A Hydbrid Quasi-ARMAX Modelibg and Identification Scheme for Nonlinear Sysytems

Hu, Jinglu
Venture Business Laboratory

Hirasawa, Kotaro
Department of Electrical and Electronic Systems Engineering, Kyushu University

Kumamaru, Kousuke
Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology

https://doi.org/10.15017/1523854

出版情報：九州大学大学院システム情報科学紀要. 2 (2), pp.213-218, 1997-09-26. 九州大学大学院システム情報科学研究科

九州大学大学院
システム情報科学研究科報告
第2巻 第2号 平成9年9月

Research Reports on Information Science and
Electrical Engineering of Kyushu University
Vol.2, No.2, September 1997

A Hybrid Quasi-ARMAX Modeling and Identification Scheme for Nonlinear Systems

Jinglu HU* , Kotaro HIRASAWA** and Kousuke KUMAMARU***

(Received June 23, 1997)

Abstract: This paper presents a hybrid quasi-ARMAX modeling and identification scheme for nonlinear systems. It is shown that a general nonlinear ARMAX system can be represented by using a constrained time-variant ARMAX model, whose coefficients are nonlinear functions of input-output variables of system. A hybrid quasi-ARMAX model are then obtained by embedding a group of fuzzy models in the coefficients of the constrained time-variant ARMAX model to describe the time-variant coefficients. One of the distinctive features of the hybrid quasi-ARMAX model is that its parameters to be estimated have useful explicit meanings which can be taken as advantages for setting initial values of parameter estimation and for improving the generalization ability of the model.

Keywords: Nonlinear systems, System modeling and identification, Time-variant ARMAX model, Fuzzy model, Convergence property, Generalization ability

1. Introduction

The key problem in system identification is to find a suitable model structure, within which a good model is to be found. When no physical insight is available or used, one usually choose black-box model structure which belongs to families that are known to have good flexibility and have been "successful in the past"[1].

Under the assumption that the unknown system is linear, linear black-box models can be chosen for the system identification. The identification based on linear approximation has been extensively and successfully handled within some well known linear black-box structures [2),3)]. If the linear assumption is relaxed, one has to use nonlinear black-box models. For nonlinear black-box modeling, the "classical" literature seems to have concentrated on global basis function expansions, such as Volterra expansions [4)]. These have apparently had limited success. Recently, some authors have suggested the use of nonlinear structures based on neural networks (NN), wavelet networks (WN), radial basis function networks (RBFN), etc, and have achieved considerable success [5),6)]. However, the latter ones emphasize only on the input-output representation ability which is realized by including a lot of parameters in the models. Since the parameters in these models do not have useful meanings, these models suffer some problems concerning their parameter estimation (e.g. local minimum problem) and their performance (e.g. generalization ability). This motivates us to develop a new modeling scheme to obtain a nonlinear black-box model whose parameters have useful meanings.

It is shown that a general nonlinear ARMAX system can be represented by using a constrained time-invariant ARMAX model, whose coefficients are nonlinear functions of input-output variables of system. In order to describe those time-variant coefficients, we first divide each coefficient into two parts: constant parameter and nonlinear term, and then represent the nonlinear term using a fuzzy model with adjustable parameters. In this way, A hybrid quasi-ARMAX model is obtained, in which a group of fuzzy models are embedded in the coefficients of the constrained time-variant ARMAX model. We will show that one of the distinctive features of such hybrid model is that its parameters to be estimated have useful explicit meanings which can be taken as advantages for setting initial values of parameter estimation and for improving the generalization ability of the model. Experimental studies using both real data and simulated data are carried out to investigate the effectiveness of the proposed modeling and identification scheme.

The paper is organized as follows: Section 2 proposes a hybrid quasi-ARMAX modeling scheme for general nonlinear systems. In Section 3, we discuss

* Venture Business Laboratory
** Department of Electrical and Electronic Systems Engineering
*** Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology

its identification problem. Numerical simulations using real data and simulated date are carried out in Section 4. Finally, Section 6 is devoted to discussions and conclusions.

2. Hybrid Quasi-ARMAX Modeling

Let us consider a SISO general nonlinear ARX (NARX) system described by

$$\mathcal{S}: \; y(t) = g(\varphi(t)) + v(t) \tag{1}$$

$$\varphi(t) = [y(t-1) \; ... \; y(t-n) \; u(t-1) \; ... \; u(t-m)]^T \tag{2}$$

where $y(t)$ is the output at time t ($t = 1, 2, ...$), $u(t)$ the input, $\varphi(t)$ the regression vector, $v(t)$ the system disturbance, and $g(\,\cdot\,)$ the unknown continuously differentiable nonlinear function.

Performing Taylor expansion to $g(\varphi(t))$ around the region $\varphi(t) = 0$ and introducing a moving average (MA) noise model to the system disturbance, we can express the system (1) in a linear ARMAX structure[7]

$$\mathcal{M}: \; y(t) = \varphi^T(t)\theta(\varphi(t)) + C(q^{-1})e(t) \tag{3}$$

$$\theta(\varphi(t)) = [a_{1,t} \; \dots \; a_{n,t} \; b_{1,t} \; \dots \; b_{m,t}]^T \tag{4}$$

$$C(q^{-1}) = 1 + c_1 q^{-1} + \; ... \; + c_l q^{-l} \tag{5}$$

where q^{-1} is backward shift operator, for instance, $q^{-1}u(t) = u(t-1)$, $e(t)$ is the white noise, and the coefficients $a_{i,t}$ and $b_{i,t}$ are nonlinear functions of $\varphi(t)$, defined by

$$a_{i,t} = a_i(\varphi(t)), \qquad b_{i,t} = b_i(\varphi(t)) \tag{6}$$

The system described by (3) can be further expressed in an ARMAX form

$$A(q^{-1},t)y(t) = B(q^{-1},t)u(t) + C(q^{-1})e(t) \tag{7}$$

$$A(q^{-1},t) = 1 - a_{1,t}q^{-1} - \dots - a_{n,t}q^{-n} \tag{8}$$

$$B(q^{-1},t) = b_{1,t} + \dots + b_{m,t}q^{-m} \tag{9}$$

(7) is a time-variant ARMAX model whose coefficients are however constrained to be nonlinear functions of $\varphi(t)$. We call it a quasi-ARMAX model.

The above results show that a general nonlinear ARMAX system can be described by using a constrained time-variant ARMAX model whose coefficients are nonlinear functions of input-output variable of system. This motivates us to introduce a hybrid quasi-ARMAX model by embedding a group of fuzzy models in the coefficients of the quasi-ARMAX model.

Suppose that the coefficients $a_{i,t}$ and $b_{i,t}$ consist of two parts: constant parameters a_i, b_i and nonlinear terms $\triangle a_{i,t}$, $\triangle b_{i,t}$

$$a_{i,t} = a_i + \triangle a_{i,t}, \qquad b_{i,t} = b_i + \triangle b_{i,t} \tag{10}$$

Then we represent each nonlinear term using a fuzzy model with adjustable parameters, which can be described explicitly by (13) [8]

$$\triangle a_{i,t} = \mathcal{Z}_i(\varphi(t)) \quad (i = 1, ..., n) \tag{11}$$

$$\triangle b_{j,t} = \mathcal{Z}_{j+n}(\varphi(t)) \quad (j = 1, ..., m) \tag{12}$$

$$\mathcal{Z}_i(\varphi_r(t)) = \frac{\sum_{j=1}^{M} \omega_{ij}(\wedge_{k=1}^{r} \mu_{A_k^j}(x_k(t)))}{\sum_{j=1}^{M}(\wedge_{k=1}^{r} \mu_{A_k^j}(x_k(t)))} = \sum_{j=1}^{M} \omega_{ij}\mathcal{N}_f(\boldsymbol{p}_{ij}, \varphi(t)) \tag{13}$$

$$r = \dim(\varphi(t)) = n + m$$

where $\wedge$ is the minimum operator, M is the number of fuzzy rules, $x_k(t)$ are the elements of $\varphi(t)$, and $\mu_{A_k^j}(\cdot)$ is the membership function of fuzzy set A_k^j. As the membership function, $\mu_{A_k^j}(\cdot)$ may simply be a triangle function or a Gaussian function defined by

$$\mu_{A_k^j}(x_k(t)) = \alpha_k^j \exp\left[-\frac{1}{2}\left(\frac{x_k(t) - \bar{x}_k^j}{\sigma_k^j}\right)^2\right] \tag{14}$$

In the latter case, the parameter vector $\boldsymbol{p}_{ij}$ is given by

$$\boldsymbol{p}_{ij} = [\alpha_i^j \; \bar{x}_i^j \; \sigma_i^j]^T \tag{15}$$

On the other hand, we can obtain another expression for the hybrid quasi-ARMAX model by using (10)-(13) in (3), which is described as

$$y(t) = \underbrace{\varphi^T(t)\theta + C(q^{-1})e(t)}_{\text{ARMAX}} + \sum_{j=1}^{M} \underbrace{\varphi^T(t)\Omega_j}_{\text{ARX}} \mathcal{N}_f(\boldsymbol{p}_{ij}, \varphi(t)) \tag{16}$$

where $\Omega_j = [\omega_{1j} \; ... \; \omega_{rj}]^T$. (16) shows that the hybrid quasi-ARMAX model is equivalent to a hybrid model combining a linear ARMAX model and a multi-ARX-model based on interpolation using the "basic function" $\mathcal{N}_f(\cdot,\cdot)$. This follows that the parameters of the hybrid quasi-ARMAX model have useful explicit meanings, where

- $\theta_e = [\theta^T, \; c_1 \; ... \; c_l]^T$ is associated with global linear ARMAX approximation,
- $\Omega = [\Omega_1 \; ...\Omega_M]^T$ is associated with M local linear ARX approximations, and
- $\boldsymbol{p} = [\boldsymbol{p}_{11} \; \cdots \; \boldsymbol{p}_{rM}]^T$ is associated with the par-

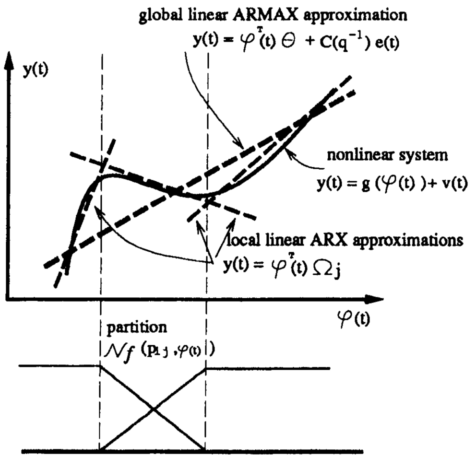


Fig.1 Useful explicit meanings of the parameter in the hybrid quasi-ARMAX modeling.

tition in the operating region of $\varphi(t)$.

Figure 1 shows the images of the global linear ARMAX approximation, the multiple local linear ARX approximations and the partition of operating region of $\varphi(t)$ in the hybrid quasi-ARMAX modeling.

3. Parameter Estimation

For nonlinear black-box models based on nonlinear structures such as neural networks, parameter estimation is usually performed by using a back propagation algorithm with random given initial values. Such scheme can certainly be applied to the hybrid quasi-ARMAX model. However, since the parameters to be estimated in our case have useful explicit meanings, a more sophisticated algorithm can be developed in order to improve the convergence properties by using a better initial value, and to improve the generalization ability of model by reducing the parameter redundancy. Such an estimation algorithm consists of the following two procedures.

3.1 Estimation Procedure 1

It is obvious form (16) that the parameters to be estimated can be divided into two groups: $\{\theta_e,\ \Omega\}$ in which the model is linear and $\{\boldsymbol{p}\}$ in which the model is nonlinear. In the Procedure 1, $\{\theta_e,\ \Omega\}$ will be adjusted while fixing $\{\boldsymbol{p}\}$ to its initial value $\boldsymbol{p}_0$.

3.1.1 Setting the initial values

(1) Since θ_e describes a linear ARMAX model

$$y(t) = \varphi^T(t)\theta + C(q^{-1})e(t), \tag{17}$$

it is naturally to set its initial value θ_{e0} by using the estimates of the linear ARMAX model for the nonlinear system.

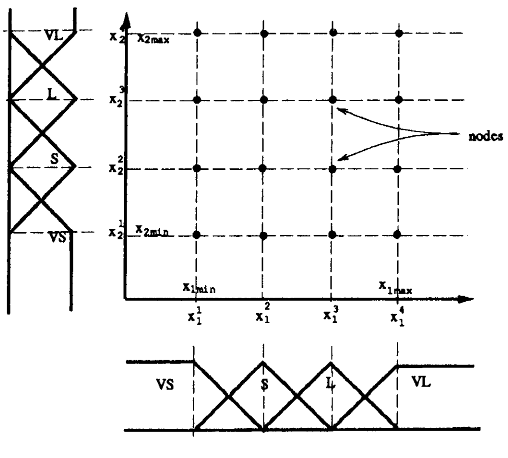


Fig.2 An example for determining $\mathbf{p}_{0j}$.

(2) The initial value Ω_0 for Ω can simply be taken as 0.

(3) The initial value $\boldsymbol{p}_0$ for $\boldsymbol{p}$ should be obtained based on an appropriate partition in the operating region of $\varphi(t)$.

Let $\boldsymbol{p}_0 = [\boldsymbol{p}_{0ij},\ i = 1,...,r;\ j = 1,...,M] = [\boldsymbol{p}_{0j},\ j = 1,...,M]$ which means that the partitions in the operating region of $\varphi(t)$ for all fuzzy models used are the same. And $\boldsymbol{p}_{0j}$ can then be determined based on the following partition.

Suppose the fuzzy model has r inputs, $X = [x_i,\ i = 1,...,r]$ and the operating region is mostly located in $X_{\min} \le X \le X_{\max}$, $X_{\min} = [x_{i\ min},\ i = 1,...,r]$, $X_{\max} = [x_{i\ max},\ i = 1,...,r]$. $X \notin [X_{\min},\ X_{\max}]$ is allowable in practice. We first partition the input hyperplane, that is, put *nodes* into the input hyperplane. As shown in **Fig.2**, if the number of nodes corresponding to x_i is denoted as n_i, the total number of the nodes (fuzzy rules) in the hyperplane will be $M = \prod_{i=1}^{r} n_i$. Next, the parameter vectors $\boldsymbol{p}_{0j}$ are chosen so that the 'basis functions' $\mathcal{N}_f(\boldsymbol{p}_{0j}, X)$ have appropriate shape and are put onto each node. Without using other knowledge information, the nodes will be uniformly assigned in the hyperplane. **Figure 2** shows an example for determining $\boldsymbol{p}_{0j}$ with $r = 2$ and $M = 4 \times 4$ in the case where triangle membership functions are used.

3.1.2 Estimation Algorithm

Introduce a parameter vector Θ and a regression vector $\varphi_{NL}(t)$ defined as

$$\Theta = [\theta_e^T,\ \Omega^T]^T \tag{18}$$

$$\varphi_{NL}(t) = [\varphi^T(t),\ e(t-1)\ ...\ e(t-1),\ \varphi^T(t) \otimes \varphi^T_{\mathcal{N}_f}(t)]^T \tag{19}$$

where $\varphi^T_{\mathcal{N}_f}(t) = [\mathcal{N}_f(\boldsymbol{p}_{ij}, \varphi(t),\ j = 1,...,M]$, and the symbol $\otimes$ denotes Kronecker production. Then the hybrid quasi-ARMAX model can be expressed as

$$y(t) = \varphi_{NL}^T(t)\Theta + e(t) \quad (20)$$

The estimation of Θ for (20) is done by minimizing the following extended criterion function

$$V_N(\Theta) = \frac{1}{2}\sum_{t=1}^{N}\left[(y(t) - \varphi_{NL}^T(t)\Theta)^2 + C_\alpha\Theta^T\Theta\right] \quad (21)$$

which can be carried out by using an existing recursive algorithm [2),3)]. The second term in the right hand of (21) is introduced to improve the generalization ability, in which C_α is a small positive value.

3.2 Estimation Procedure 2

In the Estimation Procedure 1, since the parameter vector $\boldsymbol{p}$ is kept fixed, the model is linear in the parameters $\{\theta_e, \Omega\}$ to be estimated. Therefore, the estimation algorithm can almost guarantee to find the global minimum of the criterion function (21). It have been found that a satisfying performance of the model can be obtained for many real applications by merely carrying out the Estimation Procedure 1 [7)]. However, another estimation procedure (Estimation Procedure 2) may be performed by adjusting the parameter vector $\boldsymbol{p}$ as well if higher identification accuracy is needed.

In the Estimation Procedure 2, the parameter vectors $\boldsymbol{p}$ and Ω are adjusted while the parameter vector θ_e are fixed, in which the results obtained in the Estimation Procedure 1 are used as initial values. This can be done because θ is not independent of Ω, which is obvious from (6) and (13) since, for instance, there exits $\omega'_{ij} = a_i + \omega_{ij}$ that (22) is satisfied

$$a_{i,t} = a_i + \sum_{j=1}^{M}\omega_{ij}\mathcal{N}_f(\boldsymbol{p}_{ij}, \varphi_r(t)) = \sum_{j=1}^{M}\omega'_{ij}\mathcal{N}_f(\boldsymbol{p}_{ij}, \varphi_r(t)) \quad (22)$$

Estimation of $\boldsymbol{p}$ and Ω can be carried out using a back propagation algorithm for neural networks by minimizing the following extended criterion function

$$V_N(\boldsymbol{p}, \Omega) = \frac{1}{2}\sum_{t=1}^{N}\left[\varepsilon^2(t) + C_\beta(\boldsymbol{p} - \boldsymbol{p}_0)^T(\boldsymbol{p} - \boldsymbol{p}_0)\right] \quad (23)$$

where $\varepsilon(t) = y(t) - \hat{y}(t)$ is prediction error, $\boldsymbol{p}_0$ is the initial value of $\boldsymbol{p}$, and C_β is a small positive value. The second term in the right hand of (23) is introduced to improve the generalization ability.

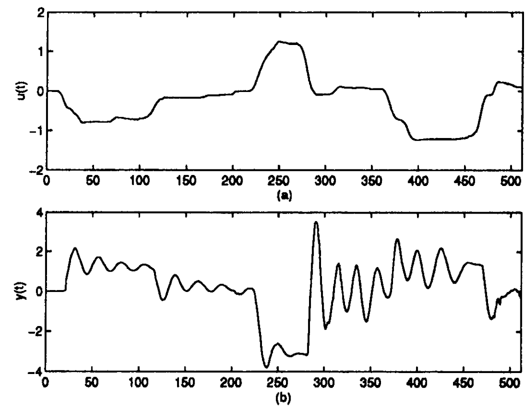


Fig.3 Estimation data: (a) the input $u(t)$, (b) the output $y(t)$.

4. Experimental Studies

In this section, experimental studies using both real data and simulated data are carried out to investigate the effectiveness of the proposed scheme. Since the systems used are not so complicated, the estimation of the hybrid quasi-ARMAX model is performed merely by using the Estimation Procedure 1. That is, $\boldsymbol{p}$ is taken as constant vector in the simulations.

4.1 Modeling a Hydraulic Robot Actuator

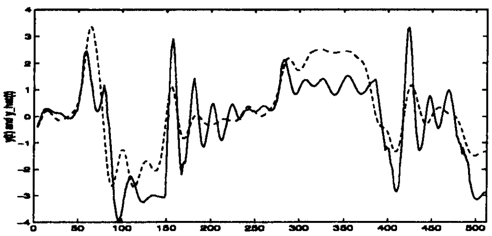
The real system considered is a hydraulic robot actuator. Let us denote by u(t) and y(t) the position of the valve and the oil pressure at time t, respectively. A sample of 1024 pairs of $\{y(t), u(t)\}$ was registered[†1]. We divide it into two equal parts for estimating and for validating our model, respectively. The estimation data is depicted in **Fig.3**.

For comparison, the simulation were carried out by using linear ARX model (ARX), one hidden layer neural networks (NN), and the proposed hybrid quasi-ARMAX model (HQAR), in which the regression vector is taken as $\varphi(t) = [y(t-1)\ y(t-2)\ y(t-3)\ u(t-1)\ u(t-2)]^T$. The results of identified models simulated on the validation data are shown in **Fig.4** and **Table 1**. We can see that the linear ARX model could not model the system well, while neural networks improves the performance quite much, and the proposed model has the best performance

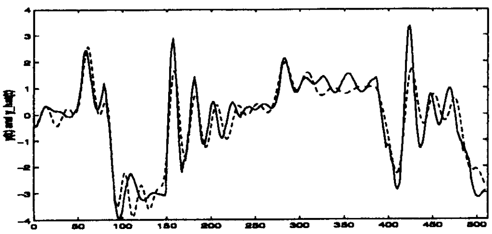
†1 The data were taken from public ftp domain. We gratefully acknowledge Linköping University for providing the data.

Table 1 Comparison of Identification Results

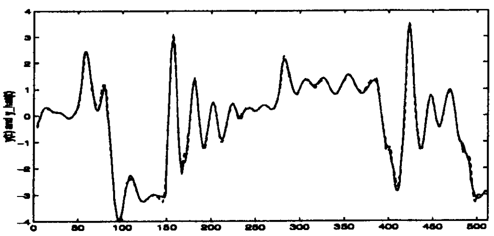
Models	Number of parameter	Mean square error
ARX	5	1.0160
NN	71	0.6170
HQAR	85	0.1360



(a) Result with linear ARX model



(b) Result with Neural Network

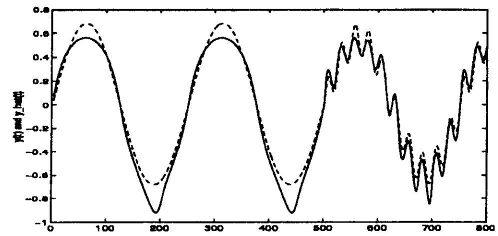


(c) Result with the proposed model

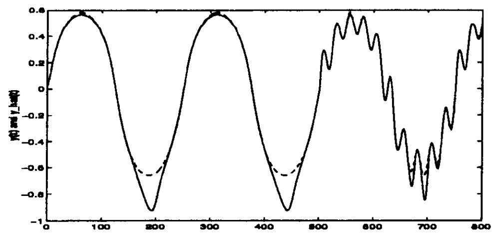
Fig.4 Simulations of identified models on validation data. The solid line shows the true oil pressure and the dashed line the simulated model output.

Table 2 Comparison of Identification Results

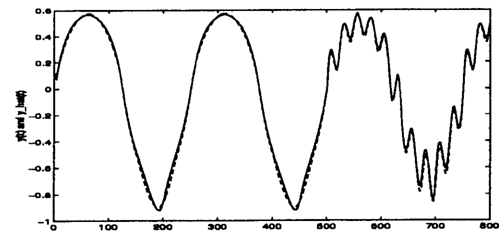
Models	Number of parameter	Mean square error
ARX	5	0.0866
NN	314	0.0678
HQAR	95	0.0270



(a) Result with linear ARX model



(b) Result with Neural Network



(c) Result with the proposed model

Fig.5 Simulations of identified models on validation data. The solid line shows the system true output and the dashed line the simulated model output.

among the three models.

4.2 Modeling a Mathematical System

The mathematical system is taken from Narendra (1990) [9], which contains rather strong nonlinearity. The system is governed by

$$y(t) = f[y(t-1), y(t-2), y(t-3), u(t-1), u(t-2)] \quad (24)$$

where

$$f[x_1, x_2, x_3, x_4, x_5] = \frac{x_1 x_2 x_3 x_5 (x_3 - 1) + x_4}{1 + x_2^2 + x_3^2}.$$

Estimation data are sampled when system is excited using random input uniformly distributed in the interval [-1, 1], while validation data are sampled from system using an input $u(t) = \sin(2\pi t/250)$ for $t \leq 500$ and $u(t) = 0.8\sin(2\pi t/250) + 0.2\sin(2\pi t/25)$ for $t > 500$.

Similar to the above one, the simulations were carried by using linear ARX model, two hidden layer neural networks and the proposed model, in which the regression vector is taken as $\varphi(t) = [y(t-1)\ y(t-2)\ y(t-3)\ u(t-1)\ u(t-2)]^T$. The results of identified models simulated on the validation data are shown in **Fig.5** and **Table 2**.

From the simulation results shown in **Fig.4** and **5**, **Table 1** and **2**, we can see that the proposed hybrid quasi-ARMAX model has superior performances to neural network models. The reason for better performances is that the estimator for the hybrid quasi-ARMAX model has better convergence properties. The two examples used in our simulations are well known in the literature where they have been used to test neural network models and other nonlinear black-box models. The interested reader is referred to the references [9),5)] for a comparison.

5. Discussions and Conclusions

We have proposed a hybrid quasi-ARMAX model, which is obtained by first representing a general nonlinear with a constrained time-variant ARMAX model, then describing the time-variant coefficients of the ARMAX model via a group of fuzzy models. One of the distinctive features of the hybrid quasi-ARMAX model is that the parameters to be estimated have useful explicit meanings which have been taken as advantages for setting initial values of parameter estimation and for improving the generalization ability of the model. The results of experimental studies using both real data and simulated data show that the identification of the model has nice convergence properties, and the identified model has better generalization ability than commonly-used neural networks.

In the hybrid quasi-ARMAX model, the model order consist of two part: the ARMAX order (n, m, l) and the order of fuzzy models (M), both of which should be determined before the parameter estimation can be carried out. The ARMAX order (n, m, l) is assumed to be known. If it is unknown, it can be determined based on the results of identifying the system using a linear ARMAX model because the hybrid quasi-ARMAX model is basically an extension of linear ARMAX model. Many existing approaches for determining the order of linear models such as Akaike criteria AIC and FPC can be applied as a reference. On the other hand, the order of fuzzy models (M), which means the number of fuzzy rules, is usually unknown and should be determined based on knowledge information about system structure. If the initial value $\boldsymbol{p}_{0j}$ is determined in a way described in the subsection 3.1.1, the order is given as $M = \prod_{i=1}^{r} n_i$, which should be chosen as small as possible in order to reduce the number of parameters to be estimated. The following hints can be used to make M small:

(1) Hint A: If the system is linear with respect to x_i, n_i may be chosen to be 1.

(2) Hint B: If no other useful information is available, n_1 and n_{n+1} corresponding to $y(t-1)$ and $u(t-1)$ are assigned with appropriate values, while all other n_i's are set to 1.

(3) Hint C: If the role of rules can be replaced by employing interpolation of other rules, those rules may be removed.

The efficient use of knowledge information plays a key role in the hybrid quasi-ARMAX modeling, for which further research are needed.

References

1) J. Sjöberg, *Non-Linear System Identification with Neural Networks*, PhD thesis, Linköping University, Sweden, 1995.

2) L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, Englewood Cliffs, N.J., 1987.

3) T. Söderström and P. Stoica, *System Identification*, Prentice Hall International, Hemel Hempstead, 1989.

4) H.W. Chen, "Modeling and identification of parallel nonlinear systems: Structural classification and parameter estimation methods", *Proceedings of the IEEE*, vol. 83, no. 1, pp. 39–66, 1995.

5) J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Deglon, P.Y. Glorennec, H. Hjalmarsson, and A. Juditsky, "Nonlinear black-box modeling in system identification: a unified overview", *Automatica*, vol. 31, no. 12, pp. 1691–1724, 1995.

6) A. Juditsky, H. Hjalmarssom, B. Delyon A. Benveniste, L. Ljung, J. Sjöberg, and Q. Zhang, "Nonlinear black-box modeling in system identification: Mathematical foundations", *Automatica*, vol. 31, no. 12, pp. 1725–1750, 1995.

7) J. Hu, *Research on Hybrid Black-Box Modeling for Nonlinear Systems and Its Applications*, PhD thesis, Kyushu Institute of Technoloy, Japan, 1997.

8) L.X. Wang and J.M. Mendel, "Back-propagation fuzzy system as nonlinear dynamic system identifiers", in *Proc. of IEEE Int. Conf. on Fuzzy Systems (San Diego, CA)*, 1992, pp. 1409–1418.

9) K.S. Narendra and K. Parthasathy, "Identification and control of dynamical systems using neural networks", *IEEE Trans. on Neural Networks*, vol. 1, no. 1, pp. 4–27, 1990.