# Feedback-Linearization-Based Neural Adaptive Control for Unknown Nonaffine Nonlinear Discrete-Time Systems

Hua Deng, Han-Xiong Li, *Senior Member, IEEE*, and Yi-Hu Wu

***Abstract*—A new feedback-linearization-based neural network (NN) adaptive control is proposed for unknown nonaffine nonlinear discrete-time systems. An equivalent model in affine-like form is first derived for the original nonaffine discrete-time systems as feedback linearization methods cannot be implemented for such systems. Then, feedback linearization adaptive control is implemented based on the affine-like equivalent model identified with neural networks. Pretraining is not required and the weights of the neural networks used in adaptive control are directly updated online based on the input–output measurement. The dead-zone technique is used to remove the requirement of persistence excitation during the adaptation. With the proposed neural network adaptive control, stability and performance of the closed-loop system are rigorously established. Illustrated examples are provided to validate the theoretical findings.**

***Index Terms*—Feedback linearization, neural networks, nonaffine nonlinear discrete-time systems, nonlinear adaptive control.**

## I. Introduction

NEURAL NETWORK (NN) adaptive control based on feedback linearization for unknown discrete-time nonlinear systems has been an active research area in the last 15 years. Most of the previous methods are elegant but limited to feedback linearizable (e.g., affine) nonlinear discrete-time systems [1]–[11]. An important feature for affine nonlinear discrete-time systems is that the output of the systems is linear with respect to the control signal. Therefore, feedback linearization for such systems can be used and their controller design is straightforward. For an unknown nonaffine nonlinear discrete-time system, however, its output depends nonlinearly on its input, and thus, feedback linearization methods cannot be implemented. Therefore, it is no longer a simple task to determine the control input of such a nonaffine nonlinear system, which results in a desired output even assuming that its model is available [12], [13].

Recently, NN adaptive control for unknown nonaffine nonlinear discrete-time systems has received considerable attention for its academic challenge and its practical interest. Noriega and Wang [14] proposed an adaptive NN control for such nonlinear systems based on online identification with backpropagation. No analytical results are given for the stability and error bound of the closed-loop system. Ahmed [15] also considered NN adaptive control but the operating point of an unknown discrete-time system must be measurable or inferred. Moreover, the measurable operating point must be independent of the system input and output [16]. In [17], NN adaptive inverse control was studied in which dynamic gradient methods [18], [19] were used to adjust the NN weights. During the adaptation, special attention must be given to the stability and convergence because they are not ensured [13], [17], [20]. To avoid instability, small learning steps are often used in gradient methods but will slow the training. In [21], a multilayer NN is employed in adaptive control with a discrete-time projection algorithm. In the control, the closed-loop system stability and tracking error are dependent on the NN node number and the learning rate of the NN weights. The dependence may increase the design difficulty of NN adaptive controllers. In [22], the linear part of an unknown nonlinear discrete-time system is first identified around a specified operating point and an NN is then used to compensate the nonlinear part. However, if the operating point is changed, the control algorithm must be reformulated. Adetona *et al.* [23] proposed a robust adaptive NN control by assuming a special affine model for the original nonaffine discrete-time system. As the special model is only valid under the requirement that the control increment is sufficiently small, the applications of the robust adaptive NN control are limited only to sufficiently slow trajectory tracking [24].

It is well known that nonaffine nonlinear plant models represent a much broader group of (both real-world and academic) systems than affine models. Actually affine nonlinear models are only a small subset of nonaffine nonlinear models. Therefore, results developed for nonaffine systems can often be directly applicable to affine systems. However, because of the complexity of nonaffine nonlinear discrete-time systems, their NN adaptive control based on feedback linearization is less well addressed. In this paper, a *new feedback-linearization-based NN adaptive*

Manuscript received January 24, 2007; revised October, 31, 2007; accepted February 4, 2008. First published July 15, 2008; current version published September 4, 2008. This work was supported in part by the National Basic Research Program of China under Grant 2006CB705400, in part by the National Natural Science Foundation of China under Grants 50276005 and 50775224, and in part by the City University of Hong Kong under SRG Project 7002114.

H. Deng is with the Key Laboratory of Modern Complex Equipment Design and Extreme Manufacturing, Ministry of Education, Changsha 410083, China and with the School of Mechanical and Electrical Engineering, Central South University, Changsha 410083, China (e-mail: hdeng@csu.edu.cn).

H.-X. Li is with the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, Hong Kong, and with the School of Mechanical and Electrical Engineering, Central South University, Changsha 410083, China (e-mail: mehxli@cityu.edu.hk).

Y.-H. Wu is with the School of Automobile and Mechanical Engineering, Changsha University of Science and Technology, Changsha 410076, China (e-mail: w2309@163.com).

Digital Object Identifier 10.1109/TNN.2008.2000804

*control* is proposed for unknown *nonaffine* nonlinear discrete-time systems. The main contributions of this paper are as follows. 1) An equivalent model in affine-like form is derived for the original nonaffine discrete-time systems. 2) Feedback-linearization-based NN adaptive control can then be determined based on the affine-like equivalent model. The weights of the NNs used in adaptive control are directly updated online based on the input–output measurement. 3) The tracking error under the proposed NN adaptive control is proportional to NN approximation error, which can be made arbitrarily small by choosing enough neurons. 4) The controller singularity problem [4] is avoided completely by using the proposed NN adaptive control law. Moreover, prelearning is not required and the dead-zone technique is used to remove the requirement of persistence excitation (PE) [25] during adaptation, as the requirement is difficult to satisfy in practice [7], [21]. With the proposed NN adaptive control and the NN weight updating law, stability and performance of the closed-loop system are rigorously established. Illustrated examples are provided to validate the theoretical findings.

## II. System Dynamics and Preliminaries

### A. Dynamics of Nonaffine Discrete-Time Systems

Consider nonaffine nonlinear discrete-time systems as follows [26]:

$$y(k+d) = f_0\left[w_k, u(k)\right] \tag{1}$$

where $w_k = [y(k), y(k-1), \ldots, y(k-n+1), u(k-1), \ldots, u(k-n+1)]$, $f_0 \in C^3$ is an unknown nonlinear function, $d$ is the relative degree [27] for system (1), $n$ is the known system order, $y(k)$ is the measured system output, and $u(k)$ is the measured system input.

It has been shown by Narendra and Mukhopadhyay [12] that, under certain conditions, the input–output representation (1) can be derived from its widely used equivalent state-space equations as follows:

$$x(k+1) = f_x\left[x(k), u(k)\right] \tag{2a}$$
$$y(k) = h\left[x(k)\right] \tag{2b}$$

where $x(k) \in R^n$ and $f_x, h \in C^\infty$. The relative degree $d$ defined in (1) represents the delay of the system (2) from input $u(k)$ to output $y(k)$. The input–output representation (1) is what is needed for later analysis and assumed to satisfy the following assumptions.

*Assumption 1:* $\partial f_0[w_k, u(k)]/\partial u(k) = \beta_k$ with $0 < |\beta_k| \le \overline{\beta}$ when $\overline{y}_k \in S_y$ and $\overline{u}_k \in S_u$, where $\overline{\beta}$ is a finite positive number, $\overline{y}_k = [y(k), y(k-1), \ldots, y(k-n+1)]$, $\overline{u}_k = [u(k), u(k-1), \ldots, u(k-n+1)]$, and $S_y \subset R$ and $S_u \subset R$ are two compact sets. It is without losing generality that positive $\beta_k$ is assumed in the following discussion.

*Assumption 2:* System (1) is invertibly stable [15] (i.e., bounded system output implies bounded system input).

Assumption 1 is a common assumption made in NN adaptive control [15], [21]. Assumption 2 means that there exists a unique and continuous (smooth) function

$$u(k) = G\left[w_k, y(k+d)\right] \tag{3a}$$

such that

$$y(k+d) = f_0\left[w_k, u(k)\right] = f_0\left[w_k, G\left[w_k, y(k+d)\right]\right]. \tag{3b}$$

Equations (3a) and (3b) imply that $G$ is the inverse mapping of $f_0$ and thus $G$ is a one-to-one mapping of $S_y$ into $S_u$ [28]. Define a reference trajectory $r(k)$ and a compact set $S_r$ such that $r(k) \in S_r \subset S_y$. The result in the following lemma is given by Ge [21] and Cabrera and Narendra [27].

*Lemma 1:* For a given reference trajectory $r(k)$, under Assumption 1, the implicit function theorem assures the existence of a unique and continuous (smooth) function

$$u^*(k) = G\left[w_k, r(k+d)\right] \tag{4a}$$

such that

$$r(k+d) = f_0\left[w_k, G\left[w_k, r(k+d)\right]\right]. \tag{4b}$$

In (4a), $u^*(k)$ can be considered as an unknown ideal control signal for system (1) to track a given reference trajectory $r(k)$. Suppose that $u^*(k) \in S_{ud}$. Equations (4a) and (4b) also imply that $G$ is a one-to-one mapping of $S_r$ into $S_{ud}$. Because $S_r \subset S_y$ and $G$ is the one-to-one mapping of $S_y$ into $S_u$, $S_{ud}$ must be a subset of $S_u$, that is, $S_{ud} \subset S_u$. As $G$ is continuous and $S_r$ is compact, then $S_{ud}$ must be compact [28]. The following properties will be used in later analysis.

*Property 1:* There exist a $\xi_{0k}$ between $r(k+d)$ and $y(k+d)$ and a $\varsigma_k$ between $u^*(k)$ and $u(k)$ that satisfy

$$f_{0u}[w_k, \varsigma_k] G_y[w_k, \xi_{0k}] \equiv 1$$

where

$$f_{0u}[w_k, \varsigma_k] = \left.\frac{\partial f_0\left[w_k, u(k)\right]}{\partial u(k)}\right|_{u(k)=\varsigma_k}$$
$$G_y[w_k, \xi_{0k}] = \left.\frac{\partial G\left[w_k, y(k+d)\right]}{\partial y(k+d)}\right|_{y(k+d)=\xi_{0k}}$$

and

$$\begin{aligned} G\left[w_k, y(k+d)\right] =& G\left[w_k, r(k+d)\right] \\ &+ G_y[w_k, \xi_{0k}]\left(y(k+d) - r(k+d)\right). \end{aligned}$$

*Proof:* See part A of the Appendix.

*Property 2:* Letting $\rho_k = (\partial f_0[w_k, u(k)]/\partial u(k))|_{u(k)=\varsigma_k}$, then $0 < \rho_k \le \overline{\beta}$ in terms of Assumption 1.

### B. Neural Networks in Control

A nonlinear controller is usually required for the control of a nonlinear system. However, the lack of a general structure makes the design of such a nonlinear controller quite difficult.

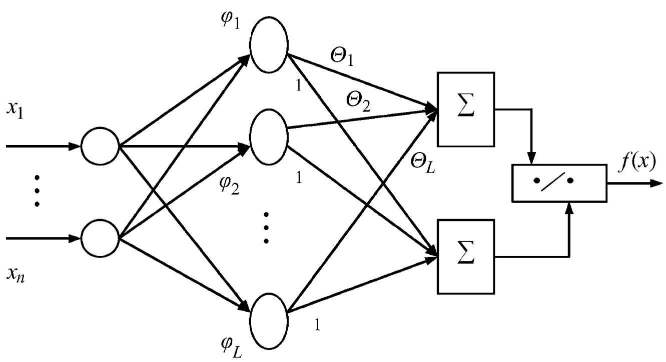


Fig. 1. Structure of an NRBFN.

Nonlinear controller design may be viewed as a nonlinear function approximation problem [15]. In recent years, there has been no doubt that NNs are quite successful in describing nonlinear functions and different NNs have been used to represent the controller nonlinearity [26], [29]–[32]. The most widely used NNs are multilayer perceptron (MLP) NNs and radial basis function (RBF) NNs because of their general approximation abilities [33]–[35]. Despite their design complexity, the function approximation properties of such NNs with compositions and superposition of simple nonlinearities equally hold provided a sufficient number of nodes are employed [15]. A nonlinear function $f(x)$ may be approximated by such one-hidden-layer networks with a linear superposition of nonlinear nodes as follows:

$$f(x) = \Theta\Phi(x) = \sum_{i=1}^{L} \theta_i \varphi_i(x) \tag{5}$$

where $x = [x_1 \ \cdots \ x_n]^T$ is the NN input vector, $\Theta = [\theta_1 \ \cdots \ \theta_L]$ is the weight vector of the NN, $L$ denotes the NN node number, and $\Phi(x) = [\varphi_1(x) \ \cdots \ \varphi_L(x)]^T$ with $\varphi_i(x)$ being a basis function corresponding to node $i$. These basis functions could be RBFs, B-splines, and so on.

Because it is a difficult and unsolved problem to find the best NN (approximator) structure, a fairly large NN is usually employed to deal with relatively complex approximation problems [13]. In this study, the normalized radial basis function network (NRBFN) [36], [37] is used in the adaptive controller as shown in Fig. 1. Evolved from RBFNNs, the NRBFN has the same structure as (5) with $\varphi_i(x)$ being a normalized activation function that is expressed as

$$\varphi_i(x) = \frac{\exp\left(-\|x - C_i\|^2/s^2\right)}{\sum_{j=1}^{L} \exp\left(-\|x - C_j\|^2/s^2\right)} \tag{6}$$

where $C_i = [c_{i1} \ \cdots \ c_{iL}]$ denotes the centroid vector and $s$ is the spread. The NRBFN can improve function approximation with a minimal number of weights [38]. Besides, the NRBFN possesses some useful properties such as $\sum_{i=1}^{L} \varphi_i(x) = 1$ and $0 < \|\Phi(x)\| \leq 1$. The parameter tuning of NRBFNs such as their centroids and spreads can be found in [39] and [40] and thus is not discussed here.

## III. New NN Adaptive Control

A new design method of feedback-linearization-based NN adaptive control is proposed here based on an affine-like equivalent model derived from the original nonaffine discrete-time systems. The method can avoid complex control development and overcome the shortcomings of other existing methods. Its control performance will be demonstrated by both rigorous analytical proof and simulation studies.

### A. *Affine-Like Input–Output Representation*

Taking the Taylor expansion of function $f_0[w_k, u(k)]$ with respect to $u(k)$ around $(w_k, u(k-1))$ and using the mean value theorem, (1) can be rewritten as

$$\begin{aligned} y(k+d) &= f_0\left[w_k, u(k)\right] \\ &= f_0\left[w_k, u(k-1)\right] + f_{0u}\left[w_k, u(k-1)\right]\Delta u(k) \\ &\quad + f_{0uu}[w_k, \mu_k]\Delta u^2(k) \end{aligned} \tag{7}$$

where

$$\begin{aligned} \Delta u(k) &= u(k) - u(k-1), \\ f_{0u}\left[w_k, u(k-1)\right] &= \left.\frac{\partial f_0\left[w_k, u(k)\right]}{\partial u(k)}\right|_{u(k)=u(k-1)} \\ f_{0uu}[w_k, \mu_k] &= \left.\frac{\partial^2 f_0\left[w_k, u(k)\right]}{2\partial u^2(k)}\right|_{u(k)=\mu_k} \end{aligned}$$

and

$$\mu_k = \kappa u(k) + (1-\kappa)u(k-1) \quad \text{with} \quad 0 \leq \kappa \leq 1.$$

Using Assumption 2 [e.g., using (3a)], one has

$$\begin{aligned} f_{0uu}[w_k, \mu_k]\Delta u^2(k) &= \left.\frac{\partial^2 f_0\left[w_k, u(k)\right]}{2\partial u^2(k)}\right|_{u(k)=\mu_k} \Delta u^2(k) \\ &= \overline{f}_{0uu}\left[w_k, y(k+d)\right] \\ &\quad \times \left(G\left[w_k, y(k+d)\right] - u(k-1)\right)^2 \end{aligned}$$

where

$$\overline{f}_{0uu}\left[w_k, y(k+d)\right] = \left.\frac{\partial^2 f_0\left[w_k, u(k)\right]}{2\partial u^2(k)}\right|_{u(k)=\eta_k}$$

and

$$\eta_k = \kappa G\left[w_k, y(k+d)\right] + (1-\kappa)u(k-1).$$

Substituting the previous equation into (7) yields

$$y(k+d) = f\left[w_k, y(k+d)\right] + g[w_k]\Delta u(k) \tag{8}$$

where

$$g[w_k] = f_{0u}\left[w_k, u(k-1)\right]$$

and

$$\begin{aligned} f\left[w_k, y(k+d)\right] &= f_0\left[w_k, u(k-1)\right] + \overline{f}_{0uu}\left[w_k, y(k+d)\right] \\ &\quad \times \left(G\left[w_k, y(k+d)\right] - u(k-1)\right)^2. \end{aligned}$$

Owing to the affine-like form of (8), the NN adaptive feedback linearization control developed for affine nonlinear discrete-time systems may be used for the nonaffine system (1). Before the development of NN adaptive control for (8), it is worth noting that (8) is equivalent to (1) only when certain conditions are met. The result is given in the following lemma.

*Lemma 2:* For a given $w_k$, if system (1) is invertible and the second-order derivative of $f_0$ with respect to $u(k)$ exists, then the affine-like system (8) is equivalent to the original nonlinear system (1).

*Proof:* We only need to prove that

$$f\left[w_k, y(k+d)\right] + g[w_k]\Delta u(k) = f_0\left[w_k, u(k)\right] = y(k+d)$$

if the given conditions hold. As $f_0 \in C^3$, the second-order derivative of $f_0[wk, u(k)]$ with respect to $u(k)$ exists. Therefore, from (8), one has

$$\begin{aligned} f\left[w_k, y(k+d)\right] &+ g[w_k]\Delta u(k) \\ &= f_0\left[w_k, u(k-1)\right] + \overline{f}_{0uu}\left[w_k, y(k+d)\right] \\ &\quad \times \left(G\left[w_k, y(k+d)\right] - u(k-1)\right)^2 \\ &\quad + f_{0u}\left[w_k, u(k-1)\right]\Delta u(k) \\ &= f_0\left[w_k, u(k-1)\right] + f_{0u}\left[w_k, u(k-1)\right]\left(u(k) - u(k-1)\right) \\ &\quad + \left.\frac{\partial^2 f_0\left[w_k, u(k)\right]}{2\partial u^2(k)}\right|_{u(k)=\eta_k} \left(G\left[w_k, y(k+d)\right] - u(k-1)\right)^2 \\ &= f_0\left[w_k, u(k)\right] + \left.\frac{\partial^2 f_0\left[w_k, u(k)\right]}{2\partial u^2(k)}\right|_{u(k)=\eta_k} \\ &\quad \times \left(G\left[w_k, y(k+d)\right] - u(k-1)\right)^2 \\ &\quad - \left.\frac{\partial^2 f_0\left[w_k, u(k)\right]}{2\partial u^2(k)}\right|_{u(k)=\mu_k} \left(u(k) - u(k-1)\right)^2. \end{aligned}$$

As system (1) is invertible, according to Assumption 2, one has $u(k) = G[w_k, y(k+d)]$ and $\eta_k = \kappa G[w_k, y(k+d)] + (1-\kappa)u(k-1) = \kappa u(k) + (1-\kappa)u(k-1) = \mu_k$. Using these results, the previous equation becomes

$$f\left[w_k, y(k+d)\right] + g[w_k]\Delta u(k) = f_0\left[w_k, u(k)\right] = y(k+d).$$

*Remark:* Lemma 2 states the conditions required for the derivation of (8) from (1), which include that system (1) must be invertible and the second-order derivative of $f_0$ with respect to $u(k)$ must exist. Actually, in the following development of NN adaptive feedback linearization control, it is required that $f_0 \in C^3$. Therefore, the proposed NN adaptive feedback linearization control can be applied to those systems that meet the aforementioned conditions.

### *B. NN Adaptive Control Development*

Similarly to affine nonlinear discrete-time systems, the ideal control signal for (8) can be given as follows:

$$u^*(k) = u(k-1) + \frac{r(k+d) - f\left[w_k, r(k+d)\right]}{g[w_k]}. \quad (9)$$

Note that (9) is equivalent to (4a) as (8) is equivalent to (1). Because $f[w_k, r(k+d)]$ and $g[w_k]$ are unknown, two NNs $\overline{W}_k\phi[w_k, r(k+d)]$ and $\overline{V}_k\sigma[w_k]$ that have the form discussed in Section II-B may be used to directly approximate them and the following NN adaptive control (10) with some updating laws of weights $\overline{W}_k$ and $\overline{V}_k$ may be achieved:

$$u(k) = u(k-1) + \frac{r(k+d) - \overline{W}_k\phi\left[w_k, r(k+d)\right]}{\overline{V}_k\sigma[w_k]} \quad (10)$$

The NN control structure (10) is similar to the feedback-linearization-based NN adaptive control of affine discrete-time nonlinear systems where two NNs are usually employed to approximate the two unknown nonlinear functions in such affine systems [2], [3], [8]–[10]. It is natural for us to use the previous results with such an understandable structure to implement feedback-linearization-based NN adaptive control for nonaffine nonlinear discrete-time systems. Although $g[w_k] \neq 0$, $\overline{V}\sigma[w_k]$ may be zero or sufficiently close to zero during weight training, which makes the controller (10) singular. The controller singularity problem is common in NN adaptive control of affine nonlinear discrete-time systems. This problem was solved excellently by Jagannathan and Lewis [7] for a class of affine nonlinear discrete-time systems but not avoided completely [4]. However, the following technique leads to an NN adaptive control avoiding the controller singularity problem completely.

Rewrite (9) as

$$u^*(k) = u(k-1) + \frac{1}{g[w_k]}r(k+d) - \frac{f\left[w_k, r(k+d)\right]}{g[w_k]}. \quad (11)$$

Because (11) is equivalent to (4a), (11) has the same properties as (4a), e.g., $u^*(k) \in S_{ud} \subset S_u$ when $r(k) \in S_r \subset S_y, \forall k \geq 0$ and $u^*(k)$ in (11) is continuous on $S_y$, $S_u$, and $S_r$. Therefore, $1/g[w_k]$ and $f[w_k, r(k+d)]/g[w_k]$ are two continuous nonlinear functions and two NRBFNs as described in Section II-B can be used to approximate them instead of $g[w_k]$ and $f[w_k, r(k+d)]$, which gives

$$u^*(k) = u(k-1) + V^*\sigma[w_k]r(k+d) - W^*\phi\left[w_k, r(k+d)\right] + \varepsilon(k) \quad (12)$$

where $W^*$ and $V^*$ are optimal weights of the two NNs

$$\begin{aligned} \frac{f\left[w_k, r(k+d)\right]}{g[w_k]} &= W^*\phi\left[w_k, r(k+d)\right] + \varepsilon_f(k) \\ \frac{1}{g[w_k]} &= V^*\sigma[w_k] + \varepsilon_g(k) \end{aligned}$$

$\varepsilon(k) = r(k+d)\varepsilon_g(k) - \varepsilon_f(k)$ is the NN approximation error, and $|\varepsilon(k)| \leq \overline{\varepsilon}$ with $\overline{\varepsilon}$ being a small positive number if $[y(k), y(k-1), \ldots, y(k-n+1)] \in S_y$, $[u(k-1), \ldots, u(k-n+1)] \in S_u$, and $\forall r(k) \in S_r$.

Based on (12), the NN adaptive control proposed here is given as

$$u(k) = u(k-1) + V_k\sigma[w_k]r(k+d) - W_k\phi\left[w_k, r(k+d)\right]. \quad (13)$$

### *C. Updating Laws of the NN Weights*

The control error dynamics for system (1) by using NN adaptive control (13) is needed for determining suitable updating laws of the NN weights. Define control error

$$e(k) = r(k) - y(k). \tag{14}$$

Thus

$$e(k+\mathrm{d}) = r(k+d) - y(k+d) = r(k+d) - f_0\left[w_k, u(k)\right]. \tag{15}$$

As $f_0 \in C^3$, using (8), (15) becomes

$$\begin{aligned} e(k+\mathrm{d}) &= r(k+d) - f\left[w_k, y(k+d)\right] - g[w_k]\Delta u(k) \\ &= r(k+d) - f\left[w_k, r(k+d)\right] - f_y[w_k, \xi_{1k}] \\ &\quad \times \left(y(k+d) - r(k+d)\right) - g[w_k]\Delta u(k) \end{aligned} \tag{16}$$

where $f_y[w_k, \xi_{1k}] = (\partial f[w_k, y(k+d)]/\partial y(k+d))|_{y(k+d)=\xi_{1k}}$ and $\xi_{1k}$ is between $r(k+d)$ and $y(k+d)$. Using the fact that (details are given in part B of the Appendix)

$$f_y[w_k, \xi_{1k}] = 1 - g[w_k]G_y[w_k, \xi_{0k}] \tag{17}$$

one has

$$\begin{aligned} e(k+\mathrm{d}) &= r(k+d) - f\left[w_k, r(k+d)\right] \\ &\quad - \left(1 - g[w_k]G_y[w_k, \xi_{0k}]\right)\left(y(k+d) - r(k+d)\right) \\ &\quad - g[w_k]\Delta u(k) \\ &= r(k+d) - f\left[w_k, r(k+d)\right] \\ &\quad + \left(1 - g[w_k]G_y[w_k, \xi_{0k}]\right)e(k+d) - g[w_k]\Delta u(k). \end{aligned} \tag{18}$$

Then, (18) can be rewritten as

$$\begin{aligned} G_y[w_k, \xi_{0k}]e(k+d) &= \frac{1}{g[w_k]} \\ &\quad \times \left(r(k+d) - f\left[w_k, r(k+d)\right] - g[w_k]\Delta u(k)\right). \end{aligned} \tag{19}$$

From (19) with Properties 1 and 2, one obtains

$$e(k+d) = \rho_k\left(\frac{1}{g[w_k]}r(k+d) - \frac{f\left[w_k, r(k+d)\right]}{g[w_k]} - \Delta u(k)\right). \tag{20}$$

Using (12), (20) becomes

$$\begin{aligned} e(k+d) &= \rho_k\left\{\left(V^*\sigma[w_k] + \varepsilon_g(k)\right)r(k+d)\right. \\ &\qquad \left. - \left(W^*\phi\left[w_k, r(k+d)\right] + \varepsilon_f(k)\right) - \Delta u(k)\right\} \\ &= \rho_k\left\{V^*\sigma[w_k]r(k+d)\right. \\ &\qquad \left. - W^*\phi\left[w_k, r(k+d)\right] - \Delta u(k)\right\} \\ &\quad + \rho_k\left(r(k+d)\varepsilon_g(k) - \varepsilon_f(k)\right) \\ &= \rho_k\left\{V^*\sigma[w_k]r(k+d)\right. \\ &\qquad \left. - W^*\phi\left[w_k, r(k+d)\right] - \Delta u(k)\right\} + \rho_k\varepsilon(k). \end{aligned} \tag{21}$$

Substituting the NN adaptive control (13) into (21), one has

$$\begin{aligned} e(k+d) &= \rho_k\left\{(V^* - V_k)\sigma[w_k]r(k+d)\right. \\ &\qquad \left. - (W^* - W_k)\phi\left[w_k, r(k+d)\right]\right\} + \rho_k\varepsilon(k) \\ &= \rho_k\left\{\widetilde{V}_k\sigma[w_k]r(k+d) - \widetilde{W}_k\phi\left[w_k, r(k+d)\right]\right\} \\ &\quad + \rho_k\varepsilon(k) \\ &= \rho_k[\widetilde{W}_k \widetilde{V}_k]\Omega(k) + \rho_k\varepsilon(k) \end{aligned} \tag{22}$$

where $\widetilde{V}_k = V^* - V_k$, $\widetilde{W}_k = W^* - W_k$, $\Omega(k) = \begin{bmatrix} -\phi[w_k, r(k+d)] \\ \sigma[w_k]r(k+d) \end{bmatrix}$, $0 < \|\Omega(k)\| \le \Omega_0$ with $\Omega_0$ being a positive number, and $|\rho_k\varepsilon(k)| \le \varepsilon_0$ with $\varepsilon_0$ being a small positive number.

According to the control error dynamics (22), the updating law for the weights in the NN adaptive control (13) is given as

$$[W_k \;\; V_k] = [W_{k-d} \;\; V_{k-d}] + \frac{\gamma}{\overline{\beta}\left\|\Omega(k-d)\right\|^2}\Omega^T(k-d)D(e(k)) \tag{23}$$

where $0 < \gamma \le 1$, $\overline{\beta}$ is defined in Assumption 1, and $D(\bullet)$ is the dead-zone function as follows:

$$D(e) = \begin{cases} 0, & \text{if } |e| \le \varepsilon_0 \\ e - \varepsilon_0, & \text{if } e > \varepsilon_0 \\ e + \varepsilon_0, & \text{if } e < -\varepsilon_0. \end{cases} \tag{24}$$

For the discrete-time NN weight updating law (23) with $\varepsilon(k) \neq 0$ and without a dead-zone function [e.g., $D(e(k)) = e(k)$], the PE of the regression vector $\Omega(k)$ is necessary for that the NN weights $[W_k \; V_k]$ remain bounded during adaptive control (further explanation is given in part C of the Appendix). However, the PE condition is difficult to guarantee in practice [7], [21]. To remove the PE condition during NN adaptive control for unknown nonaffine discrete-time systems, the discrete-time projection algorithm was employed in the weight updating law of a multilayer NN [21], and the dead-zone technique was employed in the NN adaptive control of a special affine model valid under sufficiently small control increment [23]. Here, the dead-zone technique is used in the proposed NN adaptive control of the affine-like nonlinear discrete-time model (8). The proof of the closed-loop system stability in the next section will show that the PE condition is not needed.

To use the dead-zone function (24), $\varepsilon_0$ should be chosen in advance. One common solution is to assume that an upper bound $\varepsilon_0$ is available [23]. Usually, $\varepsilon_0$ cannot be known *a priori*. It is well known that $\varepsilon_0$ can be arbitrarily small when enough nodes of neural networks are chosen [11]. Therefore, a small $\varepsilon_0$ (estimated *a priori*) is often set in practice when a fairly large NN is chosen to deal with relatively complex approximation problems.

### *D. Stability and Performance of the Closed-Loop System*

Because all assumptions are only valid under the conditions that $y(k) \in S_y$ and $u(k) \in S_u$, we need to prove that the system output $y(k)$ will track the reference trajectory $r(k)$, and $y(k+d)$ and $u(k)$ will remain in $S_y$ and $S_u$ by the proposed NN adaptive control. In this section, the tracking error, the NN weights, and the system input will be proven to be bounded.

Stability and performance of the closed-loop system with NN adaptive control (13) and weight updating law (23) are given in Theorem 1.

*Theorem 1:* For the nonaffine nonlinear discrete-time system (1) and NN adaptive control (13) with NN weight updating law (23), if there exist two compact sets $S_y$ and $S_u$ such that

i) at instant $k$, all past inputs $[u(k-1),\ldots,u(k-n+1)]$ are in $S_u$ and current output and all past outputs $[y(k),y(k-1),\ldots,y(k-n+1)]$ are in $S_y$;

ii) the initial future output sequence $y(k_0+d-1),y(k_0+d-2),\ldots y(k_0+1)$ are kept in the compact set $S_y$, where $k_0$ is the initial instant;

then the solution of error system (22) is uniformly ultimately bounded (UUB) for all $k>0$ with ultimate bound $\lim_{k\to\infty}|e(k)|\le 4\varepsilon_0$. Simultaneously, it will be guaranteed that $y(k+d)$ and $u(k)$ remain in the compact sets $S_y$ and $S_u$, respectively.

*Proof:* We will prove, one by one in the order of the NN weights, the tracking error and the system input that all the signals in the closed-loop system are bounded.

*1) Boundedness of the NN Weights:* From (23), one has

$$[\widetilde{W}_{k+d}\ \ \widetilde{V}_{k+d}]=[\widetilde{W}_k\ \ \widetilde{V}_k]-\frac{\gamma}{\overline{\beta}\,\|\Omega(k)\|^2}\Omega^T(k)D\left(e(k+d)\right). \tag{25}$$

When $|e(k+d)|\le\varepsilon_0$, then $D(e(k+d))=0$. From (25), $\widetilde{W}_k$ and $\widetilde{V}_k$ are bounded. Thus, we only need to consider the case where $|e(k+d)|>\varepsilon_0$. When $|e(k+d)|>\varepsilon_0$, we can derive that [11]

$$D\left(e(k+d)\right)=\eta_k\rho_k[\widetilde{W}_k\widetilde{V}_k]\Omega(k) \tag{26}$$

with $0<\eta_k<1$.

Choose a Lyapnuov candidate $V_{\widetilde{\theta}}(k)=\sum_{i=0}^{d-1}\widetilde{W}_{k+i}\widetilde{W}_{k+i}^T+\sum_{i=0}^{d-1}\widetilde{V}_{k+i}\widetilde{V}_{k+i}^T$. The first difference of $V_{\widetilde{\theta}}(k)$ is given as

$$\begin{aligned}\Delta V_{\widetilde{\theta}}(k)&=V_{\widetilde{\theta}}(k+1)-V_{\widetilde{\theta}}(k)\\&=\sum_{i=1}^{d}\widetilde{W}_{k+i}\widetilde{W}_{k+i}^T+\sum_{i=1}^{d}\widetilde{V}_{k+i}\widetilde{V}_{k+i}^T\\&\quad-\left(\sum_{i=0}^{d-1}\widetilde{W}_{k+i}\widetilde{W}_{k+i}^T+\sum_{i=0}^{d-1}\widetilde{V}_{k+i}\widetilde{V}_{k+i}^T\right)\\&=\widetilde{W}_{k+d}\widetilde{W}_{k+d}^T-\widetilde{W}_k\widetilde{W}_k^T+\widetilde{V}_{k+d}\widetilde{V}_{k+d}^T-\widetilde{V}_k\widetilde{V}_k^T.\end{aligned}$$

Using (25) and (26), one has

$$\begin{aligned}\Delta V_{\widetilde{\theta}}(k)=&-2\gamma[\widetilde{W}_k\widetilde{V}_k]\frac{\Omega(k)D\left(e(k+d)\right)}{\overline{\beta}\,\|\Omega(k)\|^2}\\&+\gamma^2\frac{\Omega^T(k)\Omega(k)D^2\left(e(k+d)\right)}{\overline{\beta}^2\,\|\Omega(k)\|^4}\\=&-2\gamma\frac{[\widetilde{W}_k\widetilde{V}_k]\Omega(k)\eta_k\rho_k[\widetilde{W}_k\widetilde{V}_k]\Omega(k)}{\overline{\beta}\,\|\Omega(k)\|^2}\\&+\gamma^2\frac{\left(\eta_k\rho_k[\widetilde{W}_k\widetilde{V}_k]\Omega(k)\right)^2}{\overline{\beta}^2\,\|\Omega(k)\|^2}\\=&\left(-2+\frac{\gamma\eta_k\rho_k}{\overline{\beta}}\right)\frac{\gamma\eta_k\rho_k\left([\widetilde{W}_k\widetilde{V}_k]\Omega(k)\right)^2}{\overline{\beta}\,\|\Omega(k)\|^2}.\end{aligned} \tag{27}$$

Since $0<\gamma\le 1$, $0<\rho_k\le\overline{\beta}$, and $0<\eta_k<1$, (27) becomes

$$\Delta V_{\widetilde{\theta}}(k)<-\frac{\gamma\eta_k\rho_k\left([\widetilde{W}_k\widetilde{V}_k]\Omega(k)\right)^2}{\overline{\beta}\,\|\Omega(k)\|^2}. \tag{28}$$

Therefore, the boundedness of $\widetilde{W}_k$ and $\widetilde{V}_k$ in (25) and hence of $W_k$ and $V_k$ in (23) is assured.

*2) Boundedness of the Tracking Error:* When $|e(k+d)|\le\varepsilon_0$, the boundedness of the tracking error is immediate as $|e(k+d)|\le\varepsilon_0<4\varepsilon_0$. Thus, we only need to prove the case where $|e(k+d)|>\varepsilon_0$.

Choose a Lyapnuov candidate

$$V(k)=V_s(k)+V_{\widetilde{\theta}}(k) \tag{29}$$

where $V_s(k)=(\gamma/2\overline{\beta}^2\Omega_0^2)\sum_{i=0}^{d-1}e^2(k+i)$ and

$$V_{\widetilde{\theta}}(k)=\sum_{i=0}^{d-1}\widetilde{W}_{k+i}\widetilde{W}_{k+i}^T+\sum_{i=0}^{d-1}\widetilde{V}_{k+i}\widetilde{V}_{k+i}^T.$$

The first difference of $V(k)$ is given as

$$\begin{aligned}\Delta V(k)&=V(k+1)-V(k)\\&=\frac{\gamma}{2\overline{\beta}^2\Omega_0^2}\left(\sum_{i=1}^{d}e^2(k+i)-\sum_{i=0}^{d-1}e^2(k+i)\right)\\&\quad+V_{\widetilde{\theta}}(k+1)-V_{\widetilde{\theta}}(k)\\&=\frac{\gamma}{2\overline{\beta}^2\Omega_0^2}\left(e^2(k+d)-e^2(k)\right)+\Delta V_{\widetilde{\theta}}(k).\end{aligned} \tag{30}$$

Substituting (26) and (27) into (30) yields

$$\begin{aligned}\Delta V(k)=&\frac{\gamma}{2\overline{\beta}^2\Omega_0^2}\left(e^2(k+d)-e^2(k)\right)-2\gamma\frac{D^2\left(e(k+d)\right)}{\eta_k\rho_k\overline{\beta}\,\|\Omega(k)\|^2}\\&+\gamma^2\frac{D^2\left(e(k+d)\right)}{\overline{\beta}^2\,\|\Omega(k)\|^2}\\\le&\frac{\gamma}{2\overline{\beta}^2\Omega_0^2}e^2(k+d)-\gamma\frac{D^2\left(e(k+d)\right)}{\eta_k\rho_k\overline{\beta}\,\|\Omega(k)\|^2}\\&-\gamma\frac{D^2\left(e(k+d)\right)}{\eta_k\rho_k\overline{\beta}\,\|\Omega(k)\|^2}+\gamma^2\frac{D^2\left(e(k+d)\right)}{\overline{\beta}^2\,\|\Omega(k)\|^2}.\end{aligned} \tag{31}$$

As $0<\gamma\le 1$, $0<\rho_k<\overline{\beta}$, $0<\eta_k<1$, and $0<\|\Omega(k)\|\le\Omega_0$, (31) becomes

$$\begin{aligned}\Delta V(k)&\le\frac{\gamma}{2\overline{\beta}^2\Omega_0^2}e^2(k+d)-\gamma\frac{D^2\left(e(k+d)\right)}{\eta_k\overline{\beta}^2\,\|\Omega(k)\|^2}\\&\le\frac{\gamma}{2\overline{\beta}^2\Omega_0^2}\left(e^2(k+d)-2D^2\left(e(k+d)\right)\right).\end{aligned} \tag{32}$$

When $|e(k+d)|>\varepsilon_0$, from (24), one has

$$D\left(e(k+d)\right)=e(k+d)-\varepsilon_0\text{sign}\left(e(k+d)\right).$$

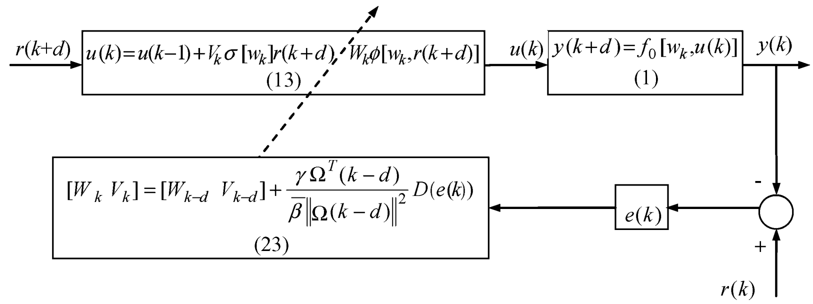


Fig. 2. Feedback-linearization-based NN adaptive control structure.

Substituting the previous equation into (32) yields

$$\begin{aligned}\Delta V(k) \leq & \frac{\gamma}{2\overline{\beta}^2\Omega_0^2} \\ & \times \left(e^2(k+d)-2\left\{e(k+d)-\varepsilon_0 \text{sign}\left(e(k+d)\right)\right\}^2\right) \\ = & \frac{\gamma}{2\overline{\beta}^2\Omega_0^2}\left(-e^2(k+d)+4\varepsilon_0 \text{sign}\left(e(k+d)\right)e(k+d)\right. \\ & \left. -2\left\{\varepsilon_0 \text{sign}\left(e(k+d)\right)\right\}^2\right) \\ \leq & \frac{\gamma}{2\overline{\beta}^2\Omega_0^2}\left(-e^2(k+d)+4\varepsilon_0\left|e(k+d)\right|\right). \end{aligned} \tag{33}$$

Define a compact set $S_e = \{e(k)||e(k)| \leq 4\varepsilon_0\}$. From (33), it is obvious that $\Delta V(k) < 0$ when $|e(k+d)|$ is out of the compact set $S_e$. Thus, one concludes that, for given reference trajectory $r(k)$, using the NN adaptive control law (13) and NN weight updating law (23), the ultimate bound on the tracking error is $\lim_{k\to\infty}|e(k)| \leq 4\varepsilon_0$. As defined in (12), $\varepsilon(k)$ is the NN approximation error and $|\rho_k\varepsilon(k)| \leq \varepsilon_0$. Therefore, $\varepsilon(k)$ can be arbitrarily small by using enough neurons in NNs and thus $\varepsilon_0$ can be set to be arbitrarily small, which leads to arbitrarily small control error. As $\Delta V(k)$ is negative and $S_r$ is only a small subset of $S_y$, there must exist a large enough compact set $S_e$ such that, for any $e(k) \in S_e$, $y(k+d) \in S_y$ is guaranteed.

*3) Boundedness of the System Input:* From (13), one has

$$\begin{aligned} u(k) &= u(k-1) + V_k\sigma[w_k]r(k+d) - W_k\phi\left[w_k, r(k+d)\right] \\ &= u(k-1) + V^*\sigma[w_k]r(k+d) - W^*\phi\left[w_k, r(k+d)\right] \\ &\quad - \widetilde{V}_k\sigma[w_k]r(k+d) + \widetilde{W}_k\phi\left[w_k, r(k+d)\right] \\ &= u(k-1) + V^*\sigma[w_k]r(k+d) - W^*\phi\left[w_k, r(k+d)\right] \\ &\quad + \varepsilon(k) - \widetilde{V}_k\sigma[w_k]r(k+d) + \widetilde{W}_k\phi\left[w_k, r(k+d)\right] \\ &\quad - \varepsilon(k). \end{aligned} \tag{34}$$

Using (12), (34) becomes

$$u(k) = u^*(k) - \widetilde{V}_k\sigma[w_k]r(k+d) + \widetilde{W}_k\phi\left[w_k, r(k+d)\right] - \varepsilon(k). \tag{35}$$

Substituting (22) into (35), one obtains

$$\begin{aligned} |u(k) - u^*(k)| &= \left|-[\widetilde{W}_k \;\; \widetilde{V}_k]\Omega(k) - \varepsilon(k)\right| \\ &= \left|-\left(\frac{e(k+d)}{\rho_k} - \varepsilon(k)\right) - \varepsilon(k)\right| \\ &= \left|\frac{e(k+d)}{\rho_k}\right|. \end{aligned} \tag{36}$$

Because $\lim_{k\to\infty}|e(k)| \leq 4\varepsilon_0$, from (36), one concludes that $\lim_{k\to\infty}|u(k) - u^*(k)| \leq 4\varepsilon_0/\rho_k$. As indicated in the proof of the boundedness of the tracking error, $\varepsilon_0$ can be set to be arbitrarily small and thus $\lim_{k\to\infty}|u(k) - u^*(k)|$ can be arbitrarily small. Because $e(k)$ is in the compact set $S_e$ and $S_{ud}$ is only a small subset of $S_u$, there must exist a large enough compact set $S_{\widetilde{u}} \subset R$ such that $u(k)$ is bounded and $u(k) \in S_u$ is guaranteed for all $|u(k) - u^*(k)| \in S_{\widetilde{u}}$.

As all signals in the system are bounded, the closed-loop system is UUB. Then, $y(k) \in S_y$ and $u(k) \in S_u$ will hold for all $k \geq 0$.

### *E. Summary*

The proposed feedback-linearization-based NN adaptive control is shown in Fig. 2 with design procedures as follows.

- Determine the NN adaptive control (13) from the equivalent affine-like model (8).
- Determine the NN weight updating law (23) with the deadzone technique (24).

It is well known that feedback linearization methods cannot be used for unknown nonaffine nonlinear discrete-time systems. However, the NN adaptive control (13) is based on feedback linearization and is implemented straightforwardly in terms of the affine-like model (8). Additionally, pretraining is not required and controller singularity problem is completely avoided. Under the proposed NN adaptive control, when enough neurons are used in the NNs, NN approximation error $\varepsilon(k) \to 0$ and thus $\varepsilon_0 \to 0$, resulting in that the tracking error $e(k) \to 0$. The result is very important because arbitrary small control error can be achieved by using enough NN nodes in the proposed NN adaptive controller. However, this point was not clearly shown in previous NN adaptive control methods for unknown nonaffine nonlinear discrete-time systems.

## IV. Simulations

The developed feedback-linearization-based NN adaptive control as shown in Fig. 2 will be simulated here and NRBFNs will be used for nonlinear function approximation in the adaptive control of unknown nonaffine nonlinear discrete-time systems. Two examples are used in the simulations. Example 1 is a first-order nonlinear discrete-time system and example 2 is a second-order system with delay. The references to be tracked are set-point step change signals. To shape the discontinuous reference signals, a first-order filter, e.g., $0.6z/(z-0.4)$, is used in the two examples. To confine the inputs of NRBFNs in $[-1, 1]$, the one-to-one mappings $\overline{y}(k) = y(k)/(a + |y(k)|)$, $\overline{u}(k) = u(k)/(b + |u(k)|)$, and $\overline{r}(k) = r(k)/(c + |r(k)|)$ are

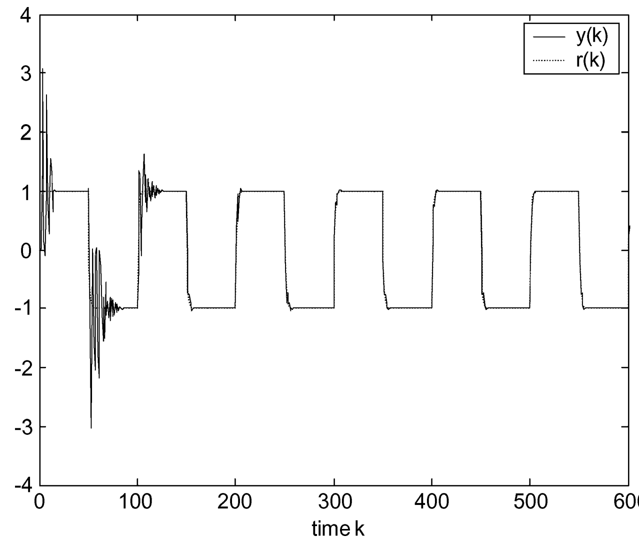


Fig. 3. Tracking performance in Example 1.

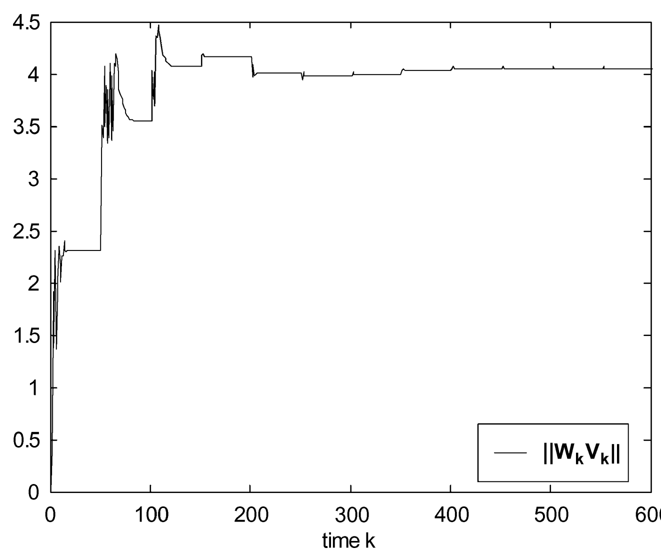


Fig. 4. Convergence of the NN weights in Example 1.

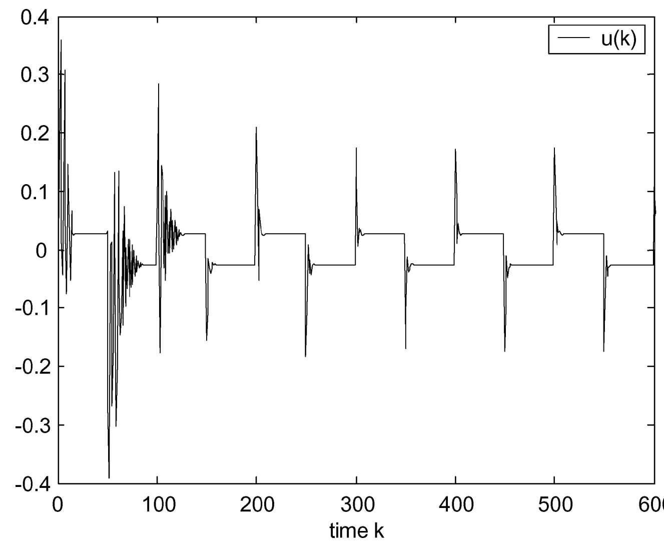


Fig. 5. Control signal in Example 1.

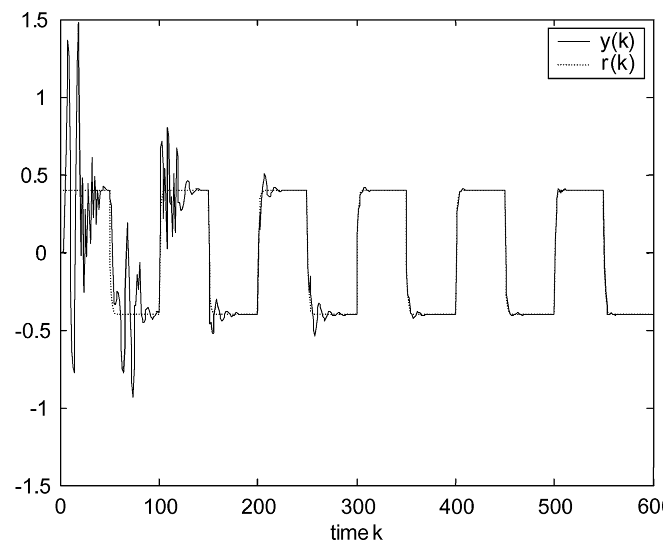


Fig. 6. Tracking performance in Example 2.

used as the NRBFN inputs instead of $y(k)$, $u(k)$, and $r(k)$. In the above mappings, $a$, $b$, and $c$ are three small positive numbers, which will be suitably chosen in terms of different examples. In all simulations, $\gamma$ in NN weight updating law (23) is chosen as $\gamma = 1$. Besides, in all NRBFNs, the spread $s$ is set to $s = 1$ and all centroids are randomly chosen in $[-1, 1]$. For simplicity, each NRBFN used in the NN adaptive control (13) has 200 nodes.

*Example 1*

A first-order system is described in the following difference equation [12], [13]:

$$y(k+1) = \sin[y(k)] + u(k)\left(5 + \cos[y(k)u(k)]\right). \quad (37)$$

The objective of the simulation is to track a set-point step change signal using the proposed NN adaptive control (13) with NN weight updating law (23). The parameters such as $a = b = c = 0.1$ and $\overline{\beta} = 4.5$ are fixed. $\varepsilon_0$ in dead-zone function is chosen as $\varepsilon_0 = 0.001$.

The tracking performance under the proposed NN adaptive control is given in Fig. 3. It is clearly shown that the tracking performance is quite good after a short self-learning period. The norm of the NN weight vector $||W_k\, V_k||$ is calculated to show the boundedness and convergence of the NN weights as in Fig. 4. The control signal is shown in Fig. 5. The simulation demonstrates that, using the proposed NN adaptive control, the control performance and stability of the closed-loop system can be guaranteed without pretraining.

*Example 2*

A second-order system with delay is described in the following difference equation:

$$\begin{aligned} y(k+2) =\, & 0.2\sin\left(0.5\left(y(k)+y(k-1)\right)\right) \\ & + 0.2\sin\left(0.5\left(y(k)+y(k-1)\right)+2u(k)+u(k-1)\right) \\ & + \left(4u(k)+u(k-1)\right) \\ & /\left(1+0.2\cos\left(0.2\left(2y(k)+y(k-1)\right)\right)\right). \end{aligned} \quad (38)$$

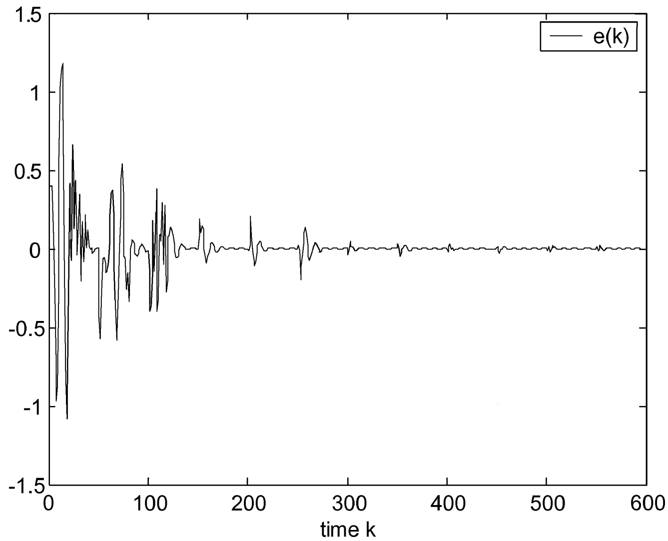


Fig. 7. Tracking error in Example 2.

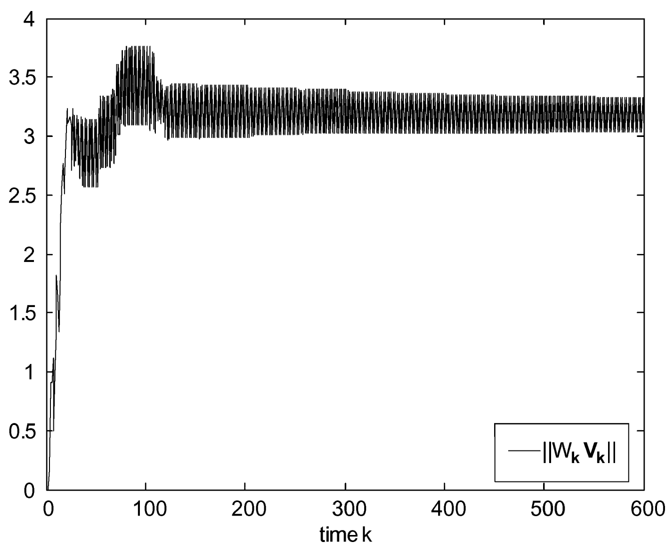


Fig. 8. Convergence of the NN weights in Example 2.

The objective of the simulation is to track a set-point step change reference signal using the proposed NN adaptive control. In the simulation, we set that $a = b = c = 0.04$, $\overline{\beta} = 5$, and $\varepsilon_0 = 0.00001$.

The control performance, the tracking error, the norm of NN weights, and the control signal are shown in Figs. 6–9. Figs. 6 and 7 demonstrate that the proposed NN adaptive control can be used to complex discrete-time nonlinear systems with delay. All signals in the closed-loop system are indeed bounded, which validate the theoretical findings. As shown in Fig. 8, the NN weights are bounded but may not converge. The small oscillation of the NN weights may be caused by the delay and further investigation for this issue is being undertaken.

## V. Conclusion

An equivalent affine-like model is derived for unknown nonaffine nonlinear discrete-time systems so that feedback linearization adaptive control with NNs can be implemented. The NN weight updating law with the dead-zone technique is

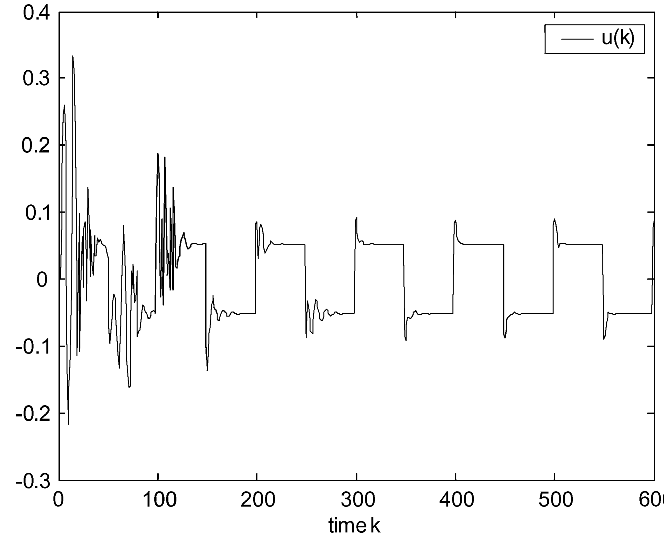


Fig. 9. Control signal in Example 2.

determined in terms of Lyapunov stability theory. Pretraining is not required for the NN adaptive control and controller singularity problem is completely avoided. The tracking error is proportional to NN approximation error, which can be made arbitrarily small by choosing enough neurons. Simulation studies demonstrate the performance of the proposed feedback-linearization-based NN adaptive control and its feasibility for unknown nonaffine nonlinear discrete-time systems.

## Appendix

### A. Proof of Property 1

When $r(k+d) = y(k+d)$, $\xi_{0k} = y(k+d)$ and $\varsigma_k = u(k)$. Therefore, Property 1 holds by taking the derivative of the both sides of (3b) with respect to $y(k+d)$. Now, we only need to consider the case where $r(k+d) \neq y(k+d)$, that is, $\xi_{0k}$ is between $r(k+d)$ and $y(k+d)$.

From (3a), taking the Taylor expansion with respect to $y(k+d)$ around $(w_k, r(k+d))$ and using the mean value theorem, one has

$$\begin{aligned} G\left[w_k, y(k+d)\right] &= G\left[w_k, r(k+d)\right] \\ &\quad + G_y[w_k, \xi_{0k}]\left(y(k+d) - r(k+d)\right) \end{aligned}$$

where $G_y[w_k, \xi_{0k}]$ is defined in Property 1. Substituting the previous equation into (3b) and taking the Taylor expansion with respect to $G[w_k, y(k+d)]$ around $(w_k, G[w_k, r(k+d)])$ and using the mean value theorem, one has

$$\begin{aligned} y(k+d) &= f_0\left[w_k, G\left[w_k, r(k+d)\right] + G_y[w_k, \xi_{0k}] \right. \\ &\qquad \left. \times \left(y(k+d) - r(k+d)\right)\right] \\ &= f_0\left[w_k, G\left[w_k, r(k+d)\right]\right] + f_{0u}[w_k, \varsigma_k] \\ &\qquad \times G_y[w_k, \xi_{0k}]\left(y(k+d) - r(k+d)\right) \end{aligned} \tag{A.1}$$

where $f_{0u}[w_k, \varsigma_k]$ is defined in Property 1. According to Lemma 1, (A.1) becomes

$$y(k+d) = r(k+d) + f_{0u}[w_k, \varsigma_k] G_y[w_k, \xi_{0k}]\left(y(k+d) - r(k+d)\right). \tag{A.2}$$

Since $y(k+d) \neq r(k+d)$, from (A.2), $f_{0u}[w_k, \varsigma_k]G_y[w_k, \xi_{0k}] \equiv 1$ must hold.

*B. Derivation of (17)*

When $r(k+d) = y(k+d)$, that is, $\xi_{0k} = \xi_{1k} = y(k+d)$, by taking the derivative of both sides of (8) with respect to $y(k+d)$, one obtains

$$f_y\left[w_k, y(k+d)\right] = 1 - g[w_k]G_y\left[w_k, y(k+d)\right] \tag{B.1}$$

where

$$f_y\left[w_k, y(k+d)\right] = \frac{\partial f\left[w_k, y(k+d)\right]}{\partial y(k+d)}$$

and

$$G_y\left[w_k, y(k+d)\right] = \frac{\partial G\left[w_k, y(k+d)\right]}{\partial y(k+d)}.$$

When $r(k+d) \neq y(k+d)$, from (8), one has

$$\begin{aligned} y(k+d) &= f\left[w_k, y(k+d)\right] + g[w_k]\left(u(k) - u(k-1)\right) \\ &= f\left[w_k, y(k+d)\right] + g[w_k] \\ &\quad \times \left(G\left[w_k, y(k+d)\right] - u(k-1)\right) \\ &= f\left[w_k, r(k+d)\right] + f_y[w_k, \xi_{1k}]\left\{y(k+d) - r(k+d)\right\} \\ &\quad + g[w_k]\left(G\left[w_k, r(k+d)\right] + G_y[w_k, \xi_{0k}]\right. \\ &\qquad \left. \times \left\{y(k+d) - r(k+d)\right\} - u(k-1)\right) \end{aligned} \tag{B.2}$$

where $f_y[w_k, \xi_{1k}]$ and $G_y[w_k, \xi_{0k}]$ are defined in (16) and in Property 1, respectively. As (8) is equivalent to (1), according to Lemma 1, one has

$$r(k+d) = f\left[w_k, r(k+d)\right] + g[w_k]\left(G\left[w_k, r(k+d)\right] - u(k-1)\right). \tag{B.3}$$

Substituting (B.3) into (B.2), together with (B.1), yields (17).

*C. PE Condition of $\Omega(k)$ Without a Dead Zone*

The following definition is needed for the explanation.

*Definition C.1 [7], [25]:* An input sequence $\Omega(k)$ is said to be PE if there exist $\lambda > 0$ and integer $k_1 \geq 1$ such that

$$\lambda_{\min}\left[\sum_{k=k_0}^{k_1} \Omega(k)\Omega^T(k)\right] > \lambda \qquad \forall k_0 \geq 0 \tag{C.1}$$

where $\lambda_{\min}[M]$ denotes the smallest eigenvalue of $M$.

From (23), with $D(e(k))$ being replaced by $e(k)$, one has

$$[\widetilde{W}_k \;\; \widetilde{V}_k] = [\widetilde{W}_{k-d} \;\; \widetilde{V}_{k-d}] - \frac{\gamma}{\overline{\beta}\left\|\Omega(k-d)\right\|^2}\Omega^T(k-d)e(k). \tag{C.2}$$

Substituting (22) into (C.2) yields

$$\begin{aligned} &[\widetilde{W}_k \;\; \widetilde{V}_k] \\ &\quad = [\widetilde{W}_{k-d} \;\; \widetilde{V}_{k-d}]\left(I - \frac{\gamma\rho_{k-d}}{\overline{\beta}\left\|\Omega(k-d)\right\|^2}\Omega(k-d)\Omega^T(k-d)\right) \\ &\qquad - \frac{\gamma\rho_{k-d}}{\overline{\beta}\left\|\Omega(k-d)\right\|^2}\Omega^T(k-d)\varepsilon(k-d). \end{aligned} \tag{C.3}$$

If (C.1) does not hold, that is, the regression vector $\Omega(k)$ is not persistently exciting, it is clearly shown in (C.3) that, although $0 < \gamma \leq 1$ and $0 < \rho_k \leq \overline{\beta}$ from Property 2, $\|I - (\gamma\rho_{k-d}/\overline{\beta}\|\Omega(k-d)\|^2)\Omega(k-d)\Omega^T(k-d)\| < 1$ cannot be guaranteed. Therefore, the boundedness of $[\widetilde{W}_k \; \widetilde{V}_k]$ in (C.2), and hence of $[W_k \; V_k]$, may not be assured as $\varepsilon(k) \neq 0$ during adaptive control.

## ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions.

## REFERENCES

[1] F. C. Chen and H. K. Khalil, "Adaptive control of nonlinear systems using neural networks," in *Proc. 29th IEEE Conf. Decision Control*, Honolulu, HI, Dec. 1990, pp. 1707–1712.

[2] F. C. Chen and H. K. Khalil, "Adaptive control of a class of nonlinear discrete-time systems using neural networks," *IEEE Trans. Autom. Control*, vol. 40, no. 5, pp. 791–807, May 1995.

[3] K. Fregene and D. Kennedy, "Stabilizing control of a high-order generator model by adaptive feedback linearization," *IEEE Trans. Energy Convers.*, vol. 18, no. 1, pp. 149–156, Mar. 2003.

[4] S. S. Ge, T. H. Lee, G. Y. Li, and J. Zhang, "Adaptive NN control for a class of discrete-time non-linear systems," *Int. J. Control*, vol. 76, no. 4, pp. 334–354, 2003.

[5] S. Jagannathan, "Control of a class of nonlinear discrete-time systems using multilayer neural networks," *IEEE Trans. Neural Netw.*, vol. 12, no. 5, pp. 1113–1120, Sep. 2001.

[6] S. Jagannathan*, Neural Network Control of Nonlinear Discrete-Time Systems*. Boca Raton, FL: CRC Press, 2006.

[7] S. Jagannathan and F. L. Lewis, "Multilayer discrete-time neural-net controller with guaranteed performance," *IEEE Trans. Neural Netw.*, vol. 7, no. 1, pp. 107–130, Jan. 1996.

[8] L. Jin, P. N. Nikiforuk, and M. M. Gupta, "Fast neural learning and control of discrete-time nonlinear systems," *IEEE Trans. Syst. Man Cybern.*, vol. 25, no. 3, pp. 478–488, Mar. 1995.

[9] D. Muñz and D. Sbarbaro, "An adaptive sliding-mode controller for discrete nonlinear systems," *IEEE Trans. Ind. Electron.*, vol. 47, no. 3, pp. 574–581, Jun. 2000.

[10] M. Norgaard, O. Ravn, N. K. Poulsen, and L. K. Hansen*, Neural Networks for Modelling and Control of Dynamic Systems*. London, U.K.: Springer-Verlag, 2000.

[11] J. T. Spooner, M. Maggiore, R. Ordonez, and K. M. Passino*, Stable Adaptive Control and Estimation for Nonlinear Systems*. New York: Wiley, 2002.

[12] K. S. Narendra and S. Mukhopadhyay, "Adaptive control using neural networks and approximate models," *IEEE Trans. Neural Netw.*, vol. 8, no. 3, pp. 475–485, May 1997.

[13] H. Deng and H.-X. Li, "A novel neural network approximate inverse control for unknown nonlinear discrete dynamical systems," *IEEE Trans. Syst. Man Cybern. B, Cybern.*, vol. 35, no. 1, pp. 115–123, Feb. 2005.

[14] J. R. Noriega and H. Wang, "A direct adaptive neural-network control for unknown nonlinear systems and its application," *IEEE Trans. Neural Netw.*, vol. 9, no. 1, pp. 27–34, Jan. 1998.

[15] M. S. Ahmed, "Neural-net-based direct adaptive control for a class of nonlinear plants," *IEEE Trans. Autom. Control*, vol. 45, no. 1, pp. 119–124, Jan. 2000.

[16] H. Wang, G. P. Liu, C. J. Harris, and M. Brown*, Advanced Adaptive Control*. New York: Elsevier Science, 1995.

[17] G. L. Plett, "Adaptive Inverse control of linear and nonlinear system using dynamic neural networks," *IEEE Trans. Neural Netw.*, vol. 14, no. 2, pp. 360–376, Mar. 2003.

[18] P. J. Werbos, "Backpropagation through time: What it does and how to do it," *Proc. IEEE*, vol. 78, no. 10, pp. 1550–1560, Oct. 1990.

[19] R. J. Williams and D. Zipser, "Experimental analysis of the real-time recurrent learning algorithm," *Connection Sci.*, vol. 1, no. 1, pp. 87–111, 1989.

[20] S. Mukhopadhyay and K. S. Narendra, "Adaptive control of nonlinear multivariable systems using neural networks," *Neural Netw.*, vol. 7, no. 5, pp. 737–752, 1994.

[21] S. S. Ge, J. Zhang, and T. H. Lee, "Adaptive MNN control for a class of non-affine NARMAX systems with disturbances," *Syst. Control Lett.*, vol. 53, pp. 1–12, 2004.

[22] Q. M. Zhu and L. Guo, "Stable adaptive neurocontrol for nonlinear discrete-time systems," *IEEE Trans. Neural Netw.*, vol. 15, no. 3, pp. 653–662, May 2004.
[23] O. Adetona, S. Sathananthan, and L. H. Keel, "Robust adaptive control of nonaffine nonlinear plants with small input signal changes," *IEEE Trans. Neural Netw.*, vol. 15, no. 2, pp. 408–416, Mar. 2004.
[24] H. Deng and H.-X. Li, "On the new method for the control of discrete nonlinear dynamic systems using neural networks," *IEEE Trans. Neural Netw.*, vol. 17, no. 2, pp. 526–529, Mar. 2006.
[25] N. Sadegh, "A perception network for functional identification and control of nonlinear systems," *IEEE Trans. Neural Netw.*, vol. 4, no. 6, pp. 982–988, Jul. 1993.
[26] H.-X. Li and H. Deng, "An approximate internal model-based neural control for unknown nonlinear discrete processes," *IEEE Trans. Neural Netw.*, vol. 17, no. 3, pp. 659–670, May 2006.
[27] J. B. D. Cabrera and K. S. Narendra, "Issues in the application of neural networks for tracking based on inverse control," *IEEE Trans. Autom. Control*, vol. 44, no. 11, pp. 2007–2027, Nov. 1999.
[28] D. Ramkrishna and N. R. Amundson, *Linear Operator Methods in Chemical Engineering*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
[29] H. Deng, H.-X. Li, and G. R. Chen, "Spectral approximation based intelligent modeling for distributed thermal processes," *IEEE Trans. Control Syst. Technol.*, vol. 13, no. 5, pp. 686–700, Sep. 2005.
[30] W. Lan and J. Huang, "Neural-network-based approximate output regulation of discrete-time nonlinear systems," *IEEE Trans. Neural Netw.*, vol. 18, no. 4, pp. 1196–1208, Jul. 2007.
[31] L. Scardovi, M. Baglietto, and T. Parisini, "Active state estimation for nonlinear systems: A neural approximation approach," *IEEE Trans. Neural Netw.*, vol. 18, no. 4, pp. 1172–1184, Jul. 2007.
[32] B.-J. Yang and A. J. Calise, "Adaptive control of a class of nonaffine systems using neural networks," *IEEE Trans. Neural Netw.*, vol. 18, no. 4, pp. 1149–1159, Jul. 2007.
[33] K. Funahashi, "On the approximate realization of continuous mapping by neural network," *Neural Netw.*, vol. 2, pp. 183–192, 1989.
[34] M. M. Gupta and D. H. Rao, *Neuro-Control Systems: Theory and Applications*. New York: IEEE Press, 1994.
[35] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feed-forward networks are universal approximators," *Neural Netw.*, vol. 2, pp. 359–366, 1989.
[36] O. Nelles, *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*. Berlin, Germany: Springer-Verlag, 2001.
[37] A. Krzyzak and D. Schäfer, "Nonparametric regression estimation by normalized radial basis function networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 1003–1010, Mar. 2005.
[38] D. Schroder, *Intelligence Observer and Control Design for Nonlinear Systems*. Berlin, Germany: Springer-Verlag, 2000.
[39] M. R. Cowper, B. Mulgrew, and C. P. Unsworth, "Nonlinear prediction of chaotic signals using a normalized radial basis function network," *Signal Process.*, vol. 82, pp. 775–789, 2002.
[40] G. Bugmann, "Normalized Gaussian radial function networks," *Neurocomputing*, vol. 20, pp. 97–110, 1998.

**Hua Deng** received the B.E. degree in electrical engineering from Nanjing Aeronautical Institute, Nanjing, China, in 1983, the M.E. degree in electrical engineering from Northwestern Polytechnical University, Xi'an, China, in 1988, and the Ph.D. degree in manufacturing engineering from City University of Hong Kong, Hong Kong, in 2005.

Currently, he is the Professor at the School of Mechanical and Electrical Engineering, Central South University, Changsha, China. His research interests include modeling and control of complex distributed parameter systems, intelligent control and learning, intelligent manufacturing, and fault diagnosis.

**Han-Xiong Li** (S'94–M'97–SM'00) received the B.E. degree in aerospace engineering from National University of Defense Technology, Changsha, China, in 1982, the M.E. degree in electrical engineering from Delft University of Technology, Delft, The Netherlands, in 1991, and the Ph.D. degree in electrical engineering from University of Auckland, Auckland, New Zealand, in January 1997.

Currently, he is the Professor at the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, Hong Kong. He is a "Chang Jiang Scholar" in Central South University. In last 20 years, he has opportunities to work in different fields including, military service, industry, and academia. He has gained industrial experience as a Senior Process Engineer from ASM—a leading supplier for semiconductor process equipment. His research interests include intelligent control and learning, process modeling and control, complex distributed parameter system, and electronics packaging process.

Dr. Li is an Associate Editor of the IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS—PART B: CYBERNETICS. He was awarded the distinguished young scholar fund by the China National Science Foundation in 2004.

**Yi-Hu Wu** received the B.E., M.E., and Ph.D. degrees in mechanical engineering from Hunan University, Changsha, China, in 1985, 1988, and 1999, respectively.

Currently, he is the Professor at the School of Automobile and Mechanical Engineering, Changsha University of Science and Technology, Changsha, China. His research interests include automotive exhaust control, automotive electronic control, and performance optimization of internal combustion engines.