

Discrete Time Nonlinear Identification via Recurrent High Order Neural Networks

Alma Y. Alanis ¹, Edgar N. Sanchez ², and Alexander G. Loukianov ²

¹Departamento de Ciencias Computacionales, CUCEI, Universidad de Guadalajara, Av. Revolucion 1500, Col. Olimpica, C.P. 44430, Guadalajara, Jalisco, Mexico. e-mail: almayalanis@gmail.com

² CINVESTAV, Unidad Guadalajara, Apartado Postal 31-438, Plaza La Luna, Guadalajara, Jalisco, C.P. 45091, Mexico, e-mail: sanchez@gdl.cinvestav.mx

(Paper received on February 13, 2008, accepted on April 15, 2008)

Abstract. This paper deals with the problem of discrete-time nonlinear system identification via Recurrent High Order Neural Networks. It includes the respective stability analysis on the basis of the Lyapunov approach for the extended Kalman filter (EKF)-based NN training algorithm, which is applied for learning. Applicability of the scheme is illustrated via simulation for a discrete-time nonlinear model of an electric induction motor.

1 Introduction

Neural networks (NN) have become a well-established methodology as exemplified by their applications to identification and control of general nonlinear and complex systems. In particular, the use of recurrent high order neural networks (RHONN) has increased recently [7]. There are recent results which illustrate that the NN technique is highly effective in the identification of a broad category of complex discrete-time nonlinear systems without requiring complete model information ([12], [13]).

Lyapunov approach can be used directly to obtain robust training algorithms for continuous-time recurrent neural networks ([7], [9]). For discrete-time systems, the problem is more complex due to the couplings among subsystems, inputs and outputs. Few results have been published in comparison with those for continuous-time domain ([12], [13]). By other hand discrete-time neural networks are more convenient for real-time applications.

For many nonlinear systems it is often difficult to obtain their accurate and faithful mathematical models, regarding their physically complex structures and hidden parameters as discussed in [1]. Therefore, system identification becomes important and even necessary before system control can be considered not only for understanding and predicting the behavior of the system, but also for obtaining an effective control law.

The identification problem consists of choosing an appropriate identification model and adjusting its parameters according to some adaptive law, such that the

response of the model to an input signal (or class of input signals), approximates the response of the real system to the same input [9].

Several training methods for discrete-time recurrent networks have been proposed in the literature as a viable alternative. New training algorithms, e.g., those based on Kalman filtering, have appeared [3], [4], [10]. In this paper, we use an Extended Kalman Filter (EKF)-based training algorithm for the RHONN, in order to identify discrete-time nonlinear systems.

2 Mathematical preliminaries

Through this paper we use k as the step sampling, $k \in 0 \cup \mathbb{Z}^+$, $|\bullet|$ for the absolute value, $\|\bullet\|$ for the Euclidian norm for vectors and for any adequate norm for matrices. For more details related to this section see [2]. Consider a MIMO nonlinear system:

$$\chi(k+1) = F(\chi(k), u(k)) \quad (1)$$

where $\chi \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $F \in \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ is nonlinear function.

Definition 1. *The solution of (1) is semiglobally uniformly ultimately bounded (SGUUB), if for any Ω , a compact subset of $\mathbb{R}^n$ and all $\chi(k_0) \in \Omega$, there exists an $\epsilon > 0$ and a number $N(\epsilon, \chi(k_0))$ such that $\|\chi(k)\| < \epsilon$ for all $k \geq k_0 + N$. In other words, the solution of (1) is said to be SGUUB if, for any apriori given (arbitrarily large) bounded set Ω and any apriori given (arbitrarily small) set Ω_0 , which contains $(0,0)$ as an interior point, there exists a control u , such that every trajectory of the closed loop system starting from Ω enters the set $\Omega_0 = \{\chi(k) \mid \|\chi(k)\| < \epsilon\}$, in a finite time and remains in it thereafter [2].*

Theorem 1 [2] Let $V(\chi(k))$ be a Lyapunov function for the discrete-time system (1), which satisfies the following properties:

$$\begin{aligned} \gamma_1(\|\chi(k)\|) &\leq V(\chi(k)) \leq \gamma_2(\|\chi(k)\|) \\ V(\chi(k+1)) - V(\chi(k)) &= \Delta V(\chi(k)) \leq -\gamma_3(\|\chi(k)\|) + \gamma_3(\zeta) \end{aligned}$$

where ζ is a positive constant, $\gamma_1(\bullet)$ and $\gamma_2(\bullet)$ are strictly increasing functions, and $\gamma_3(\bullet)$ is a continuous, nondecreasing function. Thus if

$$\Delta V(\chi) < 0 \quad \text{for} \quad \|\chi(k)\| > \zeta$$

then $\chi(k)$ is uniformly ultimately bounded, i.e. there is a time instant k_T , such that $\|\chi(k)\| < \zeta, \forall k < k_T$.

3 Discrete-time Recurrent Neural Networks

Consider the following discrete-time recurrent high order neural network (RHONN):

$$x_i(k+1) = w_i^\top z_i(x(k), u(k)), \quad i = 1, \dots, n \quad (2)$$

where x_i ($i = 1, 2, \dots, n$) is the state of the i th neuron, L_i is the respective number of higher-order connections, $\{I_1, I_2, \dots, I_{L_i}\}$ is a collection of non-ordered subsets of $\{1, 2, \dots, n\}$, n is the state dimension, w_i ($i = 1, 2, \dots, n$) is the respective on-line adapted weight vector, and $z_i(x(k), u(k))$ is given by

$$z_i(x(k), u(k)) = \begin{bmatrix} z_{i1} \\ z_{i2} \\ \vdots \\ z_{iL_i} \end{bmatrix} = \begin{bmatrix} \prod_{j \in I_1} y_{i_j}^{d_{i,j}(1)} \\ \prod_{j \in I_2} y_{i_j}^{d_{i,j}(2)} \\ \vdots \\ \prod_{j \in I_{L_i}} y_{i_j}^{d_{i,j}(L_i)} \end{bmatrix} \quad (3)$$

with $d_{i,j}(k)$ being a nonnegative integers, and y_i is defined as follows:

$$y_i = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{in} \\ y_{i_{n+1}} \\ \vdots \\ y_{i_{n+m}} \end{bmatrix} = \begin{bmatrix} S(x_1) \\ \vdots \\ S(x_n) \\ u_1 \\ \vdots \\ u_m \end{bmatrix} \quad (4)$$

In (4), $u = [u_1, u_2, \dots, u_m]^T$ is the input vector to the neural network, and $S(\bullet)$ is defined by

$$S(x) = \frac{1}{1 + \exp(-\beta x)} \quad (5)$$

Consider the problem to approximate the general discrete-time nonlinear system (1), by the following discrete-time RHONN serie-parallel representation [9]:

$$\chi_i(k+1) = w_i^*{}^T z_i(x(k), u(k)) + \epsilon_{z_i} \quad (6)$$

where χ_i is the i th plant state, ϵ_{z_i} is a bounded approximation error, which can be reduced by increasing the number of the adjustable weights [9]. Assume that there exists ideal weights vector w_i^* such that $\|\epsilon_{z_i}\|$ can be minimized on a compact set $\Omega_{z_i} \subset \mathbb{R}^{L_i}$. The ideal weight vector w_i^* is an artificial quantity required for analytical purpose [9]. In general it is assumed that this vector exists and is constant but unknown. Let us define its estimate as w_i and the estimation error as

$$\tilde{w}_i(k) = w_i^* - w_i(k) \quad (7)$$

The estimate w_i is used for stability analysis which will be discussed later. Since w_i^* is constant, then $\tilde{w}_i(k+1) - \tilde{w}_i(k) = w_i(k+1) - w_i(k)$, $\forall k \in 0 \cup \mathbb{Z}^+$.

4 The EKF Training Algorithm

Kalman filtering (KF) estimates the state of a linear system with additive state and output white noises [1], [3]. For KF-based neural network training, the network weights become the states to be estimated, with the error between the

neural network output and the desired output being considered; this error is considered as additive white noise. For identification, the desired output is information generated by the plant; in this paper, the respective state. Due to the fact that the neural network mapping is nonlinear, an extended Kalman Filtering (EKF)-type is required.

The training goal is to find the optimal weight values that minimize the prediction errors (the differences between the desired outputs and the neural network outputs). The EKF-based NN training algorithm is described by

$$\begin{aligned} K_i(k) &= P_i(k) H_i(k) M_i(k) & i &= 1, \dots, n \\ w_i(k+1) &= w_i(k) + \eta_i K_i(k) e_i(k) \\ P_i(k+1) &= P_i(k) - K_i(k) H_i^T(k) P_i(k) + Q_i(k) \end{aligned} \quad (8)$$

with

$$M_i(k) = [R_i(k) + H_i^T(k) P_i(k) H_i(k)]^{-1} \quad (9)$$

$$e_i(k) = \chi_i(k) - x_i(k) \quad (10)$$

where $e_i(k)$ is the respective identification error, $P_i(k) \in \mathbb{R}^{L_i \times L_i}$ is the prediction error covariance matrix at step k , $w_i \in \mathbb{R}^{L_i}$ is the weight (state) vector, L_i the respective number of neural network weights, χ_i is the i th plant state, x_i the i th neural network state, n is the number of states, $K_i \in \mathbb{R}^{L_i}$ is the Kalman gain vector, $Q_i \in \mathbb{R}^{L_i \times L_i}$ is the NN weight estimation noise covariance matrix, $R_i \in \mathbb{R}$ is the error noise covariance; $H_i \in \mathbb{R}^{L_i}$ is a vector, in which each entry ($H_{i,j}$) is the derivative of one of the neural network state, ($x_{i,j}$), with respect to one neural network weight, ($w_{i,j}$), as follows

$$H_{i,j}(k) = \left[\frac{\partial x_{i,j}(k)}{\partial w_{i,j}(k)} \right]_{w_{i,j}(k)=w_{i,j}(k+1)}^T \quad (11)$$

where $i = 1, \dots, n$ and $j = 1, \dots, L_i$

Usually P_i and Q_i are initialized as diagonal matrices, with entries $P_i(0)$ and $Q_i(0)$, respectively. It is important to remark that $H_i(k)$, $K_i(k)$ and $P_i(k)$ for the EKF are bounded; for a detailed explanation of this fact see [11].

Then the dynamics of the identification error (10) can be expressed as

$$e_i(k+1) = \tilde{w}_i(k) z_i(x(k), u(k)) + \epsilon_{z_i} \quad (12)$$

By the other hand the dynamics of (7) is

$$\tilde{w}_i(k+1) = \tilde{w}_i(k) - \eta_i K_i(k) e(k) \quad (13)$$

Now, we establish the main result of this paper in the following theorem.

Theorem 2: The RHONN (2) trained with the EKF-based algorithm (8) to identify the nonlinear plant (1), ensures that the identification error (10) is semiglobally uniformly ultimately bounded (SGUUB); moreover, the RHONN weights remain bounded.

Proof. Consider the Lyapunov function candidate

$$\begin{aligned} V_i(k) &= \tilde{w}_i^T(k) \tilde{w}_i(k) + e_i^2(k) \\ \Delta V_i(k) &= V(k+1) - V(k) \\ &= \tilde{w}_i^T(k+1) \tilde{w}_i(k+1) + e_i^2(k+1) - \tilde{w}_i^T(k) \tilde{w}_i(k) - e_i^2(k) \end{aligned} \quad (14)$$

Using (12) and (13) in (14)

$$\begin{aligned} \Delta V_i(k) &= [\tilde{w}_i(k) - \eta_i K_i(k) e_i(k)]^T [\tilde{w}_i(k) - \eta_i K_i(k) e_i(k)] - \tilde{w}_i(k) \tilde{w}_i(k) \\ &\quad + [\tilde{w}_i(k) z_i(x(k), u(k)) + \epsilon_{z_i}]^T [\tilde{w}_i(k) z_i(x(k), u(k)) + \epsilon_{z_i}] - e_i^2(k) \end{aligned}$$

which can be expressed as

$$\begin{aligned} \Delta V_i(k) &= \tilde{w}_i^T(k) \tilde{w}_i(k) - \tilde{w}_i^T(k) \tilde{w}_i(k) + \eta^2 e_i^2(k) K_i^T K_i(k) \\ &\quad + 2\epsilon_{z_i} \tilde{w}_i(k) z_i(x(k), u(k)) + z_i^T(x(k), u(k)) \tilde{w}_i^T(k) \tilde{w}_i(k) z_i(x(k), u(k)) \\ &\quad + \epsilon_{z_i}^2 - 2\eta_i e_i(k) \tilde{w}_i^T(k) K_i(k) - e_i^2(k) \\ \Delta V_i(k) &\leq |e_i(k)|^2 \|\eta K_i\|^2 - |e_i(k)|^2 - 2\eta_i |e_i(k)| \|\tilde{w}_i(k)\| \|K_i(k)\| + |\epsilon_{z_i}|^2 \\ &\quad + |2\epsilon_{z_i}| \|\tilde{w}_i(k)\| \|z_i(x(k), u(k))\| + \|\tilde{w}_i(k)\|^2 \|z_i(x(k), u(k))\|^2 \end{aligned}$$

Then $\Delta V_i(k) < 0$ when

$$|e_i(k)| > \frac{|\epsilon_{z_i}|^2}{1 - \|\eta K_i\|^2} \equiv \kappa_1$$

and

$$\|\tilde{w}_i(k)\| > \frac{|2\eta_i e_i \max| \|K_i(k)\|}{\|z_i(x(k), u(k))\|^2} - \frac{|2\epsilon_{z_i}|}{\|z_i(x(k), u(k))\|} \equiv \kappa_2$$

Therefore the solution of (12) and (13) is stable, hence the identification error and the RHONN weights are SGUUB [5].

5 Application

In this section we apply the above developed scheme to a three-phase induction motor model

5.1 Motor model

The six-order discrete-time induction motor model in the stator fixed reference frame (α, β) under the assumptions of equal mutual inductances and linear mag-

netic circuit is given by [6]

$$\begin{aligned}
\omega(k+1) &= \omega(k) + \frac{\mu}{\alpha} (1 - \alpha) M \left(i^\beta(k) \psi^\alpha(k) - i^\alpha(k) \psi^\beta(k) \right) - \left(\frac{T}{J} \right) T_L(k) \\
\psi^\alpha(k+1) &= \cos(n_p \theta(k+1)) \rho_1(k) - \sin(n_p \theta(k+1)) \rho_2(k) \\
\psi^\beta(k+1) &= \sin(n_p \theta(k+1)) \rho_1(k) + \cos(n_p \theta(k+1)) \rho_2(k) \\
i^\alpha(k+1) &= \varphi^\alpha(k) + \frac{T}{\sigma} u^\alpha(k) \\
i^\beta(k+1) &= \varphi^\beta(k) + \frac{T}{\sigma} u^\beta(k) \\
\theta(k+1) &= \theta(k) + \omega(k) T + \frac{\mu}{\alpha} \left[T - \frac{(1-a)}{\alpha} \right] \\
&\quad \times M \left(i^\beta(k) \psi^\alpha(k) - i^\alpha(k) \psi^\beta(k) \right) - \frac{T_L(k)}{J} T^2
\end{aligned} \tag{15}$$

with

$$\begin{aligned}
\rho_1(k) &= a \left(\cos(\phi(k)) \psi^\alpha(k) + \sin(n_p \phi(k)) \psi^\beta(k) \right) \\
&\quad + b \left(\cos(\phi(k)) i^\alpha(k) + \sin(\phi(k)) i^\beta(k) \right) \\
\rho_2(k) &= a \left(\cos(\phi(k)) \psi^\alpha(k) - \sin(\phi(k)) \psi^\beta(k) \right) \\
&\quad + b \left(\cos(\phi(k)) i^\alpha(k) - \sin(\phi(k)) i^\beta(k) \right) \\
\varphi^\alpha(k) &= i^\alpha(k) + \alpha \beta T \psi^\alpha(k) + n_p \beta T \omega(k) \psi^\alpha(k) - \gamma T i^\alpha(k) \\
\varphi^\beta(k) &= i^\beta(k) + \alpha \beta T \psi^\beta(k) + n_p \beta T \omega(k) \psi^\beta(k) - \gamma T i^\beta(k) \\
\phi(k) &= n_p \theta(k)
\end{aligned} \tag{16}$$

with $b = (1 - a) M$, $\alpha = \frac{R_r}{L_r}$, $\gamma = \frac{M^2 R_r}{\sigma L_r^2} + \frac{R_s}{\sigma}$, $\sigma = L_s - \frac{M^2}{L_r}$, $\beta = \frac{M}{\sigma L_r}$, $a = e^{-\alpha T}$ and $\mu = \frac{M n_p}{J L_r}$, besides L_s , L_r and M are the stator, rotor and mutual inductance respectively; R_s and R_r are the stator and rotor resistances respectively; n_p is the number of pole pairs; i^α and i^β represents the currents in the α and β phases, respectively; ψ^α and ψ^β represents the fluxes in the α and β phases, respectively and θ is the rotor angular displacement. Simulations are performed for the system (15), using the following parameters: $R_s = 14\Omega$; $L_s = 400mH$; $M = 377mH$; $R_r = 10.1\Omega$; $L_r = 412.8mH$; $n_p = 2$; $J = 0.01Kg m^2$; $T = 0.001s$.

5.2 Neural network identification

The RHONN proposed for this application is as follows:

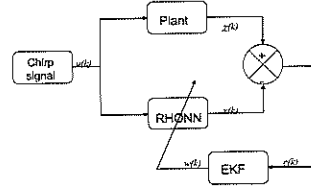


Fig.1. Identification scheme

$$\begin{aligned}
 x_1(k+1) &= w_{11}(k) S(\omega(k)) + w_{12}(k) S(\omega) S(\psi^\beta(k)) i^\alpha(k) \\
 &\quad + w_{13}(k) S(\omega) S(\psi^\alpha(k)) i^\beta(k) \\
 x_2(k+1) &= w_{21}(k) S(\omega(k)) S(\psi^\beta(k)) + w_{22}(k) i^\beta(k) \\
 x_3(k+1) &= w_{31}(k) S(\omega(k)) S(\psi^\alpha(k)) + w_{32}(k) i^\alpha(k) \\
 x_4(k+1) &= w_{41}(k) S(\psi^\alpha(k)) + w_{42}(k) S(\psi^\beta(k)) \\
 &\quad + w_{43}(k) S(i^\alpha(k)) + w_{44}(k) u^\alpha(k) \\
 x_5(k+1) &= w_{51}(k) S(\psi^\alpha(k)) + w_{52}(k) S(\psi^\beta(k)) \\
 &\quad + w_{53}(k) S(i^\beta(k)) + w_{54}(k) u^\beta(k)
 \end{aligned}$$

The training is performed on-line, using a series-parallel configuration as illustrated in Fig. 1. During the identification process the plant and the NN operates in open-loop. Both of them (plant and NN) have the same input vector $[u_\alpha \ u_\beta]^T$; u_α and u_β are chirps functions with 170volts of maximal amplitude and incremental frequencies from 0Hz to 250Hz and 0Hz to 200Hz respectively. All the NN states are initialized in a random way as well as the weights vectors. It is important remark that the initial conditions of the plant are completely different from the initial conditions for the NN. The identification is performed using (8) with $i = 1, 2, \dots, n$ with n the dimension of plant states ($n = 6$).

5.3 Simulation results

The results of the simulation are presented in Figs. 2-8. Fig. 2 shows the identification of rotor angular displacement; Fig. 3 displays the identification performance for the speed rotor; Fig. 4 and Fig. 5 present the identification performance for the fluxes in phase α and β , respectively. Figs 6 and 7 portray the identification performance for currents in phase α and β , respectively. Finally, the weights evolution are presented in Fig. 8.

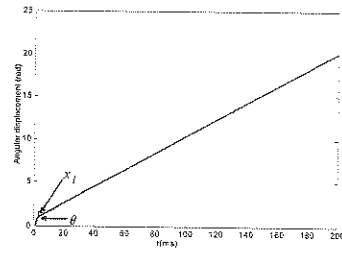


Fig. 2. Angular displacement identification

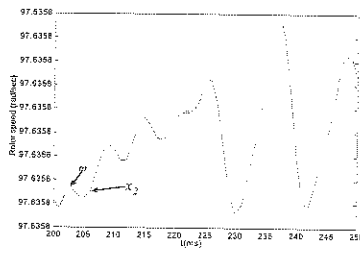


Fig. 3. Rotor speed identification

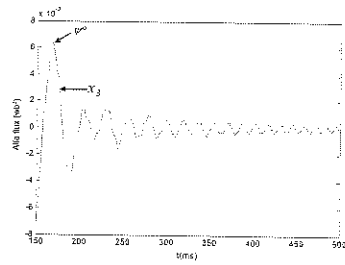


Fig. 4. ψ^α identification

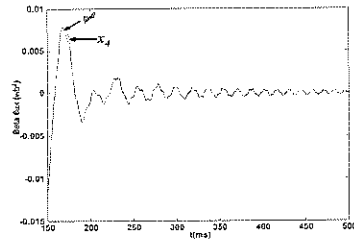


Fig. 5. ψ^β identification

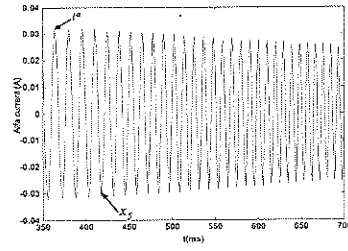


Fig. 6. i^α identification

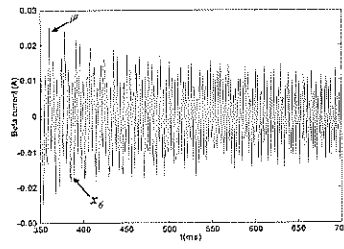


Fig. 7. i^β identification

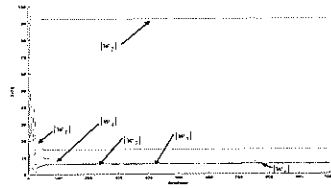


Fig. 8. Weights evolution

6 Conclusions

This paper has presented the application of recurrent high order neural networks to identification of discrete-time nonlinear systems. The training of the neural networks was performed on-line using an extended Kalman filter. The boundedness of the identification error was established on the basis of the Lyapunov approach. Simulation results illustrate the applicability of the proposed identification methodology. Researches are being pursued to develop new discrete-time nonlinear adaptive control based on the discussed identification scheme.

Acknowledgement: The authors thank the support of CONACYT Mexico, through Projects 46069Y and 57801Y. The first author also thanks the support of "CONACYT Fondo Institucional".

References

1. C. K. Chui and G. Chen, *Kalman Filtering with Real-Time Applications*, 3rd ed., Springer-Verlag, New York, USA, 1998.
2. S. S. Ge, J. Zhang and T.H. Lee, "Adaptive neural network control for a class of MIMO nonlinear systems with disturbances in discrete-time", *IEEE Transactions on Systems, Man and Cybernetics*, Part B, Vol. 34, No. 4, August, 2004.
3. R. Grover and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, 2nd ed., John Wiley and Sons, New York, USA, 1992.
4. S. Haykin, *Neural Networks. A comprehensive foundation*, 2nd ed., Prentice Hall, New Jersey, USA, 1999.
5. Y. H. Kim and F. L. Lewis, *High-Level Feedback Control with Neural Networks*, World Scientific, Singapore, 1998.
6. A. G. Loukianov, J. Rivera and J. M. Cañedo, "Discrete-time sliding mode control of an induction motor", *Proceedings IFAC'02*, Barcelone, Spain, July, 2002.
7. E. N. Sanchez and J. L. Ricalde, "Trajectory tracking via adaptive recurrent neural control with input saturation", *Proceedings of International Joint Conference on Neural Networks'03*, Portland, Oregon, USA, July, 2003.
8. E. N. Sanchez, A. Y. Alanis and G. Chen, "Recurrent neural networks trained with Kalman filtering for discrete chaos reconstruction", *Proceedings of Asian-Pacific Workshop on Chaos Control and Synchronization'04*, Melbourne, Australia, July, 2004.
9. G. A. Rovithakis and M. A. Chistodoulou, *Adaptive Control with Recurrent High-Order Neural Networks*, Springer Verlag, New York, USA, 2000.
10. S. Singhal and L. Wu, *Training multilayer perceptrons with the extended Kalman algorithm*, in D. S. Touretzky (ed), *Advances in Neural Information Processing Systems 1*, pp. 133-140, Morgan Kaufmann, San Mateo, CA, USA, 1989.
11. Y. Song and J. W. Grizzle, *The extended Kalman Filter as Local Asymptotic Observer for Discrete-Time Nonlinear Systems*, *Journal of Mathematical systems, Estimation and Control*, Vol. 5, No. 1, pp. 59-78, Birkhauser-Boston, 1995.
12. W. Yu and X. Li, "Discrete-time neuro identification without robust modification", *IEE Proc-Control Theory Appl.*, Vol. 150, No. 3, May, 2003.
13. W. Yu and X. Li, "Nonlinear system identification using discrete-time recurrent neural networks with stable learning algorithms", *Information Sciences*, Vol. 158, pp. 131-147, 2004.