Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Nonlinear dynamic system control using wavelet neural network based on sampling theory

Ehsan Hossaini-asl
Department of Automation and Instrumentation
Petroleum University of technology
Tehran,Iran
e.hossaini.asl@gmail.com

Mehdi Shahbazian
Department of Automation and Instrumentation
Petroleum University of technology
Tehran,Iran
shahbazian_m@yahoo.com

***Abstract*— Wavelet neural network based on sampling theory has been found to have a good performance in function approximation. In this paper, this type of wavelet neural network is applied to modeling and control of a nonlinear dynamic system and some methods are employed to optimize the structure of wavelet neural network to prevent a large number of nodes. The direct inverse control technique is employed for investigating the ability of this network in control application. A variety of simulations are conducted for demonstrating the performance of the direct inverse control using wavelet neural network. The performance of this approach is compared with direct inverse control using multilayer perceptron neural network (MLP). Simulation results show that our proposed method reveals better stability and performance in reference tracking and control action.**

Keywords—*wavelet, wavelet neural network, sampling theory, direct inverse control, nonlinear dynamic control*

I. INTRODUCTION

Although the use of linear control methods have been prevalent in the chemical process industries, they have their limitations especially when dealing with nonlinear plants in a wide operating region as commonly found in these industries. Many economically important units, such as reactors and high-purity distillation columns, can be very nonlinear and very difficult to control adequately with linear controllers [9]. Complex steady state and dynamic behaviors create challenges that are tough for traditional linear controllers to handle.

Analysis of nonlinear dynamics and design of globally stabilizing controllers for nonlinear systems are addressed in many textbooks. A universal requirement of the behavior of the closed loop system is that it should be stable. Unfortunately the stability issue is most often a complicated matter for nonlinear systems.

When neural networks originally were proposed for controlling unknown nonlinear systems, one of the first methods being reported was on training a network to act as the inverse of the system and use this as a controller.

The major characteristic of the direct inverse control is that they are intuitively simple and also simple to implement. One of the main problems of this technique is that they don't work for systems with an unstable inverse, which often occur when using a high sampling frequency. This paper try to solve this problem using wavelet neural network based on sampling theory.

As described in [1, 2], the wavelet neural network based on sampling theory, acts as a good approximator for any target function with properties of global convergence and avoiding overfitting. In this paper, this new wavelet neural network is applied as a direct inverse controller for investigation of its performance in identification and control of a nonlinear dynamic system. We compare its performance with commonly used MLP with Levenberg-Marquardt training.

This work is divided in 4 sections. The first section briefly reviews the theory of multiresolution analysis and wavelet neural network based on sampling theory. The second section describes the procedure of designing the direct inverse control (DIC) based on wavelet neural network. In the third section, some methods are suggested for decreasing the number of nodes and improving the structure of wavelet network. In the forth section, simulation proves the performance of this new DIC controller for reference tracking and appropriate control action in comparison with MLP.

II. THEORETICAL DESCRIPTION

A. Review of Multiresolution Analysis

In neural network learning, in order to take the full advantage of orthonormality of basis function, with localized learning, we need a set of basis functions which are local and orthogonal.

Wavelets are new family of localized basis functions that have found many applications in large areas of science and engineering [4, 5]. Wavelets are universal approximator which can be used to approximate any arbitrary multidimensional nonlinear function. They have many powerful mathematical properties such as orthonormality, locality in time and frequency domains, different degrees of smoothness, fast implementations, and effective compact support. Wavelets are introduced in a multiresolution framework developed by Mallat [5]. We focus on the wavelet networks constructed from a multiresolution analysis (MRA) [5]. Consider a function f(x) in $L^2(\mathbb{R})$, where $L^2(\mathbb{R})$ denotes the vector space of all measurable, square integrable one dimensional functions. In

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

addition, assume V_j be the vector space containing all possible approximations of $f(x)$ at the resolution m. Then, the ladder of spaces V_j, $j \in \mathbb{Z}$ represents the successive resolution levels for $f(x)$. The properties of these spaces are as follows:

1. (Nested) $V_j \subseteq V_{j+1}, \forall j \in \mathbb{Z}$ (1)
2. $f(x) \in V_j \Leftrightarrow f(x-k) \in V_j, \forall (j,k) \in \mathbb{Z}^2$ (2)
3. (Density) $\text{close}_{L^2}\left(\bigcup_{j\in\mathbb{Z}} V_j\right) = L^2(\mathbb{R})$ (3)
4. (Separation) $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$ (4)
5. (Scaling) the function $f(x)$ belongs to V_j if and only if the function $f(2^{-j}x)$ belongs to V_0
6. (Basis) There exists a function $\phi \in L^2$ (called a scaling function or a father wavelet), with $\phi_{j,k} = 2^{j/2}\phi(2^j x - k)$, such that $\{\phi_{0,k}; k \in \mathbb{Z}\}$ is a basis for V_0.

The function ϕ is called a scaling function of the multiresolution analysis (MRA). A family of scaling functions of the MRA is expressed as:

$$\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k), \; j,k \in \mathbb{Z} \quad (5)$$

Where 2^j and k correspond to the dilation and translation factors of the scaling function respectively while $2^{j/2}$ is an energy normalisation factor. Let W_j be the orthogonal complement of V_j to V_{j+1} ($V_j \oplus W_j = V_{j+1}$). Then the orthonormal basis functions corresponding to W_j's named wavelets and denoted by $\psi_{j,k}$'s can be easily obtained from $\phi_{j,k}$'s [3]. A family of wavelets may be represented as:

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k), \; j,k \in \mathbb{Z} \quad (6)$$

With 2^j, k and $2^{j/2}$ being the dilation, translation, and normalisation factor of the wavelets, respectively. Next $L^2(\mathbb{R})$ can be expressed as:

$$L^2(\mathbb{R}) = \bigcup_{j\in\mathbb{Z}} V_j = \cdots W_{-1} \oplus W_0 \oplus W_1 \ldots = \bigoplus_{j\in\mathbb{Z}} W_j \quad (7)$$

Where $W_j \perp W_m$ for $j \neq m$. Fig.1 illustrates the relation between V_j and W_j spaces in MRA. Equation (7) indicates that the wavelet basis generates decomposition of the L^2 space. It shows that any L^2 function is uniformly approximated using a wavelet series:

$$f(x) = \sum_{j=-\infty}^{j=+\infty} \sum_{k=-\infty}^{k=+\infty} d_{j,k}\psi_{j,k}(x) \quad (8)$$

If we start from the approximation of the function at resolution $j = 0$, then:

$$f(x) = f_0(x) + \sum_{j=0}^{j=+\infty} \sum_{k=-\infty}^{k=+\infty} d_{j,k}\psi_{j,k}(x) \quad (9)$$

Where

$$f_0(x) = \sum_{k=-\infty}^{k=+\infty} a_{0,k}\phi_{0,k}(x) \quad (10)$$

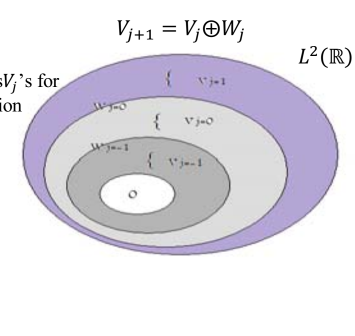

Figure 1. Embedded spaces V_j's for multiresolution representation of $L^2(\mathbb{R})$.

We can conclude that any function $f(x) \in L^2$ can be written as a unique linear combination of wavelets of different resolutions. This means that $f(x) = \cdots + g_{-1}(x) + g_0(x) + g_1(x) + \cdots$, where $g_j(x) \in W_j$ is unique. Since $V_j = W_j + W_{j-1} + \cdots$ and spaces V_j can be generated by the scaling function $\phi(x) \in L^2$, there exists:

$$f_{ne}(x) = \sum_{k=-\infty}^{\infty} c_{j,k}\phi(2^j x - k) = \sum_{k=-\infty}^{\infty} c_{j,k}\phi_{j,k} \quad (11)$$

Such that $\|f(x) - f_{ne}(x)\| \longrightarrow 0$ when $j \to \infty$. In fact (11) is just the presentation of wavelet networks with three layers. In an impact interval of interest, (11) can be written as:

$$f_{ne}(x) = \sum_{k=I_0}^{I_1} c_{j,k}\phi_{j,k}(x) \quad (12)$$

Where $\phi_{j,k} = \phi(2^j x - k)$. A wavelet network is realized by taking $c_{j,k}$'s as the output weights, 2^j's as the input weights and $\phi(x-k)$ as the activation function.

Variety of approaches have been proposed for determining wavelet network parameters such as input weights 2^j and also output weights $c_{j,k}$'s. Here we use the approach based on sampling theory proposed by Zhang [2] for specifying appropriate resolution j.

B. Wavelet Neural Network Based on Sampling Theory

The wavelet neural network based on sampling theory first proposed by Zhang [2] and found to be a good approximator of band limited function with ability of global convergence and avoiding local minimum. But this wavelet neural network appeared to have some disabilities in avoiding of overfitting in training of non uniform noisy data. Then this neural network was modified by Hossaini asl [1] for this type of training data and become a good tools for modeling of nonlinear dynamic systems. In this part, this modified wavelet neural network is briefly described. For more details on this theory, please refer to [1, 2].

Suppose that the band limited function $f(t)$ is sampled with sample time T. Then the function can be written as follows:

$$f(nT) = f(nT)\delta(x - nT) \quad (13)$$

The Fourier transform of the discrete signal obtained by sampling f at intervals T is:

$$\hat{f}_d(w) = \frac{1}{T}\sum_{k=-\infty}^{+\infty} \hat{f}\left(w - \frac{2k\pi}{T}\right) \quad (14)$$

If the support of $\hat{f}$ is included in $[-\pi/T, \pi/T]$ then

$$f(x) = \sum_{k=-\infty}^{+\infty} f(kT)((x - kT)/T) \quad (15)$$

On the other hand, the frequency band of wavelet network that described in previous section is obtained as follows:

$$\int_{-\infty}^{\infty} \left|\tilde{f}_{ne}(w)\right|^2 dw \leq \int_{-2^j b_w}^{2^j b_w} \left|\tilde{f}_{ne}(w)\right|^2 dw + 2^{-j}\varepsilon \sum_k \left|c_{j,k}\right|^2 \quad (16)$$

So the energy of wavelet network is concentrated well in the following frequency band:

$$[-2^j b_w, 2^j b_w] \quad (17)$$

The parameter b_w only depends on scaling function. Equation (17) means that the frequency band of wavelet network can be controlled by input weights.

Suppose $\{kT, f_s(kT)\}$'s are training data with $\sum_k |f_s(kT)|^2 < +\infty$ then, by the sampling theorem, there exists a unique function $f(x)$ to interpolate all training data. On the other hand, a wavelet network is a function in $L^2(\mathbb{R})$ so a wavelet network represents a function in P_T if its Fourier transform has a support included in $[-\pi/T, \pi/T]$.

This means that the network $f_{ne}(x)$ whose Fourier transform has a support in $[-\pi/T, \pi/T]$ is complex enough to recover a band-limited function. Therefore according to (16), the input weights can be calculated using following equation:

$$2^j = \pi/(b_w \times T) \tag{18}$$

For constructing the structure of wavelet network, the property of energy concentration of wavelet in time domain should be employed. In wavelet network, the k^{th} node has the following input-output function:

$$S_{out} = \phi(2^j S_{in} - k) \tag{19}$$

Where S_{in} is the input, 2^j's are the input weights, k is the k^{th} threshold and $\phi(\circ)$ is the scaling function. If the support of scaling function is limited to $[0, N_\phi]$, then the k^{th} node of network has the following support:

$$\left[2^{-j}k, 2^{-j}(N_\phi + k)\right] \tag{20}$$

Assume the domain of interest for estimation of function is the interval $[a, b]$, then the translations are found as follows:

$$2^j a - N_\phi \le k \le 2^j b \tag{21}$$

The output weights are found based on minimization of the following cost function:

$$J(f_s, f_{ne}) = \sum_{i=1}^{N} |f_s(x_i) - f_{ne}(x_i)|^2 \tag{22}$$

Where $\{x_i, f_s(x_i)\}$'s are samples and $f_{ne}(x_i)$'s are output of approximator. Without any additive term, this cost function is widely used in the training of networks because of convenient implementation.

Three commonly used methods are direct solution method, iterative method and inner product method. In [1], the iterative method is employed for training the output weights. In this method the output weights can be calculated as follows:

$$E^{(k+1)} = F_s - \phi_{m\times n} C^{(k+1)} \tag{23}$$

$$C^{(k+1)} = C^{(k)} + A_f E^{(k)} \tag{24}$$

The column vector $E^{(k)}$ denotes the error of interpolation by the wavelet network at k^{th} iteration, the column vector $C^{(k)}$ represents the output weights at k^{th} iteration and the matrix A_f is the feedback matrix. The values of elements in the feedback matrix indicate that how much the errors in each data point would affect on output weights. The $\phi_{m\times n}$ matrix is:

$$\phi_{m\times n} = \begin{bmatrix} \phi_{J,I_0}(x_1) & \cdots & \phi_{J,k}(x_1) & \cdots & \phi_{J,I_1}(x_m) \\ \vdots & \cdots & & \cdots & \vdots \\ \phi_{J,I_0}(x_m) & \cdots & \phi_{J,k}(x_m) & \cdots & \phi_{J,I_1}(x_m) \end{bmatrix} \tag{25}$$

Where the subscript $I_0 = \left[2^J a - N_\phi\right]$ and the subscript $I_1 = [2^J b]$, which denote respectively, the minimum and maximum of translation k obtained from (21).

In [1], an intuitive approach for finding the appropriate feedback matrix was proposed. In this method the feedback matrix is constructed based on the ϕ matrix. This method uses the receptive field of each node or scaling function in wavelet network. For detail information about calculating the appropriate feedback matrix, refer to [1].

For uniform sampled data, training wavelet neural network based on sampling theory shows quite acceptable results. However, for non uniform data, the algorithm encounters severe problems such as high overfitting error and deviation of estimated function from the actual target function. One of the reasons is that in the training of non uniform noisy data, the number of sub wavelet network that data are uniform within it, are intensively large because of equation that was proposed in [2]. To overcome this problem, a modified equation for calculating the number of sub wavelet network is proposed in [1]. By using this method, the number of sub wavelet network is decreased.

Another technique that is suggested in [1], is the applying of wavelet thresholding into the training procedure. This technique causes the removing of nodes that represent the noise in the wavelet neural network.

The final technique that is suggested in [1], is applying early stopping in the training procedure. This technique greatly improves the training procedure. By using this technique the profile of convergence of algorithm is monitored and therefore the stopping time is decided.

III. DESIGNING OF DIC CONTROLLER

In this section, the method of designing a DIC controller using the wavelet neural network based on sampling theory for nonlinear dynamic system is described. This theory of control is described in [3, 7, 8].

The scheme of direct inverse control with neural network is characterized by the controller being a network trained as the inverse of the system. "Inverse" was understood in the sense that the transfer function for the closed-loop system, consisting of controller and system equaled the time delay of the system. Assuming that the nonlinear dynamic system can be described by:

$$y(t+1) = g[y(t), \dots, y(t-n+1), u(t), \dots, u(t-m)] \tag{26}$$

The desired network is then the one that isolate the most recent control input, $u(t)$ as shown below:

$$\hat{u}(t) = \hat{g}^{-1}[y(t+1), y(t), \dots, y(t-n+1), u(t), \dots, u(t-m)] \tag{27}$$

Assuming such a network has somehow been obtained, it can be used for controlling the system by substituting the output at time $t+1$ by the desired output, the reference, $r(t+1)$. If the network represents the exact inverse, the control input produced by it will thus drive the system output at time $t+1$ to $r(t+1)$.

There are two methods for establishing the inverse model, namely the general training and the on-line method, called the specialized training.

Since this is the first time that the wavelet network based on sampling theory is employed as a controller of nonlinear dynamic system, we use off-line method for reasonable assessment of its performance. The most straightforward way of training a network as the inverse of a system is to approach the problem as a system identification problem as described below:

a) Identification experiment is performed (generate training data).

b) Wavelet neural network architecture is selected according to inverse of system.

c) The wavelet neural network is trained offline.

The wavelet neural network is then trained to minimize the criterion:

$$J(\theta, Z^N) = \frac{1}{2N}\sum_{t=1}^{N}[u(t) - \hat{u}(t,\theta)]^2 \tag{28}$$

As described in Part B of Section II, the iterative method is employed for training of wavelet neural network that leads to minimization of the objective function.

IV. Improving the Structure of DIC Based on Wavelet Neural Network

One of the main problems that may encounter in training of nonlinear dynamic system using wavelet neural network based on sampling theory, is generation of large number of nodes in each sub wavelet network. The main reason of this problem is the variation of density of data in each multi dimensional sub wavelet network. The sub networks that contains more dense training data, affect the other sub networks in case of resolution of scaling function. In this paper, three methods are suggested for solution of this problem. These methods are explained below:

A. Decreasing Nodes using Internal Thresholding

In this method, we first cluster the training data into sub wavelet network and determine the scales and threshold of each sub network. Then by looking at the number of data and number of nodes in each sub wavelet network, a balance between these parameters could be set. In this approach, the balance is set by defining a threshold level and applying it on ϕ matrix of each sub wavelet network.

B. Decreasing nodes using feedback matrix

After construction of the feedback matrix, there will be some rows that all of its elements are zeros. This means that none of training data contribute in training of the corresponding node. By this explanation, we could neglect that node in estimation of wavelet neural network.

C. Decreasing nodes using wavelet thresholding

After computing wavelet coefficients, the number of nodes is decreased using wavelet coefficient thresholding. This method is explained in [1] and includes two kind of hard thresholding and soft thresholding. In this paper, the hard thresholding is used due to its simplicity.

In the next section, these methods are applied to training procedure and their effects on estimation of wavelet network will be investigated.

V. Simulation

In this section, the performance of wavelet neural network as a direct inverse controller is investigated and also compared to MLP with Levenberg-Marquardt training algorithm. A system that is chosen to be controlled is the model of spring that is an open-loop stable system with smooth nonlinearities [3]. The challenging part of this system is that the inverse of this system is also unstable. This system is described as shown below:

$$\ddot{y}(t) + \dot{y}(t) + y(t) + y^3(t) = u(t) \tag{29}$$

For finding the architecture of the DIC controller, first the structure of the system should be obtained. For this objective, the continuous model of the system should be discretized as shown below:

$$y(t) - 0.7859 * y(t-1) + 0.3679 * y(t-2) = 0.3403 * u(t-1) + 0.2417 * u(t-2) \tag{30}$$

For approaching this model, the nonlinear term $y^3(t)$ in continuous model is neglected. Therefore the structure of system can be shown as Fig.2. According to the structure of the system and also description in section III, the difference equation model of the DIC controller is calculated as follows:

$$y(t+1) = g[y(t), y(t-1), u(t), u(t-1)] \tag{31}$$

$$\hat{u}(t) = \hat{g}^{-1}[y(t+1), y(t), y(t-1), u(t-1)] \tag{32}$$

According to (32), the architecture of DIC controller can be depicted as Fig.3. According to Fig.3, the structure of wavelet neural network is found. The next step is to execute a suitable experiment to generate training data that could represent all the states of the nonlinear dynamic system. After a lot of experiments, the best input signals for system that could excites approximately all the dynamic of system is found as multifrequency sinusoidal signal and uniform random signal. In this paper, the uniform random signal is employed as the identification signal for system. The input signal and the response of the system are shown in Fig.4. In this experiment, the number of training data is chosen as 400 with amplitude value between [-2, 2]. As clearly depicted in Fig.4, the amplitude value of the system response is between [-1, 1]. The training procedure is first executed without using the three methods that were suggested in section IV. The reason is for demonstrating the performance of these methods in training of the network. During the training procedure, after each iteration of training, the wavelet neural network is cross validated using 3 different response of the actual system to three signals. The results of training are shown in Fig.5 and Fig.6 as shown below:

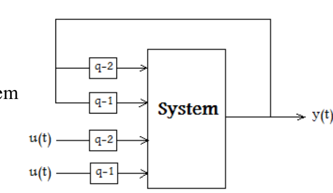

Figure 2. Discretized model of the system

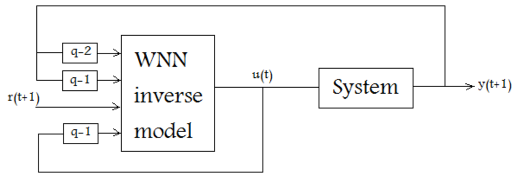

Figure 3. Architecture of DIC controller

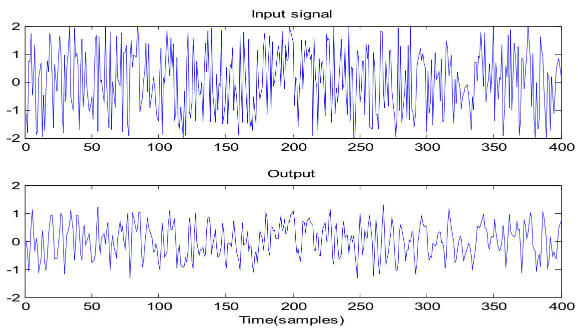

Figure 4. Training data set

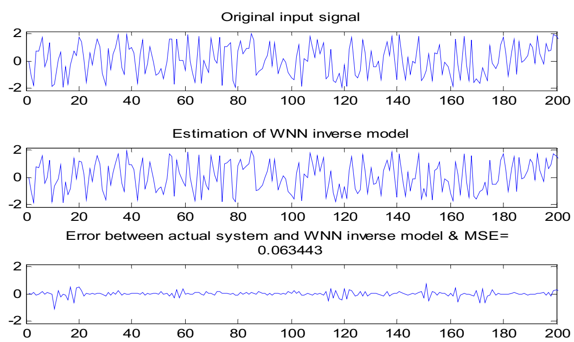

Figure 5. Comparison between actual system and WNN inverse model

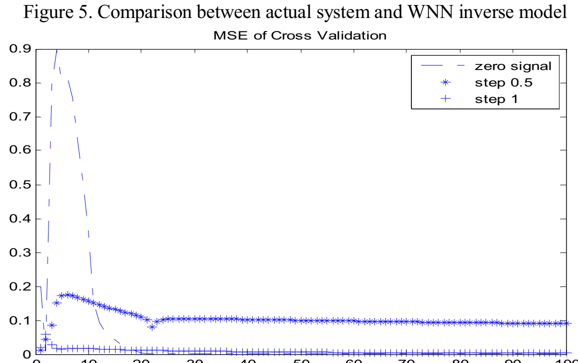

Figure 6. Comparison of MSE of cross validation between actual system and WNN inverse model

TABLE I. STATISTICS OF TRAINING

Name	Value
Root mean square error	0.25188
Mean absolute error	0.1423
Maximum absolute error	2.4129
Number of nodes	56400
Number of iterations	100

As mentioned before, the WNN inverse model is validated with the response of the actual system to three input signals. These input signals are zero signal, step signal with amplitude value of 0.5 and also step with amplitude value of 1. Fig.6 depicts the profile of variation of mean square error between WNN inverse model estimation and the actual system response of these three signals during training procedure. As it is clear, the MSE for three signals are decreased during the training procedure. This wavelet neural network has 4 inputs and one output. It also contains two sub wavelet networks for $y(t-1)$, two sub wavelet networks for $y(t)$, two sub wavelet networks for $r(t+1)$ and also four sub wavelet networks for $u(t-1)$. The results of training are shown in Table I as below:

In this training, the methods that were proposed in section IV are not used and the results of training show that the number of nodes that is generated exceeds 56400. Therefore, the three methods are added to the training procedure. After training, the total number of nodes decreases from 56400 to 366 nodes. More detailed information about these methods is shown in Table II. According to this Table, by using these methods, the number of nodes is gradually decreased while there are just small increases in statistical errors while the structure of the network is greatly reduced.

Now the reference signal should be designed according to training data. For determining of the domain that the reference can change within it, the density of training data should be considered. One can interpret that the reference to WNN is better to vary in the domain of [-1,1] and the control action from WNN controller varies in the domain of [-2, 2]. Now the WNN inverse model is ready for taking control the system. In this simulation, the WNN inverse model of third type in Table II (method I+II+III) is used for comparison with MLP with Levenberg-Marquardt learning algorithm. This neural network has two layers of nodes that use "tanh" as the activation function. This network is trained using Levenberg-Marquardt method in 500 iterations. Two experiments with two different references are applied to wavelet neural network and MLP. The results of control are shown in Fig.7 to Fig.10 as below:

TABLE II. STATISTIC ERRORS OF THREE METHODS

Name	Method I	Method I+II	Method III
Root mean square error	0.27251	0.30252	0.31277
Mean absolute error	0.16052	0.20126	0.20493
Maximum absolute error	1.9703	1.9852	1.9867
Number of nodes	2032	458	366

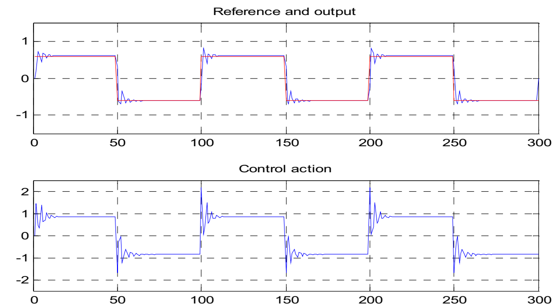

Figure 7. System response to wavelet neural network as inverse controller

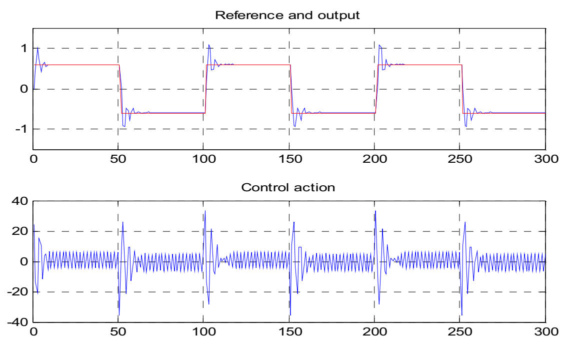

Figure 8. System response to MLP as inverse controller

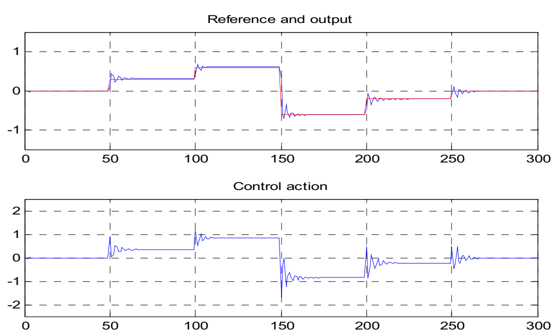

Figure 9. System response to wavelet neural network as inverse controller

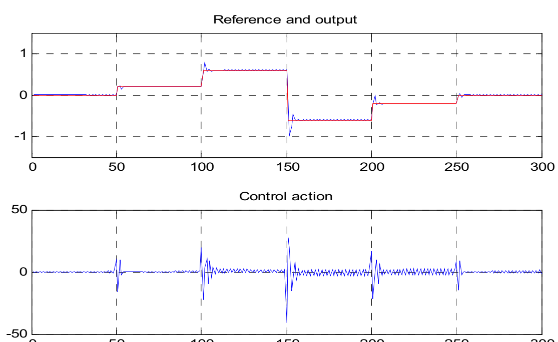

Figure 10. System response to two layer Levenberg-Marquardt neural network as inverse controller

According to Fig.7 to Fig.10, MLP controller has strong control action with high fluctuation resulting in larger overshoot. This is a quite common outcome when using this type of controller. This problem could be overcome by using a low-pass filter for cancellation of zeros that is described in [3]. On the other hand, the WNN controller has a reasonable control action without any oscillation and also similar to reference signal that results in more appropriate overshoot. It can be inferred from this comparison that the WNN inverse model is robustly stable system while MLP controller is poorly stable (because of existence of zero near the unit circle in transfer function of the discretized model) in dealing with nonlinear dynamic systems. Since the frequency band of WNN controller is controlled by its input weights that are determined using the sampling theory, this property of wavelet neural network based of sampling theory prevents the WNN controller to result in an unstable or poorly stable system.

VI. CONCLUSION

This work investigates the performance of wavelet neural network based on sampling theory in modeling and control of nonlinear dynamic systems and compares it with MLP neural network. For this reason, the technique of direct inverse control of an inversely unstable nonlinear system is employed. First some methods are suggested to prevent the generation of high number of nodes in wavelet neural network. The simulation results prove that the poor stability of the closed loop system that is one of the main problems in using the technique of DIC has been overcame. This is because the frequency band of wavelet neural network can be effectively controlled using sampling theory. Therefore we recommend the application of this powerful property of wavelet neural network based on sampling theory in implementation of other nonlinear dynamic control techniques.

REFERENCES

[1] E. Hossaini asl, M. Shahbazian, K. Salahshoor, "Wavelet neural network based on sampling theory for non uniform noisy data". Selected Papers from the WSEAS Conferences in Spain, pp.51-56, 2008.

[2] Z. Zhang, "Learning algorithm of wavelet network based on sampling theory." Neuro computing, Vol.71, Issue 1-3, pp.244-269, 2007.

[3] M. Norgaard, O.H. Ravn, N.K. Poulsen, L.K. Hansen, "Neural Networks for modeling and control of Dynamic Systems" (Springer, 2003).

[4] I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets", Com. Pure Appl. Math., Vol.41, Issue 7, pp.909-996, 1988.

[5] S.G. Mallat, "A theory for multi-resolution signal decomposition: the wavelet representation", IEEE Trans. Pattern Analysis Mach. Int., Vol.11, No.7, pp.674-693, 1989.

[6] B.R. Bakshi, G. Stephanopoulos, "Wave-net: a multiresolution hierarchical neural network with localized learning", AIChE Journal, Vol.39, No.1, pp.57-81, 1993.

[7] M. Norgaard, "System Identification and Control with Neural Networks". PhD thesis, Department of Automation, Technical University of Denmark, Kgs.Lyngby, Denmark.

[8] O. Sorensen, "Neural Networks in Control Applications", PhD thesis, Aalborg University, Department of Control Engineering, Aalborg, Denmark, 1994.

[9] M.A. Hussain, L.S. Kershenbaum, "Implementation of an inverse-model-based control strategy using neural networks on a partially simulated exothermic reactor" Chemical Engineering Research and Design, Vol.78, pp.299-311, 2000.