15th Triennial World Congress, Barcelona, Spain

ROBUST BLACK-BOX IDENTIFICATION WITH DISCRETE-TIME NEURAL NETWORKS

Wen Yu

CINVESTAV-IPN, Departamento de Control Automatico, Av.IPN 2508, A.P.14-740, Mexico D.F., 07000, Mexico, e-mail: yuw@ctrl.cinvestav.mx

Abstract: In general, neural networks cannot match nonlinear systems exactly, neuro identifier has to include robust modification in order to guarantee Lyapunov stability. In this paper input-to-state stability approach is applied to access robust training algorithms of discrete-time neural networks. We conclude that the gradient descent law and the backpropagation-like algorithm for the weights adjustment are stable in an L_∞ sense and robust to any bounded uncertainties.

Keywords: robust identification, neural networks

1. INTRODUCTION

Resent results show that neural network technique seems to be very effective to identify a wide class of complex nonlinear systems when we have no complete model information. There are many results on robust identification of continuous-time nonlinear systems(Kosmatopoulos *el al.*, 1995), (Suykens *el al.*, 1999) and (Yu *el al.*, 2001). Lyapunov approach can be used directly to obtain robust training algorithms for continuous-time neural networks (Poznyak *el al.*, 2001). On the other hand, discrete-time neural networks are more convenient for real applications. The dynamic behavior and stability of discrete-time neural networks have been widely studied. The absolute stability of discrete-time recurrent neural networks were analyzed in (Feng and Michel , 1999)(Jin and Gupta, 1999). Identification via discrete-time neural networks has also received much attention over the past decade. (Polycarpou and Ioannou, 1992) assumed neural networks could represent nonlinear systems exactly, and concluded that backpropagation-type algorithm guaranteed exact convergence. (Jagannathan and Lewis, 1996) and (Song, 1998) considered bounded modeling errors, robust modifications were introduced in the weight tuning algorithms. Since unmodeled dynamic will cause parameter drift, and leads to instability solutions (Ioannou and Sun, 1996), robust modification terms have to be added to assure that the adaptive learning processes are stable, for example, $\sigma-$modification in (Kosmatopoulos *el al.*, 1995), modified $\delta-$rule in (Jagannathan and Lewis, 1996) and dead-zone in (Song, 1998).

It is well known that neuro identification is in sense of black-box approximation. All of the uncertainties can be considered as parts of the black-box, i.e., unmodeled dynamics are inside of the model, not as structure uncertainties. Hence the commonly-used robustifying techniques are not necessary. By using passivity theory, we successfully proved that for continuous-time recurrent neural networks, gradient descent algorithms without robust modification were stable and robust to any bounded uncertainties (Yu and Li, 2001), and for continuous-time identification they were also robustly stable (Yu and Li, 2001a). But for discrete-time system identification, do discrete-time neural networks have the similar properties? That is the motivation of this paper. To the best of our knowledge, identification without robust modification via discrete-time neural networks has not yet been established in the literature.

Input-to-state stability is another elegant approach to analyze stability compared with Lyapunov method. It can lead to general conclusions on the stability by using only input-state characteristics. We will use input-to-state stability approach to obtain learning laws without robust modifications in this paper. A simple simulation gives the effectiveness of the proposed algorithm.

2. PRELIMINARIES

The main concern of this section is to understand some concepts of input-to-state stability (ISS). Consider following discrete-time nonlinear system

$$\begin{array}{c} x(k+1) = f\left[x\left(k\right), u\left(k\right)\right] \\ y(k) = h\left[x\left(k\right)\right] \end{array}, \tag{1}$$

where $u\left(k\right) \in \Re^m$ is the input vector, $x\left(k\right) \in \Re^n$ is a state vector, and $y\left(k\right) \in \Re^l$ is the output vector. f and h are general nonlinear smooth function $f, h \in C^\infty$. Following (Jiang and Wang, 2001), let us now recall following definitions.

Definition 1. A system (1) is said to be globally *input-to-state stability* if there exists a $\mathcal{K}$-function $\gamma(\cdot)$ (continuous and strictly increasing $\gamma(0) = 0$) and $\mathcal{KL}$ -function $\beta\left(\cdot\right)$ ($\mathcal{K}$-function and $\lim\limits_{s_k \to \infty} \beta\left(s_k\right) = 0$) , such that, for each $u \in L_\infty$ $\left(\sup\left\{\left\|u(k)\right\|\right\} < \infty\right)$ and each initial state $x^0 \in R^n$, it holds that

$$\left\|x\left(k, x^0, u\left(k\right)\right)\right\| \leq \beta\left(\left\|x^0\right\|, k\right) + \gamma\left(\left\|u\left(k\right)\right\|\right)$$

Definition 2. A smooth function $V : \Re^n \to \Re \geq 0$ is called a smooth ISS-Lyapunov function for system (1) if: (I) there exists a $\mathcal{K}_\infty$-function $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$ ($\mathcal{K}$-function and $\lim\limits_{s_k \to \infty} \beta\left(s_k\right) = \infty$) such that

$$\alpha_1(s) \leq V\left(s\right) \leq \alpha_2(s), \quad \forall s \in \Re^n$$

(2) There exist a $\mathcal{K}_\infty$-function $\alpha_3(\cdot)$ and a $\mathcal{K}$-function $\alpha_4(\cdot)$ such that

$$V_{k+1} - V_k \leq -\alpha_3(\left\|x\left(k\right)\right\|) + \alpha_4(\left\|u\left(k\right)\right\|)$$

for all $x\left(k\right) \in \Re^n, u\left(k\right) \in \Re^m$

Theorem 1. (Jiang and Wang, 2001) Consider a system (1), the following are equivalent

- It is input-to-state stability (ISS).
- It is robustly stable.
- It admit a smooth ISS-Lyapunov function.

Property. If a system is input-to-state stability, the behavior of the system should remain bounded when its inputs are bounded.

Since $f, h \in C^\infty$, the state equation (1) can be expressed in a NARMA model (Narendra and Mukhopadhyay, 1997)

$$\begin{array}{c} y(k+1) = F[y\left(k\right), y\left(k-1\right), \cdots \\ u\left(k\right), u\left(k-1\right), \cdots] \end{array} \tag{2}$$

In this paper, we consider following nonlinear discrete-time single-input, single-output plant

$$\begin{array}{c} y(k) = f[y\left(k-1\right), y\left(k-2\right), \cdots \\ u\left(k-d\right), u\left(k-d-1\right), \cdots] \end{array} \tag{3}$$

where $f\left(\cdot\right)$ is an unknown nonlinear difference equation representing the plant dynamics, $u\left(k\right)$ and $y\left(k\right)$ are measurable scalar input and output, d is time delay. Since the nonlinear systems (3) and (1) are equivalent, ISS approach can be applied in both of them.

3. DISCRETE-TIME NEURO IDENTIFICATION WITHOUT ROBUST MODIFICATION

First, we consider a single-layer neural network which is represented as

$$\widehat{y}\left(k\right) = \phi\left[W_k x\left(k\right)\right] \tag{4}$$

where $\widehat{y}\left(k\right)$ is the output of neural network, $W_k \in R^{1\times n}$ is weight matrix, $x\left(k\right) \in R^{n\times 1}$ is a input vector, $\phi(\cdot)$ is a sigmoidal activation function.

The identified nonlinear system (3)can be represented as

$$y(k) = f\left[x\left(k\right)\right] \tag{5}$$

where

$$\begin{array}{c} x\left(k\right) = \left[y\left(k-1\right), y\left(k-2\right), \cdots \right. \\ \left. u\left(k-d\right), u\left(k-d-1\right), \cdots\right] \end{array}$$

According to the Stone-Weierstrass theorem (Cybenko, 1998), this general nonlinear smooth function can be written as

$$y\left(k\right) = \phi\left[W^* x\left(k\right)\right] - \mu\left(k\right) \tag{6}$$

where W^* is optimal weight, $\mu\left(k\right)$ is the modeling error. Since ϕ is bounded function and the output of the plant is assumed bounded, $\mu\left(k\right)$ is bounded as $\mu^2\left(k\right) \leq \overline{\mu}$, $\overline{\mu}$ is an unknown positive constant. The neuro identification error is defined as

$$e\left(k\right) = \widehat{y}\left(k\right) - y\left(k\right) \tag{7}$$

Using Taylor series, the identification error can be represented as

$$\begin{aligned} e(k) &= \phi\left[W_k x(k)\right] - \phi\left[W^* x(k)\right] + \mu(k) \\ &= \phi^{'}\left[W_k x(k)\right] \widetilde{W}_k x(k) + \zeta(k) \end{aligned} \tag{8}$$

where $\widetilde{W}_k = W_k - W^*$, $\zeta(k) = \varepsilon(k) + \mu(k)$, $\varepsilon(k)$ is second order approximation error. $\phi^{'}$ is the derivative of nonlinear activation function $\phi(\cdot)$. Since ϕ is a sigmoidal activation function, $\varepsilon(k)$ is bounded as $\varepsilon^2(k) \leq \overline{\varepsilon}$, $\overline{\varepsilon}$ is an unknown positive constant. The following theorem gives a stable learning algorithm of discrete-time single-layer neural network.

Theorem 2. If we use the single-layer neural network (4) to identify nonlinear plant (5), the following gradient updating law without robust modification can make identification error $e(k)$ bounded (stable in an L_∞ sense)

$$W_{k+1} = W_k - \eta_k e(k) \phi^{'} x^T(k) \tag{9}$$

where $\eta_k = \dfrac{\eta}{1 + \left\| \phi^{'} x^T(k) \right\|^2}$, $0 < \eta \leq 1$

Proof. We select Lyapunov function as

$$V_k = \left\| \widetilde{W}_k \right\|^2 = \sum_{i=1}^{n} \widetilde{w}_i^2 = tr\left\{ \widetilde{W}^T \widetilde{W} \right\}$$

From the updating law (9)

$$\begin{aligned} &\widetilde{W}_{k+1} = \widetilde{W}_k - \eta_k e(k) \phi^{'} x^T(k) \\ &\Delta V_k = V_{k+1} - V_k \\ &= \left\| \widetilde{W}_k - \eta_k e(k) \phi^{'} x^T(k) \right\|^2 - \left\| \widetilde{W}_k \right\|^2 \\ &= \eta_k^2 e^2(k) \left\| \phi^{'} x^T(k) \right\|^2 - 2\eta_k \left\| e(k) \phi^{'} \widetilde{W}_k x(k) \right\| \end{aligned}$$

Using (8) and $0 < \eta \leq 1$, $0 \leq \eta_k \leq \eta \leq 1$,

$$\begin{aligned} &\Delta V_k = \eta_k^2 e^2(k) \left\| \phi^{'} x^T(k) \right\|^2 \\ &-2\eta_k \left\| e(k) \left[e(k) - \zeta(k) \right] \right\| \\ &\leq \eta_k^2 e^2(k) \left\| \phi^{'} x^T(k) \right\|^2 \\ &-2\eta_k e^2(k) + \eta_k e^2(k) + \eta_k \zeta^2(k) \\ &= -\eta_k \left[1 - \eta \frac{\left\| \phi^{'} x^T(k) \right\|^2}{1 + \left\| \phi^{'} x^T(k) \right\|^2} \right] e^2(k) \\ &+\eta_k \zeta^2(k) \\ &\leq -\pi e^2(k) + \eta \zeta^2(k) \end{aligned}$$

where

$$\begin{aligned} \pi &= \frac{\eta}{1+\kappa} \left[1 - \frac{\kappa}{1+\kappa} \right] > 0 \\ \kappa &= \max_k \left\| \phi^{'} x^T(k) \right\|^2 \end{aligned}$$

Since

$$n \min\left(\widetilde{w}_i^2\right) \leq V_k \leq n \max\left(\widetilde{w}_i^2\right)$$

where $n \times \min\left(\widetilde{w}_i^2\right)$ and $n \times \max\left(\widetilde{w}_i^2\right)$ are $\mathcal{K}_\infty$-functions, and $\pi e^2(k)$ is an $\mathcal{K}_\infty$-function, $\eta \zeta^2(k)$ is a $\mathcal{K}$-function, so V_k admits the smooth ISS-Lyapunov function. From Theorem 1, the identification process is ISS. The input is the approximation error $\zeta(k) = \varepsilon(k) + \mu(k)$, the state is the identification error $e(k)$. Because the input $\zeta(k)$ is bounded, the state $e(k)$ is bounded. ■

Remark 1. (9) is the gradient descent algorithm, the normalizing learning rate η_k is time-varying in order to assure the identification process is stable. This learning law is easier to use, because we do not need to care about how to select a better learning rate to assure both fast convergence and stability. No any prior information is required. If we select η as dead-zone function,

$$\begin{cases} \eta = 0 & \text{if } e(k) \leq \overline{\varepsilon} + \overline{\mu} \\ \eta = \eta_0 & \text{if } e(k) > \overline{\varepsilon} + \overline{\mu} \end{cases}$$

(9) is the same as (Song, 1998) and (Yu *el al.*, 2001). If a more $\sigma-$modification term or modified $\delta-$rule term are added in (9), it becomes that of (Jagannathan and Lewis, 1996) or that of (Lewis *el al.*, 1996). But all of them need additional information of the modeling error. And the identification error is enlarged by the robust modifications (Ioannou and Sun, 1996).

Remark 2. If we add a fixed matrix $A \in R^{1 \times m}$ in (4), ϕ is m-dimension vector function

$$\widehat{y}(k) = A\phi\left[W_k x(k)\right], \quad W_k \in R^{m \times n} \tag{10}$$

The learning law (9) becomes

$$\begin{aligned} &W_{k+1} = W_k - \\ &\frac{\eta}{1 + \left\| \phi^{'} A^T x^T(k) \right\|^2} e(k) \phi^{'} A^T x^T(k) \end{aligned}$$

This is the same as the Equ.20 in (Polycarpou and Ioannou, 1992), but they assumed the neural networks (10) can match nonlinear system (5) exactly. In our case, modeling errors $\mu(k)$ and $\varepsilon(k)$ are allowed.

Now, we consider multilayer neural network which is represented as (Lewis *el al.*, 1996)

$$\widehat{y}(k) = V_k \phi\left[W_k x(k)\right] \tag{11}$$

where the scalar output $\widehat{y}(k)$ and vector state $x(k) \in R^{n \times 1}$ are the same as (4), the weights in output layer are $V_k \in R^{1 \times m}$, the weights in hidden layer are $W_k \in R^{m \times n}$, ϕ is m-dimension vector function. Similar as (6), the nonlinear plant (5) may be expressed as

$$y(k) = V^* \phi[W^* x(k)] - \mu(k)$$

where V^* and W^* are unknown optimal values of V_k and W_k. The nonlinear plant (5) may be also expressed as

$$y(k) = V^0 \phi[W^* x(k)] - \delta(k) \tag{12}$$

where V^0 is an initial value of V_k. In general, $\|\delta(k)\| \geq \|\mu(k)\|$. The identification error (7) for multilayer neural networks is changed as

$$\begin{aligned} &e(k) = V_k \phi[W_k x(k)] - V^0 \phi[W^* x(k)] + \delta(k) \\ &= V_k \phi[W_k x(k)] - V^0 \phi[W_k x(k)] \\ &+ V^0 \phi[W_k x(k)] - V^0 \phi[W^* x(k)] + \delta(k) \\ &= \widetilde{V}_k \phi + V^0 \phi' \widetilde{W}_k x(k) + \zeta_1(k) \end{aligned}$$

where $\widetilde{V}_k = V_k - V^0$, $\zeta_1(k) = V^0 \varepsilon(k) + \delta(k)$. The following theorem gives a stable backpropagation-like algorithm for discrete-time multilayer neural network.

Theorem 3. If we use the multilayer neural network (11) to identify nonlinear plant (5), the following gradient updating law without robust modification can make identification error $e(k)$ bounded

$$\begin{aligned} W_{k+1} &= W_k - \eta_k e(k) \phi'[W_k x(k)] V^{0T} x^T(k) \\ V_{k+1} &= V_k - \eta_k e(k) \phi^T[W_k x(k)] \end{aligned} \tag{13}$$

where

$$\eta_k = \frac{\eta}{1 + \left\|\phi' V^{0T} x^T(k)\right\|^2 + \|\phi\|^2}$$

$0 < \eta \leq 1$

Proof. Lyapunov function is $L_k = \left\|\widetilde{W}_k\right\|^2 + \left\|\widetilde{V}_k\right\|^2$. From the updating law (13), we have

$$\begin{aligned} \widetilde{W}_{k+1} &= \widetilde{W}_k - \eta_k e(k) \phi' V^{0T} x^T(k) \\ \widetilde{V}_{k+1} &= \widetilde{V}_k - \eta_k e(k) \phi^T \end{aligned}$$

Since ϕ' is diagonal matrix,

$$\begin{aligned} &\Delta L_k = \left\|\widetilde{W}_k - \eta_k e(k) \phi' V^{0T} x^T(k)\right\|^2 \\ &+ \left\|\widetilde{V}_k - \eta_k e(k) \phi^T\right\|^2 - \left\|\widetilde{W}_k\right\|^2 - \left\|\widetilde{V}_k\right\|^2 \\ &= \eta_k^2 e^2(k) \left(\left\|\phi' V^{0T} x^T(k)\right\|^2 + \|\phi\|^2\right) \\ &-2\eta_k \|e(k)\| \left\|V^0 \phi' \widetilde{W}_k x(k) + \widetilde{V}_k \phi\right\| \\ &= \eta_k^2 e^2(k) \left(\left\|\phi' V^{0T} x^T(k)\right\|^2 + \|\phi\|^2\right) \\ &-2\eta_k \|e(k)[e(k) - \zeta_1(k)]\| \\ &\leq -\eta_k e^2(k) \left[1 - \eta_k \left(\left\|\phi' V^{0T} x^T(k)\right\|^2 + \|\phi\|^2\right)\right] \\ &+ \eta \zeta_1^2(k) \\ &\leq -\pi_1 e^2(k) + \eta \zeta_1^2(k) \end{aligned}$$

where $\pi_1 = \dfrac{\eta}{1+\kappa_1}\left[1 - \dfrac{\kappa_1}{1+\kappa_1}\right] > 0$, $\kappa_1 = \max\limits_k \left(\left\|\phi' V^{0T} x^T(k)\right\|^2 + \|\phi\|^2\right)$. Since

$$\begin{aligned} &n\left[\min\left(\widetilde{w}_i^2\right) + \min\left(\widetilde{v}_i^2\right)\right] \\ &\leq L_k \leq n\left[\max\left(\widetilde{w}_i^2\right) + \max\left(\widetilde{v}_i^2\right)\right] \end{aligned}$$

Since $\pi_1 e^2(k)$ is an $\mathcal{K}_\infty$-function, $\eta \zeta_1^2(k)$ is an $\mathcal{K}$-function, L_k admits the smooth ISS-Lyapunov function. From Theorem 1, the identification process is input-to-state stability. The input is the approximation error $\zeta_1(k) = V^0 \varepsilon(k) + \delta(k)$, the state is the identification error $e(k)$. Because the input is bounded as $\overline{\varepsilon}$ and $\overline{\mu}$, the state $e(k)$ is bounded. ■

Remark 3. The identification error will converge to the ball radium the upper bounded of ζ_1, and it is influenced by the initial matrix V^0. Since V^0 does not influence the stability property, we may select any value for V^0 at first. The algorithm (13) can make the identification error convergent, i.e. V_k will make the identification error smaller than that of V^0. V^0 may be selected by following off-line steps:

(1) Start from any initial value for V^0.
(2) Do identification with this V^0.until $k = T$.
(3) If the $\|e(T)\| < \|e(0)\|$, let V_T as the new initial condition, i..e, $V^0 = V_T$, repeat the identification process, go to 2
(4) If the $\|e(T)\| \geq \|e(0)\|$, stop this off-line identification, now V_T is the final value for V^0.

With this V^0 we may start on-line identification.

4. SIMULATION

We will use a very simple example to illustrate the algorithm and the stable issue proposed in this paper. The identified plant (5) is

$$\begin{aligned} &y(k) = -0.12y(k-1) \\ &+0.7y(k-2) + u(k-2) \end{aligned} \tag{14}$$

The single layer neural network (4) is

$$\widehat{y}(k) = 25 \times \tanh[W_k x(k)] \tag{15}$$

where

$$\begin{aligned} W_k &= [w_{1,k}, w_{2,k}, w_{3,k}] \\ x(k) &= [y(k-1), y(k-2), u(k-2)]^T \end{aligned}$$

The learning algorithm (9) is

$$W_{k+1} = W_k - \eta \frac{[\widehat{y}(k) - y(k)] S(x) x^T(k)}{1 + 25 S(x) Y(y,u)} \tag{16}$$

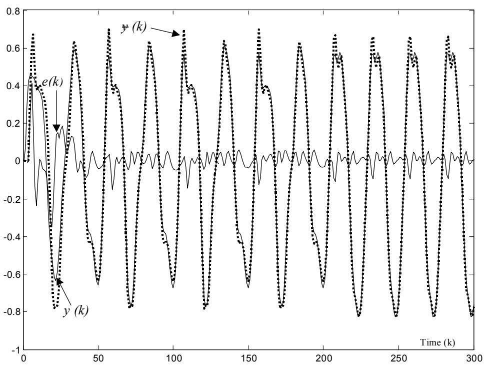

Fig. 1. Neuro identification with $\eta = 1$

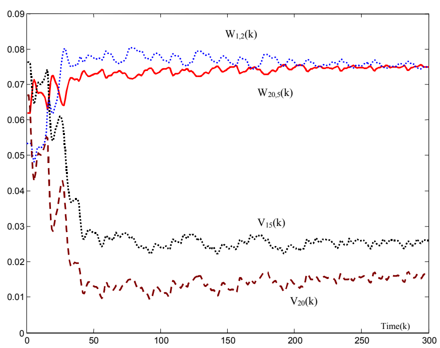

Fig. 2. Nonlinear system identification

where

$$S(x) = \sec h^2 \left[W_k x(k) \right]$$
$$Y(y,u) = \left[y(k-1)^2 + y(k-2)^2 + u(k-2)^2 \right]$$

First we select η in (16) as $\eta = 1$. In the learning phase we use a square ware $u(k)$ =round$(\frac{k}{30})$ to obtain the optimal weight of the neural network. Then we fix the weight, and put the input

$$u_1(k) = \sin(\frac{k}{30}) + \cos(\frac{k}{20}) + 2\sin(\frac{k}{20}) \quad (17)$$

into the plant (14) and neural network (15), the outputs are shown in Fig.1. But if we select $\eta = 2$, the learning process becomes unstable. Theorem 2 gives a necessary condition for stable learning, i.e., if $\eta \leq 1$ the learning process is stable. $\eta > 1$ may make the learning process unstable, in this example $\eta = 2$.Now we test a nonlinear system. We use a single-link robot model as in (Song *el al.*, 1999)

$$y(k+2) = 2(1-T)y(k+1) + \\ (2T-1)y(k) + 10T^2 \sin y(k) + u(k)$$

where $T = 0.01$ is sampleing time. We use the control input as in (17), the learning procedure is shown in Fig.2

5. CONCLUSION

In this paper we study nonlinear system identification by the discrete-time single layer and multilayer neural networks. By using ISS approach, we conclude that the commonly-used robustifying techniques, such as dead-zone and $\sigma-$modification, are not necessary for the gradient descent law and the backpropagation-like algorithm. Further works will be done on discrete-time recurrent neural networks and neuro control based on ISS approach.

6. REFERENCES

[1]

G.Cybenko (1998), Approximation by Superposition of Sigmoidal Activation Function, *Math.Control, Sig Syst*, Vol.2, 303-314.

Z.Feng and A.N.Michel (1999), Robustness Analysis of a Class of Discrete-Time Systems with Applications to Neural Networks, *Proc. of American Control Conference*, 3479-3483, San Deigo

P.A.Ioannou and J.Sun (1996), *Robust Adaptive Control*, Prentice-Hall, Inc, Upper Saddle River: NJ

S.Jagannathan and F.L.Lewis (1996), Identification of Nonlinear Dynamical Systems Using Multilayered Neural Networks, *Automatica*, Vol.32, No.12, 1707-1712 .

L.Jin and M.M.Gupta (1999), Stable Dynamic Backpropagation Learning in Recurrent Neural Networks, *IEEE Trans. Neural Networks*, Vol.10, No.6, 1321-1334 .

Z.P.Jiang and Y.Wang (2001), Input-to-State Stability for Discrete-Time Nonlinear Systems, *Automatica*, Vol.37, No.2, 857-869.

E.B.Kosmatopoulos, M.M.Polycarpou, M.A.Christodoulou and P.A.Ioannpu (1995), High-Order Neural Network Structures for Identification of Dynamical Systems, *IEEE Trans. on Neural Networks*, Vol.6, No.2, 442-431.

F.L.Lewis, A.Yesildirek and K.Liu, Multilayer Neural-Net Robot Controller with Guaranteed Tracking Performance, *IEEE Trans. Neural Networks*, Vol.7, No.2, 388-399, 1996.

K.S.Narendra
and S.Mukhopadhyay (1997), Adaptive Control Using Neural Networks and Approximate Models, *IEEE Trans. Neural Networks*, Vol.8, No.3, 475-485 .

M.M.Polycarpou and P.A.Ioannou (1992), Learning and Convergence Analysis of Neural-Type Structured Networks, *IEEE Trans. Neural Networks*, Vol.3, No.1, 39-50

A.S.Poznyak, E.N.Sanchez and W.Yu (2001), *Differential Neural Networks for Robust Nonlinear Control-Identification, State Estima-*

tion and Trajectory Tracking, World Scientific Publishing Pte Ltd

Q.Song (1998), Robust Training Algorithm of Multilayered Neural Networks for Identification of Nonlinear Dynamic Systems, *IEE Proceedings - Control Theory and Applications*, Vol.145, No.1, 41-46

Q.Song, J.Xiao and Y.C.Soh (1999), Robust Backpropagation Training Algorithm for Multilayered Neural Tracking Controller, *IEEE Trans. Neural Networks*, Vol.10, No.5, 1133-1141.

J.A.K.Suykens, J.Vandewalle and B.De Moor (1999), Lur's Systems with Multilayer Perceptron and Recurrent Neural Networks; Absolute Stability and Dissipativity, *IEEE Trans. on Automatic Control,* Vol.44, 770-774.

W.Yu and X. Li (2001), Some Stability Properties of Dynamic Neural Networks, *IEEE Trans. Circuits and Systems, Part I*, Vol.48, No.1, 256-259.

W.Yu and X. Li (2001), Some New Results on System Identification with Dynamic Neural Networks, *IEEE Trans. Neural Networks*, Vol.12, No.2, 412-417 .

W.Yu, A.S. Poznyak and X.Li, Multilayer Dynamic Neural Networks for Nonlinear System On-line Identification, *International Journal of Control*, accepted for publication.