NEURAL NETWORK BASED VARIABLE STRUCTURE CONTROL FOR NONLINEAR DISCRETE SYSTEMS

G.P. Liu, **V. Kadirkamanathan** and **S.A. Billings**

Department of Automatic Control & Systems Engineering
University of Sheffield, PO Box 600, Mappin Street, Sheffield S1 3JD, UK
Email: guoping/visakan/steve@acse.sheffield.ac.uk

Research Report No. **626**

Department of Automatic Control & Systems Engineering
The University of Sheffield

May 1996

Neural Network Based Variable Structure Control for Nonlinear Discrete Systems

G. P. Liu, V. Kadirkamanathan and S. A. Billings

Department of Automatic Control and Systems Engineering
University of Sheffield, PO Box 600, Mappin Street, Sheffield S1 3JD, U.K.
Tel: +44 1142825231 Fax: +44 1142731729
Email: guoping/visakan/steve@acse.sheffield.ac.uk

Abstract

Neural network based variable structure control is proposed for the design of nonlinear discrete systems. Sliding mode control is used to provide good stability and robustness performance for nonlinear systems. An affine nonlinear neural predictor is introduced to predict the outputs of the nonlinear process and to make the variable structure control algorithm simple and easy to implement. When the predictor model is inaccurate, variable structure control with sliding modes is used to improve the stability of the system. A recursive weight learning algorithm for the neural networks based affine nonlinear predictor is also developed and the convergence of both the weights and the estimation error is analysed.

Keywords: Neural networks, nonlinear discrete systems, variable structure control, sliding mode control.

1 Introduction

Variable structure control with sliding modes was first proposed in the early 1950's [4] [9] [21] and has subsequently been used in the design of a wide spectrum of system types including linear and nonlinear systems, large-scale and infinite-dimensional systems, and stochastic systems. It has also been applied to a wide variety of engineering systems. The most distinguished feature of variable structure control based on sliding modes is the ability to improve the robustness of systems which are subject to uncertainty. If however the uncertainty exceeds the values allowed for the design, the sliding mode can not be attained and this results in an undesirable response [21]. In the continuous-time case this problem was solved by combining variable structure and adaptive control [20], but this requires that all the system variables are available and can be measured. This case has also been discussed for linear discrete systems using input-output plant model [7] [18]. But, as far as authors are aware the above problem has not been solved for nonlinear discrete systems where the input-output model is unknown.

Recently, neural networks have become an attractive tool which can be used to construct the model of complex nonlinear processes. This is because neural networks have an inherent ability of learning and approximating a nonlinear function arbitrarily well. This therefore provides a possible way of modelling complex nonlinear processes effectively. A large number of identification and control structures have been proposed on the basis of neural networks in recent years (see, for example, [1] [3] [5] [15] [17] [19]).

This paper presents a neural network based variable structure controller design procedure for unknown nonlinear discrete systems. A neural network based affine nonlinear predictor is introduced so that the control algorithm is simple and easy to implement. Two cases are considered for variable structure neural control. First, a performance function which is only concerned with minimization of the prediction error is considered. Second, a performance function which includes the minimization of the prediction error and the control input is studied. A recursive weight learning algorithm of a neural networks for the neural network affine nonlinear predictor is also developed. This algorithm can be used for both on-line and off-line weight training. It is shown that both the weights of the neural networks and the estimation error converge.

The paper begins in Section 2 with the structure of the affine nonlinear predictors which is based on neural networks. The variable structure neural control is given in Section 3. The generalized variable structure neural control is discussed in Section 4. Section 5 develops the recursive weight learning algorithm for the neural networks used for the d-step ahead predictor and the properties of the algorithm are analysed. Finally, simulation results are shown in Section 6.

2 Neural Network Based Predictors

This paper considers a discrete-time affine nonlinear control system which has been described by

$$y_t = F(\mathbf{x}_t) + G(\mathbf{x}_t)u_{t-d} \tag{1}$$

where $F(.)$ and $G(.)$ are nonlinear functions, y is the output and u the control input, respectively, the vector $\mathbf{x}_t = [\, y_{t-1} \quad y_{t-2} \quad \ldots \quad y_{t-n} \,]$, n is the order of $y(t)$ and d is the time-delay of the system. It is assumed that the order n and the time delay d are known the nonlinear functions $F(.)$ and $G(.)$ are smooth but unknown, and $G(.)$ is bounded away from zero.

Based on the affine nonlinear system described by Eq.(1), we present a d-step ahead affine nonlinear predictor to compensate for the influence of the time-delay. This predictor uses sequences of both past inputs and outputs of the precess upto the sampling time t to construct the predictive models, which are of the following form:

$$\hat{y}_{t+d} = \hat{F}(\mathbf{x}_t) + \hat{G}(\mathbf{x}_t)u_t \tag{2}$$

where $\hat{F}(\mathbf{x}_t)$ and $\hat{G}(\mathbf{x}_t)$ are nonlinear functions of the vector $\mathbf{x}_t$ which are to be estimated.

Due to the arbitrary approximation property of neural networks, the nonlinear functions $\hat{F}(\mathbf{x}_t)$ and $\hat{G}(\mathbf{x}_t)$ can be approximated by single hidden layer networks. This is expressed by

$$\hat{F}(\mathbf{x}_t) = \sum_{k=1}^{N_0} f_k \phi_k(\mathbf{x}_t) \tag{3}$$

$$\hat{G}(\mathbf{x}_t) = \sum_{k=1}^{N_1} g_k \gamma_k(\mathbf{x}_t) \tag{4}$$

where $\phi_k(\mathbf{x}_t)$ and $\phi_k(\mathbf{x}_t)$ are the basis functions of the networks, N_0 and N_1 denote the size of the networks. Define the weight and basis function vectors of the neural networks as

$$\bar{F} = [\, f_1 \quad f_2 \quad \ldots \quad f_{N_0} \,]^T \tag{5}$$

$$\bar{G} = [\, g_1 \quad g_2 \quad \ldots \quad g_{N_1} \,]^T \tag{6}$$

$$\bar{\Phi}_t = [\, \phi_1(\mathbf{x}_t) \quad \phi_2(\mathbf{x}_t) \quad \ldots \quad \phi_{N_0}(\mathbf{x}_t) \,]^T \tag{7}$$

$$\bar{\Gamma}_t = [\, \gamma_1(\mathbf{x}_t) \quad \gamma_2(\mathbf{x}_t) \quad \ldots \quad \gamma_{N_1}(\mathbf{x}_t) \,]^T \tag{8}$$

Then the neural network based predictors can be rewritten by

$$\hat{y}_{t+d} = \bar{F}^T \bar{\Phi}_t + \bar{G}^T \bar{\Gamma}_t u_{t+d} \tag{9}$$

It is well known from the universal approximation theory for neural networks that the modelling error of the predictor can be reduced arbitrarily by properly choosing the basis functions and adjusting the weights. There are many basis functions available, e.g., radial functions, sigmoidal functions, polynomial functions and so on. This paper does not intend to discuss how to choose between these. But a recursive learning algorithm for the weight adjustment of the networks used in the predictors will be presented in a later section.

3 Variable Structure Neural Control

Based on the d-step-ahead affine nonlinear predictor modelled using the neural networks described in the previous section, this section considers variable structure neural control using sliding modes. It will be assumed that all the basis functions in the neural network predictor are given but the weights of the network are unknown.

The objective of the control is to minimize the following performance function.

$$J_s = \frac{1}{2}(\hat{y}^*_{t+d} - r_{t+d})^2 \tag{10}$$

where r is the reference input and $\hat{y}^*_{t+d}$ is the optimal d-step-ahead prediction of the output y_t.

For the given neural network structure, the optimal d-step ahead predictor is given by

$$\hat{y}^*_{t+d} = (\bar{F}^*)^T\bar{\Phi}_t + (\bar{G}^*)^T\bar{\Gamma}_t u_t \tag{11}$$

where $\bar{F}^*$ and $\bar{G}^*$ are the optimal estimates of the weights which yield a prediction error within the required accuracy.

Based on the optimal d-step ahead predictor given by Eq.(11), the control input to minimize J_s can be solved analytically and is expressed by

$$u_t = ((\bar{G}^*)^T\bar{\Gamma}_t)^{-1}(r_{t+d} - (\bar{F}^*)^T\bar{\Phi}_t) \tag{12}$$

In practice , it is very difficult to know the optimal weight vectors $\bar{F}^*$ and $\bar{G}^*$ in the affine nonlinear predictor if some uncertainties or disturbances exist in the system. This section considers the use of the above neural predictor and the variable structure controller based on the set Ξ defined below. Let

$$s_{t+d} = \hat{y}^*_{t+d} - r_{t+d} \tag{13}$$
$$\Xi(\varepsilon) = \{\xi : |\xi| \leq \varepsilon\} \tag{14}$$
$$\tilde{F}(\mathbf{x}_t) = (\bar{F}^* - \bar{F})^T\bar{\Phi}_t = \tilde{F}^T\bar{\Phi}_t \tag{15}$$
$$\tilde{G}(\mathbf{x}_t) = (\bar{G}^* - \bar{G})^T\bar{\Gamma}_t = \tilde{G}^T\bar{\Gamma}_t \tag{16}$$

where ε is a positive number. Thus, the following control input is considered:

$$u_t = (\bar{G}^T\bar{\Gamma}_t)^{-1}(r_{t+d} - \bar{F}^T\bar{\Phi}_t + s_t + v_t) \tag{17}$$

From now on, it is assumed that the weight vector $\bar{G}$ is such that $\bar{G}^T\bar{\Gamma}_t$ is bounded away from zero. This reasonable assumption is based on that the nonlinear function $G(.)$ of the system is assumed to be bounded away from zero. The auxiliary control input v_t is chosen to achieve $s_{t+d} = s_t$ and v_t is chosen as the output feedback

$$v_t = \left(\sum_{k=1}^{N_0} a_k\phi_k(\mathbf{x}_t) + \sum_{k=1}^{N_1} b_k\gamma_k(\mathbf{x}_t)\psi_t\right) \tag{18}$$

where

$$\psi_t = (\bar{G}^T\bar{\Gamma}_t)^{-1}(r_{t+d} + s_t - \bar{F}^T\bar{\Phi}_t) \tag{19}$$

a_k and b_k are the coefficients which are to be designed.

The problem is to choose the coefficients a_k and b_k to guarantee the stability of the system. To solve this problem, the following theorem is obtained.

Theorem 1: If the the coefficients a_k and b_k of v_t are chosen as

$$a_k = \begin{cases} -d_0\text{sign}(\phi_k s_t) & s_t \notin \Xi(\sigma_t) \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

$$b_k = \begin{cases} -d_0\text{sign}(\gamma_k \psi_t s_t) & s_t \notin \Xi(\sigma_t) \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

$d_0 > \kappa_t + \mu - 1 > 0$ and $\tau > 1$, where

$$\sigma_t = \frac{\tau(d_0\kappa_t + \mu)^2}{2(d_0 + 1 - \kappa_t - \mu)}\left(\sum_{k=1}^{N_0}|\phi_k(\mathbf{x}_t)| + \sum_{k=1}^{N_1}|\gamma_k(\mathbf{x}_t)\psi_t|\right) \tag{22}$$

$$\kappa_t = \sum_{k=1}^{N_1}|\gamma_k(\mathbf{x}_t)(\bar{G}^T\bar{\Gamma}_t)^{-1}|\mu + 1 \tag{23}$$

$$\mu = \max_{i=1,2,\dots,N_0, j=1,2,\dots,N_1}\left\{|f_i^* - f_i|, |g_j^* - g_j|\right\} \tag{24}$$

then s_{t+d} converges to $\Xi(\sigma_t)$.

Proof: To ensure the stability of the system, the Lyapunov technique is used. Choose the Lyapunov function as

$$V_{t+d} = s_{t+d}^2 \tag{25}$$

The difference of the Lyapunov function can be expressed by the following form:

$$\Delta V_{t+d} = V_{t+d} - V_t = (s_t + \Delta s_{t+d})^2 - s_t^2 = 2s_t\Delta s_{t+d} + (\Delta s_{t+d})^2 \tag{26}$$

Using the control input given by Eq.(17), it can be shown form Eq.(13) that

$$\begin{aligned} s_{t+d} &= F^*(\mathbf{x}_t) + G^*(\mathbf{x}_t)u_t - r_{t+d} \\ &= (\tilde{F} + \bar{F})^T\bar{\Phi}_t + (1 + \tilde{G}^T\bar{\Gamma}_t(\bar{G}^T\bar{\Gamma}_t)^{-1})(r_{t+d} - \bar{F}^T\bar{\Phi}_t + s_t + v_t) - r_{t+d} \\ &= \tilde{F}^T\bar{\Phi}_t + \tilde{G}^T\bar{\Gamma}_t\psi_t + (1 + \tilde{G}^T\bar{\Gamma}_t(\bar{G}^T\bar{\Gamma}_t)^{-1})v_t + s_t \end{aligned} \tag{27}$$

and it then follows that

$$\Delta s_{t+d} = \tilde{F}^T\bar{\Phi}_t + \tilde{G}^T\bar{\Gamma}_t\psi_t + \rho_t v_t \tag{28}$$

where

$$\rho_t = 1 + \tilde{G}^T\bar{\Gamma}_t(\bar{G}^T\bar{\Gamma}_t)^{-1} \tag{29}$$

Substituting Eq.(18) into Eq.(28) results in

$$
\begin{aligned}
|\Delta s_{t+d}| &\leq \sum_{k=1}^{N_0} |(\tilde{f}_k + a_k\rho_t)\phi_k(\mathbf{x}_t)| + \sum_{k=1}^{N_1} |(\tilde{g}_k + b_k\rho_t)\gamma_k(\mathbf{x}_t)\psi_t| \\
&\leq (d_0\kappa_t + \mu)\left(\sum_{k=1}^{N_0} |\phi_k(\mathbf{x}_t)| + \sum_{k=1}^{N_1} |\gamma_k(\mathbf{x}_t)\psi_t|\right) \qquad (30)
\end{aligned}
$$

Using Eqs.(20) and (21) yields

$$
\begin{aligned}
2s_t\Delta s_{t+d} &= 2s_t\left(\tilde{F}^T\bar{\Phi}_t + (\tilde{G}^T\bar{\Gamma}_t)^{-1})\psi_t + \rho_t v_t\right) \\
&= 2s_t\left(\sum_{k=1}^{N_0}(\tilde{f}_k + a_k\rho_t)\phi_k(\mathbf{x}_t) + \sum_{k=1}^{N_1}(\tilde{g}_k + b_k\rho_t)\gamma_k(\mathbf{x}_t)\psi_t\right) \\
&\leq -2(d_0 + 1 - \kappa_t - \mu)\left(\sum_{k=1}^{N_0}|\phi_k(\mathbf{x}_t)s_t| + \sum_{k=1}^{N_1}|\gamma_k(\mathbf{x}_t)\psi_t s_t|\right) \\
&\leq -\tau\left((d_0\kappa_t + \mu)\left(\sum_{k=1}^{N_0}|\phi_k(\mathbf{x}_t)| + \sum_{k=1}^{N_1}|\gamma_k(\mathbf{x}_t)\psi_t|\right)\right)^2 \\
&\leq -\tau(\Delta s_{t+d})^2 \qquad (31)
\end{aligned}
$$

As a result, the following relation is derived

$$\Delta V_{t+d} < -(\tau - 1)(\Delta s_{t+d})^2 \qquad (32)$$

The above relation implies that Δs_{t+d} converges to zero as t approaches infinity. This shows that s_{t+d} is brought into the inside of the set $\Xi(\sigma_t)$.

4 Generalized Variable Structure Neural Control

In the previous section, the performance function J_s only involves the difference between the reference and the optimal prediction. For many practical systems, the control input of the system should be taken into account in the performance function. Thus, the objective of the control in this section is to minimize the following performance function, which includes the control input.

$$J_g = \frac{1}{2}(\hat{y}^*_{t+d} - r_{t+d})^2 + \frac{1}{2}\alpha(\Delta u_t)^2 \qquad (33)$$

where $\Delta u_t = u_t - u_{t-1}$, α is a positive number.

Using the neural network based d-step-ahead affine nonlinear predictor, the control input to minimize J_g is given by

$$u_t = \left(((\bar{G}^*)^T\bar{\Gamma}_t)^2 + \alpha\right)^{-1}\left((\bar{G}^*)^T\bar{\Gamma}_t(r_{t+d} - (\bar{F}^*)^T\bar{\Phi}_t) + \alpha u_{t-1}\right) \qquad (34)$$

To avoid the difficulty of finding the optimal weight vectors $\bar{F}^*$ and $\bar{G}^*$ in the affine nonlinear predictor, this section considers the use of a predictive neural controller and variable structure control. Similar to the previous section, the following control input is used:

$$u_t = ((\bar{G}^T\bar{\Gamma}_t)^2 + \alpha)^{-1}\left(\bar{G}^T\bar{\Gamma}_t(r_{t+d} - \bar{F}^T\bar{\Phi}_t + s_t + v_t) + \alpha u_{t-1}\right) \tag{35}$$

The auxiliary control input v_t is chosen as

$$v_t = (\bar{G}^T\bar{\Gamma}_t)^{-2}\left(\sum_{k=1}^{N_0} a_k\eta_t^{-1}\phi_k + \sum_{k=1}^{N_1} b_k\gamma_k\psi_t + c_1\omega_t\right) \tag{36}$$

where

$$\psi_t = \bar{G}^T\bar{\Gamma}_t(r_{t+d} + s_t - \bar{F}^T\bar{\Phi}_t) + \alpha u_{t-1} \tag{37}$$

$$\omega_t = \alpha\mu^{-1}(r_{t+d} + s_t - \bar{F}^T\bar{\Phi}_t - \alpha\bar{G}^T\bar{\Gamma}_t u_{t-1}) \tag{38}$$

$$\eta_t = ((\bar{G}^T\bar{\Gamma}_t)^2 + \alpha)^{-1} \tag{39}$$

For this case, the following theorem gives the design of the auxiliary control input v_t so that s_{t+d} converges from the outside to the inside of the set Ξ.

Theorem 2: If the the coefficients a_k and b_k of the v_t are given by

$$a_k = \begin{cases} -d_0\mathrm{sign}(\phi_k s_t) & s_t \notin \Xi(\tau_t) \\ 0 & \text{otherwise} \end{cases} \tag{40}$$

$$b_k = \begin{cases} -d_0\mathrm{sign}(\gamma_k\psi_t s_t) & s_t \notin \Xi(\tau_t) \\ 0 & \text{otherwise} \end{cases} \tag{41}$$

$$c_1 = \begin{cases} -d_0\mathrm{sign}(\omega_t s_t) & s_t \notin \Xi(\tau_t) \\ 0 & \text{otherwise} \end{cases} \tag{42}$$

$d_0 > \kappa_t + \mu - 1 > 0$ and $\tau > 1$, where

$$\tau_t = \frac{\tau(d_0\kappa_t + \mu)^2}{2(d_0 + 1 - \kappa_t - \mu)}\left(\sum_{k=1}^{N_0}|\phi_k| + \sum_{k=1}^{N_1}|\gamma_k\eta_t\psi_t| + |\eta_t\omega_t|\right) \tag{43}$$

$$\kappa_t = \sum_{k=1}^{N_1}|\gamma_k(\mathbf{x}_t)(\bar{G}^T\bar{\Gamma}_t)^{-1}|\mu + 1 \tag{44}$$

$$\mu = \max_{i=1,2,\ldots,N_0, j=1,2,\ldots,N_1}\left\{|f_i^* - f_i|, |g_j^* - g_j|\right\} \tag{45}$$

ζ and d_0 are positive numbers, then s_{t+d} converges to the set $\Xi(\sigma_t)$.

Proof: Choose the Lyapunov function as

$$V_{t+d} = s_{t+d}^2. \tag{46}$$

It is known from the proof of Theorem 1 that the difference of the above Lyapunov function can be given by Eq.(26). With Eqs.(13) and (35), s_{t+d} can be expressed by

$$\begin{aligned}
s_{t+d} &= F^*(\mathbf{x}_t) + G^*(\mathbf{x}_t)u_t - r_{t+d} \\
&= (\tilde{F}+\bar{F})^T\bar{\Phi}_t + (\tilde{G}+\bar{G})^T\bar{\Gamma}_t\eta_t\left(\bar{G}^T\bar{\Gamma}_t(r_{t+d} - \bar{F}^T\bar{\Phi}_t + s_t + v_t) + \alpha u_{t-1}\right) \\
&= \tilde{F}^T\bar{\Phi}_t + \tilde{G}^T\bar{\Gamma}_t\bar{G}^T\bar{\Gamma}_t\eta_t(r_{t+d} - \bar{F}^T\bar{\Phi}_t + s_t + \alpha(\bar{G}^T\bar{\Gamma}_t)^{-1}u_{t-1}) \\
&\quad -\alpha\eta_t(r_{t+d} - \bar{F}^T\bar{\Phi}_t + s_t - \bar{G}^T\bar{\Gamma}_t u_{t-1}) + (\tilde{G}^T\bar{\Gamma}_t\bar{G}^T\bar{\Gamma}_t + (\bar{G}^T\bar{\Gamma}_t)^2)\eta_t v_t + s_t
\end{aligned} \tag{47}$$

Moving the term s_t from the right side to the left side in the equation above gives

$$\Delta s_{t+d} = \tilde{F}^T\bar{\Phi}_t + \tilde{G}^T\bar{\Gamma}_t\psi_t\eta_t - \mu\eta_t\omega_t + \eta_t(\bar{G}^T\bar{\Gamma}_t)^2\zeta_t v_t \tag{48}$$

where

$$\zeta_t = \tilde{G}^T\bar{\Gamma}_t(\bar{G}^T\bar{\Gamma}_t)^{-1} + 1 \tag{49}$$

The upper bound of the absolute value of Δs_{t+d} is estimated by

$$\begin{aligned}
|\Delta s_{t+d}| &\leq \sum_{k=1}^{N_0}|(\tilde{f}_k + a_k\zeta_t)\phi_k(\mathbf{x}_t)| + \sum_{k=1}^{N_1}|(\tilde{g}_k + b_k\zeta_t)\eta_t\gamma_k(\mathbf{x}_t)\psi_t| + |(c_k\zeta_t - \mu)\eta_t\omega_t| \\
&\leq (\mu + d_0\kappa_t)\left(\sum_{k=1}^{N_0}|\phi_k(\mathbf{x}_t)| + \sum_{k=1}^{N_1}|\gamma_k(\mathbf{x}_t)\eta_t\psi_t| + |\eta_t\omega_t|\right)
\end{aligned} \tag{50}$$

Using Eqs (40)-(42) leads to

$$\begin{aligned}
2s_t\Delta s_{t+d} &= 2s_t\left(\tilde{F}^T\bar{\Phi}_t + \tilde{G}^T\bar{\Gamma}_t\psi_t\eta_t - \mu\eta_t\omega_t + \eta_t\zeta_t v_t\right) \\
&= 2s_t\left(\sum_{k=1}^{N_0}(\tilde{f}_k + a_k\zeta_t)\phi_k(\mathbf{x}_t) + \sum_{k=1}^{N_1}(\tilde{g}_k + b_k\zeta_t)\eta_t\gamma_k(\mathbf{x}_t)\psi_t + (c_k\zeta - \mu)\eta_t\omega_t\right) \\
&\leq -2(d_0 + 1 - \kappa_t - \mu)\left(\sum_{k=1}^{N_0}|\phi_k(\mathbf{x}_t)s_t| + \sum_{k=1}^{N_1}|\gamma_k(\mathbf{x}_t)\eta_t\psi_t s_t| + |\eta_t\omega_t s_t|\right) \\
&\leq -\tau\left((d_0\kappa_t + \mu)\left(\sum_{k=1}^{N_0}|\phi_k(\mathbf{x}_t)| + \sum_{k=1}^{N_1}|\gamma_k(\mathbf{x}_t)\eta_t\psi_t| + |\eta_t\omega_t|\right)\right)^2 \\
&\leq -\tau(\Delta s_{t+d})^2
\end{aligned} \tag{51}$$

Thus, the following relation is derived

$$\Delta V_{t+d} < -(\tau - 1)(\Delta s_{t+d})^2 \tag{52}$$

The above relation shows that s_{t+d} converges to the inside of the set $\Xi(\sigma_t)$ as t approaches infinity. Thus, this proves the theorem.

5 Recursive Weight Learning of Neural Networks

In practice uncertainties and/or disturbances will always exist and recursive weight learning of the neural networks used to construct the d-step ahead affine nonlinear predictor becomes necessary. Here, we consider the recursive adjustment algorithm of the weights of the d-step ahead predictor. The algorithm can be used for both on-line and off-line weight training.

Using the available output data $y_{t-d-1}, ..., y_{t-d-n}$ and the input data u_{t-d}, the d-step ahead predictor is given by

$$y_t = (\bar{F}^*)^T \bar{\Phi}_{t-d} + (\bar{G}^*)^T \bar{\Gamma}_{t-d} u_{t-d} + \varepsilon_t \tag{53}$$

where $\bar{F}^*$ and $\bar{G}^*$ are the optimal estimates of the weights, ε_t is the approximation error of the predictor which is assumed to be bounded, i.e., $\max |\varepsilon_t| \leq \delta_L$, but the upper bound δ_L is not known exactly.

The estimated d-step ahead predictor can be compactly written as

$$\hat{y}_t = W_{t-1}^T \Phi_{t-1} \tag{54}$$

where the weight vector W_{t-1} and the basis function vector Φ_{t-1} are

$$W_{t-1} = [\, f_1 \quad f_2 \quad \ldots \quad f_{N_0} \quad g_1 \quad g_2 \quad \ldots \quad g_{N_1} \,]^T \tag{55}$$

$$\Phi_{t-1} = [\, \phi_1(\mathbf{x}_{t-d}) \quad \ldots \quad \phi_{N_0}(\mathbf{x}_{t-d}) \quad \gamma_1(\mathbf{x}_{t-d})u_{t-d} \quad \ldots \quad \gamma_{N_1}(\mathbf{x}_{t-d})u_{t-d} \,]^T \tag{56}$$

Based on the recursive least squares algorithm for a bounded noise [22] [23], the recursive weight learning algorithm for the neural network is proposed as

$$W_t = W_{t-1} + \lambda_t P_t \Phi_{t-1} e_t \tag{57}$$

$$P_t = P_{t-1} - \lambda_t \gamma_t P_{t-1} \Phi_{t-1} \Phi_{t-1}^T P_{t-1} \tag{58}$$

$$\lambda_t = \beta_t (\delta_e \Phi_{t-1}^T P_{t-1} \Phi_{t-1})^{-1} (|e_t| - \delta_e) \tag{59}$$

$$\gamma_t = \delta_e |e_t|^{-1} \tag{60}$$

$$e_t = y_t - \hat{y}_t \tag{61}$$

$$\beta_t = \begin{cases} 0 & |e_t| \leq \delta_e \\ 1 & |e_t| > \delta_e \end{cases} \tag{62}$$

where the positive number δ_e is assumed not to be less than the upper bound δ_L of the approximation, $P(0)$ is a positive finite matrix and $\lambda_{max}(.)$ is the maximum eigenvalue of its argument matrix.

Consider the Lyapunov function

$$V_t = \tilde{W}_t^T P_t^{-1} \tilde{W}_t \tag{63}$$

where $\tilde{W}_t = W_t - W_t^*$, W_t^* is the optimal estimate of the weight vector W_t. Combining Eqs. (57)-(61), the Lyapunov function can be expressed by

$$\begin{aligned} V_t &= (\tilde{W}_{t-1} + \lambda_t P_t \Phi_{t-1} e_t)^T P_t^{-1} (\tilde{W}_{t-1} + \lambda_t P_t \Phi_{t-1} e_t) & (64) \\ &= \tilde{W}_{t-1}^T P_{t-1}^{-1} \tilde{W}_{t-t} + \lambda_t \varepsilon_t^2 - \lambda_t \gamma_t e_t^2 & (65) \end{aligned}$$

Since $\max|\varepsilon_t| \leq \delta_L$,

$$V_t \leq \tilde{W}_{t-1}^T P_{t-1}^{-1} \tilde{W}_{t-t} + \lambda_t \delta_L^2 - \lambda_t \gamma_t e_t^2 \tag{66}$$

Using Eqs.(59)-(60) gives

$$V_t \leq V_{t-1} - \beta_t (\Phi_{t-1}^T P_{t-1} \Phi_{t-1})^{-1} (|e_t| - \delta_e)(|e_t| - \delta_e^{-1} \delta_L^2) \tag{67}$$

Thus,

$$\Delta V_t \leq V_t - V_{t-1} \leq -f(\delta_e) \tag{68}$$

where

$$f(\delta_e) = \beta_t (\Phi_{t-1}^T P_{t-1} \Phi_{t-1})^{-1} (|e_t| - \delta_e)(|e_t| - \delta_e^{-1} \delta_L^2) \tag{69}$$

Since the bound δ_L of the estimation error $\varepsilon_L(t)$ is assumed not to be greater than δ_e, it is easy to show that $f(\delta_e) > 0$ until the error $|e_t| = \delta_e$. So, the error $|e_t|$ converges to δ_e. On the other hand, if $|e_t| < \delta_e$, it is possible that $\Delta V_t > 0$. This implies that the weight vector W_t may drift away over time. In this case, set the $\beta_t = 0$ in the weight learning algorithm given by Eqs.(57)-(62) to avoid divergence of the weight vector. Thus the error $|e_t|$ always converges to the range $[0, \delta_e]$.

The analysis above shows that if the upper bound δ_L is known, then the error $|e_t|$ will converge to δ_L by simply setting $\delta_e = \delta_L$. In the case where the upper bound δ_L of the estimation error $\varepsilon_L(t)$ is not known exactly, the error $|e_t|$ still converges to δ_e if δ_e is set to be greater than δ_L. Thus, the closer the number δ_e is chosen to the upper bound δ_L, the more accurate the estimation of the predictor is.

6 An Example

In this section, consider the following affine nonlinear system [2]:

$$y_t = \frac{2.5 y_{t-1} y_{t-2}}{1 + y_{t-1}^2 + y_{t-2}^2} + 0.3cos(0.5(y_{t-1} + y_{t-2})) + 1.2 u_{t-1} \tag{70}$$

The initial condition of the plant is $(y_{-1}, y_{-2}) = (0, 0)$ and the reference input

$$r(t) = \begin{cases} 6cos(\pi t/80) & 0 < t \leq 160 \\ 0 & t > 160 \end{cases} \tag{71}$$

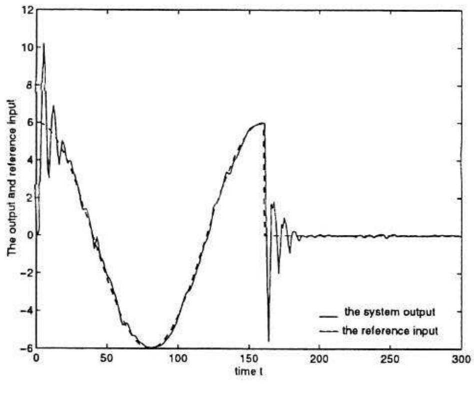

Figure 1: The output y_t and the reference input r_t of the system without using variable structure control.

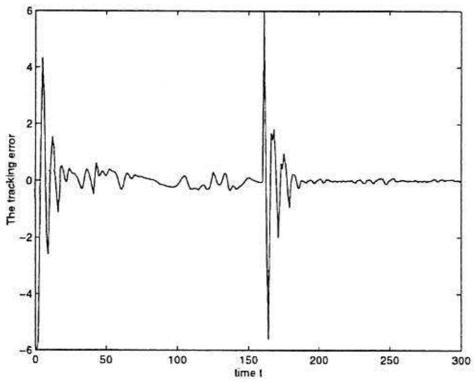

Figure 2: The tracking error $y_t - r_t$ of the system without using variable structure control.

Since the structure and parameters of the functions $F(\mathbf{x}_t)$ and $G(\mathbf{x}_t)$ in the affine nonlinear system are assumed to be unknown. Growing Gaussian radial basis function (GRBF) neural network was used to approximate the functions. The growing GRBF network was initialised with no basis function units. As observations are received the network grows by adding new units. The decision to add a new unit depends on the observation novelty for which two conditions must be satisfied. The first condition states that the approximation error between the real output and the estimated output must be significant. The second condition states that the new centre of the GRBF must be far away from existing centres. The more details about the growing GRBF neural network are given in [16]. In this way, the the approximation accuracy of the functions $F(\mathbf{x}_t)$ and $G(\mathbf{x}_t)$ will converge to the required bound.

In the simulation, the recursive weight algorithm was used for off-line training of the growing GRBF network. When the variable structure neural control was applied, the recursive weight algorithm was then used for on-line training of the growing GRBF network. The generalized variable structure neural control strategy was used. The parameters were $\alpha = 0.5$, $\tau = 1.1$, $\mu = 0.1$, $d_0 = 0.5$. The performance of the system is shown in Figs.1 and 2 without the neural network based variable structure control and in Figs. 3 and 4 with neural network based variable structure control. Figs.1 and 3 show the output y_t and the reference input r_t of the system. The tracking error $r_t - y_t$ is shown in Figs.2 and 4.

The results of the simulation shows that the tracking error of the system using variable structure control is smaller and converges faster than one of the system without variable structure control. Thus, it is clear that the difference with the variable structure control is significant.

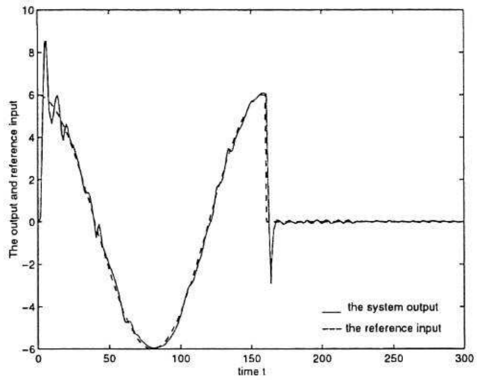

Figure 3: The output y_t and the reference input r_t of the system using variable structure control.

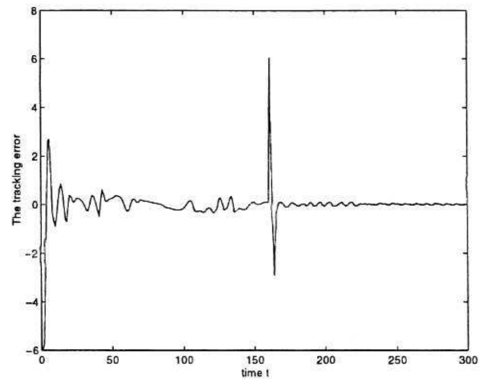

Figure 4: The tracking error $y_t - r_t$ of the system using variable structure control.

7 Conclusions

This paper has presented a novel neural network based variable structure controller design for unknown nonlinear discrete systems. A neural network based affine nonlinear predictor was introduced to predict the outputs of the nonlinear process, and a variable structure control algorithm was developed which is simple and easy to implement. In order to improve the stability and robustness performance of the system, discrete sliding mode control technique was applied. Two cases were considered for the variable structure neural control. The first was based on the minimization of the square prediction error. The second was based on combined minimization of both the squared prediction error and the squared control input. A recursive weight learning algorithm for the affine nonlinear predictors was also developed which can be used for both on-line and off-line weight training. The analysis of the weight learning algorithm demonstrated that both the weights of the neural networks and the estimation errors converge.

8 Acknowledgement

The authors are grateful for the support of EPSRC under the grant GR/J46661.

References

[1] S. A. Billings and S. Chen, "Neural Networks and System Identification", In K. Warwick *et al.*, (eds.) *Neural networks for systems and control*, pp.181-205, 1992.

[2] F.C. Chen and H.K. Khalil, "Adaptive control of a class of nonlinear discrete-time systems using neural networks," *IEEE Trans. on Automatic Control*, vol. 40, no. 5, pp. 791-801, 1995.

[3] S. Chen, S. A. Billings and P. M. Grant, "Nonlinear system identification using neural networks", *International Journal of Control*, Vol.51, No.6, pp.1191-1214, 1990.

[4] S.V. Emelyanov, Variable Structure Control Systems, Moscow: Nauka (in Russian), 1967.

[5] S. Fabri, and V. Kadirkamanathan, "Dynamic structure neural networks for stable adaptive control of nonlinear systems", to appear in the *IEEE Trans. on Neural Networks*, 1996.

[6] K. Furuta, "Sliding mode control of a discrete system," Systems and Control Letter, vol. 14, pp. 145-152, 1990.

[7] K. Furuta, "VSS type self-tuning control," IEEE Trans. on Industrial Electronics, vol. 40, no. 1, pp. 37-44, 1993.

[8] J.Y. Hung, W. Gao and J.C. Hung, "Variable structure control: A survey," IEEE Trans. on Industrial Electronics, vol. 40, no. 1, pp.2-22, 1993.

[9] Y. Itkis, Control Systems of Variable Structure, New York: Wiley, 1976.

[10] V. Kadirkamanathan, *Sequential Learning in Artificial Neural Networks, Ph.D Thesis*, University of Cambridge, UK, 1991.

[11] V. Kadirkamanathan and M. Niranjan. "Application of an architecturally dynamic network for speech pattern classification", In *Proceedings of the Institute of Acoustics*, Vol.14, Part 6, pp.343-350, 1992.

[12] V. Kadirkamanathan and M. Niranjan, "A function estimation approach to sequential learning with neural networks", *Neural Computation*, Vol.5, pp.954-957, 1993.

[13] G.P. Liu and V. Kadirkamanathan, "Multiobjective criteria for nonlinear model selection and identification with neural networks," to appear in the *IEEE Transactions on Neural Networks.*

[14] G.P. Liu, V. Kadirkamanathan and S.A. Billings, "Sequential identification of nonlinear systems by neural networks," *Proceedings of the 3rd European Control Conference*, Rome, pp. 2408-2413, 1995.

[15] G.P. Liu, V. Kadirkamanathan and S.A. Billings, "Variable neural networks for nonlinear adaptive control," to be presented in the *13th IFAC Congress*, USA, 1996.

[16] G.P. Liu, V. Kadirkamanathan and S.A. Billings, "Stable sequential identification of continuous nonlinear dynamical systems by growing RBF networks," to appear in the *International Journal of Control.*

[17] K. S. Narendra and K. Parthasarathy. "Identification and control of dynamical systems using neural networks", *IEEE Trans. on Neural Networks*, Vol.1, No.1, pp.4-27, 1990.

[18] Y. Pan and K. Furuta, "VSS approach to the design of robust digital controller using transfer equation," European Journal of Control, vol. 1, no. 2, pp.166-173, 1995.

[19] R. M. Sanner and J. J. E. Slotine, "Gaussian networks for direct adaptive control", *IEEE Trans. on Neural Networks*, Vol.3, No.6, pp.837-863, 1992.

[20] J. J. E. Slotine and W. Li, *Applied Nonlinear Control.* Prentice-Hall International, Englewood Cliffs, New Jersey 1991.

[21] V.I. Utkin, "Variable structure system with sliding mode: A survey," *IEEE Trans. on Automatic Control*, vol. 22, no. 2, pp. 212-222, 1964.

[22] H. Wang, G.P. Liu, C.J. Harris and M. Brown, Advanced Adaptive Control, *Pergamon Press Ltd*, Oxford, 1995.

[23] J.F. Whidborne and G.P. Liu, Critical Control Systems: Theory, Design and Applications, *Research Studies Press*, UK., 1993.