2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Structured Wavelet-based Neural Network for Control of Nonlinear Systems

A. Karami and M.J. Yazdanpanah

Abstract— **In this paper, a wavelet-based neural network is proposed for the control of nonlinear systems. Activation functions of neural network nodes are determined based on the wavelet transform. The controller can efficiently compensate for the undesired effects of hard nonlinearities such as saturation and/or dead zone of control input. Compared with standard neuro-controllers, the structure of the controller is definite and simple. The proposed controller is localizable and has a systematically chosen structure, which improves the close-loop performance. An off-line algorithm determines the number of nodes. In addition, an on-line algorithm adjusts the parameters of wavelet bases and network weights. Back propagation algorithm with a momentum term is used for updating the weights and parameters of activation functions. This controller reduces the quantity of network parameters, calculation cost and convergence time of online algorithms with respect to the conventional neural network. Also, the controller is able to control unstable and MIMO systems. To illustrate the capability and performance superiority of the proposed controller, two nonlinear systems are controlled and the corresponding results are compared.**

Index Terms— **Wavelet transform, neural network, nonlinear system control, adaptive activation functions.**

I. INTRODUCTION

CONTROL of nonlinear systems is an important branch in control engineering. Control methods based on NN (neural networks) are effective tools for identification and control of nonlinear uncertain systems [1]-[3] since classic control methods need a mathematical model of plant and performance of control is reduced with modeling error of nonlinear system. NN can be used for determining an appropriate model by utilizing input-output data of the plant; however, neural networks are not able to localize and modeling error converge slowly. Using an appropriately chosen adaptive mechanism, the parameters of NN can be tuned online in order to approximate the unknown nonlinear model. The application of the Lyapunov's direct method shows that the adaptive neural control can guarantee the stability of closed loop system [4]. Tracking error in the neural controller is a direct result of the function approximation error [5]. It is proven that a neural network is able to approximate any nonlinear function with the perspecified accuracy, but the selected network structure is very important [6]. Nevertheless, the selected structure is very important to meet the desired accuracy. A trial and error method is often used for structure selection.

Neural networks have several important disadvantages: no localizability, lack of a systematic method for structure selection and slow convergence of approximation. To overcome the lack of localization and global approximation of neural networks, the global activation function is substituted with a localized and compact support function such as radial base functions. However, in these local neural networks, only small spaces of input space can map to the output space. A wavelet network can overcome the above mentioned disadvantages because wavelet functions are compact-support and able to be localized in both time and space domains. Also, since different functions are used for the node's activation functions, wavelet networks can map any input space to the output space. Different activation functions can be determined based on wavelet decomposition techniques such as multi-resolution approximation, cut-down discrete wavelet transform and continuous wavelet transform.

Recently, various studies have used wavelet transform and wavelet networks for the identification and control of nonlinear systems [7]-[11]. It has been mentioned that wavelet network is a universal approximator and that it can approximate any nonlinear function with prespecified accuracy using a linear combination of wavelet functions [4]. In [4], an adaptive wavelet controller (AWC) is used to control only SISO systems with a partially known structure or input-to-state-stable; in other words, L^2 and L^∞ stable systems. In this controller, activation functions are determined based on the multi-resolution analysis and are not adjustable. In addition, the convergence of tracking error is very slow.

In this paper, a certain class of wavelet network is proposed for control of the nonlinear systems. This controller is a wavelet network in three layers, but the connection weights of only one layer will be adjusted. In this controller, the scaled and translated mother wavelet functions are utilized as the activation function of nodes based on the continuous wavelet transform. At first, the network structure is definite and only the number of nodes is unknown. An algorithm is used for determining the node's number, which found the number of wavelet bases and the network structure proper for the identification of the nonlinear system. The supervised gradient descent method and on-line back propagation (BP) algorithm is applied in

The authors are with the Control & Intelligent Processing Center of Excellence, School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran (e-mail: alikarami@ut.ac.ir).

order to adjust the scale and translation of mother wavelet functions (shape of wavelet functions) and the connection weights of the network. The proposed controller has several advantages. Due to the ability of wavelet functions in localization in both time and space domains, any arbitrary functions can be approximated more effectively than those of the conventional neural networks, especially functions with sharp changes [12], [13]. These wavelet networks have fewer coefficients to be adjusted than NN so that these parameters can converge much faster with less error in order to control or identify nonlinear systems [14]. In this control scheme, the network structure is definite and a proper number of wavelet bases (number of hidden layers' nodes) can be obtained. In contrast, NN structures need to be determined by trial and error methods. The suggested controller can be used to control unstable and MIMO systems. In general, this wavelet network can be an appropriate candidate for replacing neural networks in various control schemes.

The paper is organized as follows. In Section II, the preliminaries on the wavelet theory for neural networks is presented. In Section III, network structure and learning algorithms for the proposed controller is depicted. The illustrative examples and comparative results are given in Section IV and, in Section V, the conclusion is provided.of nonlinear systems is an important branch in control engineering. Control methods based on NN (neural networks) are effective tools for identification and control of nonlinear uncertain systems [1]-[3].

II. Wavelet Theory for Neural Networks

From [15], we have following discussion:

"Choose the function $\varphi(x) \in L^2(\Re)$ which is also called the mother wavelet and satisfies the following conditions:

- The finite energy condition:

$$\int_{-\infty}^{\infty} \left|\varphi(t)\right|^2 dt < \infty \tag{1}$$

- The admissibility condition:

$$\int_R \frac{\left|\hat{\varphi}(\omega)\right|^2}{|\omega|} d\omega < \infty \tag{2}$$

where $\hat{\varphi}(w)$ is the Fourier transform of $\varphi(x)$. Then, the corresponding family of dilated and translated wavelets can be defined by:

$$\{\varphi_{j,k}(x) = a^{-2/j}\ \varphi(a^{-j}\ x - kb),\ (j,k) \in Z^2\} \tag{3}$$

where a and b are, respectively, the dilation and translation parameters. By selecting a and b properly, $\{\varphi_{j,k}(x)\}$ can be called affine wavelet which may constitute a frame of $L^2(\Re)$, i.e.

$$A\|f\|^2 \leq \sum_{(j,k)\in Z^2} \left|\langle \varphi_{j,k}, f\rangle\right|^2 \leq B\|f\|^2 \tag{4}$$

where $f \in L^2(\Re)$, $\langle \varphi_{j,k}, f\rangle = \int_R \varphi_{j,k}(t) f(t)\,dt$ is inner product and $A>0$ and $B>0$ are frame bounds. If $A=B$, $\{\varphi_{j,k}(x), (j,k) \in Z^2\}$ is the tight frame. In this case, it leads to:

$$f(x) = A^{-1} \sum_{(j,k)\in Z^2} \langle \varphi_{j,k}, f\rangle . \varphi_{j,k}(x) \tag{5}$$

while $A=B=1$, $\{\varphi_{m,n}(x),), (m,n) \in Z^2\}$ becomes an orthonormal basis. In such a case,

$$f(x) = \sum_{(j,k)\in Z^2} \langle \varphi_{j,k}, f\rangle . \varphi_{j,k}(x) \tag{6}$$

It should be emphasized that the wavelet transform holds the advantage of variable time-frequency localization. The lattice points of the mother wavelets $\{\varphi_{j,k}(x)\}$ are located on (kbaj,±a-jωo); therefore, the width of the time-window of $\varphi_{j,k}(x)$ can be changed with the variation of the frequency. Thus, this property is very useful for the analysis of non-stationary signals and the nonlinear function learning. An often-quoted example of a wavelet is the second derivative of a Gaussian function:

$$\varphi(x) = (1 - x^2)e^{-\frac{x^2}{2}} \Leftrightarrow \varphi(w) = \sqrt{2\pi\omega^2}\, e^{-\frac{\omega^2}{2}} \tag{7}$$

This function has the excellent localization in time and frequency and satisfies the admissibility condition.

In terms of the results shown above, the so-called wavelet basis function (WBF) neural networks can be defined as follows:

$$f(x) = \sum_{j=1}^{K} w_j \varphi_j(x) = \sum_{j=1}^{K} w_j \varphi_j(a_j x - b_j) \tag{8}$$

where $w_i \in \Re$, $a_i \in \Re^d$, $b_i \in \Re^d$, d is the dimension of input and K is the number of wavelet bases. Moreover, the values of a_i and b_i construct a regular lattice in wavelet bases and frames" [14], [15].

For a multi-dimensional case, tensor product of one-dimension wavelet can be used:

$$\psi(\mathbf{x}) = \psi(x_1, x_2, \cdots x_n) = \prod_{i=1}^{n} \varphi(x_i) \tag{9}$$

The network input is $\mathbf{x} = [x_1, x_2, \ldots x_n]^T$, where n is the input dimension. In general, any function can be approximated by a wavelet network whose activation function nodes are the scaled and translated mother wavelet, $\varphi_{ab}(x)$:

$$\varphi_{ab}(x) = \sqrt{a}\ \varphi\ (ax - b), \quad a \in R^+, b \in R \tag{10}$$

The normalization factor $\sqrt{a}$ is introduced so that the energy of $\varphi_{ab}(x)$ be the same as that of $\varphi(x)$.

$$\varphi_{a_j b_j}(x_i) = \sqrt{a_{ij}}\,\varphi\left(a_{ij} x_i - b_{ij}\right) \tag{11}$$

$$\psi_j(\mathbf{x}) = \prod_{i=1}^{n} \varphi_{a_j b_j}(x_i) \tag{12}$$

$$f(x)=\sum_{j=1}^{K} w_j \psi_j(\boldsymbol{x}) \tag{13}$$

If system output is multi-dimensional $(\mathbf{y} \in \Re^m)$:

$$y_i = \sum_{j=1}^{K} w_{ij}\psi_j(\boldsymbol{x}) + e_i, \quad i = 1,2,\cdots,m \tag{14}$$

$$y_i \cong \sum_{j=1}^{K} w_{ij}\psi_j(\boldsymbol{x}), \quad i = 1,2,\cdots,m \tag{15}$$

Therefore, wavelet network has three layers: wavelet layer for computing wavelet function as activation functions of nodes (11), product layer for computing wavelet bases (12) and output layer for computing outputs (15).

The control system configuration is shown in Fig. 1.

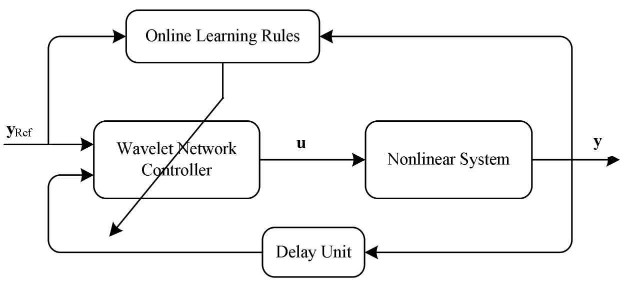

Fig. 1. The proposed control system configuration

III. Network Structure and Learning Algorithms

In this section, an algorithm is introduced to determine the network size; then learning rules are presented for adjusting network parameters.

A. Determination of Number of Wavelet Bases

Regarding the approximation function (15), wavelet network structure is clear and consists of three layers: two hidden layers (wavelet layer and product layer) and one output layer. Only the number of wavelet bases in the wavelet layer and product layer (K) is unknown. So, the wavelet network has n inputs, $n*K$ nodes in wavelet layer, K nodes in product layer and m nodes in output layer, where n is input dimension and m is plant output dimension. K must be defined off-line and then network parameters will be adjusted on-line. The following algorithm determines K. At first, $K=1$ is assumed. A sequence of random input (u) is applied to plant and plant output is obtained. The corresponding input (x) is applied to the network and all wavelet bases (ψ_j) are computed. If the network has proper wavelet bases, at least a base is would be fired and new bases would not be required. Otherwise, a wavelet would be added and translation and scaling factors for the new wavelet would be randomly considered. In other words, compact support property of wavelets, if the input is occurred in the dead zone of wavelet functions, no wavelet functions would be fired, the number of wavelet bases would not be appropriate and new wavelets would be required. This algorithm would continue until all the sequence input is applied. In this algorithm, the input and output vectors are normalized. For more details, see the illustrated algorithm in Fig. 2.

B. The Parameters Learning Rules

In Section A, the number of wavelet bases is defined and, in this section, adjusting rules of network parameters are introduced which include connection weights (W) for output layer and wavelet functions parameters (a, b). The error back-propagation (EBP) with momentum term algorithm is used for the online update of controller parameters.

Error is defined as

$$\mathbf{e} = \mathbf{y}(\mathbf{x}) - \mathbf{y}_d(\mathbf{x}) \tag{16}$$

where $\mathbf{y}(\mathbf{x})$ is plant output and $\mathbf{y}_d(\mathbf{x}) = \mathbf{y}_{ref}(\mathbf{x})$. The cost function E can be defined as

$$E = \frac{1}{2}\sum_{i=1}^{m} e_i^{\,2} \tag{17}$$

where $\mathbf{e} = [e_1 \ \cdots \ e_i \ \cdots \ e_m]^T$.

The adjusting rules for the network parameter are

$$\mathbf{W}(k+1) = \mathbf{W}(k) + \Delta\mathbf{W}(k), \tag{18}$$

$$\Delta\mathbf{W}(\mathrm{k}) = -\eta_{\mathrm{w}} \frac{\partial \mathrm{E}}{\partial \mathbf{W}(\mathrm{k})} + \alpha\Delta\mathbf{W}(\mathrm{k}-1) \tag{19}$$

where η_w is learning rate.

$$w_{ij}(k+1) = w_{ij}(k) + \Delta w_{ij}(k) \tag{20}$$

So, connection weights can be updated as

$$\Delta w_{ij}(k) = -\eta_w \frac{\partial E}{\partial w_{ij}} \tag{21}$$

$$\begin{aligned}\frac{\partial E}{\partial w_{ij}} &= \frac{\partial}{\partial w_{ij}}\left(\frac{1}{2}\sum_{k=1}^{m} e_k^{\,2}\right) = \frac{\partial}{\partial w_{ij}}\left(\frac{1}{2}e_i^{\,2}\right) \\ &= e_i.\frac{\partial}{\partial w_{ij}}(y_i) = e_i.\frac{\partial}{\partial w_{ij}}(\hat{y}_i)\end{aligned} \tag{22}$$

$$\frac{\partial E}{\partial w_{ij}} = e_i.\left(\sum_{j=1}^{N} w_{ij}\psi_j(\mathbf{x})\right) = e_i.\psi_j(\mathbf{x}) \tag{23}$$

$$\Delta w_{ij}(k) = -\eta_w.\ e_i.\ \psi_j(\mathbf{x}) \tag{24}$$

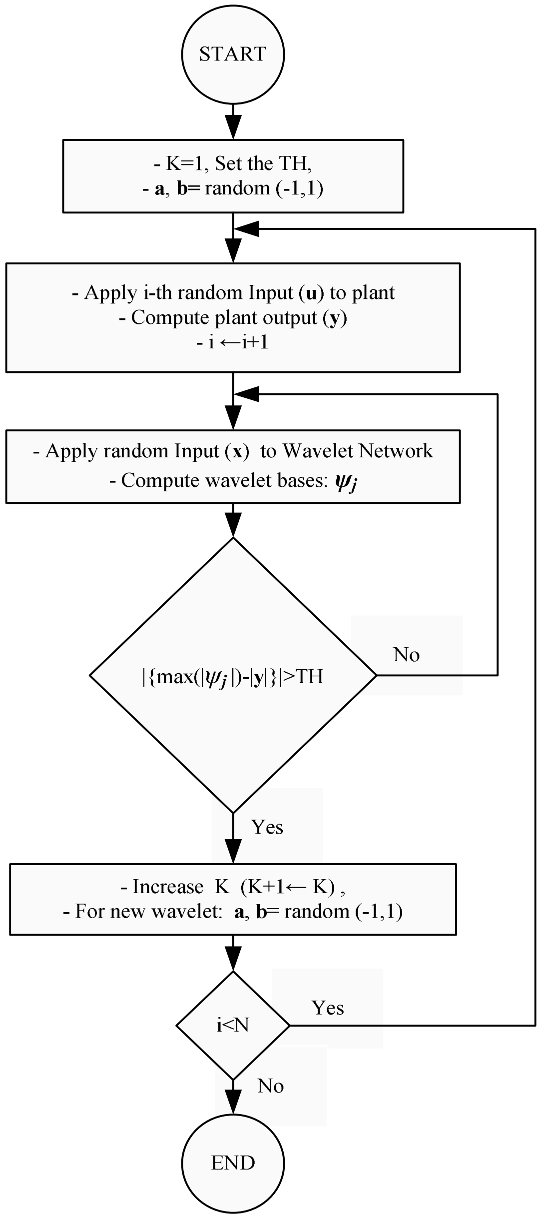

Fig. 2. The algorithm for determining the number of wavelet bases and network size

Similarities, the updating rules for a_{ij} and b_{ij} are as follows:

$$a_{ij}(k+1) = a_{ij}(k) + \Delta a_{ij}(k) \quad (25)$$

$$b_{ij}(k+1) = b_{ij}(k) + \Delta b_{ij}(k) \quad (26)$$

where

$$\Delta a_{ij}(k) = -\eta_a \frac{\partial E}{\partial a_{ij}(k)} + \alpha . \Delta a_{ij}(k-1) \quad (27)$$

$$\Delta b_{ij}(k) = -\eta_b \frac{\partial E}{\partial b_{ij}(k)} + \alpha . \Delta b_{ij}(k-1) \quad (28)$$

and

$$\begin{aligned}\Delta a_{ij}(k) &= -\eta_a \frac{\partial E}{\partial a_{ij}} = -\eta_a e_i . \frac{\partial}{\partial a_{ij}}(\hat{y}_i) \\ &= -\eta_a e_i w_{ij} (\frac{1}{2} a_{ij}^{-1/2} \varphi(x'_{ij}) + \\ &\quad a_{ij}^{1/2} x'_{ij} (x'^{2}_{ij} - 3) \exp(-\frac{1}{2} x'^{2}_{ij}))\end{aligned} \quad (29)$$

$$\begin{aligned}\Delta b_{ij}(k) &= -\eta_b \frac{\partial E}{\partial b_{ij}} = -\eta_a e_i . \frac{\partial}{\partial b_{ij}}(\hat{y}_i) \\ &= -\eta_b e_i w_{ij} a_{ij}^{1/2} x'_{ij} (3 - x'^{2}_{ij}) \exp(-\frac{1}{2} x'^{2}_{ij})\end{aligned} \quad (30)$$

where $x'_{ij} = a_{ij} x_i - b_{ij}$.

IV. ILLUSTRATIVE EXAMPLE

In this section, two examples are given which illustrate the advantages of the proposed controller over the neural network controller and adaptive wavelet controller (AWC) in [4].

A. Comparison of Wavelet based and Conventional Neural Network Controllers

Here, an example is given that is controlled by the proposed controller and conventional neural network. Then, these controllers are compared. This example demonstrates the benefits of wavelet network over conventional neural network for nonlinear control systems.

Consider a temperature control system as

$$\begin{aligned} y(k+1) = {} & A(T_s) y(k) + \\ & \frac{B(T_s)}{1 + e^{0.5 y(k) - \gamma}} u(k) + (1 - A(T_s)) y_0 \end{aligned} \quad (31)$$

where $A(T_s) = e^{-\alpha T_s}$ and $B(T_s) = (b/a)(1 - e^{-\alpha T_s})$. The parameters in this simulation are $a = 1.00151\, e^{-4}$, $b = 8.67973\, e^{-3}$, $\gamma = 40.0$, $T_s = 30$ and the initial temperature is 25 ^{o}C [16]. In this control system, input u is saturated as:

This open loop system is unstable. The control system is shown in Fig. 3. In the proposed control scheme, the plant was controlled successfully by implementing the off-line algorithm to determine the number of wavelet bases and on-line algorithms for adjusting the controller parameters. For demonstrating the ability and benefit of the wavelet controller, the plant is controlled by a four-layer neural network. In this controller, a bipolar sigmoid function is used as the activation function and error back-propagation algorithm and a momentum term is used for updating connection weights for both controllers. In order to compare the performance of these controllers, a criterion is used as follows:

$$J = \int_0^{T_f} |e(t)| dt \quad (32)$$

For discrete cases:

$$J = \sum_{k=1}^{K_f} \left| e(k) \right| \quad (33)$$

The performance of two controllers is compared and the results are presented in Figs. 4, 5 and Table I.

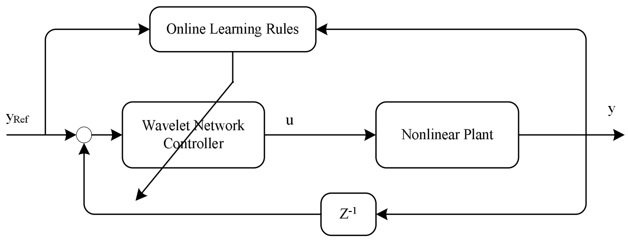

Fig. 3. The control scheme for temperature control problem

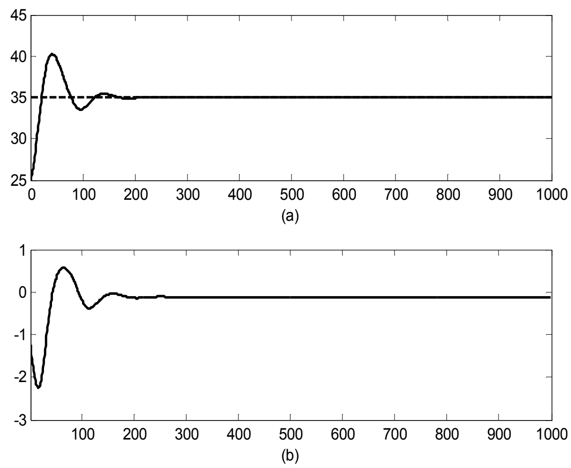

Fig. 4. Performance of the proposed controller: (a) Step response, (b) Control signal (TH=0.8, *K*=2, structure: 1-2-2-1)

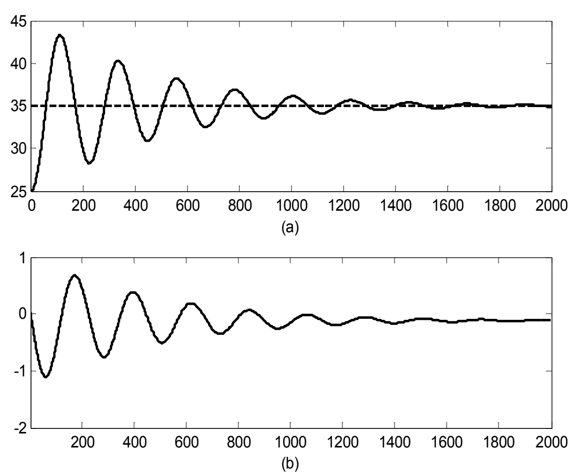

Fig. 5. Performance of the four-layer neural network controller: (a) Step response, (b) Control signal (NN structure: 2-5-5-1)

TABLE I
PERFORMANCE COMPARISON OF THE PROPOSED CONTROLLER WITH THE CONVENTIONAL NEURAL NETWORK

Controller Type	TH [a]	Controller structure	Number of Parameters	Control Performance (J)
Wavelet Network	0.8	1-2-2-1	6	15.4
	0.6	1-3-3-1	9	11.2
	0.4	1-7-7-1	21	8.5
Neural Network		2-5-5-1	40	115.3
		2-10-10-1	130	69.1
		2-15-15-1	270	47.6

In both controllers, all initial values are considered uniformly random in (-1, 1) only a ∈ (0,1).
The results are as mean 10 independent run.
[a] TH is threshold value in the algorithm Fig. 2.

A. Comparison of Wavelet based Nero-controller and AWC

In order to compare the proposed controller and the adaptive wavelet controller (AWC) proposed in [4], a nonlinear system which was formerly controlled by AWC is controlled by the wavelet-based controller. The nonlinear non-affine dynamic system is constructed as follows:

$$\dot{x} = (5 + x)\sin(x) +$$

(35) $$[2u + \cos(u)]\,\frac{1}{1 + e^{-x^2}} \quad (34)$$

where $x(0) = 0$. The tacking error profile by AWC is shown in Fig. 6 [4]. This nonlinear system is controlled by the proposed controller and the tracking error curve with unit reference is shown in Fig. 7.

V. CONCLUSION

In this study, a wavelet network was proposed for controlling nonlinear systems with hard non-linearity. In this control scheme, neural network disadvantages such as network structure selection as trial-and-error, lack of localization and slow convergence are avoided. This controller improves the control performance and reduces the calculation cost of updating parameters by reducing the number of adjustable parameters. An off-line algorithm was proposed to determine the size of network and an on-line updating rule based on the error back propagation with a momentum term was used for adjusting the controller parameters including the shape of wavelet bases and output layer weights.

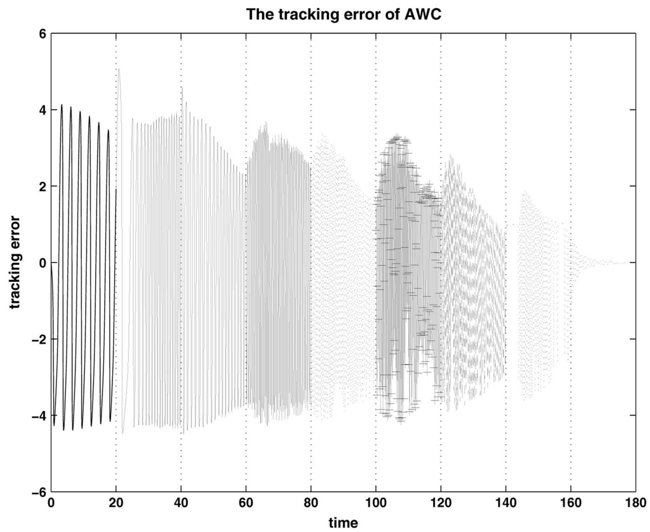

Fig. 6. Performance of the adaptive wavelet controller (AWC) for a step response: Tracking error [4]

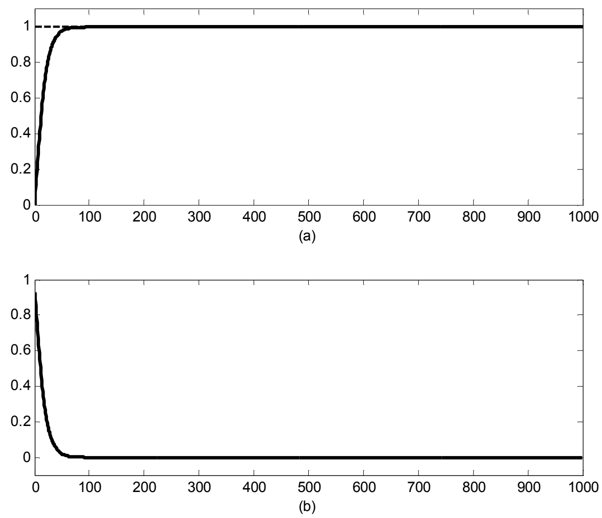

Fig. 7. Performance of the proposed controller (a) Step response, (b) Error profile. (K=5, structure: 2-5-5-1)

The suggested controller may be used to control unstable and MIMO systems. To illustrate the performance and effectiveness of the proposed controller, an unstable nonlinear system was controlled as an example. In addition, this system was controlled by a conventional neural network. The comparison of the control performance between the proposed controller and neural network controller in a computer simulation demonstrated that the wavelet-based controller could result in a better control performance, quicker convergence and fewer parameters than the NN controller for controlling nonlinear systems, especially for the systems with hard nonlinearity. In general, this wavelet network can be an appropriate candidate and can replace neural networks in various control schemes. In addition, another nonaffine nonlinear system was controlled by the proposed controller and simulation results illustrated the advantages of the controller over AWC which were proposed in [4].

REFERENCES

[1] A. Poznyak, W. Yu, E. Sanchez, and J. Perez, "Nonlinear adaptive trajectory tracking using dynamic neural network," *IEEE Trans. Neural Network*, vol. 10, no. 6, pp. 1402–1411, 1999.

[2] F. Sun, Z. Sun, and P. Woo, "Stable neural-network-based adaptive control for sampled-data nonlinear systems," *IEEE Trans. Neural Network*, vol. 9, no. 5, pp. 956–968, 1998.

[3] S. Z. Ge and C. Wang, "Adaptive neural control of uncertain MIMO nonlinear systems," *IEEE Trans. Neural Network*, vol. 15, no. 4, pp. 674–692, 2004.

[4] J. X. Xu and Y. Tan, "Nonlinear adaptive wavelet control using constructive wavelet networks," *IEEE Trans. on Neural Networks*, vol. 18, no. 1, pp. 115-127, 2007

[5] J. Farrell, "On performance evaluation in on-line approximation for control," *IEEE Trans. Neural Network*, vol. 9, no. 5, pp. 1001–1007, 1998.

[6] K. Hornik, "Multilayer feedforward networks are universal approximator," *Neural Network*, vol. 2, no. 5, pp. 359–366, 1989.

[7] Ch. Hsu, Ch. Lin and T. Lee, "Wavelet adaptive back stepping control for a class of nonlinear systems," *IEEE Trans. Neural networks*, vol. 17, no. 5, pp. 1175-1183, 2006.

[8] Ch. Sh. Chen and W. S. Lin, "Adaptive wavelet sliding mode control of uncertain multivariable nonlinear systems," *Proceedings of the American Control Conference, Denver, Colorado*, pp. 180-185, 2003

[9] A. Karami, H. R. Karimi, B. Moshiri, and P. J. Maralani, "Wavelet-based intelligent optimal control of robotic manipulators," *Proceedings of the European Control Conference, Kos, Greece*, pp. 3777-3782, July 2-5, 2007.

[10] R. J. Wai, "Robust control for nonlinear motor-mechanism coupling system using wavelet neural network," *IEEE Trans. Systems, Man and Cybernetics*, vol. 33, no. 3, pp. 489-497, 2003.

[11] S. Billings and H.Wei, "A new class of wavelet networks for nonlinear system identification," *IEEE Trans. Neural Network.*, vol. 16, no. 4, pp. 862–874, 2005.

[12] A. Karami, H. R. Karimi, B. Moshiri, and P. J. Maralani, "Wavelet-based adaptive collocation method for the resolution of nonlinear PDEs," *International Journal of Wavelets, Multiresolution and Information Processing (IJWMIP)* , vol. 5, no. 6 , pp. 957–973, 2007.

[13] A. Karami, H.R. Karimi, B. Moshiri and P.J. Maralani, "Investigation of interpolating wavelets theory for the resolution of PDEs," *International Journal of Contemporary Mathematical Sciences (IJCMS)*, vol. 3, no. 21, pp. 1017 - 1029, 2008.

[14] A. Karami, H. R. Karimi, B. Moshiri, and P. J. Maralani, "Intelligent optimal control of robatic manipulators using wavelets," *International Journal of Wavelets, Multiresolution and Information Processing (IJWMIP)*, vol. 6, no. 4, pp. 575–592, 2008.

[15] Y. Tan, X. Dang, F. Liang and C. Y. Su., "Dynamic wavelet neural network for nonlinear dynamic system identification," *Int. Conference Control Appl. in Proc. 9th IEEE (CCA*), Anchorage, Alaska, USA, pp. 214–219, 2000.

[16] Ch.T. Lin, Ch.F. Juang and, Ch.P. Li, "Temperature control with a neural fuzzy inference network," *IEEE Trans. Systems, Man and Cybernetics,* vol. 29, no. 3, pp. 440-451, 1999.