Different Approach to Designing Neural Network for Similar Handwritten Chinese Character Recognition

Devil H.F. YIP and William W.H. YU
Department of Industrial and Manufacturing Systems Engineering
The University of Hong Kong
Pokfulam Road, Hong Kong

ABSTRACT

The input images of Chinese characters are normally preprocessed using different image processing techniques before the main classification in the handwritten Chinese character recognition. The authors proposed a different approach to the system philosophy of solving the handwritten Chinese character recognition problem where no preprocessing is necessary. The Chinese characters are treat as ideographs. The proposed system consists of a Rough Classifier which trigger the different Fine Classifiers. Each classifier is an optimized artificial neural network using genetic algorithms. A reduced system has been implemented. The result shows that the proposed system has higher recognition rate than the similar systems reported and is more efficiency.

Keywords: Handwritten Chinese Character Recognition, Neural Networks, ETL9B

1. INTRODUCTION

It is much more difficult to learn Chinese characters than the characters of the western languages. The reason is the huge set of Chinese character with up to forty-eight strokes and a lot of Chinese characters have the same radical by vary in number of strokes. As learning the Chinese characters is a formidable task for a person, it is better to develop the computer system for Chinese character recognition (CCR). The handwritten Chinese characters are much more difficult than the automatic recognition of the printed Chinese characters because of the different written styles. An automatic CCR system, which is fast and has a high success rate, will greatly simply and accelerate converting a handwritten Chinese document into a printed document.

In 1966, Casey and Nagy [1] at IBM reported one of the first attempts at CCR for a set of about 1,000 printed Chinese characters. Until now, a large number of papers and reports have been published on this topic. The commercial system for handwritten and printed English character recognition systems, off-line and on-line printed CCR systems, and on-line handwritten CCR systems have been well developed. The research of off-line handwritten CCR systems is still going on.

Generally, the techniques of Character recognition are grouped into three classes. The first class adopts global transformation and series expansion. Fourier transform [2], Hadamard and Harr transform [3], Walsh transform [4], Hough transform and code transform [5] are the examples of this class. Although, these transformations are generally computationally demanding but easy to implement. The second class compares features obtained from the spatial distribution of pixels. The examples are cross counts, loci features [6] and moments [7]. These methods are fast but difficult to implement.

Third class matches the geometrical or topological features such as strokes and bays in various directions, end points and loops [8]. These techniques are complex in their implementations requiring long processing time but powerful. Extractions of the geometrical or topological features of the character are time consuming. Artificial neural networks (ANNs) for character recognition have been done by some researchers [9-11]. Nevertheless, most of the tests are completed in small scales only. It is hard to achieve high recognition rate to handwritten Chinese characters with these approaches.

Part of the SPIE Conference on Applications and Science of Neural Networks, Fuzzy Systems, and Evolutionary Computation • San Diego, California • July 1998
SPIE Vol. 3455 • 0277-786X/98/$10.00

The development of effective handwritten Chinese character recognition (HCCR) system for about 50,000 Chinese characters accurately is still a challenge. A reasonable commercial system should be able to recognize at least 5,000 common characters. Several researches have developed the high recognition rate HCCR systems [12-16]. ANNs are implemented as the rare end of the systems. The front end of the systems comprises of some computer vision systems preprocessing techniques. The results of the preprocessing stages are quantified and fed to the input nodes of the ANN. Senda et al., [12] used a pattern recognition method, minimum distance classifier, to calculate the eigenvectors of the prototypes.

Kato and Nemoto [13] use smoothing, noise reduction and normalization in their preprocessing stage. Guo et al., [14] include non-linear normalization, 3 x 3 mask smoothing and extraction of the contour in the preprocessing stage. Saruta et al., developed the HCCR systems, ELNET [15] and ELNET-II [16], where input data first undergo non-linear normalization and outline extraction prior to the training of ANN. The preprocessing of data may corrupt information vital for differentiating between two Chinese characters with similar strokes and thus render the recognition network less effective. Furthermore, time must be spent to process every character presented to the system before the feature extraction process.

In this research, the authors proposed approach [17] a new system for HCCR which differs significantly from the methodology investigated by the other researchers. The proposed system does not use the preprocessing stages the others used. Genetic algorithms (GAs) are adopted to determine the optimal parameters of the ANN topology of the adaptive back-propagation network for similar Chinese character recognition under the specified conditions. The ETL9B database provided by the Electro-Technical Laboratory from Japan is used to evaluate the performance of the proposed system. Compare with the work done by Saruta et al. Section 2 consists of the system philosophy of the proposed HCCR system. Section 3 presents the network architecture and the training algorithm. Section 4 describes the training and test results of the proposed system. In section 5, conclusions and future works are given.

2. SYSTEM PHILOSOPHY

In another paper, the authors explained the philosophy of the proposed system and have implemented successfully a reduced system to learn and recognized a specific group of Chinese characters. This paper reports the experiment and results obtained from the same reduced system with a different group of Chinese characters. For the sake of the completeness of the paper, the system philosophy is reiterated below.

An ANN mimics the ability of a person to recognize complex patterns without having to perform complex analysis of the pattern. A child learning to recognize complex patterns such as Chinese characters generally do not rely on any high level analytical skills. A child has this ability before acquiring any formal lessons or schooling. The authors argue that using image processing technique at the front end of a HCCR system somewhat prevents the ANN to mimic a person's full ability to recognize complex patterns. When a child goes to school, the analytical skills are very important. However, to implement a person's analytical skills in a HCCR system means enormous efforts in deriving complex equations, algorithms and other sophisticated and time consuming techniques. The results may not always be encouraging in proportion to the efforts made.

The proposed system mimics how child learns Chinese characters and does not use the above preprocessing techniques. The system treats the Chinese characters as ideographs. The authors also argue that when a child learns the characters, he or she somewhat classifies the characters into the simpler and the more complex categories without invoking any critical and analytical procedures. This primitive mental classification process facilitates an efficient recognition process. The mental search for matches can be directed towards the simpler or the more complex set of characters as the case may be. Thus the system philosophy comprises of the following components.

1. A person has the ability to learn and recognize a large set of patterns such as Chinese characters without using more complex analytical skills.

2. A person tends to classify the patterns into simple and complex categories of various degrees. The boundary between any two categories varies from person to person and is not analytically derived.
3. A person tends to classify a large set of Chinese characters according to various degrees of similarities in terms of general outlines, shapes, portions of the characters, or radicals. Again these classifications need not be the result of some critical examinations of the characters. The results from such classification exercises may differ tremendously with the results produced by a learned individual.

This paper investigates the proposed philosophy using a reduced system. The reduced system assigned a limited function to the Rough Classifier. It classifies the data set in terms of their radicals. The characters to be classified are those having the character "言" on the left side of the characters. Examples of these set of characters are: { 譕, 謶, 託, 詑, 誌, 認 }.

Figure 1 illustrates the conceptual architecture of the ANN HCCR system based on the system philosophy above.

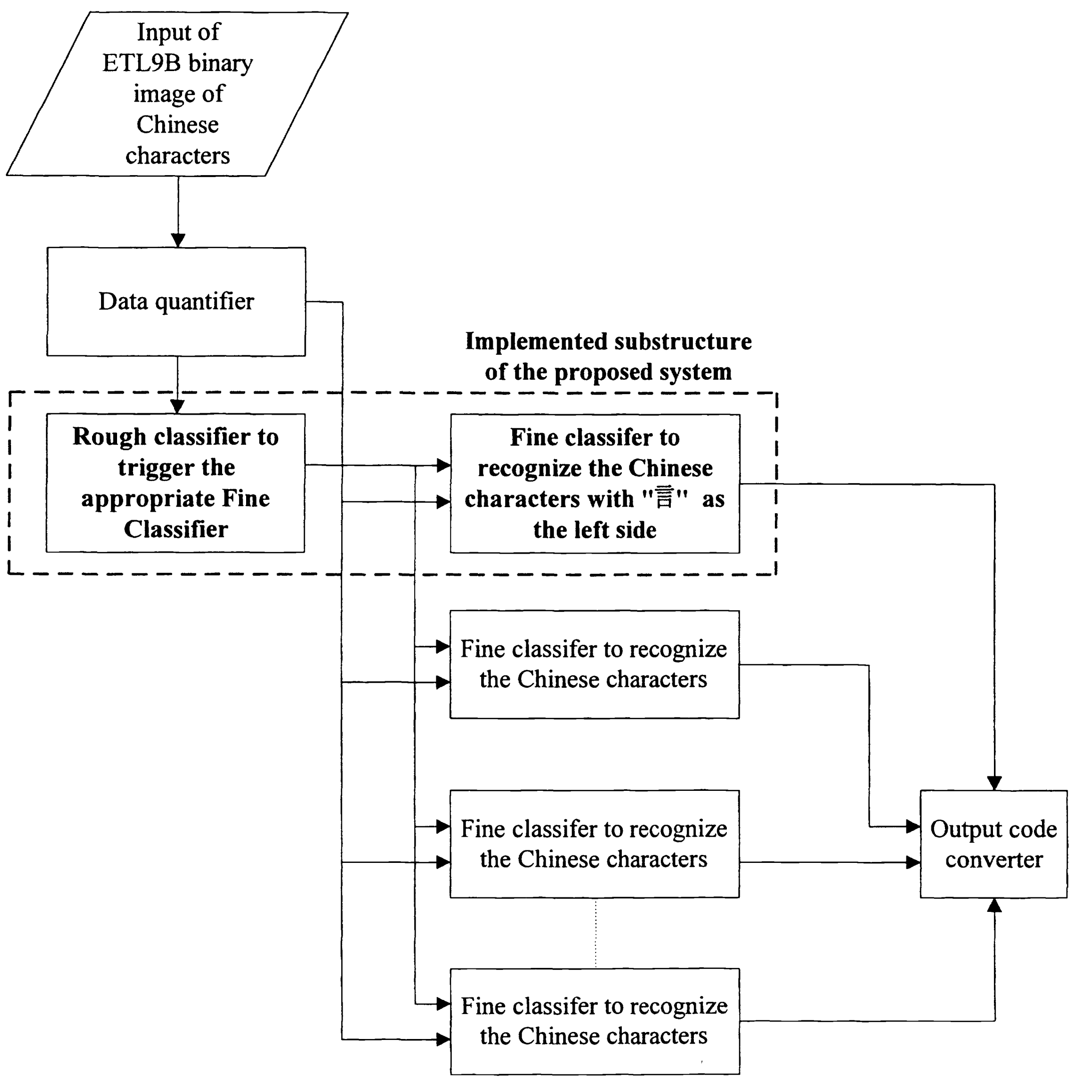

Figure 1. The proposed HCCR system

3. NETWORK ARCHITECTURE AND TRAINING ALGORITHM

To implement and test a substructure of the proposed ANN HCCR system, the test system comprises of a Rough Classifier and a number of Fine Classifiers. The input data are the pixel binary output of the character presented to the input nodes of all the classifier networks simultaneously. Each of these classifiers is a stand-alone ANN and comprises of input nodes, hidden layers and output nodes. In addition to the input data the Fine Classifiers input nodes also receive signals from the Rough Classifier. The Rough Classifier first groups the similar characters into the same class according to their pixel density distribution.

The output data of the Rough Classifier represent complexity of strokes, the lexicographical radicals, the spatial distribution of strokes and the overall shape of the character. The weighted output from the Rough Classifier to the Fine Classifiers trigger or inhibit the Fine Classifiers responses to the input data. Then, the triggered Fine Classifier categorizes the character and the output will produce the relevant Chinese character code specified by the system (such as Big 5) through an appropriate converter. Each Fine Classifier learns and recognizes a subset of the character set. The workloads of the Fine Classifiers are thus divided (not necessarily evenly) amongst them. This approach greatly improves the efficiency of the system.

One of the critical problems with using ANNs is in designing the network topology for a given problem such as character recognition. Typically the network topology (number of hidden units, number of hidden layers and the interconnects) and the parameters in the network (learning rate, momentum, weight decay, initial weight, etc.,) are determined through trials and experiments which could often be long and tedious in design considerations. The temptation is to accept any solutions which may not be optimal. In some situations, the problem may appear insolvable after a long process of trials and experiments for the right combination of parameters has not been selected.

The authors have shown that genetic algorithms (GA) are suitable tools for determining the optimal topology and parameters of ANN [17-19]. GAs have the advantage of being less likely to get struck in a local optima. When there is no a prior knowledge about the parameter space and the search space is large, the optimization process may take a long time to converge.

The back-propagation network has been adopted because it is capable of approximating arbitrary mapping given a set of examples and adaptive back-propagation (ABP) technique [20] is used to achieve fast convergence during the training process. Since the authors are using the ABP for training, the ANNs are restricted to a feed-forward multi-layer perception (MLP). The MLP may be randomly or fully connected. In fully connected MLP, every node in a layer is connected only to all the nodes in a subsequent layer.

GAs have been used to determine the most effective number of hidden layers, number of hidden nodes per layer and interconnects. For each j^{th} genetic string or chromosome, the training of the network produces a fitness value given by:

$$f_j = \frac{1}{\sum_{i=1}^{8} W_i F_i(X_{ij})}$$

Where W_i is the significance of weighting factor for the i^{th} parameter, F_i is a function for the i^{th} parameter and, X_{ij} is the i^{th} parameter value.

The parameters of the fitness values included are: the overall error, the maximum number of nodes in the network, the number of links in the network, the number of epochs required to train the network, maximum number of epochs specified, the number of misclassified patterns during training/recall, the number of nodes out of the tolerance limit during training/recall, and the total number of training/recall patterns. Each gene represents a specific network topology and the fitness value indicates its performance.

Crossover and mutation follows each training process and the genes compete for survival based on their fitness values. The optimization process is subject to some specified conditions. The network of the system has a maximum of 2 hidden layers each with a maximum of 30 hidden nodes; 50 generations with population size 50; uniform crossover probability 0.6 and mutation probability 0.01 (based on the recommendation of DeJong [21]).

4. TRAINING AND TEST RESULTS

The training and testing of this two-stage classifier were conducted on a Pentium based personal computer. The character database ETL9B, which was prepared by Electro-Technical Laboratory from Japan, contains 3036 categories (2965 handwritten Chinese characters and 71 handwritten Japanese characters) in 200 difference patterns per character. In principle, in larger the training set is, the more robust the network. The authors used the same training and testing conditions of the experiment done by ELNET [15] and ELNET-II [16], and cross validation in the experiment. The data available are in set of 200 characters. The experiment used 257 sets of characters. All the characters have the radical "言" on the left side. Each character set comprises of 200 samples and 180 samples per character set were used for cross validation. The input data were 64 x 63 binary patterns. Each was subdivided into 8 x 8 square matrices. Each square matrix was a 10 x 10 pixels array overlapping its neighing square by 2 pixels width.

Group	Unknown test data sets	Training data sets	Recognition rate
1	Samples 1 to 20	Remaining 180 sets	99.06 %
2	Samples 21 to 40	Remaining 180 sets	98.58 %
3	Samples 41 to 60	Remaining 180 sets	98.94 %
4	Samples 61 to 80	Remaining 180 sets	98.65 %
5	Samples 81 to 100	Remaining 180 sets	98.66 %
6	Samples 101 to 120	Remaining 180 sets	98.50 %
7	Samples 121 to 140	Remaining 180 sets	98.40 %
8	Samples 141 to 160	Remaining 180 sets	98.59 %
9	Samples 161 to 180	Remaining 180 sets	98.52 %
10	Samples 181 to 200	Remaining 180 sets	98.47 %
Average recognition rate			98.64 %

Table 1 Recognition results of the proposed system with 180 sets data in training

The average recognition rate of the proposed system is 98.64% with a fully connected optimal topology of 64 input nodes, 22 hidden nodes in the first hidden layer, 18 hidden nodes in the second hidden layer and 16 binary output under the specified conditions. The result is better than the ELNET (94.54%) and ELNET-II (95.84%) with the same character database. The authors have also conducted an experiment using only 100 samples per character set for training and 100 for tests. The average recognition rate is 96.12% which compares well against ELNET and ELNET-II which used a larger training sample set of 180.

5. CONCLUSIONS AND FUTURE WORKS

Though the experiments were conducted on a reduced system, the result shows that the proposed system philosophy is implementable, effective and of high performance. The authors in another paper have reported the results of recognition of another group of characters on the same reduced system. All the results show that the full-scale systems can be implemented and equally successful results can be expected. The implementation of the full-scale systems is currently underway.

REFERENCES

1. R. Casey and G. Nary, "Recognition of Printed Chinese Characters", *IEEE Transactions on Electronic Computers*, **15,** pp. 91 - 101, 1966.
2. S. Wendling and G. Stamon, "Hadamard and Harr Transforms and Their Power Spectrum in Character Recognition", *Joint Workshop on Pattern Recognition and Artificial Intelligence*, Hyannis, USA, pp. 103-112, 1976.
3. J.S. Huang and M.L. Chung, "Separating Similar Complex Chinese Characters by Walsh Transform", *Pattern Recognition*, **20**, pp. 425-428, 1987.
4. Y.S. Cheung and C.H. Leung, "Chain-Code Transform for Chinese Character Recognition", *IEEE Proceedings of International Conference on Cybernetics Society*, Tusuon, USA, pp. 42-45, 1985
5. A.L. Knoll, "Experiments with Characteristic Loci for Recognition of Handprinted Characters", *IEEE Transactions on Computer*, **18**, pp. 366-372, 1969
6. N.D. Tucker and F.C. Evans, "A Two Step Strategy for Character Recognition using Geometrical Moments", *Second International Joint Conference on Pattern Recognition*, pp. 223-225, 1974.
7. M. Shridar and A. Bareldin, "High Accuracy Character Recognition Algorithm using Fourier and Topological Descriptors", *Pattern Recognition*, **17**, pp. 515-523, 1984.
8. S. Mori, K. Yamamoto, H. Yamada and T. Saito, "An a Handprinted KYOIKU-KANJI Character Data Base", *Bull. Electrotech. Lab. (Japan)*, **43**, pp. 752-773, 1979.
9. I. Guyon, P. Albrecht, Y. LeCun, J. Denker and W. Hubbard, "Design of a Neural Network Character Recognizer for a Touch Terminal", *Pattern Recognition*, **24**, pp. 105-120, 1991.
10. I.C. Jon, S.S. Yu and S.C. Tsay, "A New Feature Extraction Method by Neural Networks", *IEEE International Symposium on Circuits and Systems*, **4**, pp. 3249-3252, 1990.
11. A. Khotanzad and J.H. Lu, "Distortion Invariant Character Recognition by a Multi-Layer Perceptron and Backpropagation", *IEEE International Conference on Neural Networks*, **1**, pp. 625-632, 1988.
12. S. Senda, M. Minoh and I. Katsuo, "A Fast Algorithm for the Minimum Distance Classifier and Its application to Kanji Character Recognition", *The Third Conference on Document Analysis and Recognition*, Montreal, Canada, **1**, pp. 283 - 286, 1995.
13. N, Kato and Y. Nemoto, "Large Scale Hand-Written Character Recognition System using Subspace Method", *IEEE Conference on Systems Man and Cybernetics*, Beijing , China, **1**, pp. 432 - 437, 1996.
14. J. Guo, R. Sato, N. Sun and Y. Nemoto, "Recognition of Handwritten Character Database ETL9B using Pattern Transformation Method", *Systems and Computers in Japan*, **25**, pp. 58 - 67, 1994.
15. K. Saruta, N. Sun, M. Abe and Y. Nemoto, "Recognition of ETL9B using Exclusive Learning Neural Network (ELNET)", *IEICE Technical Report*, **PRU94-93**, 1994.
16. K. Saruta, N. Kato, M. Abe and Y. Nemoto, "High Accuracy Recognition of ETL9B using Exclusive Learning Neural Network - II (ELNET - II)", *IEICE Transactions on Information and Systems*, **79D**, pp. 516 - 522, 1996.
17. D.H.F. Yip and W.W.H. Yu, "Novel Design of Neural Networks for Handwritten Chinese Character Recognition", to be published on the *Conference of Geometry VII*, SPIE's International Symposium on Optical Science, Engineering, and Instrumentation, 1998.
18. D.H.F. Yip and W.W.H. Yu, "Classification of Coffee using Artificial Neural Network", *IEEE International Conference on Evolutionary Computation*, Nagoya, Japan, pp. 655 - 658, 1996.
19. D.H.F. Yip, E.L. Hines and W.W.H. Yu, "Application of Artificial Neural Networks in Sales Forecasting", *IEEE International Conference on Neural Networks*, Houston, USA, **4**, pp. 2121 - 2124, 1997.
20. F.M Silva and L.B. Almeida, "Acceleration Techniques for the Backpropagation Algorithm", *Lecture Notes in Computer Science: Part 412*, Springer-Verlag, 1990.
21. K.A. DeJong, "Analysis of the Behavior of a Class of Genetic Adaptive Systems", Ph.D. Dissertation, Department of Computer and Communication Sciences, University of Michigan, Ann Arbor, MI, 1975.