectively. Any tool feeding for cutting F_i or F_j does not interfere with the other.
- *Attachment:* F_i attaches to only the surface of F_j as a protrusion or depression (it is possible that F_i embraced or overlapped with other feature entity). Profile cutting of F_i feeds along the top or bottom *(drive face)* of F_j, in other words, the *part face* is just the profile surface of F_i self, and its host face is a control variable of avoiding hurting F_j.
- *Compound:* F_i joints F_j, and they share a common attachment host feature, so that F_i and F_j group a compound feature. When F_i is compounded with F_j, and both are protrusions, profile cutting of any one of them certainly destroys the shape of another feature.

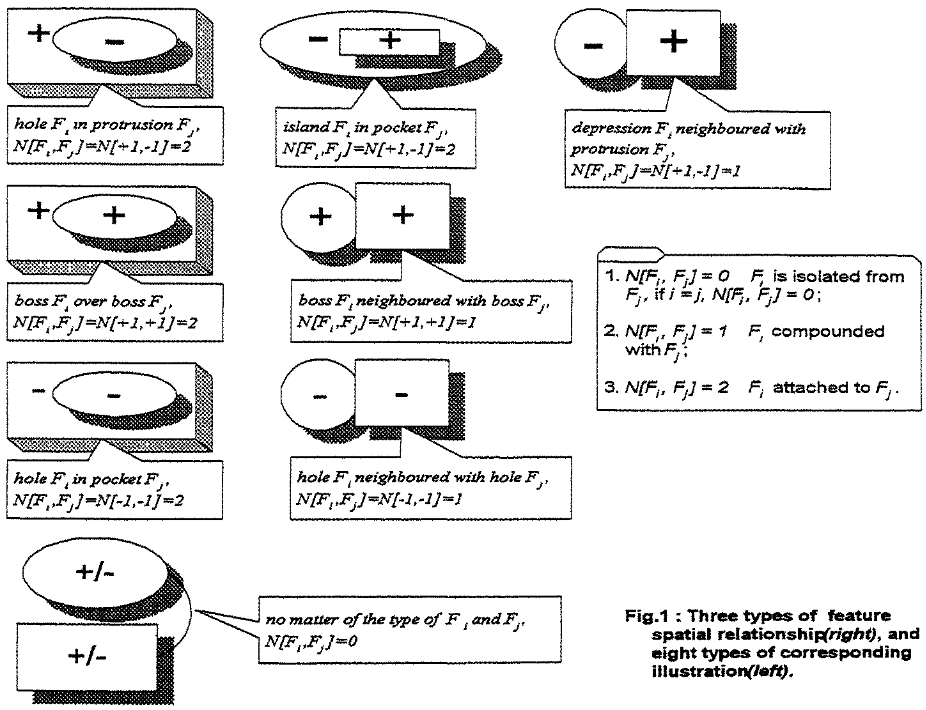


Fig.1 : Three types of feature spatial relationship*(right)*, and eight types of corresponding illustration*(left)*.

According to the above three definitions, eight possible neighboring styles of two arbitrary features can be concluded as *Fig. 1*. We define feature neighboring symbol *N[X, Y]* to represent the neighboring relationship of feature X with feature Y, and $T[X]$ to represent the type of feature X, while '+1' stands for protrusion and '-1' for depression.

Without loss of generality, 2D profile projections of 3D feature entities are used as the illustration to discuss compounding mechanism. Rectangular and ellipse represent F_i and F_j, respectively, denotation '+' stands for protrusion feature, and '–' for depression feature. The corresponding product models to the eight neighboring types are those of islands or holes within pockets, protrusions or depressions in open-ended face, and those of overlapping, etc.

It is very important to understand that, even though F_i and F_j are isolated from each other, the two features are also possible to be compounded together due to the translation of a third common feature with which F_i, F_j neighbored individually. For $N[F_i, F_j] = 2$, if $i < j$, F_j is the son feature of F_i. (i and j are positive integers that represent the sequence of feature modeling.)

2.2 Translation of Feature Compounding Relationship

Multi-features compounding is resulted from the translating of neighboring of two arbitrary features. Translating of neighboring must comply with: (1) Only the features of identical type can be considered, i.e., any compound feature group is either protrusion or depression, and (2) Translation media are only those features that shared one common host face, in other word, feature attaching can not join neighboring translation.

- *Translation of feature neighboring:* Suppose $N[F_i, F_k]=1$, $N[F_k, F_j]=1$,and $T[F_i]=T[F_k]=T[F_j]$ (this assumption embodies the concept of *homogenous compounding* that originated from the axiom of *DFM*), i.e., F_i neighbors with F_k and F_k neighbors with F_j. Logically, F_i is isolated from F_j, but because of the translating of the common feature F_k; F_i, F_j and F_k group together and construct a triple-degree compound feature. In this situation, F_i and F_j should be recognized as indirectly neighboring compound. Given feature entities X and Y meeting $N[X,Y]=0$, $\tilde{N}[X,Y]$ stands for the indirect neighboring relationship between X and Y. In order to maintain the consistent of feature neighboring, it is stipulated that for any two features F_i and F_j meeting $N[F_i, F_j]\neq 0$, $\tilde{N}[F_x, F_y]$ must be 0; or vice versa.

- *Depth of Compound Feature:* For any compound feature X, $D[X]$ is the total number of the feature entities grouped in X. $D[X] = 1$ means that X is a single feature.

- *Recursive definition of feature compounding:* If $\tilde{N}[F_x, F_y]=1$, F_x is indirectly compounded with F_y, there must exist at least one common feature F_z meeting $N[F_x, F_z]=1$, $N[F_z, F_y]=1$. Theoretically, F_z can be a single feature, i.e. $D[F_z]=1$; or can be a compound feature itself, $D[F_z]\neq 1$, which implies the recursive definition of compound feature.

3. FEATURE ADJACENCY MATRIX ELIMINATION

3.1 Symbol Defining & Feature Attribute Matrix

For product model $P = B \pm F_1 \pm F_2 \pm F_3 \pm \Lambda \pm F_n = B + \sum_{i=1}^{n} (\pm F_i)$; $\pm F_i$ denote the feature type corresponding with protrusion or depression (Boolean add or subtract), there totally exist n feature entities. F_i, $i=1,\dots,n$. is the sequence of feature modeling. Marking modeling stock B as F_0 unifies the representation of feature entity. Defining the feature type T symbolized as $T[F_{i, i=0, n}]$ = '+1' or '-1' represents the characteristics of positive or negative of every entity, and creating feature attribute matrix C as follows, practically, stock B (F_0) must be a protrusion.

$$C = [T[F_0], T[F_1], T[F_2], \cdots, T[F_n]] = [+1, \pm 1, \pm 1, \cdots, \pm 1]$$

3.2 Creating Feature Adjacency Matrix

1st. Defining matrix $A_{(n+1)(n+1)}$, set all nodes 0;

2nd. Creating feature adjacency matrix A: for any given $F_{i,\ i=0,\ n}$ judge if it neighbors with the other entities $F_{j,\ j=0,n\ |\ j\neq i}$, and let $A[i, j] = N[F_i, F_j]$, obviously, the current A is a symmetrical matrix;

3rd. Revising A: let $A[i,j] = A[i,j]\times C[j]$. A changed to be non-symmetrical which provides criteria for homogenous compounding searching of the downstream judging.

3.3 Feature Neighboring Translating & Feature Adjacency Matrix Eliminating

Translating of feature neighboring is associated with the compound depth D. For any row i of A, recurring the following steps can get a whole series of compound feature.

1st. Traversing all element of row i, if $T[F_i] = A[i,p]$ (homogeneous feature), let $A[i,p] = A[p,i] = 0$, record F_p as a compound component, then turn to row p of A;

2nd. When all $A[i,j] \neq T[F_i]$, end the searching in current row for compound component. $A[i,j]$ indicates the element of the current row recurred to.

While this cycle of recurring finishes, all features compounded with F_i are obtained.

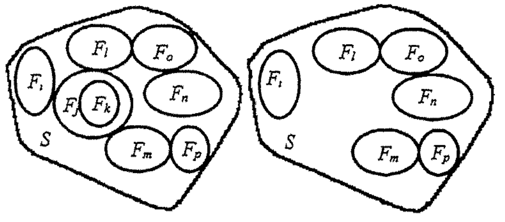


Fig. 2: Translating of feature neighboring relationship
Fig. 3: Feature compounding of homogeneous type

See *Fig.2*, F_i is the compounding basis, assume $N[F_i, F_j] = N[F_j, F_l] = N[F_j, F_m] = 1$, and $N[F_l, F_o] = N[F_m, F_p] = N[F_o, F_n] = 1$. According to *Sec. 2.2*, F_i,F_j, F_l,F_m,F_n,F_o,F_p are compounded together as a whole feature CF_1 and CF_1 attaches to host face S (suppose all entities are protrusions), however $N[F_j, F_k] = 2$, F_k will not join in translating neighboring relationship. Even the host of F_k is shifted to this new compound feature group from the origin alone F_j. From the point of view of machining, this kind of shifting of host face from feature compounding to attaching conforms to the realities of NC programming.

The two obtained compound groups are $CF_1(F_i, F_j, F_l, F_m, F_n, F_o, F_p)$ and $CF_2(F_k)$, also CF_2 attaches to CF_1, $N[CF_1, CF_2]$=2. So called element elimination evinces that the two neighbored FAM node are instantaneously dealt with and set to zero, which simplifies the future calculating and ensures every feature to be the component of only one compound feature, so the most recursive depth is n. For the rest of rows, the same procedures are executed until the base node F_0, meanwhile, the element elimination finished. Every cycle of elimination can determine one compound feature group. Finally, a whole compound feature series $CF_{i,\ i=0,\ ,m\ |\ m\leq n}$ is obtained.

The recursive searching is independent from the modeling methods and sequences, the initial recurring entry node of every cycle is determined in the reverse order of modeling sequence, i.e., initial node of every searching cycle is from F_n to F_0. Thus, the final abstracted compound series complies with the real process and manufacture planning, and every CF_i coincides with one series of tool cutting trajectories.

Also for *Fig. 2*, F_i, $\cdots$, F_p attaches to S, they are more apt to be compounded together. In the real cutting, S is just the machining face, however when $T[F_j] = -1$, i.e., F_j is a hole or pocket, the compound feature series are substituted with: $CF_1(F_i)$, $CF_2(F_l, F_o, F_n)$ and $CF_3(F_m, F_p)$ as three protrusions on S. Positive feature F_k will, although it is located in negative feature F_j, affect the calculating of tool trajectory of CF_1, CF_2 and CF_3, so it is necessary to pay special attention to this kind of isolated protrusion feature entity like F_k to avoid interfering (see *Fig.3*).

4. IMPLEMENTATION

Fig. 4 shows a rather complex compound feature-based part. There are total 13 feature entities, and the corresponding attribute matrix is $C = [1,-1,-1,1,1,1,1,-1,-1,-1,1,-1,-1]$. Only referring to the final part model, without considering the modeling procedures and sequence, can be generated the feature direct-adjacency matrix A, and its revised form (multiplied by C) depicted by *(a)* and *(b)* respectively. Selecting some typical compounding types, we execute the recursive eliminating to demonstrate the detail procedures of the feature adjacency matrix elimination algorithm.

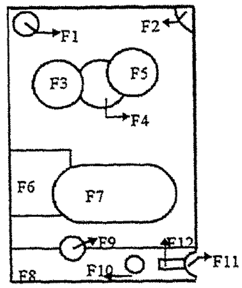


Fig. 4-a: Top view

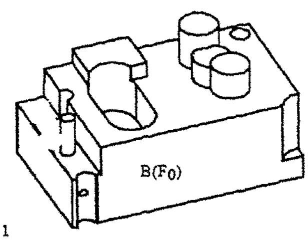


Fig. 4-b: Axonometrical view

Beginning searching with the first non-single node F_{11} in the reverse order of modeling:

Step 1: F_{11} neighbors with F_8, $T[F_{11}] = A[11,8] = -1$, let $A[11,8] = A[8,11]=0$, marking F_8 as a compound element Then, entering into recurring, turning to row F_8 that neighbored with F_9, and both of them are negative types, so assigning $A[8,9]=A[9,8]=0$, coming into row F_9, all elements, at the moment, in this row are 0 and ending examining in this row. Returning to row F_8 that includes a positive feature F_{10}, the attachment relationship of these two nodes does not join compound determining, so F_{10} is bypassed. The next node $A[8,11]$ is already changed to 0, hence we end the searching into the row F_8 searching, and back to row F_{11}. Because F_{11} is the host of F_{12}, this cycle of recursive searching for the whole compound feature involving the entry node F_{11} comes to a halt. We obtain $CF_1(F_{11}, F_8, F_9)$, and the current A is shown as *(c)* F_{10} originally attached to only F_8, but owing to F_8, F_9 and F_{11} compounding together into the negative feature CF_1, F_{10} alters to be an island and attaches to CF_1, which will affect the tool path planning of CF_1. F_{12} also is an inner hole of CF_1.

Step 2: F_6 direct neighbors with F_7, but $T[F_6] \neq T[F_7]$, from the point of view of machining, F_6 and F_7 can be regarded as two independent features, i.e., tool feeding of either cutting outer profile of F_6 or that of inner profile cutting of F_7 will not gnawed another feature. Hence let $A[6, 7] = A[7, 6]=0$, stopping searching of these two rows.

Step 3: F_3, F_4 and F_5 are positive features and construct a whole compound positive feature $CF_2(F_5,F_4,F_3)$. Because positive features compounding will result in profile overlapping, CF_2 can not direct be used to generate its NC tool trajectory until its final hybrid-profile is re-constructed. After the symmetrical eliminating, A changed to *(d)*, the italics stand for all eliminated elements.

Step 4: While all compound feature series are determined, any two features maybe isolate from or attach to each other. Referring to *(d)*, for those left features, still applying symmetrical eliminating, in the reverse order of modeling: (1) If $A[i, j] = \pm 2$, assign $A[i, j] = A[j, i] = 0$, and (2) if $i < j$, then F_j attaches to F_i, on the contrast, F_i attached to F_j. In the final, A degenerated to 0 matrix.

5. CONCLUTION & RELATED TOPICS

Eliminating of FAM and judging of neighboring relationship N for two arbitrary features in the CAD model of a product are based on the feature-oriented exact BRep/CSG hybrid model[4] that is applied mainly into the polyhedron modeling. The recognized compound features coincide with the practical arranging of cutting group and that of cutting order. Every compound feature series attached to machining face is either protrusion or depression, which unifies the representation of machining cells (including machining feature and machining domain) in manufacturing stage. After the depth and type of feature compounding are determined, the corresponding geometry data should be, from BRep/CSG hybrid model, restored and re-constructed according the compound topological type[4]. Especially, the profile re-constructing of the positive compound feature group is more of realism to avoid the possible mutual overcutting. Finally, we should transform this type of compound feature into machining domain adopting the form of either original design data for negative feature or supplementary set of geometry data in the minimum solid hull of positive feature, and provide more minute and exact calculating model for NC programming.

$$A = \begin{array}{c} F_0 \\ F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \\ F_7 \\ F_8 \\ F_9 \\ F_{10} \\ F_{11} \\ F_{12} \end{array} \begin{vmatrix} 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 1 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \end{vmatrix} \text{(a)}$$

$$A = \begin{vmatrix} 0 & -2 & -2 & 2 & 2 & 2 & 2 & -2 & -2 & -2 & 0 & -2 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 \end{vmatrix} \text{(b)}$$

$$A = \begin{vmatrix} 0 & -2 & -2 & 2 & 2 & 2 & 2 & -2 & -2 & -2 & 0 & -2 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 \end{vmatrix} \text{(c)}$$

$$A = \begin{vmatrix} 0 & -2 & -2 & 2 & 2 & 2 & 2 & -2 & -2 & -2 & 0 & -2 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 \end{vmatrix} \text{(d)}$$

6. ACKNOWLEDGEMENTS

The authors would like to thank the anonymous referees of this paper for their laborious evaluating. This research was supported by Special Scientific Fund of Doctoral Subject, P. R. China.

7. REFERENCES

1. Perng, D.B., C.Z. and Li, R.K., "Automatic 3D Machining Feature Extraction from 3D CSG Solid Input". CAD, 22(5): pp.285-295, 1990.
2. Joshi, S. and Chang, T.C., "Graph-Based Heuristics for Recognition of Machined Features from a 3D Solid Model". CAD, 20(2): pp.58-66, 1988.
3. Fields, M.C. and Anderson, D.C., "Fast Feature Extraction for Machining Applications". CAD, 20(11): pp.803-812, 1994.
4. Y. Yu, W. Peng, T. C. Zhang and P. J. Yang, "Reliability of Feature Modeling And its Searching & Restoring", Journal of Northwestern Polytechnical University, (forthcoming).

Towards adaptable product development

Y. Gardan, JP. Jung, Y. Lanuel, S. Leinen, B. Martin,
C. Minich, E. Perrin, C. Poinsignon, L Stemart
LRIM
University of Metz
France

ABSTRACT :

We are developing an experimental system, whose goal is to offer tools to define adaptable CAD applications. We then evaluate some new concepts in man-machine interface and in modelling. Considering the conclusion of our ideas in a system called SACADO [1], we are now implementing a new generation (REGAIN), more complete. In this paper, we give a brief overview of the essential new concepts we are implementing.

KEYWORDS :
CAD, man-machine interface, features, constraints

1. INTRODUCTION

CAD systems are more and more complex, because they try to take into account new tasks, such as : constraints management, simultaneous engineering ... If you can consider that a certain number of basic tools, such as geometric modelling, are operational (even if some problems remain, due, for example to numerical errors and approximations), actual systems lack tools to give the end-user, or at least applications programmers, easy to use tools to develop applications oriented systems. It is probable that future systems will be more adaptable and flexible.

We are developing an experimental system, whose goal is to offer tools to define adaptable CAD applications. We then evaluate some new concepts in man-machine interface and in modelling. Considering the conclusion of our ideas in a system called SACADO [1], we are now implementing a new generation (REGAIN), more complete. In this paper, we give a brief overview of the essential new concepts we are implementing. They deal with :

- man-machine interface : it is an important part, because it gives the basic tools to dialogue with an operational system, but also ways of defining a new tool.
- modelling : it is the heart of CAD systems. We need to dispose of a features based modeller (associated with a solid modeller and its operations) and we try to use the internal semantic to facilitate operations.
- constraints management : this is becoming a basic tool for every CAD system.

2. MAN-MACHINE INTERFACE

In the last few years, the man-machine interface development became a real problem for all the intervening parts in the CAD systems life cycle. Indeed, interfaces aim to be the central part of any CAD systems development kernel. Their complexity increases according to the requirements: graphical, user-friendly, portable ... In particular, since the interfaces are intensively interactive, the dialogue, which manages the communication between the end-users and the system, becomes complex and difficult to develop.

Facing such a development, the trend is to consider an application not as a monolithic one, but as a set of linkable components provided by various mechanisms : OLE Automation, CORBA, Open DOC ... Our approach based on a CAD field seems to be more specialised. From an external point of view, a CAD system can be perceived as manipulations and models. We propose to design dialogue throughout these two components described below.

Manipulations

When an end-user interacts, the system must supply various tools in order to help him in its manipulations. As in CAD systems a lot of manipulations are proposed, a lot of tools appear. We can notice that tools are not based on a standard description. To avoid a combinatorial explosion of manipulation tools specifications, we introduce a unique concept to describe any manipulations : the manipulators. Our approach aims to have a better hold on the interactions of an end-user. By the manipulators, we extend the concept of interaction introduced in the system SACADO [2] which defines an intervention of an end-user. More precisely, to describe a manipulation tool, the manipulator includes two parts : one reserved for definition *domains* and one reserved for *helps* which characterise the manipulation.

We distinguish two definition domains : a domain of use and an application one. By domain of use, we mean the definition of all animated object or end-user that can use the manipulation tool. As for the application domain, it allows the manipulator to know the set of modelled objects or end-users that may be influenced by the manipulation. In order to include contextual manipulations, each domain can receive conditions on its components, what makes the manipulator more flexible.

Helps that compose the description of a manipulator represent all the appearances of the manipulation tool and all the mechanisms linked to the processing cycle of the tool. Various appearances can occur during a manipulation : it may be interaction styles (menus, dialogue boxes ...) or modelled objects dedicated to manipulations (lever ...). Moreover, the mechanisms defined allow the system to manage complex manipulations that can get more relevant information, in particular deduced ones.

With manipulators, the system centralises knowledge about helps mode that it can provide. As a result, it's easier for the system to control the manipulations of end-users and to propose at one time, only adapted manipulations. To realise this, we introduce a coherence control mechanism based on compatibilities applied on manipulators according to a goal seek for an end-user. A compatibility is contextual and conditioned, and expresses the impact of using the manipulator on the goal sought for. The system can detect three compatibilities : *local,* when the manipulator allows to get closer to the goal ; *immediate,* when the manipulator suspends a goal ; *deferred,* when the manipulator changes the goal. The system may then detect that the end-user is not coherent in its manipulations and proposes helps. There is not only one kind of manipulator appearance from which the system can deduce a compatibility : for example, from the use of a menu, the system can deduce a compatibility, but the system can also determine one from the use of a lever.

Thanks to the concept of manipulators, the system increases its control on end-users. Moreover, the modular decomposition in the definition of a manipulator provides so much flexibility that many manipulation tools can be defined ; we can even introduce new manipulators without disrupting the system and we can distinguish dialogue proper to manipulation from dialogue dedicated to application models update. Besides, defining adapted tools with manipulators allows the system to manage multi-users environments.

Generic model

The description of a manipulator is based on knowledge of objects and end-users. In CAD systems, models are created to get different kinds of knowledges within different representations. According to this philosophy, a generic model is defined as the base of other models. It describes components in an external way, throughout their possibilities called behaviours. Components are extended to all what contains knowledge useful to modelling i.e. application objects (cube ...), properties (tangency ...) and even end-users. It seems to be natural to model end-users not only because they can be instigators of modelling but also because multi-users systems expand.

Each component of the generic model is described according to the same approach : a couple composed by a canonical representation and behaviours. Three behaviours complete the description in order to get more knowledge : *reactors, transmutors* and *productors.*

A reactor contains knowledge to interpret the component into datum for a given context. It becomes extrinsic when the datum is dedicated to animated objects or end-users. It aims to supply them helps. So, this datum is used as a link between the component and manipulators in order to manage these helps. The reactor is assumed to be intrinsic when its datum is dedicated to more static components (properties ...). In this case, the knowledge induced by the reactor provides semantic views required by some computational parts.

A transmutor includes knowledge about links between components. According to the canonical representation of a component, it describes the evolution towards a new component. As a result, the transmutor maintains the coherence of models.

A productor represents all creation modes beyond the canonical representation. The association with others behaviours guarantees interactive dialogues like those encountered in CAD systems (helps, anticipation ...).

Introduced behaviours give a suitable basis for an independent description with respect to the implementation. In addition, they can encapsulate more specialised knowledge like manufacturing one [3].

Synthesis

In view of the growing complexity of the graphic man-machine interfaces (amount of code, difficulty to design and maintain ...), we uphold the idea of an open architecture to break the monolithic aspect of classical systems. This point of view is particularly right in CAD applications. We propose a distribution of knowledge around a kernel. The design is then distributed among two concepts : the generic model and the manipulators. Their description is close to a declarative approach so that the dialogue is updated according to the generic model components and the manipulators described.

3. THE *REGAIN* FORM FEATURES MODELLER

3.1. The modeller

We call form features all the physical components of a part. There are primary features, which are primitive volumes such as blocks or spheres, and secondary features, which locally modify the shape of a part, such as slots or drillings. The REGAIN features based modeller [4] is designed in order to allow an efficient and intuitive handling of form features. We have three major goals.

We first aim to provide a model with a complete design history, including the parameters of all the final part components, their constraints and their nature (primary or secondary feature, such as blocks and slots). The model we developed is a two levels structure : the first level is called the semantical level ; it contains the whole design history. The second level is a B-rep evaluation, tightly linked by bi-directional links to the semantical level to warrant their consistency. As form features are stored in an explicit way, with all their parameters and constraints, they can be very easily modified (in a dialogue point of view) or accessed (see 3.2. for their use in the visualisation tool).

We also want the user to place the features and to give their dimensions in an intuitive and flexible way. The dialogue we propose is essentially based on constraints : instead of giving the position and orientation of the local coordinate system of the feature, the user gives constraints such as : "this face of the feature must be parallel to this face of the part". The local coordinate system can then be automatically computed and the constraints stored for further updates. Moreover, this technique does not depend on the feature's shape. It can be applied to every form feature, in particular to the ones that were not pre-defined in the CAD system library and were created by the user (see next paragraph). Another dialogue facility is given by the automatic extension behaviour of secondary form features. This is detailed with the U-shaped through slot. Traditionally, this object is modelled in the generic form feature library as a block to be subtracted to the part. The block has three parameters and when the slot is instanced, the user has to place it and to give its width, length and height. In our software, the user just has to place the feature (with constraints) and to give the width. As the slot automatically grows to the "front", to the "rear" and "upwards" until it gets out of the material, it is no longer necessary to give the last two parameters. This brings two interesting advantages : user interventions are less numerous and it is possible to insert slots whose height and length are not constant, which is impossible if these two quantities must be provided by the user. The automatic extension behaviour is achieved by describing the generic secondary feature as a half space instead of a volume. For example, the generic slot is modelled by a U-shaped half space, which is infinitely long and high. To compute the exact amount of material to remove from the part (or to add if the feature adds material), the intersection (or the difference) of the part and the half space is first computed. One of the resulting volumes is chosen and called the *useful volume*. The useful volume is then subtracted from (or added to) the part. Note that this involves Boolean operations on volumes with adjacent faces, that's why these operations have been carefully studied (see 3.2.). Also note that, when a form feature parameter changes, it is often necessary to delete the form feature and to re-instance it with the modified value (other solutions - variational geometry and restart design from beginning - are not satisfying). In that context, the knowledge of the useful volumes makes the feature's deletion very easy : it is only necessary to recombine the part and the useful volume with the opposite operator.

The third goal of the REGAIN form feature modeller is to offer the end-user a tool to define his own generic form features and to add them into the form feature library, at the same level as the pre-defined generic features. As the latter are described by half spaces, this means that the tool must allow the user to describe a half space. This is easily achieved by using the interactive environment of the

CAD system : the user first designs a volume corresponding to the feature. This means, for example, that he builds a podium like volume for the T-slot. To tell the software in which directions the volume must be extended to become a half space, he just shows the faces to remove. For the T-slot, he shows the front, back and top faces. During the half space construction, the system spies the user and when this stage is over, the shape is recorded and added to the generic features library. Of course, other methods can be proposed to compute the half space, specially for simple shapes ; for instance, we use a sweep operator for surfaces that can be generated with a 2D profile.

Design by feature is a first step towards a real integration of semantical information in CAD/CAM systems : it provides the end-user with a familiar language but it also gives the software developer and the application developer a powerful way to associate a behaviour to objects. Hence, features appear as a way to improve the interactivity and functionalities of CAD systems. In the current version of our system, the behaviour of a feature is limited to the shape (with its geometric constraints and some validity rules) and its positioning method. It will be completed in such a way that a feature becomes able to detect its instances that appeared without explicit instanciation (due to Boolean operations for example), to give interpretations of itself according to the application (viewpoint notion) or to handle its semantical modifications (i.e. a through hole becomes a blind hole). If the features intelligence keeps on growing, it will even be possible that a feature determines itself ; the CAD softwares will then really allow a functional design.

3.2. Adapted operations

If you have a features based modeller at your disposal, you need to transform its semantic in order to define the associated solid modeller or to visualise it. We tried to implement Boolean operators well adapted to our approach (basing them on faces) and we are also aiming to describe adapted visualisation algorithms.

When using features, such as described in 3.1., it is easy to understand that they can be based on faces to faces Boolean operations. Moreover, 3D Boolean operations between solids with boundary representation generally suffer from a lack of calculation precision, a lack of robustness and computational times that grow with particular cases.

Our algorithm for 3D Boolean operations is based on 2D Boolean operations which reduce complexity. It is divided into two steps. First, each face of the two objects is processed in order to define which parts of it belong to the final object. This is made by a 2D Boolean operation between the face and a section of the other object computed from the face which is processed. The selected sub-faces are then joined under B-Rep formalism to obtain the final object.

All the steps of this algorithm have been rigorously formalised [5]. The formalization allows us to avoid too many particular cases and it has been shown there was finally only one. This particular case happens when faces of the first object are coplanar with faces of the second one. Calculations are rejected until it isn't possible to decide with topological information.

The main interests of this algorithm are the formalization in order to obtain a robust algorithm, the processing of each face taking into account dangling faces and a uniform algorithm even if the type of objects changes (polyhedral, non-planar faces and non-manifold).

In the framework of the REGAIN system, one wishes to be able to visualise an object without using a geometrical approximation and, possibly, with realism. Form features are high-level information containing important notions of coherence. A study under the visualisation point of view, allows on the one hand an optimised evaluation of the object geometry, and, on the other hand, the implementation of particular algorithms. For example, a blind hole in a plane will be visible only if the drilling direction and the view direction have the same orientation. If it is not visible, its geometrical evaluation will not be necessary. The visualisation of the object will then be faster.

From this principle, the accurate display of an object modelled by form features comes not down to implement one single algorithm using a single low-level geometric model (B-Rep or CSG). It is an operation composed of a set of relevant means to process the display at best from realism and efficiency view points. Consequently, an enhancement of the form features library would have to be logically associated with an expansion of these display means.

4. Constraints management

4.1. Variational geometry and variational design

Variational design can be defined as a methodology which is based on graph theory and which uses robust techniques for numerical solutions in order to provide a constraints-driven system. This system also manages both geometric and engineering constraints [6]. The difference between variational systems and other systems (parametric for instance) depends on the kind of satisfaction constraints algorithms used by these systems [7]. Variational systems solve constraints by constructing a system of equations representing the constraints, and solving all constraints of the system simultaneously on the basis of a numerical equation-solving procedure or some equivalent method.

Variational systems can generally realise the necessary changes as soon as a parameter has been modified or as soon as a constraint has been added in order to get a new consistent solution. Moreover, as the designer may decide at any moment to reformulate his problem and as these changes may be done easily, the end-user is encouraged to test every design approach.

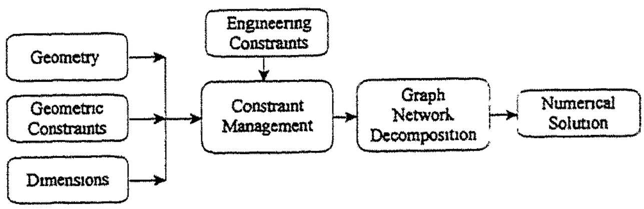


Figure 1 : *Schematic diagram of a variational design system*

Additional characteristics make a variational scheme attractive :

- most intentions are modelled by the geometric constraints, the engineering constraints and the links between the parameters. This allows other users to modify a model and preserve, after all, the intentions of the designer ;
- a variational model contains a complete mathematical description of a design. The tolerances analysis, the mechanisms analysis and the design optimisation can therefore be taken into account.

4.2. Models comparison

The variational design corresponds to the criteria we can expect from a CAD system. Therefore, we developed such a system, in integrating a graph to represent the different elements (the model) and the relations between these elements (the constraints).

So, the search for similarities on common pieces of different parts consists in the comparison of their representation graphs and more precisely in the search for equality of graphs or sub-graphs.

In order to prove the equality or the inclusion of part components, the only correspondence of graphs is not always possible. In particular, this is the case when two design graphs representing the same part have been built in a different way. A complete normalised graph maintain all the possible constraints -the ones explicitly requested by the end-user and the ones deduced by the system-. This graph represents a part in a unique manner.

However, this complete graph can be reduced because of the large number of redundant information. The main idea is to define a set of addition/replacement rules in order to get a unique limited normalised graph. This one has to be minimal and must represent a given situation in a unique way.

So, a sequence of constraints sets will be built and, from the convergence of this sequence, the ending of the normalisation process and the unicity of the solution can be proved.

The comparison of two parts is equivalent to the comparison of the models representing these parts. The problem of the equality of two graphs G1 and G2 (figure 2) may be represented by CSP (Constraint Satisfaction Problems). For each node in the CSP, we define a domain of possible values ; at the initial step, this set corresponds to the nodes of the second graph. An edge between two nodes exists if a binary constraint (perpendicular, parallel for instance) exists between these nodes ; on each arc of the CSP, the same constraints of the second graph are marked.

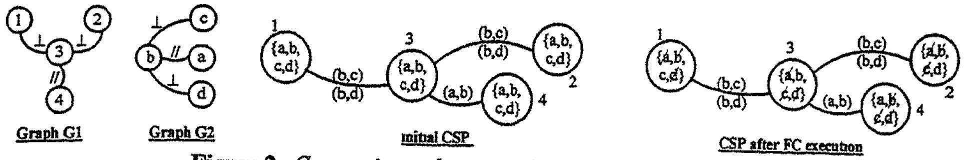


Figure 2. *Comparison of two graphs with CSP representation*

During the CSP-resolution process, the equality of two graphs will be proved if a different value can be assigned to each variable of the graph. On the other hand, if only a subset of the variables can be assigned, the sub-graph represented by these nodes is included in the second graph.

The modelling of graphs equality or graphs inclusion by CSP leads to a unique and concise representation of the problem ; moreover, robust and efficient algorithms can be used for its resolution. Among these algorithms, the *forward-checking* can be noticed because of its good efficiency compared to the cost. This is a prospective method whose aim is to fail when detecting inconsistency in the research tree as soon as possible.

5. CONCLUSION

This paper has briefly presented three basic elements of our approach. We have emphasised on some original proposals. We think that our man-machine approach is promising, giving us concepts facilitating the development of CAD systems and providing us with tools to extend the scope of the dialogue tools. Our features based modeller and the associated operators have to be considered as a basis for further implementation. The idea of comparing models should become a kernel within re-using.

6. REFERENCES

1 Y. Gardan. "Defining CAD systems adaptable to design process", ETFA'95 INRIA/IEEE proceedings conference on emerging technologies and factory automation, Paris, october, 1993.

2 Y. Gardan, J.P. Jung and B. Martin. "An end-user oriented approach to design man-machine interface for CAD/CAM". Proceedings of IEEE International Conference on Systems, Man -and Cybernetics, France, Le Touquet, pages 525-530, 17-20 october, 1993.

3 Y. Gardan, B. Martin and I. Stemart. "Behaviors : from man-machine interface to design for manufacturing". Proceedings of CESA'96 IMACS Multiconference, Computational Engineering in Systems Applications, symposium on Discrete Events and Manufacturing Systems, IEEE-SMC, P. Borne, J. C. Gentina, E. Craye and S. El Khattabi editors, France, Lille, pages 530-535, 9-12 july, 1996.

4 Y. Gardan, C. Minich, C. Poinsignon - "Proposals for a product model", in Selected papers of IDMME'96, Nantes, France, April 15-17, 1996 P. Chedmail, J.C. Bocquet, D Dornfeld (eds), Kluwer academic publishers

5 Y. Gardan, E. Perrin - "An algorithm reducing 3D Boolean operations to a 2D problem : concepts and results" - Computer Aided Design - Vol 29 - n° 4 - pp 277/287 - 1996

6 Shah J.J., Mantyla M., Parametric and feature-based CAD/CAM, John Wiley & Sons, Inc., 1995.

7 Chung J.C.H, Schussel M.D., Technical evaluation of variational and parametric design, Computers in Engineerig 1990, p.289-298.

AUTOMATIC SELECTION OF DIE STRUCTURE BASED ON PRODUCT FEATURES OF DEEP-DRAWN COMPONENTS

W.M. Sing and K.P. Rao
Department of Manufacturing Engineering & Engineering Management
City University of Hong Kong, Hong Kong

ABSTARCT

Recently, emphasis has been placed on developing intelligent computer-aided processes for determining die structure and provide detailed designs in order to realise reduction in design time, increase in productivity and consistent designs without relying too much on human expert. In the proposed computer-aided system for deep drawing, the required finished product shape and a set of highest feasible forming process parameters would be used as input, and selection of most suitable die structure is the output. According to the type of die structure and the product specification, full details of the other die components can then be determined, such as the size and length of screws, dowels, kicker pins, knockout pin, lifter pin and stopper pin, punch size and length and, die size and thickness.

KEYWORDS

Deep Drawing, Die Structure, Computer-aided Die Design, Decision Tables, Group Technology

1. INTRODUCTION

Presswork is an economical production method frequently used for mass production of discrete cylindrical shells. The quality, production cost, delivery time, production rate, maintenance cost and accuracy of a component is affected by the die structure employed in producing that component. A shell can be produced in a single drawing operation, a number of operations (transfer or progressive die), or a combination of two operations in a single stage (combination or compound die). The die can be constructed following conventional or inverted die structure with a movable blank holder or a fixed stock guide. A knock-out device can also be incorporated. Though numerous combinations are possible, the actual design depends on the experience of the designer and the facilities available in the workshop. There are several standard die structures, and the determinant factors for the selection of a proper die structure capable of drawing a product economically and reliably are the features of the drawn cup. Mass production and extensive use of deep drawing process for axisymmetrical sheet metal products has made it possible to use standard structures as well as standard components.

Useful technological data and heuristic rules on die structure design are readily available in the published technical papers and die design handbooks as well as through developed in-house experience. Numerous techniques can be used to computerise such design knowledge. Decision tables can accommodate rules as well as technological data for the evaluation of a specific target, and can be easily understood and absorbed by the users [1]. They have been used extensively in this study to represent design rules and data on die structure. However, only a limited number of rules are included in each decision table so that they are comprehensible and can be easily edited in future. In practice, dividing the technological data and rules into different tables can allow the developers to implement the program easily, develop the software accurately and change the table content to suit actual situation. After translating the decision tables into computer programs, each table can be regarded as a small module inside the whole program. Accumulating the recommended actions from each of these decision tables, a generative die structure design for the required drawn product can be produced.

2. DRAWING DIES AND THEIR CLASSIFICATION

A die assembly can be simply divided into two sections; one is the working area which is composed of functional components such as punch and die, and the other one is the non-working area which is composed of supporting components such as die holder, guide posts and bushes, spring, stock guide etc. The functions of the die components, the equipment being used and the specification of the product determine the position and shape of each die component inside an assembly. However, there are several possible arrangements of the position of the components to form a workable die. Some important types of dies, which have their respective advantages and disadvantages are classified as : (i) Movable blankholder die, (ii) Fixed blankholder die, (iii) Drop-through die, (iv) Return-type die, (v) Compound die (Combination die), (vi) Progressive die, and (vii) Transfer die. They can be further classified into conventional or inverted die type based on the position of the punch and die. In the conventional die, the punch is fixed to the upper die set and the die is mounted on the lower die set. For inverted die type, the punch is mounted on the lower die and the die is fixed to the upper die.

2.1 Types of Die Construction

As mentioned earlier, a cup can be drawn using a single operation, a number of operations (progressive or transfer die) in different stages, or a combination of two operations in a single stage (combination or compound die), and the actual design depends on the experience of the designer and the facilities available in the workshop [2-5]. Fig. 1 shows a collection of various types of constructions that are commonly used for the axisymmetrical deep drawing.

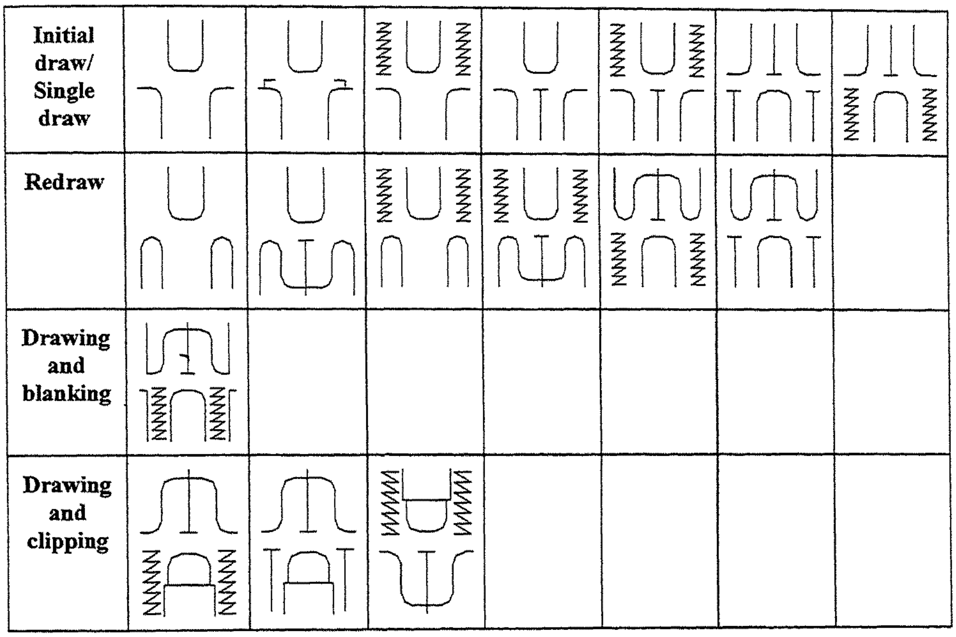
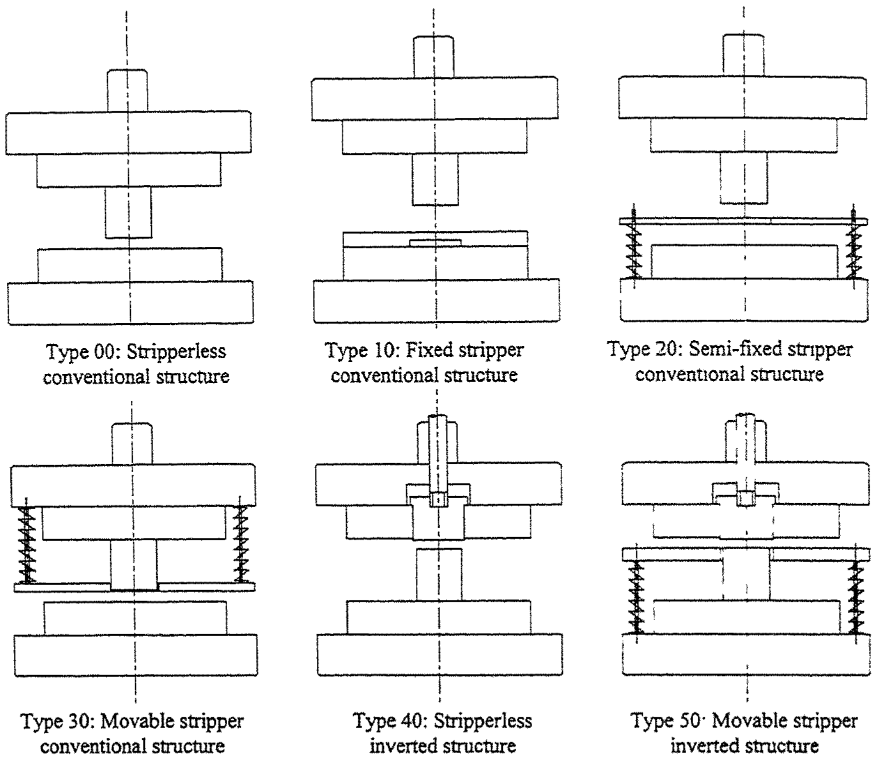


Fig. 1 : Schematic representation of various die structures that are frequently used.

2.2 Classification using Group Technology

It will be an advantage if a CAD/CAM system can retrieve a tentative solution from the records of an electronic database to solve the problem on hand. Group technology (GT) gives the power to utilize the recorded solution over and over again wherever the particular characteristics or attributes are shared by other products. GT can serve to focus our technology so that best practices can be consistently applied leading to increase in efficiency and productivity as well as possible elimination of prototype stage. Design standardization for families of products is obviously more easily achieved through a GT system since it can quickly collect similar parts and examine the group as whole. Correct application of GT to die design and making establishes a coherent system which originates with the design engineering activity and proceeds through subsequent stages of process layout planning, die design, die making and the production control of the product.

Though numerous types of die construction are possible, they can be simply classified into six groups based on their construction characteristics and applications [3]. A two digit numerical system s used to classify the various die constructions; the first digit indicates the framework type whether it s a single operation die, combination die or progressive die, the position of punch and die and the ype of blankholding method. The second digit indicates the application of presswork such as iercing, bending, blanking, deep drawing etc. Different die constructions can be grouped using such lassification, and the resultant basic groups are schematically presented in Fig. 2.



Type 00: Stripperless conventional structure

Type 10: Fixed stripper conventional structure

Type 20: Semi-fixed stripper conventional structure

Type 30: Movable stripper conventional structure

Type 40: Stripperless inverted structure

Type 50: Movable stripper inverted structure

Fig. 2 : Schematic representation of basic groups of die construction.

3. SELECTION OF THE TYPE OF DIE CONSTRUCTION

The selection of the type of die depends on the production quantity, size of the product, equipment used for production and the experience of the die designer. Also, prior to deciding the overall structure, the sub-construction details such as knockout device and blankholder type are first determined on the basis of the product specification and the data provided from process planning. A blank ejection device is required to eject the workpiece from the die cavity after it has been formed, especially, when ironing is involved and/or flanged cup is being formed. The installation of such devices is usually based on the experience of die designer. As such, these practical experiences can be easily converted into decision tables and preserved for future use. Tables 1 and 2 illustrate the decision tables that can be used in the selection of blank holder and knockout and the type of deep drawing method respectively. The aim of this study is to select the most suitable die structure automatically from among the six basic die constructions illustrated in Fig. 2. As different die structures offer different levels of tolerances on the product drawn with them, a precision grade is assigned based on the decision table illustrated in Table 3. Then the decision matrix given in Table 4 aids the selection of the type of die construction, which is simply based on the precision of the product and the production quantity. Alternatively, when the user input provides the product collection method, product precision and the requirement of blankholder, the selection of die structure can be made using the decision criteria presented in Table 5.

Table 1 : Decision table for the determination of blankholding and work ejection methods

CONDITION STUBS	CONDITION ENTRIES (Rules)							
	1	2	3	4	5	6	7	8
(t/D)x100	> 2	> 2	< 2	< 2	< 1	< 1	>1.5	>1.
Drawing rate ($m = d/D$)	> 0.6	> 0.6	< 0.6	< 0.6	< 0.8	< 0.8	>0.8	>0.8
Ironing	N	Y	N	Y	N	Y	N	Y
ACTION STUBS	**ACTION ENTRIES**							
Knockout device	N	Y	N	Y	N	Y	N	Y
Springed blankholder	N	N	Y	N	Y	Y	N	N

t : Thickness of sheet material D : Diameter of the blank d : Diameter of the drawn cup

Table 2 : Decision table for the determination of process method for deep drawing

CONDITION STUBS	CONDITION ENTRIES						
Drawing rate	-	-	>0.6	>0.4	>0.8	<0.4	-
(t/D)x100	-	-	>2.0	>0.2 ~ 2.0	-	<2.0	-
Product collection method	Drop-through type	Ejection type	-	-	-	-	-
H/D	-	-	-	-	<0.5	-	-
Production quantity	<1,000,000	<1,000,000	<1,000,000	<1,000,000	>1,000,000	>1,000,000	>5,000,000
Product diameter (mm)	-	-	-	-	>50	<20	-
Number of stages	≤3	≤3	≤3	≤3	-	-	-
ACTION STUBS	**ACTION ENTRIES**						
Single operation: Conventional die	Y	N	Y	N	N	N	N
Single operation: Inverted die	N	Y	N	Y	N	N	N
Combination die	N	N	N	N	Y	N	N
Progressive die	N	N	N	N	N	Y	N
Transfer die	N	N	N	N	N	N	Y

H : Height of the target product

Table 3 : Decision table for the determination of accuracy class of drawn product

CONDITION STUBS	CONDITION ENTRIES		
Tolerance specification: Diameter	> ± 0.250	± 0.250 ~ ± 0.125	< ± 0.125
Depth	> ± 0.500	± 0.500 ~ ± 0.250	< ± 0.250
ACTION STUBS	**ACTION ENTRIES**		
Accuracy	Low	Moderate	High

Table 4 : Decision matrix for selecting the type of die construction based on the precision requirements of product

Accuracy	Precision/High			General/Moderate			Rough/Low		
Quantity*	Large	Medium	Small	Large	Medium	Small	Large	Medium	Small
Die Structure									00
					10	10	10	10	
		20	20	20					
	30								
						40		40	40
	50	50	50	50	50		50		

* Large : >1,000,000; Medium : 500,000 - 1,000,000; Small : <500,000

Table 5 : An alternative decision table for selecting the type of die construction

CONDITION STUBS	CONDITION ENTRIES															
Drop-through type die	Y	Y	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y	N	N
High precision	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	N	N	N	N
Large quantity	Y	Y	Y	Y	N	N	N	N	Y	Y	Y	Y	N	N	N	N
Blankholder	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N
ACTION STUBS	**ACTION ENTRIES**															
Die structure group	30	20	50	40	10	00	50	40	30	20	50	40	10	00	50	40

4. SELECTION OF DIE-SET

The use of a die-set not only shortens the manufacturing time of the forming die but also assists in the smooth proceeding of the forming process. In general, ready-made die-sets are used for smaller dies and larger dies are specially designed according to the requirements of the product and the press available. Substantial amount of practical and tested information is available from commercial sector, trade standards, national standards as well as international standards that can be readily used. The advantage of such information in the present context is that all the standardised data and information, usually presented in table format, can be converted readily as decision tables. Such a representation of die structure as decision tables enables users to easily access and modify them as per their own practical data, without any changes in the software of the CAD/CAM system. Standardization of the size and thickness of the plates can speed up the die-set selection process. The size of different plates (in planar dimensions) are constrained by the size of the product, and the thickness depends on the planar dimensions, which can be determined following the design data presented in Table 6.

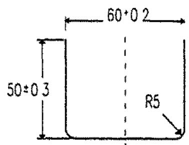
Table 6 : Relationship between the thickness of various plates and their planar dimensions

	Planar dimensions of the plate (in mm):				
Thickness (mm):	up to 100×100	150×150	180×180	250×250	350×350
Punch Holder	25 ~ 32			30 ~ 40	40 ~ 50
Punch Plate	15 ~ 25	20 ~ 25		25 ~ 28	25 ~ 32
Backing Plates	5	8 ~ 16			10
Stripper Plate	10 ~ 16			16 ~ 20	
Guiding Stripper	16 ~ 25	20 ~ 25		25 ~ 28	
Die Plate	16 ~ 25	25 ~ 28		25 ~ 32	
Die Holder	28 ~ 40			40 ~ 50	

5. CASE STUDY

The general operation of the proposed system is illustrated through a case study. The details of the chosen product are given in Fig. 5. Application of Rule 3 in Table 7 shows that the product can only be drawn using two operations with $m_1 = 0.58$ and $m_2 = 0.79$. Relevant columns in Tables 1 to 5 indicate the following results for the first draw operation:

- Column 5 of Table 1 → Springed blankholder for $m_1 = 0.58$.
- Column 4 of Table 2 → Single operation and inverted die.
- For moderate precision grade and medium production quantity of the product, Table 4 suggests "Fixed stripper conventional structure" or " Movable stripper inverted structure". Since inverted die has been chosen from Table 2, the latter should be the solution.
- Alternatively, as the specifications of the product meet the condition entries in Column 15 of Table 5, " Movable stripper inverted structure" is the resultant selection, which is same as above.



All dimensions in mm

Material = Deep drawing steel
Production quantity = 600,000 pcs.
Thickness of the sheet metal = 0.7 mm
Relative sheet thickness, $t/D \times 100$ (%) = 0.57
Height of the cup to diameter of the blank, $H/D = 0.41$
Overall drawing rate, $d/D = 0.492$

Fig. 5 : Product specifications used in the case study.

Table 7 : Determination of limit drawing rate for full cup drawing

	CONDITION ENTRIES (Rules)					
Sheet thickness range	1	2	3	4	5	6
t/D (%)	0.08~0.15	0.15~0.30	0.30~0.60	0.60~1.0	1.0~1.5	1.5~2.0
Drawing rate	ACTION ENTRIES					
m_1	0.63	0.60	0.58	0.55	0.53	0.50
m_2	0.82	0.80	0.79	0.78	0.76	0.75
m_3	0.84	0.82	0.81	0.80	0.79	0.78
m_4	0.86	0.85	0.83	0.82	0.81	0.80
m_5	0.88	0.87	0.86	0.85	0/84	0.82

6. SUMMARY

Standard die structures used for producing hollow sheet production are standardised based on group technology concepts. The selection of the die structure is based on the process layout and the product features, namely the production quantity, precision, product shape and other special requirements. In consideration of achieving quick die design and manufacturing, international and commercial standards are followed for the design of die components. In the determination of the type of die structure, its size and related attributes, decision table method has been adopted extensively.

7. REFERENCES

1. Decision Table Task Group Report, Beitz, E. Henry (Editor), Conference on Data Systems Languages (CODASYL), Association for Computing Machinery, New York, 1982.
2. Nakagawa, T., Press Ka Ko Data Handbook, pp. 227-259, Nikkan Kogyo Shin Bun Sha (Japanese) (Presswork data book, Nikkan Ind. Publisher, Tokyo, Japan), 1986.
3. Yoshida, H. and Yamaguchi, F., Kana Gata Sekki Ki Jun Manual, Shin Gi Ju Tsu Kai Hatsu Center (Japanese) (Die Design Standard Manual, New Technique Development Centre, Tokyo, Japan), 1986.
4. Yamaguchi, F., Fundamental Press Die Practice Text, Part 2, Press Technology, Vol. 29, pp.7-47, 1991.
5. Nakagawa, T., New Presswork Data - Deep Drawing, Press Technology, Vol. 30, No. 5, pp.1-96, 1992.

IDENTIFICATION OF GRAIN GROWTH IN WOOD MACHINING USING ARTIFICIAL NEURAL NETWORKS

V. Karri
Dept. of Mechanical Engineering
University of Tasmania, Hobart, Australia - 7001

ABSTRACT

The anisotropic properties and different directions of grain growth make wood a complex work material in automated machining. In an automated machining of wood, in order to keep a uniform surface finish and associated vibrations along different directions of wood, a selection of appropriate cutting tools is necessary. In the first instance, there is a need for an intelligent system to identify the grain growth before the appropriate selection of cutting tool. In this paper single edged machining operations are carried out for Australian grown Huon Pine wood. Experiments are carried out along and across the grain growth over a comprehensive range of tool geometrical combinations and speeds. A neural network architecture is developed, based on the extensive experimental data, to identify the direction of grain growth for automatic selection of cutting tools. The network is trained over 350 sets of tool geometrical combinations and extensively tested. The testing of the network has shown a very high reliability of the developed neural network model.

KEYWORDS

Neural Networks, Intelligent Machines, Machining, Manufacturing Processes

1. INTRODUCTION

Research in wood machining has progressed from investigation of a particular problem of a certain machining process to study of the problems basic to a particular process, as noted by the work of Franz [1] in planing, Reineke [2] in sawing, and Leney [3] in veneer cutting. It is apparent that the more fundamental such work becomes, the less it is concerned with the details of the machine or the tool, and the more it is concerned with the action of a highly simplified cutter on a workpiece of wood. On the other hand the practical problems associated with the anisotropy of wood have complicated the cutting tool selection in an automated machining operation. The deformation and type of chip formation in machining of wood depends entirely on the direction of cutting and the grain growth. The performance of the cutting tool varies along the different directions of the grain growth. In an automated machining there is a need for an intelligent controller to identify the appropriate area of grain growth in wood and then select appropriate cutting tool and cutting conditions to carry out machining. The intelligent controllor has two major tasks; one is to identify the area of working i.e. if the grain growth is along, across or perpendicular and the second task is to select appropriate cutting tool. In this paper an intelligent neural network model is developed which when given certain cutting conditions and some measured parameters can estimate the area of grain growth.

This investigation is based on the premise that the diverse wood machining processes have vital problems in common, and that the most basic of these concerns the action of a wedge-shaped cutter on a wooden workpiece, in the simplest possible cutting situation such as the orthogonal cutting operation [4-9]. This is a term applied where a simple wedge-shaped cutter, with a straight edge wider than the workpice, moves with a uniform velocity in a plane parallel to the surface of a semi-infinite workpiece. A Quasi-equilibrium is assumed [4,9] in the process, which forms a static problem in mechanics. To study the general oblique or orthogonal cutting situation [5,6], it is evident that the greatest obstacle to generally analyse is the extreme anisotropy of wood [6,7,10-16]. The complexity to visualize and establish relationships between events for cutting situations differing widely with respect to grain, has restricted a few workers [10-12] to study a situation in which both the cutter edge

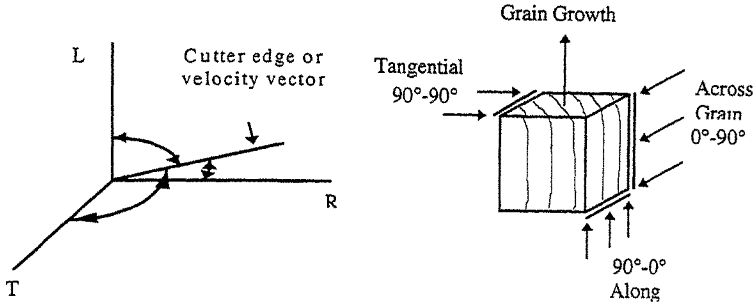


Fig. 1 Three Axis of Wood and Direction of Grain Growth

and its velocity vector are perpendicular to the grain growth. This was done in the hope that principles elicited and analytical results obtained might be extended to other grain angles. This possibility was briefly investigated, with limited success, on American grown *Pinus strobus* (white pine), *Liriodendron Tulipifera* (yellow poplar) and *Acer Saccharum* (sugar maple) [14].

As applied to metals, orthogonal cutting is regarded as a plane problem in mechanics. Because of the anisotropy of wood, this is not the general case, but holds only when one of the axes of symmetry is parallel to the cutting edge. The preliminary investigation carried out by the chief investigator on Australian grown Huon pine has resulted in the qualitative comparability, of the effect of major process variables on the forces, power, friction and shear stress, when machined along the grain growth [15,16]. In the general situation in wood, in which the grain direction is arbitrary, it can be characterized to be orthotropic, with the longitudinal (L), tangential (T) and radial (R) directions mutually orthogonal, then the cutting situation is represented by the direction angles of the cutter edge and the velocity vector relative to these three axes (see fig.1 below). Figure 1 also shows a typical grain growth and grain direction.

A traditional approach to identification of the grain type using the forces, power and vibrations data needs complex modelling. The intricate modelling and associated complications in 'traditional' approach have necessitated empirical approach in modelling the process. To understand the effect of various major process variables on forces, power and vibrations in wood machining, by empirical approach, require an extensive and often prohibitive experimental investigation. Further the quantitative validity of the empirical equations resulted from complex statistical curve fitting techniques are of doubtful accuracy due to the limited data. Hence neural networks are sought for estimating the performance predictions and also able to decide on the grain type when automating wood working based on the experimental data. In other words the neural network model developed in this work caters for both the predictive capabilities of the model as well as making intelligent decision to identify the grain growth type on a wooden workpiece.

This information is extremely critical in developing the automated neural network controller for identification of various regions of an anisotropic wooden material so that depending on the region, cutting conditions can be used to maintain the uniform surface finish. Hence a neural network architecture is developed, based on the extensive experimental data, to identify the direction of grain growth for automatic selection of cutting tools. Based on the data if the direction of grain growth is determined, then the associated tool geometrical combinations can be used for machining wood.

2. DEVELOPMENT OF EXPERIMENTAL RIG AND DATA ACQUISITION

The experimental setup involves the measurement of both the force components and the vibrations during the cutting process. The concentric dynamometer with strain gauges is connected and mounted round on the head of a milling machine. The strain gauges were connected to the strain gauge amplifier for recording the force components. A data acquisition system was used together with a PC Pentium computer for recording the force components. The cutting tools were mounted on the dynamometer for the measurement of three force components (Fp (power force), Fq (thrust force) and Fr (radial force)) in all three directions during cutting [15] as shown in the experimental set up in figure 2 below. The Huon pine wood was machined along and across the grains (see fig.1). The vibrations were also simultaneously measured using an accelerometer mounted at the back of the cutting tool. The analog signal indicates the real-time, instantaneous acceleration of the object on which the accelerometer is mounted. The selection of the accelerometer is made based on the thickness of the cutting tool selected. The output from the accelerometer is sent into frequency display digital oscilloscope and subsequenty through a data acquisition system into the computer for storage.

In carrying out simpler wedge typed orthogonal and oblique cutting operations 3 normal rake angles (10°,20°30°) and 3 angles of inclination (0°,10°,20°) were selected to cover a comprehensive tool geometry. For each tool geometrical combination 9 cut thicknesses were taken so that a set of 81 cuts were taken in this part of investigation. Further to check the effect of speed on the process parameters, three speeds were run for two angles of inclination amounting to 56 cuts. Hence considering two directions (along and across the grain growth) a total of 272 (2 x (56+81)) cuts were taken in this part of the investigation. It is hoped that the quantitative reliability of the neural network model will be very high with the training being carried out on a comprehensive experimental data.

The preliminary investigation carried out on this material [12,13] has shown that the qualitative trends of the major process variables on the forces, power and vibrations are similiar to that of metals. This is particularly encouraging since the well established predictive models for machining of metals can be readily transported to machining Australian grown Huon pine wood.

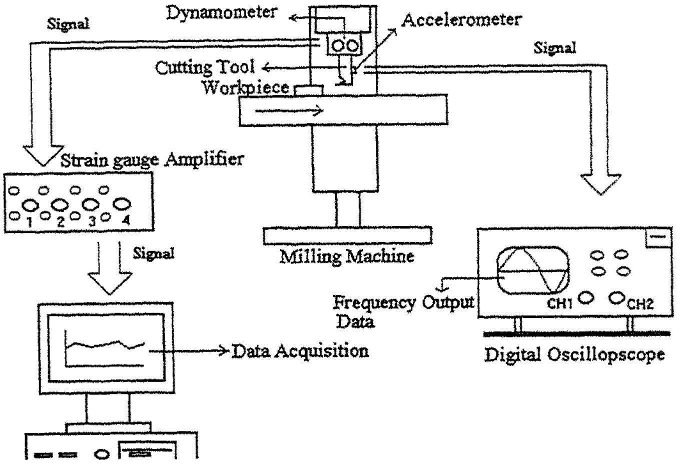


Fig.2 Experimental Setup for Forces and Vibration Measurements

3. NEURAL NETWORK ARCHITECTURE FOR IDENTIFICATION OF GRAIN GROWTH

A multi-layer perceptron network with backpropagation [14] is chosen as shown in fig.3. The back propagation follows gradient descent on the error surface to minimize the network error. Although adding momentum to back-propagation can decrease training times and the probability that the network will get stuck in a shallow minimum in the error surface, in this particular application, no momemtum is added. Back propagation networks are multi-layered perceptrons feed forward neural networks that apply the error back-propagation procedure for learning.

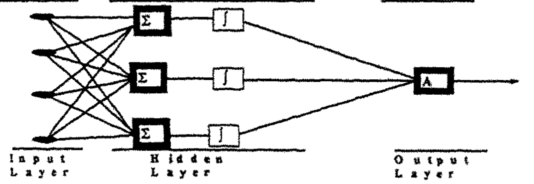


Fig 3 Neural Network for automatic grain growth detection.

The back propagation procedure uses a gradient descent method which adjusts the weight in its original and simplest form by an amount proportion to the partial derivative of the error function with respect to the given weight. The learning procedure of a supervised multi-layered perceptron backpropagation network is to initialise the weights of the network at small random values and start the learning cycle by exposing the network to a certain input pattern paired with the desired output.

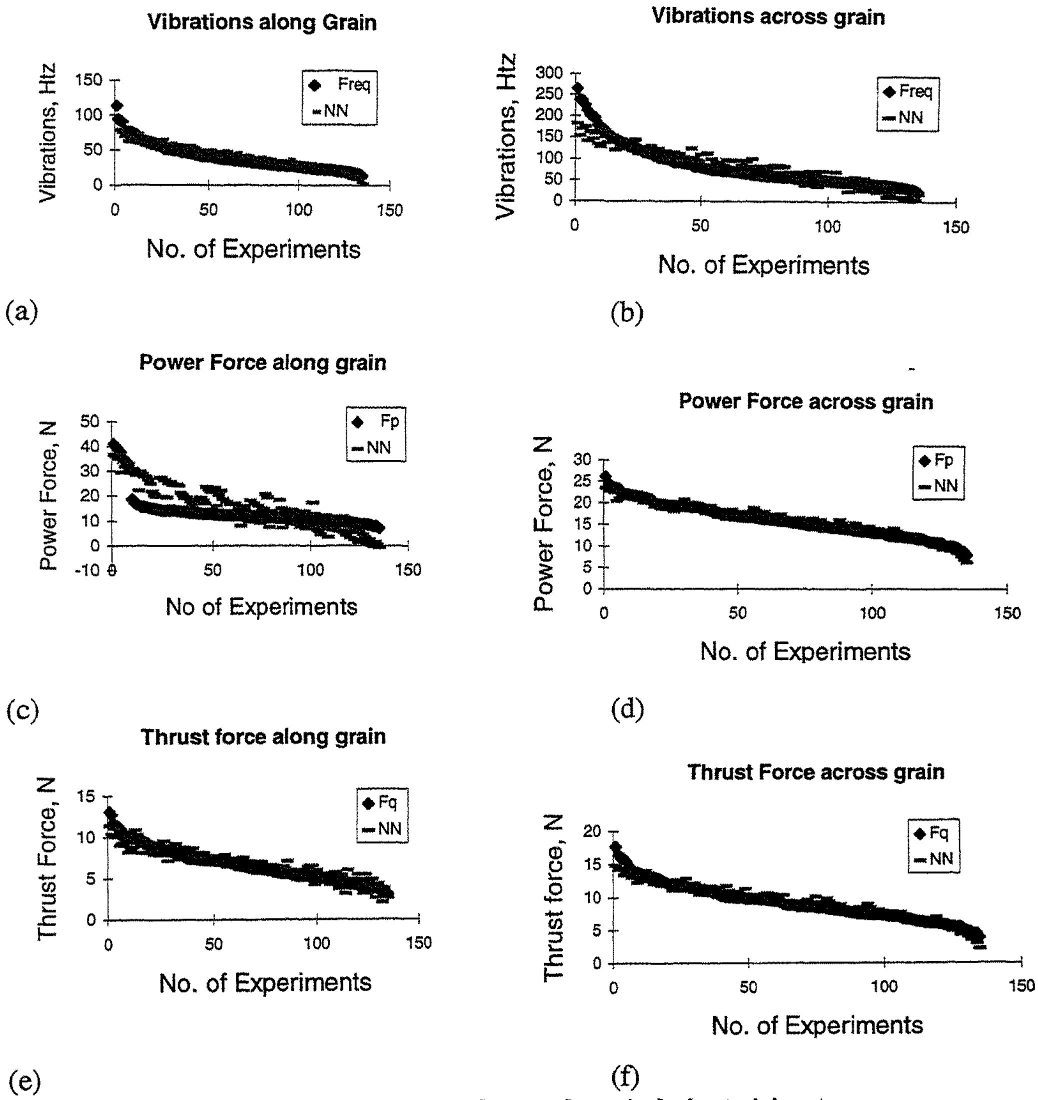


Fig.4 Vibrations and Forces along and across the grain during training stage

The network then computes the output so that the error can be calculated following by adjustment of weights of the network using the error backpropagation algorithm so that a certain amount of the detected error is removed. The process is repeated until the cumulative error is within the tolerable range. The most common difficulty in using the back-propagation for training neural network models is to choose both adequate learning rate and initial weights for each application. It should be noted here that both learning rate and initial weight factors play critical roles in the asymptotic accuracy of chosen neural network models for system representation and their convergence rate in estimating the suitable weight vectors.

After designing the architecture of the neural networks, the number of inputs needed are to be decided to identify the grain growth. The number of variables tested are the normal rake angle, inclination angle, cut thickness and the velocity so that given a set of these variables the forces, power and vibrations can be predicted along and across the grain sizes. The neural network is trained over 136 cutting conditions for each type of grain growth separately. This also helps to estimate the performance along and across the grain growth independently. There are 4 inputs at the input layer stage, with 5 neurons in the hidden layer and 1 output for each force component and vibration for each of the cases along and across the grain. Back propagation can train multi-layered feed forward networks with differentiable transfer functions to perform function approximation, pattern association, and pattern classification.

Fig.4 shows the quantiative comparability of neural network to experimental values of vibrations, power and thrust forces along and across the grain growth. Figs.4a,c and e show the vibrations, power and thrust forces during the training stage when machining along the grain. It can be seen that the neural network has exhibited a good correlation matching the experimental results. The average percentage deviation defined as ((NN output - experimental)/experimental))*100 was found to be 6% for vibrations and 5%, 4% for power and thrust forces respectively highlighting the excellent training of neural network architecture. Similarly figs.4b,e and f show the neural network predictions for vibrations, power and thrust forces across the grain growth. The average percentage deviations for vibrations, power and thrust forces were found to be 4%, 7% and 5% respectively.

To use the network as an intelligent system to detect the direction of the grain growth based on the input information all the 4 inputs together with the vibrations for both along and across the grain directions are used at the training stage. The vibrations along the grain, for the purpose of neural networks, are given a value of 0 and that across the grain have been given a value of 1. The network has been well trained within an error of ± 1% error. Figure 5a shows the training stage of the network. It can be seen that the network architecture, based on the patterns, recognises and identifies the areas along and across the grain growth. It has been found that for $0 \leq X \leq .5$, the output indicates that the cutting tool is along the grain growth. On the other hand if $.5 \leq X \leq 1$, the cutting tool is across the grain growth, where X is the vibrations output from the neural network architecture. The network is tested over 36 different cutting conditions within the domain of the trained cutting conditions. Figure5b shows the output of the neural network for the testing conditions. It can be seen that the network identifies the regions along and across the grain growth. For along the grain growth (0 identification), all the values estimated by the neural networks are between 0 and 5. On the other hand the values predicted for across the grain growth (1 identification) are well with in the limits of 5 and 10. This confirms the intelligent behaviour of the networks to be able to identify different regions of anisotropic material based on the experimental data incorporated into the controller for automated wood machining.

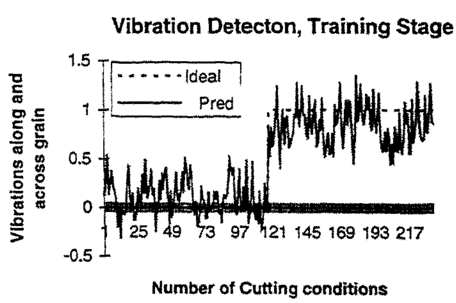


Fig. 5a. Vibration detection, Training Stage

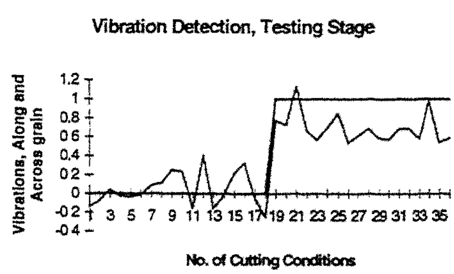


Fig. 5b. Vibration Detection, Testing Stage

4. CONCLUSIONS

Wood as an anisotropic material is discussed and the need for developing reliable performance estimation models is highlighted. It has been shown that with the advent of neural networks the performance features such as forces, power and vibrations can be estimated. A neural network architecture is developed and tested for estimating performance for machining wood along and across the grain growth. The quantitative estimation of performance predictions using have been found to be with ± 5 % average percentage deviation. Further using the multi-layer perceptron with back progagation without momentum, the machining data for both along and across the grains have been trained for identification of the type of surface. The machining surface along the grain of the wooden piece has been identified as zero and that across the grain has been identified as 1. The network has been well trained with an average percentage deviation of 1.8% in the error and while testing with 36 cutting conditions, the machining surface was well detected by the neural network model.

5. REFERENCES

1. Franz, N.C., "An Analysis of the Wood Cutting Process", Univ. of Michigan, Ann Arbor, Mich. 1957.

2. Reineke, L.H., "Sawteeth in Action"., Proc. for Prod. Res. Soc., 4: pp.36-51, 1950.

3. Leney, L., "A Photographic Study of Veneer Formation", For. Prod. J., V.10 (3), pp.133., 1960.

4. Armarego, E.J.A., "Practical Implications of Classical Thin Shear Zone Analysis", UNESCO/CIRP Seminar on Manuf.Tech., Singapore, p167. 1982.

5. Karri,V, Fundamental Studies of Rotary Tool Cutting Processes, Ph.D. Thesis, The Uni.of Melbourne, 1991.

6. Karri,V., "Computer-Aided Predictive Cutting model for Simulated Keyway Broaching Operation", 3rd IASTED Int.Conf. Robotics and Manufacture, June14-17, Cancun, Mexico, pp.410, 1995.

7. Karri,V., "Forces, Power, Stress and Displacements in Orthogonal Keyway Broaching Operations". Int. Conf. on Intelligent Manufacturing Systems, Vienna - Austria, June.13-15, p.549-554, 1994,

8. Merchant, M.E., "Mechanics of Metal Cutting Process", J.Appl.Phy., V.16/5, P.267, V.66/6. p. 318.

9. Armarego, E.J.A., Brown, R.H., The Machining of Metals, Prentice Hall Inc., New-Jersey, 1969.

10. Freudenthal, A.M., The inelastic behaviour of engiering materials and structures, Wiley and Sons, Newyork., 1950.

11. Norris, C.B., "Strength of orthotropic materials subjected to combined stresses", F.P.L. Madison Report No. 1816.

12. Kivimaa, E., The Cutting Forces in Woodworking, State Institute for Technical Research, Helsinki, Finland, Publ.No.18, 1950.

13. AS01 Glossary or terms used in timber standards.

14.Bootle, K.R., Wood in Australia, Mc.Graw Hill Company, Sydney, 1993.

15. Karri, V., "Orthogonal Cutting Characteristics of Huon Pine - An Experimental Investigation", 14th Int. Conf. on Mech. Struc. & Mat., Dec.,Hobart,p 664. 1995.

16. Karri, V., 1995, "Oblique Cutting Characteristics of Huon Pine", 14th Int. Conf. on Mech. Struc. & Mat., Dec.,Hobart,p 669, 1995.

17. Rumelhart,D.E., McCleland,J.L., "Parallel Distributed Processing", MIT press, Cambridge, 1986.

Author/Session/Page Index

Abachi H. 1B/72
Abrantes Maria Joao 6A/612
Acar B.S. 7A/697
Ajmal A. 11C/1179
Al-Kindi G.A. 9D/1033
Albrecht R. 4B/389
Arnold J. 4B/395
Au C.K. 4D/486
Awan Khalil Aslam 7B/733
Azevedo J.L. 11A/1137
Babu N. Ramesh 5B/554
Babu N. Ramesh 8B/845
Bachmann Bernd 9A/941
Bahns O. 4B/389
Bao Li 1C/122
Bazzi M. 7B/743
Beck S. 4B/389
Bell D.A. 11A/1143
Bell R. 9B/959
Bellalem Lotfi 9A/934
Bender K. 9C/1001
Bernardi Ansgar 9A/941
Beyer C. 3C/294
Bindbeutel K. 9C/1001
Birke C 3C/294
Blackmore D. 11B/1161
Blackmore D. 11B/1167
Bohn Jan Helge 8C/878
Boudebous Dalila 3B/266
Cai Hegao 2A/164
Cai Lilong 9D/1027
Cameron Stephen 10A/1057
Campbell Thomas 7A/703
Cao J.D. 11D/1219
Cao Ming 5C/560
Carpenter I.D. 5C/583
Case K. 7A/697
Chan C.C. 5A/529
Chan F.T.S. 4B/407
Chan F.T.S. 5B/535
Chan F.T.S. 9A/947
Chan Felix T.S. 1B/77
Chan Felix T.S. 8D/915
Chan K.C. 8A/803
Chan K.W. 4A/341
Chan K.W. 4D/492
Chan K.W. 8D/921
Chan L.C.F. 9A/953
Chang Jiazhong 3B/255
Chen Bingkui 1A/56
Chen Bingkui 11C/1197
Chen Calvin C. 8C/894
Chen Chi-Hui 5A/523
Chen F.L. 8D/928
Chen Hui-tang 1D/146
Chen JiaDong 9C/1015
Chen Jihong 9D/1027
Chen Jiuqiang 8A/821
Chen Jyun-Ming 5A/523
Chen Ke-Zhang 9C/1007
Chen Ke-Zhang 10C/1079
Chen L.C. 8D/903
Chen W.J. 10D/1106
Chen X.H. 10B/1075
Chen Y.H. 3A/231
Chen Y.H. 4D/492
Chen Y.H. 7C/761
Chen Y.H. 9D/1039
Chen Y.H. 10A/1045
Chen Y.H. 11C/1185
Cheng Fuhua Frank 11D/1213
Cheng J.John 11B/1155
Cheng K. 10B/1075
Cheng S.K. 7C/755
Chiang Y.M. 8D/928
Chiu K.F. 9B/984
Choi S.H. 3C/300
Chu L.K. 9B/984
Chu T.H. 10B/1069
Clark D.E.R. 7D/797
Condom J.M. 3D/323
Corney J. 7D/797
Cui N. 1B/89
Cui Yaodong 5C/560
Daho O. Ben Ahmed 7B/743
Dangelmaier W. 3B/272
Deng Chao 5C/565
Detand J. 2C/188
Dewar R.G. 5C/583
Ding Xinquan 1B/83
Dormann S. 8C/884
Du H. 6B/629
Du Mike 2D/220
Du R. 2C/196
Du Zhaohui 3C/284
Durr H. 4C/426
Ellis Ray 9B/965
Es-Said O. 8C/884
Eshwaraiah P. 4A/353
Estima de Oliveira J.P. 11A/1137
Fan I.S. 11C/1191
Fan Q.J. 1B/89
Ferreira J.A. 11A/1137
Fielding T. 1D/134

Fritz B. 8C/884
Fu Yili 3D/311
Fukuda Junko 4A/360
Galantucci L.M. 7C/773
Gao Bin 3B/255
Gao J.X. 11C/1173
Gao Jian 4D/461
Gao Juhong 4D/467
Garden Y. 5D/594
Garden Yvon 6A/623
Garden Yvon 9A/934
Gerencher P 8C/884
Gibson I. 4C/455
Gindy N.N. 8A/826
Gindy N.N. K/9
Glander M. 9C/1001
Goh Angela 8B/851
Gongbin 1A/60
Gorelik A.G. 2D/208
Grote K.H. 3C/294
Grote Karl-H. 8C/868
Gu Guanqun 1B/83
Guan J.W. 11A/1143
Guan Youjiang 6D/665
Gui Jin-Kang 6C/659
Guo Haibin 1C/111
Guo Haibin 1C/117
Guo Qun 4D/474
Guo Y. 8A/803
Han Ming 1C/105
Han Z.Y. 8B/863
He Huacan 6C/647
Herscovici Daniel 8C/894
Hill Simon D. 6A/612
Hoffmann J. 8D/909
Holm Hans 3D/317
Holm Hans 3D/335
Hsu Hung-Yao 4B/383
Hu Xiaobing 8A/821
Hu Y.N. 10A/1045
Hu Y.N. 11C/1185
Hu Yujin 3A/249
Huang Chao 2A/158
Huang G.P. 1A/50
Huang Jin 3C/290
Huang N.K. 1B/89
Huang Peizhi 4D/480
Huang Qiming 8C/890
Huang Shuhuai 1C/105
Huang Shuhuai 3C/290
Hui K.C. 3A/237
Hui K.C. 3A/243
Hull R.S. 3B/278
Hull R.S. 9B/978
Huo Chuanhue 4D/480
Ibrahim R. 1B/72
Ip A. 11C/1173
Ip R.W.L 4B/407
Ip R.W.L. 5B/535
Ip R.W.L. 9A/947
Ip Ralph W.L. 1B/77
Ip Ralph W.L. 8D/915
Ip W.H. 5B/542
Ip W.H. 8B/838
Ip W.H. 9B/971
Jacquel D. 10C/1087
Jia Y.Y. 8A/815
Jiang Ping 1D/146
Jin Aiming 5C/560
Jokinen Harri 6C/659
Joneja Ajay 7D/791
Jones R. 4C/420
Jung J.P. 5D/594
Jung Yoong Ho 5A/510
Kadhim M.J. 9D/1033
Kang Jichang 6C/647
Kang Min-Hyoung 5C/577
Kao Y.C. 1A/44
Karcher A. 9C/1001
Karri V. 5D/606
Karri V. 6A/618
Kaschka U. 4C/426
Kennedy R.D. 1B/95
Keung K.W. 5B/542
Kim J.S. 10A/1051
Kirkwood D. 10B/1075
Kishore A. Krishna 5B/554
Kong Tom 11B/1155
Krishna Bala N. 7C/767
Kruth J.P. 2C/188
Kuhn A. 3B/272
Kunii Tosiyasu L. K/21
Kurdthongmee W. 11D/1207
Kuttner Brian C. 10C/1093
Kwok M.C.K. 5B/548
Kwon K.S. 10A/1051
La Pierre B. 4B/389
Lai C.H. 5A/529
Lai C.H. 11B/1149
Laliberte M. 1D/140
Lam F.W. 4B/414
Lam S.M. 7D/779
Lam T.W. 7A/685
Langemann T. 3B/272

Author/Session/Page Index

Lanuel Y.	5D/594
Lau Henry	1B/77
Lau Ka-Lok	8B/851
Lau T.L.	7B/721
Law S.	11B/1149
Lee C.W.	10B/1063
Lee Eng Wah	10D/1114
Lee Kwan H.	4C/449
Lee T.M.	7B/727
Lee W.B.	4B/414
Leinen S.	5D/594
Leng Y.N.	8A/815
Leu M.C.	11B/1161
Leu M.C.	11B/1167
Leung C.B.	7D/785
Li C.L.	4A/341
Li Deren	4D/480
Li H.T.	9D/1039
Li J.	8B/857
Li Jian	6B/641
Li Peigen	5C/565
Li Runfang	1A/56
Li Xueqing	1A/60
Li Y.	5B/542
Li Y.D	3A/237
Li Yan	3C/290
Li Yuguang	4D/480
Liang M.	2C/202
Liang Zhongming	7A/709
Lim Leslie K.C.	8C/874
Lin Chao	1A/56
Lin G.C.I.	3A/225
Lin G.C.I.	6B/629
Lin G.C.I.	8D/903
Lin Grier C.I.	1A/44
Lin Grier C.I.	1D/128
Lin Grier C.I.	4B/383
Lin Jianping	2B/170
Lin Tengjiao	1A/56
Lin Tengjiao	11C/1197
Little G.	7D/797
Liu Bin	1C/105
Liu C.K.	11C/1191
Liu C.Y.	9D/1039
Liu Ding Yuin	11D/1201
Liu Dong	8A/821
Liu J.R.	9D/1039
Liu Pei Ling	10D/1114
Liu Shenquan	1A/60
Liu Shiping	2D/220
Liu W.H.	2B/182
Liu X.M.	9D/1039
Lu Qingping	1C/122
Lu Tien-Fu	1D/128
Lu Yiping	2B/170
Lui C.K.	4A/374
Luo Bin	5C/565
Luo Yan	2A/152
Luong L.H.S.	9A/947
Ma M.C.	3A/243
Ma Matthew Jia	2C/196
Ma Weiyin	9D/1021
Ma Y.J.	4B/414
Ma Yulin	3D/311
Madsen Ole	3D/317
Mai Chungfan	6C/652
Martin B.	5D/594
Mathews P.G.	11A/1131
Medland A.J.	4A/366
Mendelson M.	8C/884
Meng Xiangxu	1A/60
Mensing G.	4C/455
Mertins K.	4B/389
Mill F.G.	10C/1087
Miller Jeffrey L.	8C/868
Mills J.K.	1D/134
Mills J.K.	1D/140
Ming Tang	6C/652
Minich C.	5D/594
Mitchell S.R.	4C/420
Mo Ming	2A/152
Mosemann H.	3D/329
Mueller U.	7C/749
Nagasawa Isao	5A/504
Nagasawa Isao	7A/691
Naqasha R.R.	9D/1033
Narendran T.T.	7C/767
Ng Johnny K.C.	8B/838
Ng K.H.	5B/548
Ng Micky W.M.	4C/434
Noorani R.	8C/884
Norrish John	4C/443
Ouriachi K.	3D/323
Owen S.	8A/826
Park Ji-Hyung	5C/577
Park Myon-Woong	5C/577
Perrin E.	5D/594
Petuelli G.	7C/749
Pickett N.J.	8A/826
Pirron J.	4B/395
Poinsignon C.	5D/594
Poiraudeau J.F.	7B/743
Porter J.M.	7A/697
Price J.W.H.	1B/95

Author/Session/Page Index

Qin Caigong	10A/1057
Qin Kaihuai	6D/665
Qin Zheng	6C/647
Rao K.P.	5D/600
Rao K.P.	7C/755
Rautenstrauch Claus	4B/401
Ravi Raju K.	4A/353
Raviart D.	9C/990
Reddy S.V. Bhaskara	7C/767
Reisch O.	4B/395
Ribeiro Fernando	4C/443
Ritchie J.M.	5C/583
Rohrdanz F.	3D/329
Saad S.M.	K/9
Salmon J.C.	10C/1087
Sarat Babu A.L.	4A/353
Schoeller Anna-Maria	9A/941
Schoene Chr.	8D/909
Senechal O.	9C/990
Shaw M.H.	10B/1063
Shi C.	8B/857
Shi Cheng Fu	10D/1114
Shunmugam M.S.	7C/767
Shunmugam M.S.	11A/1131
Simmons J.E.L.	5C/583
Sing W.M.	5D/600
Spina R.	7C/773
Stemart I.	5D/594
Subrahmanyam V.H.R.	8B/845
Sui Guanghua	1C/117
Sun Jian	4D/461
Susanto S.	1B/95
Swaelens Bart	9D/1021
Szczerbicki Edward	8A/809
Szczerbicki Edward	11A/1119
Sze W.S.	4A/374
Sze W.S.	10D/1097
Tahon C.	9C/990
Tan Jianrong	2A/158
Tan K.H.	8A/832
Tan S.T.	4A/341
Tan S.T.	4A/374
Tan S.T.	4C/434
Tan S.T.	8D/921
Tan S.T.	10D/1097
Tan S.T.	10D/1106
Tan Xiaochun	9C/1015
Tang L.	5A/517
Tang Qian	1A/56
Tang Xiaoqi	9D/1027
Tang Zhiwen	2B/176
Tewari N.K.	4A/353
Timo Yla-Mononen	6C/659
Tricarico L.	7C/773
Triouleyre Jean	9C/996
Tsang N.K.	8D/921
Tse W.C.	10A/1045
Tse W.C.	11C/1185
Tseng Henry C.	9B/971
Tu Yiliu	11A/1125
Tuttle R.	7D/797
Umeda Masanobu	5A/504
Umeda Masanobu	7A/691
van den Berg Bert	6B/635
Van Zeir G.	2C/188
Vancraen Wilfried	9D/1021
Voelcker H.B.	K/30
Wahl F.M.	3D/329
Wallbank J.	10B/1069
Walo Michael L.	8C/868
Walton D.J.	2D/214
Wan Li	4D/474
Wang C.C.	10B/1063
Wang C.H.	1A/50
Wang Chenpei	6B/641
Wang D.W.	5B/542
Wang Dehua	11C/1197
Wang Dingwei	2B/176
Wang Fengyin	6D/680
Wang Jia Ye	10D/1114
Wang Jia Ye	11D/1201
Wang Jianguo	1C/122
Wang Jiaye	1A/60
Wang Joshua F.	8B/857
Wang Kai	6D/672
Wang L.	3A/225
Wang L.P.	11B/1161
Wang L.P.	11B/1167
Wang Pingjiang	9D/1027
Wang W.	6D/665
Wang Wei	3D/311
Wang Wenping	11D/1201
Wang Xiankui	4D/461
Wang Xinhua	4D/480
Wang Xuefu	11D/1213
Wang Y.G.	8C/874
Wang Y.Z.	7C/761
Wang Yongjie	2A/164
Wang Yue-juan	1D/146
Wang Yungan	8C/890
Wei Baogang	6C/647
Wei Xiuting	2A/158
Wohlers Terry T.	K/1
Wong T.	10D/1097

Wong T.N. 5B/548
Wong T.N. 7D/779
Wong T.N. 7D/785
Wong T.N. 9A/953
Wong T.N. 10D/1106
Wu B. 3B/278
Wu B. 9B/978
Wu Bin 9B/965
Wu Bo 6D/680
Wu F 2B/182
Wu Hsien-Jung 3C/306
Wu Jieyi 1B/83
Wu Liangwei 1C/111
Wu M.J. 10B/1063
Wu Yunxing 2D/220
Wu Z. 1A/66
Xiao D. 7A/697
Xiao T.Y. 2B/182
Xiao X. 8B/857
Xiao Yuejia 1C/105
Xie Lin 5C/571
Xiong M.H. 8B/863
Xu Baodong ·3B/255
Xu Bo 3B/255
Xu J.X. 1B/89
Xu Shixin 2A/164
Xu Yanshen 4D/467
Xue Hongyuan 3B/260
Yamaguchi Hideyuki 5A/504
Yamaguchi Hideyuki 7A/691
Yan Jianxin 6C/652
Yan Yongnian 1C/111
Yan Yongnian 1C/117
Yan Yongnian 1C/122
Yan Yongnian 3C/284
Yang Daniel C.H. 11B/1155
Yang Jingbin 7A/703
Yang M.Y. 7B/727
Yang Qinghao 7B/721
Yang Shuzi 6D/680
Yang Zhi-Xin 7D/791
Yeap T. 2C/202
Yeung Eddie H.H. 11A/1125
Yeung M. 2D/214
Yin Guofu 8A/821
Yip Y.C. 7A/685
Yoo Il-sang 4C/449
Yu K.M. 7A/685
Yu Kai-Ming 4A/347
Yu Q. 1A/50
Yu Yong 5D/588
Yuan M.L. 1A/50
Yuen M.M.F. 1A/66
Yuen M.M.F 3A/249
Yuen M.M.F 4D/486
Yuen M.M.F. 8A/832
Zakari Abdelouahed 9A/934
Zeng Guang 1C/111
Zhan Yong 7B/715
Zhang C. 4D/492
Zhang Caiming 6D/672
Zhang Jun 11A/1125
Zhang Keshu 9C/1015
Zhang P. 9B/959
Zhang Ping 4D/461
Zhang Renji 1C/111
Zhang Renji 1C/117
Zhang Renji 3C/284
Zhang S. 4B/414
Zhang S.G. 11C/1179
Zhang Tie-Chang 5D/588
Zhang Wei 1C/111
Zhang Wenzu 6D/680
Zhang Xianglin 3C/290
Zhang Z.M. 1B/89
Zhang Zhihua 5A/504
Zhang Zhihua 7A/691
Zhao Y.F. 10D/1106
Zheng Detao 4D/461
Zhong Y.F. 8B/863
Zhong Yifang 3B/260
Zhong Yifang 4D/474
Zhou J. 1A/50
Zhou J. 8B/863
Zhou Jack G. 8C/894
Zhou Ji 2B/170
Zhou Ji 3A/249
Zhou Ji 3B/260
Zhou Ji 4D/474
Zhou Ji 6B/641
Zhou Ji 7B/715
Zhou Ji 7D/791
Zhou Ji 9D/1027
Zhou Xinjian 2A/152
Zhou Yanhong 6B/641
Zhou Yanhong 7B/715
Zhu Guowang 4D/461
Zhu Jun 1C/122
Zhu Lian 5C/560
Zhu Wei-Ming 4A/347
Zuo M. 2C/202

X08316248

UNIVERSITY OF HONG KONG

670.427 I61 p97
International Conference on
Manufacturing Automation (1997
: Hong Kong)
Proceedings of the
International Conference on
Manufacturing Automation,

ISBN: 962-85138-1-8 (1 Set)
ISBN: 962-85138-2-6 (Vol. 1)
ISBN: 962-85138-3-4 (Vol. 2)