Proceedings of the American Control Conference
Philadelphia, Pennsylvania • June 1998

Nonlinear Observer Design with Unknown Nonlinearity via B-spline Network Approach

H. Y. Zhang†, C. W. Chan‡, K. C. Cheung‡, H. Jin†
† Department of Automatic Control, Beijing University of Aeronautics and Astronautics
Beijing 100083, P. R. China, Email: buaa301@mimi.cnc.ac.cn
‡ Department of Mechanical Engineering, University of Hong Kong
Hong Kong Pokfulam road, Email: mechan@hkucc.hku.hk

Abstract: A novel approach is proposed to the state estimation of a class of nonlinear systems which consist of known linear part and unknown nonlinear part. A linear observer is first designed then a nonlinear compensation term in the nonlinear observer is determined using the proposed "deconvolution method". The B-spline neural network is used to model the estimated compensation term. Three simulation examples are given to compare the effectiveness of the proposed approach and some analytical approaches.

Keywords: nonlinear system, observer, neural network

1. Introduction

Estimating the states of nonlinear systems is important to control, supervisory and fault diagnosis problems. There are some research results on the design of observer of nonlinear system using analytical approach in the publications [1~4]. The Canonical Observer Design developed by Bestle, Zeitz [1], Krener and Respondek [2] uses Lie-algebraic methods to transform the nonlinear plant into observer canonical form from which the design of an observer is facilitated. However to find an appropriate nonlinear one to one transformation is highly nontrivial. Moreover, the nonlinear model of the plant must be known exactly. Baumann and Rugh developed another approach [3] in which an extended linearization technique is utilized to produce an observer. Here again the exact knowledge of the nonlinearities and, further, the first derivative of these nonlinearities must be known in order to calculate the gain function of the observer. Walcott and Zak proposed a variable structure method [4] which only uses the bounds of nonlinearity of the plant in observer dynamics. However the bounds may not be known exactly in practice. Comparative study of above methods is given in [5]. Many problems of designing nonlinear observer remain to be solved. Especially, when the model of the nonlinearity is unknown or does not satisfy the conditions required by the above methods.

Neural network has proven to be an universal approximator. It can successfully approximate nonlinear function. However, only few papers have applied neural network approach to the state estimation of nonlinear systems [6], [7] as compared with the neural approach to the system modeling [8]~[11]. Although they are similar in terms of the function approximation, the state estimation problem is more difficult than the modeling problem. It requires the observability of nonlinear system which is hard to check.

In this paper, the B-spline neural network is applied to the state estimation of nonlinear systems with unknown nonlinearity and known linear part. The nonlinear observer consists of a linear observer and a nonlinear compensation term. The compensation term is determined using a proposed "deconvolution method". Then, it is modeled using B-spline neural network. Simulation examples of the proposed approach are given to illustrate the effectiveness.

The rest of the paper is organized as follows: In second section, a novel design procedure of nonlinear observer based on neural network is presented; In third section, the estimation of unknown nonlinear compensation term in the nonlinear observer using "deconvolution method" is developed; In forth section, B-spline neural network is used to model the nonlinear compensation term; In fifth section estimation error of the Neural Network observer is analyzed; In sixth section simulation examples are given, the results of the proposed approach and some analytical approaches are compared; The last section is the conclusion.

2. New design procedure of nonlinear neural network based observers

(1) System

Consider a class of nonlinear systems described by

$$x(k+1) = Ax(k) + f(x(k)) \qquad (1)$$

$$y(k) = Cx(k) \qquad (2)$$

where A is the linear part of the system; $f(x)$ is nonlinear part of the system. The measurement equation is linear. We assume the linear part of the system is known and observable, i.e., (A,C) is a observable pair. The nonlinear part $f(x)$ is unknown and not necessary smooth.

(2) Design procedure

(a) Firstly, we design a linear observer for the linear part of the system as follows:

$$\hat{x}_l(k+1) = A_0\hat{x}_l(k) + Ky(k) \qquad (3)$$

$$A_0 = A - KC \qquad (4)$$

where K is observer gain, which can be designed using existing methods, and $\hat{x}_l$ is the state of linear observer. The nonlinear observer has the following form

$$\hat{x}(k+1) = A_0\hat{x}(k) + Ky(k) + S(k) \qquad (5)$$

$$\hat{y}(k) = C\hat{x}(k) \qquad (6)$$

where S is an unknown nonlinear compensation term which is due to the existence of nonlinear part in (1).

(b) Secondly, we estimate the unknown nonlinear compensation term S using a " deconvolution procedure". Since the nonlinear observer output $\hat{y}$ can be expressed as a convolution relation with the nonlinear compensation term, by forcing the observer output $\hat{y}$ to equal the system output y, the unknown nonlinear compensation term can be determined.

(c) Finally, using neural network mapping ability we can model the nonlinear compensation term as a nonlinear function of some inputs selected based on the knowledge of the system nonlinearity, or by trial and errors.

3. Determination of the nonlinear compensation term using "deconvolution procedure"

The idea is to force $\hat{y}(k)$ in (6) to track y(k), in doing so we can determine the compensation term. From (5) and (6) we have

$$\hat{y}(k) = C\hat{x}(k) = C[A_0\hat{x}(k-1) + Ky(k-1) + S(k-1)] = \cdots$$
$$= CA_0^k\hat{x}(0) + \sum_{i=1}^{k} CA_0^{i-1}Ky(k-i) + \sum_{i=1}^{k} CA_0^{i-1}S(k-i) \quad (7)$$

$$\tilde{y}(k) = y(k) - \hat{y}(k) = y^*(k) - \sum_{i=1}^{k} CA_0^{i-1}S(k-i) \quad (8)$$

$$y^*(k) = y(k) - CA_0^k\hat{x}(0) - \sum_{i=1}^{k} CA_0^{i-1}Ky(k-i) \quad (9)$$

$\hat{x}(0)$ can be assigned arbitrary, except to satisfy $C\hat{x}(0) = \hat{y}(0) = y(0)$, therefore $y^*(k)$ is known. Let $\tilde{y}(k) = 0$ and $C_i = CA_0^{i-1}$, *k=1,2,...,M+1*. From (8) we can obtain

$$\begin{cases} C_1S(0) = y^*(1) \\ C_1S(1) + C_2S(0) = y^*(2) \\ \vdots \\ C_1S(M) + \cdots + C_{M+1}S(0) = y^*(M+1) \end{cases} \quad (10)$$

when $rank(C) = n$, and $p \geq n$ ($C \in R^{p\times n}$), the solution of *S(k)* is

$$\begin{cases} S(0) = C^+y^*(1) \\ S(k) = C^+\left[y^*(k+1) - \sum_{i=0}^{k-1} C_{k+1-i}S(i)\right] \end{cases} \quad (11)$$

where $C^+ = (C'C)^{-1}C'$, $k = 1, \cdots, M$

When *S(k)* can be expressed as

$$S(k) = \begin{bmatrix} 0 \\ \bar{S}(k) \end{bmatrix}$$

where $\bar{S}(k) \in R^{d\times n}$. C can be partitioned accordingly as $C = [\tilde{C} \quad \bar{C}]$, where $\bar{C} \in R^{p\times d}$.

When $rank(\bar{C}) = d$, $p \geq d$, we can obtain a solution for $\bar{S}(k)$ as

$$\begin{cases} \bar{S}(0) = \bar{C}^+y^*(1) \\ \bar{S}(k) = \bar{C}^+\left[y^*(k+1) - \sum_{i=0}^{k-1} C_{k+1-i}S(i)\right] \end{cases} \quad (11)'$$

where $\bar{C}^+ = (\bar{C}'\bar{C})^{-1}\bar{C}'$, $k = 1, \cdots, M$

4. Modeling nonlinear compensation term using neural network

In this paper , we will adopt the B-Spline network to model the nonlinear compensation term. The advantages of using such network are as follows. In the feedforward network (BP network), the output is a complicated nonlinear function of weights w_{ij}^s. The performance index for training the network is

$$I = \sum_k \| y(k) - d(k) \|^2 = \sum_k \| e(k) \|^2 \quad (12)$$

which is highly nonlinear. That means there may be several local minimums associated with the index function *I*. In contract, the output of a B-spline network[12] is given by

$$y(k) = \sum_{i=1}^{p} a_i(k)w_i \quad (13)$$

where w_i is the weight corresponding to the ith basis function $a_i(k)$, its input is $\hat{x}(k)$, p is the number of basis functions. (12) is quadratic in the unknown weights, hence a global minimum exists.

If nonlinear function has several augments, multivariate basis function is needed for approximation. Multivariate basis functions are formed by taking the tensor product of the univariate basis functions. The network output is linearly dependent on these multivariate basis functions.

The estimated nonlinear compensation term is taken as the desired network output while the network input can be selected based on the knowledge of the system nonlinearity or by try and error method .

Assuming input variable is $\hat{x}(k)$, and compensation term *S(k)* is a scalar function(if *S(k)* is a vector, we can model its elements similarly), the network modelling can be expressed as

$$\begin{cases} S(1) = \sum_{i=1}^{p} a_i(\hat{x}(1))w_i + \varepsilon(1) \\ \vdots \\ S(N) = \sum_{i=1}^{p} a_i(\hat{x}(N))w_i + \varepsilon(N) \end{cases} \quad (14)$$

$\varepsilon(k)$ is the modelling error of network, *k=1,...,N*. (14) can be written in vector-matrix form:

$$S = \bar{A}W + E \quad (15)$$

where $S' = [S(1) \ \cdots \ S(N)]$, $W' = [w_1 \ \cdots \ w_p]$,

$$\bar{A} = \begin{bmatrix} a_1(1) & \cdots & a_p(1) \\ \cdots & \cdots & \cdots \\ a_1(N) & \cdots & a_p(N) \end{bmatrix}, \quad E' = [\varepsilon(1) \ \cdots \ \varepsilon(N)]$$

Normally, the number of equations is greater than the number of unknown weights, i.e., *N>p*, and assuming $\bar{A}$ has full column rank, therefore Least-Squares method can be used to determine the weights as follows:

$$W = (\bar{A}'\bar{A})^{-1}\bar{A}'S \quad (16)$$

5. Estimation error analysis

The equation of neural network estimator can be expressed as

$$\hat{x}(k+1) = A_0\hat{x}(k) + Ky(k) + \hat{f}(\hat{x}(k)) \quad (17)$$

where $\hat{f}(\hat{x}(k))$ is the output of network. From (5) and (24) we have

$$e(k+1) = \hat{x}(k+1) - x(k+1)$$
$$= A_0e(k) + [\hat{f}(\hat{x}) - f(x)] \quad (18)$$
$$= A_0e(k) + \Delta\hat{f}(k) + \varepsilon(k)$$

where $\Delta\hat{f}(k) = \hat{f}[x(k) + e(k)] - \hat{f}[x(k)]$,

$\varepsilon(k) = [\hat{f}(x) - f(x)]$

Since A_0 is a Hurwitz matrix, then given any positive definite real symmetric matrix Q, there exists a positive definite real symmetric matrix P such that

$$A_0' P A_0 - P = -Q \qquad (19)$$

Next, consider the following Lyapunov function

$$V(k) = e'(k) P e(k)$$

For the time being, we assume that the approximation of network is perfect, i.e., $\varepsilon(k) \equiv 0$. From (18) we get

$$\begin{aligned} V(k+1) &= e'(k+1) P e(k+1) = e'(k) A_0' P A_0 e(k) \\ &\quad + 2\Delta\hat{f}'(k) P A_0 e(k) + \Delta\hat{f}'(k) P \Delta\hat{f}(k) \end{aligned} \qquad (20)$$

$$\begin{aligned} \Delta V(k) &= V(k+1) - V(k) = -e'(k) Q e(k) \\ &\quad + 2\Delta\hat{f}'(k) P A_0 e(k) + \Delta\hat{f}'(k) P \Delta\hat{f}(k) \end{aligned} \qquad (21)$$

We assume the output function $\hat{f}$ of the B-spline network satisfies Lipschitz condition, i.e.,

$$\|\Delta\hat{f}(k)\| = \|\hat{f}(x(k)+e(k)) - \hat{f}(x(k))\| \le L\|e(k)\| \qquad (22)$$

L is a positive number. Then we have

$$\begin{aligned} \Delta V(k) &\le -\lambda_{q\min}\|e(k)\|^2 + 2L\lambda_{p\max}\sigma_{\max}\|e(k)\|^2 \\ &\quad + L^2\lambda_{p\max}\|e(k)\|^2 \end{aligned} \qquad (23)$$

where $\lambda_{q\min}$ is the minimum eigenvalue of Q, $\lambda_{p\max}$ is the maximum eigenvalue of P, $\sigma_{\max}$ is the largest singular value of A_0. Therefore, if

$$-\lambda_{q\min} + 2L\lambda_{p\max}\sigma_{\max} + L^2\lambda_{p\max} < 0 \qquad (24)$$

then $\quad \Delta V(k) < 0$

which means $e(k)$ can asymptotically approach zero.

In practice, network approximation will never be perfect, i.e., $\|\varepsilon(k)\| \ne 0$ for all k. In this case, we can give an estimate of upper bound of estimate error.

From (18), we have

$$\begin{aligned} \|e(k+1)\| &\le \|A_0 e(k)\| + \|\Delta\hat{f}(k)\| + \|\varepsilon(k)\| \\ &\le (\sigma_{\max} + L)\|e(k)\| + \|\varepsilon(k)\| \end{aligned} \qquad (25)$$

When $(\sigma_{\max} + L) < 1$, the above difference equation is stable, therefore for sufficient large k, we have

$$\|e(k+1)\| \approx \|e(k)\| \qquad (26)$$

Substituting (26) into (25), we obtain

$$\|e(k)\| \le \frac{\|\varepsilon(k)\|}{1 - (\sigma_{\max} + L)} \quad \text{(for large } k) \qquad (27)$$

Although the above estimation may be very rough and conservative, it does give some physical insight into the problem. From (27), we can see that $\|e(k)\|$ is proportional to $\|\varepsilon(k)\|$ and inversely proportional to $\sigma_{\max}$ and L. It means that smaller network approximation error, smaller $\sigma_{\max}$ and smaller L will give a smaller estimation error.

6. Simulation examples

In this section, nonlinear observers are first given by analytical approaches and then by the proposed approach. The results are compared. Later , the proposed approach is applied to some examples of nonlinear systems to which some analytical approaches can not be applied.

(1) **Example 1**.

System and measurement equations

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -\sin x_1 \end{bmatrix} \qquad (28)$$

$$y = x_1 + x_2 \qquad (29)$$

This example is taken from [5] where four analytical approaches were used. We list the designed observers using these approaches below and compare the results with that of proposed approach.

(a) The Lie-algebraic methods

The designed observer has the following form

$$\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{bmatrix} = \begin{bmatrix} \hat{x}_2 \\ -\sin \hat{x}_1 \end{bmatrix} + \begin{bmatrix} g_1(\hat{x}) \\ g_2(\hat{x}) \end{bmatrix} (y - \hat{x}_1 - \hat{x}_2)$$

where

$$\begin{aligned} g_1(\hat{x}) &= (\cos\hat{x}_1 + 1)^{-3}[\cos^3\hat{x}_1(\hat{x}_2^2 + 3)\cos^2\hat{x}_1 \\ &\quad + (3 + \hat{x}_2^2 - 2\hat{x}_2\sin\hat{x}_1 - \sin^2\hat{x}_1)\cos\hat{x}_1 \\ &\quad + (2\hat{x}_2^2 - 1)\sin^2\hat{x}_1 - 2\hat{x}_2^2\sin\hat{x}_1 + 1] \end{aligned}$$

$$\begin{aligned} g_2(\hat{x}) &= (\cos\hat{x}_1 + 1)^{-3}[\cos^3\hat{x}_1 + (3 + \hat{x}_2^2 + \hat{x}_2\sin\hat{x}_1)\cos^2\hat{x}_1 \\ &\quad + (3 + 4\hat{x}_2\sin\hat{x}_1 + \sin^2\hat{x}_1 + \hat{x}_2^2)\cos\hat{x}_1 \\ &\quad + (2\hat{x}_2^2 - 1)\sin^2\hat{x}_1 + 3\hat{x}_2\sin\hat{x}_1 + 1] \end{aligned}$$

(b) The methods of extended linearization and Tau

The designed observers have the same form as follows

$$\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -\sin\hat{x}_1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} y$$

(c) The Variable Structure System (VSS) technique

The designed observer has the following form:

$$\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \hat{S}(\hat{x}_1, \hat{x}_2, y) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} y$$

where $\hat{s}(\hat{x}_1, \hat{x}_2, y) = -2\,\mathrm{sgn}(\hat{x}_1 + \hat{x}_2 - y)$

The simulation responses of these observers for some initial conditions are similar. In [5], it is claimed that VSS observer gives the best performance. We depict the VSS observer response in Fig. 1. The estimate $\hat{x}_2$ exhibits obvious chattering which is typical for VSS design . Notice that this chattering was not shown in [5]

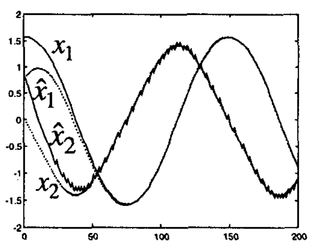

Fig.1: VSS observer design for *f(x)=sinx*

(d). The proposed approach

First, design the linear observer which has the same form as in (b) and (c), then discretize it at a sampling interval 0.05 seconds, and add a nonlinear compensation term $S(k)$ to form a nonlinear observer. The designed discrete nonlinear observer has the following form

$$\begin{bmatrix} \hat{x}_1(k+1) \\ \hat{x}_2(k+1) \end{bmatrix} = \begin{bmatrix} 0.9512 & 0 \\ -0.0476 & 0.9512 \end{bmatrix} \begin{bmatrix} \hat{x}_1(k) \\ \hat{x}_2(k) \end{bmatrix} + \begin{bmatrix} 0.0488 \\ 0.0476 \end{bmatrix} y(k) + \begin{bmatrix} 0 \\ S(k) \end{bmatrix}$$

where the nonlinear compensation term $S(k)$ is determined using "decovolution method", then modeled using B-spline neural network. Based on the knowledge of system's nonlinearity the input of network is taken as $\hat{x}_1$, i.e., $S(k)$ is approximated by network output $\hat{S}(\hat{x}_1(k))$ which is a function of $\hat{x}_1(k)$.If we do not have the knowledge we can try to take $\hat{x}_1$,or $\hat{x}_2$ as input or their suitable combination as input. In the latter case multivariate B-spline function may be needed.

In the simulation, we used 9 univariate second order (linear)B-spline with even supports. The input range of the network is[-2,2]. The simulation responses are shown in Fig. 2. Where the $x_1(k)$, $\hat{x}_1(k)$, $x_2(k)$, $\hat{x}_2(k)$ and $S(k)$, $\hat{S}(\hat{x}_1(k))$ are compared. $\hat{y}(k)$ is not shown, since $\hat{y}(k) = \hat{x}_1(k) + \hat{x}_2(k)$.

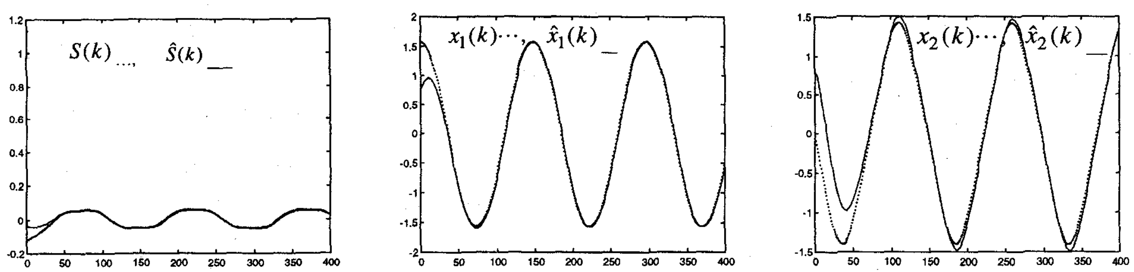

Fig.2: Observer designed by proposed approach for *f(x)=sinx*

As can see from these responses, good estimates are obtained. We should point out that further improvement in estimation accuracy is possible by refinement of the network structure and parameters.

(2) **Example 2**

The linear part is the same as in example 1. the nonlinear part now is a deadzone function

$$f(x) = \begin{cases} k(x-b_2) & x \geq b_2 \\ 0 & b_1 < x < b_2 \\ k(x-b_1) & x \leq b_1 \end{cases} \qquad (30)$$

In the simulation , $k=1, b_1 = -0.8, b_2 = 0.8$.

The Lie-algebraic method and extended linearization method can not be applied to design observer for this kind of system nonlinearity. Although VSS method can be applied to this case principally, it needs to know the upper bound of *f(x)*. In example 1, the nonlinear function is $\sin(x_1)$, therefore $\|\sin x_1\| < 1$, the upper bound is 1. For the deadzone function, the upper bound depends on the range of x_1 Since this range is not known before the design, different values of the upper bound in the VSS design are used. Fig. 3 shows that if the upper bound is set too low, $\hat{s}(\hat{x}_1, \hat{x}_2, y) = -0.1\,\mathrm{sgn}(\hat{x}_1 + \hat{x}_2 - y)$, the $\hat{x}_2$ is very sluggish to track x_2, and the estimation error is large. Fig. 4 shows that if the upper bound is set too high, $\hat{s}(\hat{x}_1, \hat{x}_2, y) = -8\,\mathrm{sgn}(\hat{x}_1 + \hat{x}_2 - y)$. though $\hat{x}_2$ is very quick to track x_2, the magnitude of chattering is very large, and the estimation error is also large. Fig.5 shows the results of the proposed neural network approach, the estimation errors are small.

(3) **Example 3**

The linear part is the same as in example 1, the

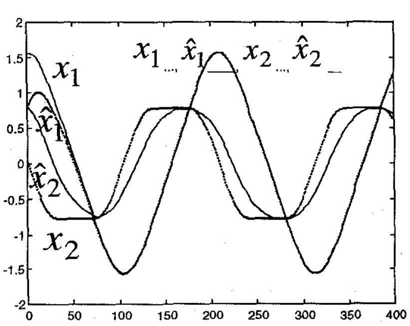

Fig.3: VSS observer design, $\hat{S}$ is too low;

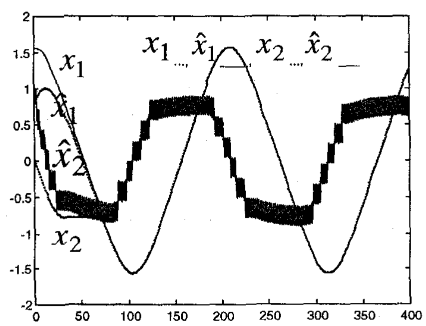

Fig.4: VSS observer design, $\hat{S}$ is too high;

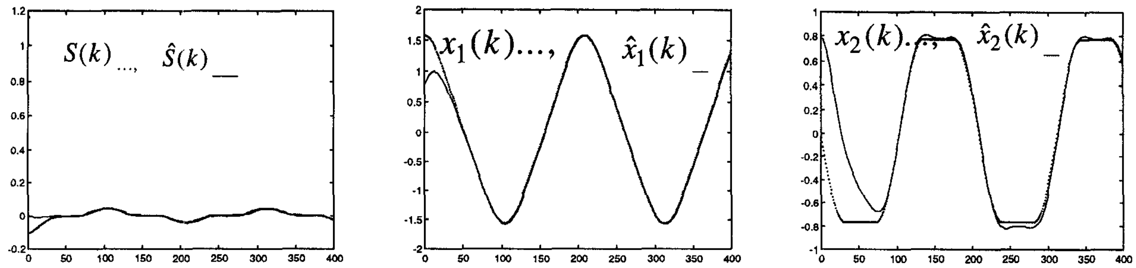

Fig.5: Observer design by proposed approach for $f(x)$ being a deadzone function

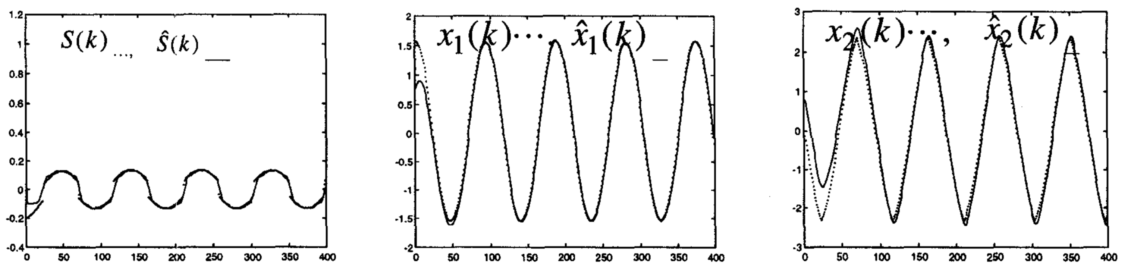

Fig.6: Observer design by proposed approach for $f(x)$ being a Coulombic friction function

nonlinear part now is a Coulombic friction function described as

$$f(x)=\begin{cases} b+kx & x>0 \\ -b+kx & x<0 \end{cases} \qquad (31)$$

Again, the Lie-algebraic method and extended linearization method can not applied to this example. Fig.6 shows the results of proposed method, the estimation accuracy is quite good.

7. Conclusion

The design of nonlinear observer is not easy. Some existing analytical approaches can give systematic design procedures , however their applications are limited. The proposed approach can be applied to a class of nonlinear systems described in (1) and (2). The linear part of the system is assumed known however the nonlinear part can be unknown and there is no restriction on its type.

The proposed approach consists of three steps: (a) Design a linear observer for the linear part of system. The nonlinear observer is the combination of linear observer and a unknown nonlinear compensation term as described in (5). (b) Estimate the unknown nonlinear compensation term using " Deconvolution method" in (11) or (11)'. (c) Model the estimated nonlinear compensation term $S(k)$ using B-spline neural network as in (16).

Simulation examples have shown that the proposed approach is very effective and can apply to cases where analytical approaches fail to apply.

Acknowledgment

The work was supported by the Research Grants Council of Hong Kong.

References:

[1]. Bestle, D., and Zeitz, M., Canonical form observer design for non-linear time-variable systems, Int. J. Control, 38, pp. 419-431, 1983.

[2]. Krener, A. J.,and Respondek, W., Nonlinear observer with linearizable error dynamics, SIAM J. Control and Optimize., 23, pp 197-216, 1985.

[3]. Baumann, W. T., and Rugh, W. J., Feedback control of nonlinear systems by extended linearization, IEEE Trans. AC-31, pp. 40-46, 1986.

[4]. Walcott, B. L., and Zak, S. H., State observation of nonlinear uncertain dynamical systems, IEEE Trans. AC-32, pp. 166-170, 1987.

[5]. Walcott, B. L., Corless, M. J., and Zak, S. H., Comparative study of non-linear state-observation techniques, Int. J. Control, 45, pp. 2109-2132, 1987.

[6]. Levin, A. U., and Narendra, K. S., Control of nonlinear dynamical systems using neural network - part 2: Observability, Identification, and Control, IEEE Trans. on Neural Networks, Vol. 7, No. 1, January 1996.

[7]. Li, Y. T., Qin, Z. X., and Zhang, H. Y., Nonlinear observer based on the neural network, Proceeding of The First Chinese World Congress on Intelligent Control and Intelligent Automation, Beijing, China, August, 1993.

[8]. Narendra, K. S., and Pathasarathy, K., Identification and control of dynamical systems using neural network, IEEE Trans. Neural Network, Vol. 1, pp. 4-27, Mar. 1990.

[9] Chen, S., and Billings, S. A., Neural networks for nonlinear dynamic system modelling and identification, Int. J. Control, Vol. 56, No. 2, pp. 319-346, 1992.

[10] Sjoberg, J., et. al., Nonlinear black-box modeling in system identification: a unified overview, Automatica, Vol. 31, No. 12, pp. 1691-1724, 1995.

[11] Harris, C. J., et al., Advances in neurofuzzy algorithms for real-time modelling and control, (invited paper), Engng. Applic. Artif. Intell., Vol. 9, No. 1, pp. 1-16, 1996.

[12]. Brown, M., and Harris, C., Neurofuzzy Adaptive Modelling and Control, Prentice Hall, 1994.