Nonlinear Dynamics and Systems Theory, **12** (3) (2012) 279–288

Wavelet Neural Network Based Adaptive Tracking Control for a Class of Uncertain Nonlinear Systems Using Reinforcement Learning

M. Sharma [1*] and A. Verma [2]

[1] *MediCaps Institute of Tech. and Mgmt., Indore, India*
[2] *Institute of Engineering and Technology, Indore, India*

Received: May 6, 2011; Revised: March 5, 2012

Abstract: In this paper an adaptive critic based wavelet neural network (WNN) based tracking control strategy for a class of uncertain systems in continous time is proposed. The adaptive critic WNN controller comprises two WNNs: critic WNN and action WNN. The critic WNN is approximating the strategic utility function, whereas the action WNN is minimizing both the strategic utility function and the unknown nonlinear dynamic estimation errors. Adaptation laws are developed for the online tuning of wavelets parameters. The uniformly ultimate boundedness of the closed-loop tracking error is verified even in the presence of WNN approximation errors and bounded unknown disturbances, using the Lyapunov approach and with novel weight updating rules. Finally some simulations are performed to verify the effectiveness and performance of the theoretical development.

Keywords: *wavelet neural networks; optimal control; adaptive control; reinforcement learning; Lyapunov functional.*

Mathematics Subject Classification (2010): 49J15, 93C40.

* Corresponding author: mailto:er.mann24@gmail.com

1 Introduction

Typical control strategies are based on a mathematical model that captures as much information as possible about the plant to be controlled. The ultimate objective is not to design the best controller for the plant model, but for the real time plant. This objective is addressed by robust control theory by including in the model a set of uncertainties. Robust control techniques are applied to the plant model, augmented with uncertainties and candidate controllers, to analyze the stability of the actual system. This is a powerful tool for practical controller design, but designing a controller that remains stable in the presence of uncertainties limits the aggressiveness of the resulting controller, and can result in suboptimal control performance [1, 2].

In this paper, the robust control techniques are combined with a reinforcement learning algorithm to improve the performance of robust controller while maintaining the stability of the system. Reinforcement learning is a class of algorithms for solving multi-step, sequential decision problems by finding a policy for choosing sequences of actions that optimize the sum of some performance criterion over time [3]– [7].

In recent years, learning-based control methodology using Neural networks (NNs) has become an alternative to adaptive control since NNs are considered as general tools for modeling nonlinear systems [17]. Work on adaptive NN control using the universal NN approximation property is now pursued by several groups of researchers. By using neural network (NN) as an approximation tool, the assumptions on linear parameterized nonlinearities in adaptive controller designing aspects have greatly been relaxed. It also broadens the class of the uncertain nonlinear systems which can be effectively dealt by adaptive controllers. However there are some difficulties associated with NN based controller. The basis functions are generally not orthogonal or redundant; i.e., the network representation is not unique and is probably not the most efficient one and the convergence of neural networks may not be guaranteed. Also the training procedure for NN may be trapped in some local minima depending on the initial settings. Wavelet neural networks are the modified form of the NN having the properties of space and frequency localization properties leading to a superior learning capabilities and fast convergence. Thus WNN based control systems can achieve better control performance than NN based control systems [8]– [10].

Adaptive actor-critic WNN-based control has emerged as a promising WNN approach due to its potential to find approximate solutions to dynamic programming [11]– [14]. In the actor-critic WNN based control a long-term system performance measure can be optimized, in contrast to the short-term performance measure used in classical adaptive and WNN control. While the role of the actor is to select actions, the role of the critic is to evaluate the performance of the actor. This evaluation is used to provide the actor with a signal that allows it to improve its performance, typically by updating its parameters along an estimate of the gradient of some measure of performance, with respect to the actor's parameters. The critic WNN approximates a certain strategics utility function that is similar to a standard Bellman equation, which is taken as the long-term performance measure of the system. The weights of action WNN are tuned online by both the critic WNN signal and the filtered tracking error. It minimizes the strategic utility function and uncertain system dynamic estimation errors so that the optimal control signal can be generated [3].

This paper deals with the designing of reinforcement learning WNN based adaptive tracking controller for a class of uncertain nonlinear systems. WNN are used for approximating the system uncertainty as well as to optimize the performance of the control strategy.

The paper is organized as follows. Section 2 deals with the system preliminaries, system description is given in Section 3. WNN based controller designing aspects are discussed in Section 4. Section 5 describes the tuning algorithm for actor-critic wavelets. The stability analysis of the proposed control scheme is given in Section 6. Effectiveness of the proposed strategy is illustrated through an example in Section 7 while Section 8 concludes the paper.

2 System Preliminaries

2.1 Wavelet neural network

Wavelet network is a type of building block for function approximation. The building block is obtained by translating and dilating the mother wavelet function. Corresponding to certain countable family ofa_m and $\mathrm{b_n}$, wavelet function can be expressed as

$$\left\{a_m^{-d/2}\psi\left(\frac{x-b_n}{a_m}\right) : m \in Z, n \in Z^d\right\}. \tag{1}$$

Consider

$$a_m = a_0^m, b_n = na_0^{-m}b_0, m \in Z, n \in Z^d. \tag{2}$$

The wavelet in (1) can be expressed as

$$\psi_{mn} = \left\{a_0^{-md/2}\psi\left(a_0^{-m}x - nb_0\right) : m \in Z, n \in Z^d\right\}, \tag{3}$$

where the scalar parameters a_0 and $\mathrm{b_0}$ define the step size of dilation and translation discretizations (typically a_0=2 and $\mathrm{b_0}$=1) and $x = [x_1, x_2, ..., x_n]^T \in R^n$ is the input vector.

Output of an n dimensional WNN with m wavelet nodes is

$$f = \sum_{m\in Z}\sum_{n\in Z^d}\alpha_{mn}\psi_{mn}. \tag{4}$$

3 System Description

Consider a nonlinear system of the form

$$\begin{aligned}
\dot{x}_1 &= x_2,\\
\dot{x}_2 &= x_3,\\
&\vdots\\
\dot{x}_n &= f(x) + u,\\
y &= x_1,
\end{aligned} \tag{5}$$

where $x = [x_1, x_2, ..., x_n]^T, u, y$ are state variable, control input and output respectively. $f = [f_1, f_2, ..., f_n]^T : \Re^{n+1} \to \Re^n$ are smooth unknown, nonlinear functions of state variables.

Rewriting the system (5) as

$$\begin{aligned} \dot{x} &= Ax + B(f(x) + u(t)), \\ y &= Cx, \end{aligned} \tag{6}$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

The objective is to formulate a state feedback control law to achieve the desired tracking performance. The control law is formulated using the transformed system (6). Let $\bar{y}_d = [y_d, \dot{y}_d, \dots, \overset{n-1}{y_d}]^T$ be the vector of desired tracking trajectory. The following assumptions are taken for the systems under consideration.

Assumption 3.1 Desired trajectory $y_d\,(t)$ is assumed to be smooth, continuous C^n and available for measurement.

4 Basic Controller Design Using Filtered Tracking Error

Define the state tracking error vector $\hat{e}(t)$ as

$$\hat{e}(t) = \hat{x}(t) - \bar{y}_d(t). \tag{7}$$

The filter tracking error is defined as

$$\hat{r} = K\hat{e},$$

where $K = [k_1, k_2, \dots k_{n-1}, 1]$ is an appropriately chosen coefficient vector such that $\hat{e} \to 0$ exponentially as $\Re \to 0$. Differentiating it along the trajectory of the systems, we get

$$\dot{\hat{r}} = K_e\hat{e} + f(x) + u - \overset{n}{y_d}\,. \tag{8}$$

Applying the feedback linearization method, the control law is defined as

$$u = (\overset{n}{y_d} - \hat{f}(x) - K_e\hat{e} - \hat{r}), \tag{9}$$

where $K_e = [0, k_1, k_2, \dots, k_{n-1}]$. Substituting (9) in (8),

$$\dot{\hat{r}} = -\hat{r} + \tilde{f}(x). \tag{10}$$

Stability of the system (6) with the proposed control strategy will be analyzed in the subsequent section.

5 Adaptive WNN Controller Design

A novel strategic utility function is defined as the long-term performance measure for the system. It is approximated by the WNN critic signal. The action WNN signal is constructed to minimize this strategic utility function by using a quadratic optimization function. The critic WNN and action WNN weight tuning laws are derived. Stability analysis using the Lyapunov direct method is carried out for the closed-loop system (6) with novel weight tuning updates.

5.1 Strategic utility function

The utility function $p(k) = [p_i(k)]_{i=1}^{m} \in \Re^m$ is defined on the basis of the filtered tracking error $\hat{r}$ and is given by [3]:

$$\begin{array}{ll} p_i(k) = 0, & \text{if } \hat{r}_i^2 \leq \eta, \\ \mathrm{p_i(k)} = 1, & \text{if } \hat{r}_i^2 > \eta, \end{array} \tag{11}$$

where $p_i(k) \in \Re, i = 1, 2, \dots, m$ and $\eta \in \Re$ is the predefined threshold, $p(k)$ can be considered as the current performance index. The long term system performance measure $Q'(k) \in \Re^m$ can be defined using the binary utility function as

$$Q'(k) = \alpha^N p(k+1) + \alpha^{N-1} p(k+2) + \ldots + \alpha^{k+1} p(N) + \ldots, \tag{12}$$

where $\alpha \in \Re$ and $0 < \alpha < 1$ and N is the horizon. Above equation may be rewritten as

$$Q(k) = \min_{u(k)} \{\alpha Q(k-1) - \alpha^{N+1} p(k)\}.$$

This measure is similar to standard Bellman's equation [15].

5.1.1 Critic WNN

$Q'(k)$is approximated by the critic WNN by defining the prediction error as

$$e_c(k) = \hat{Q}(k) - \alpha(\hat{Q}(k-1) - \alpha^N p(k)), \tag{13}$$

where $\hat{Q}(k) = \hat{w}_1^T(k)\phi_1(v_1^T x(k)) = \hat{w}_1^T(k)\phi_1(k)$, $e_c(k) \in \Re^m$, $\hat{Q}(k) \in \Re^m$ is the critic signal, $w_1(k) \in \Re^{n_1 \times m}$ and $v_1 \in \Re^{nm \times n_1}$ represent the weight estimates, $\phi_1(k) \in \Re^{n1}$is the wavelet activation function and $\mathrm{n_1}$ is the number of nodes in the wavelet layer. The objective function to be minimized by the critic NN is defined as:

$$E_c(k) = \frac{1}{2} e_c^T(k) e_c(k). \tag{14}$$

The weight update rule for the critic NN is a gradient-based adaptation, which is given by [3]

$$\hat{w}_1(k+1) = \hat{w}_1(k) + \Delta\hat{w}_1(k),$$

where

$$\Delta\hat{w}_1(k) = \alpha_1 \left[-\frac{\partial E_c(k)}{\partial \hat{w}_1(k)} \right] \tag{15}$$

or

$$\hat{w}_1(k+1) = \hat{w}_1(k) - \alpha_1 \phi_1(k) \times (\hat{w}_1^T(k)\phi_1(k) + \alpha^{N+1} p(k) - \alpha \hat{w}_1^T(k-1)\phi_1(k-1))^T, \tag{16}$$

where $\alpha_1 \in \Re$ is the WNN adaptation gain. The critic WNN weights are tuned by the reinforcement learning signal and discounted values of critic WNN past outputs.

5.1.2 Action WNN

The output of the action NN is to approximate the unknown nonlinear function $f(x(k))$ and to provide an optimal control signal to be the part of the overall input $u(k)$ as

$$\hat{f}(k) = \hat{w}_2^T(k)\phi_2(v_2^T x(k)) = \hat{w}_2^T(k)\phi_2(k), \tag{17}$$

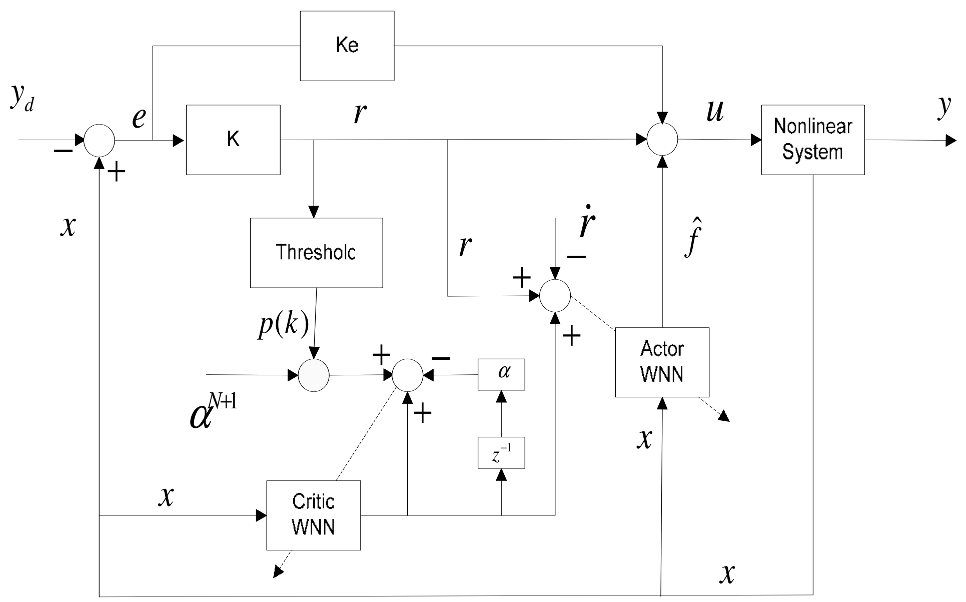

Figure 1: Block diagram of the closed loop system.

where $\hat{w}_2(k) \in \Re^{n_2 \times m}$ and $v_2 \in \Re^{nm \times n_2}$ represent the matrix of weight estimate, $\phi_2(k) \in \Re^{n_2}$ is the activation function, n_2 is the number of nodes in the hidden layer. Suppose that the unknown target output-layer weight for the action WNN is w_2 then we have

$$f(k) = w_2^T(k)\phi_2(v_2^T x(k))\varepsilon_2(x(k)) = w_2^T(k)\phi_2(k)\varepsilon_2(x(k)), \tag{18}$$

where $\varepsilon_2(x(k)) \in \Re^m$ is the WNN approximation error. Combining (17) and (18),

$$\tilde{f}(k) = \hat{f}(k) - f(k) = (\hat{w}_2(k) - w_2)^T \phi_2(k) - \varepsilon_2(x(k)), \tag{19}$$

where $\tilde{f}(k) \in \Re^m$ is the functional estimation error. The action WNN weights are tuned by using the functional estimation error$\tilde{f}(k)$and the error between the desired strategic utility function$Q_d(k) \in \Re^m$ and the critic signal $\hat{Q}(k)$ as shown in figure 2. Define

$$e_a(k) = \tilde{f}(k) + (\hat{Q}(k) - Q_d(k)). \tag{20}$$

The objective is to make the utility function $Q_d(k)$ zero at every step. Thus (20) becomes

$$e_a(k) = \tilde{f}(k) + \hat{Q}(k). \tag{21}$$

The objective function to be minimized by the action NN is given by

$$E_a(k) = \frac{1}{2} e_a^T(k) e_a(k). \tag{22}$$

The weight update rule for the action NN is also a gradient based adaptation, which is defined as

$$\hat{w}_2(k+1) = \hat{w}_2(k) + \Delta\hat{w}_2(k),$$

where

$$\Delta\hat{w}_2(k) = \alpha_2 \left[-\frac{\partial E_a(k)}{\partial \hat{w}_2(k)} \right] \tag{23}$$

or

$$\hat{w}_2(k+1) = \hat{w}_2(k) - \alpha_2 \phi_2(k)(\hat{Q}(k) + \tilde{f}(k))^T, \tag{24}$$

where $\alpha_2 \in \Re$ is the WNN adaptation gain.
The WNN weight updating rule in (24) cannot be implemented in practice since the nonlinear function $f(x(k))$ is unknown. However, using (10), the functional estimation error is given by

$$\tilde{f}(k) = \dot{\hat{r}} + \hat{r}. \tag{25}$$

Substituting (25) into (24), $\hat{w}_2(k+1) = \hat{w}_2(k) - \alpha_2 \phi_2(k)(\hat{Q}(k) + \dot{\hat{r}} - \hat{r})^T$.
Here the weight update for the action WNN is tuned by the critic WNN output, current filtered tracking error, and a conventional outer-loop signal as shown in Figure 2.

6 Stability Analysis

Consider a Lyapunov functional of the form

$$V = \frac{1}{2}\hat{r}^2. \tag{26}$$

Differentiating it along the trajectories of the system, we have

$$\dot{V} = \hat{r}(K_e \hat{e} + K(\hat{f}(\hat{x}) + u(t) - v_r - \overset{n}{\underset{d}{y}}).$$

By the substitution of control law $u(t)$ in the above equation,

$$\dot{V} = \hat{r}(-K\hat{r} + \tilde{f}(\hat{x}) - v_r).$$

$$\dot{V} \leq -K\hat{r}^2 + |\hat{r}| \left| \tilde{f}(\hat{x}) \right| - \hat{r}v_r).$$

Substituting the robust control term $v_r = -\frac{(\rho^2+1)\hat{r}}{2\rho^2}$ in the above equation, we get

$$\dot{V} \leq -s_1\hat{r}^2 + s_2(|\hat{r}| \left| \tilde{f}(\hat{x}) \right|)^2,$$

where $s_1 = (K + \frac{K}{2})$ and $s_2 = \frac{K\rho^2}{2}$. The system is stable as long as

$$s_1\hat{r}^2 \geq s_2(|\hat{r}| \left| \tilde{f}(\hat{x}) \right|)^2. \tag{27}$$

7 Simulation Results

Simulation is performed to verify the effectiveness of proposed reinforcement learning WNN based control strategy. Consider a system of the form

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= 0.01 x_1 \sin x_2 + u, \\ y &= x_1. \end{aligned} \tag{28}$$

System belongs to the class of uncertain nonlinear systems defined by (5) with $n = 2$. The proposed controller strategy is applied to this system with an objective to solve the tracking problem of system. The desired trajectory is taken as $y_d = 0.5 \sin t + 0.1 \cos 0.5t + 0.3$. Initial conditions are taken as $x(0) = [0.5, 0]^T$. Attenuation levels for robust controller are taken as 0.01. Controller gain vector is taken as $k = [35, 5]$. Wavelet networks with discrete Shannon's wavelet as the mother wavelet is used for approximating the unknown system dynamics. Wavelet parameters for these wavelet networks are tuned online using the proposed adaptation laws. Initial conditions for all the wavelet parameters are set to zero. Simulation results are shown in Figure 1. As observed from the figures, system response tracks the desired trajectory rapidly.

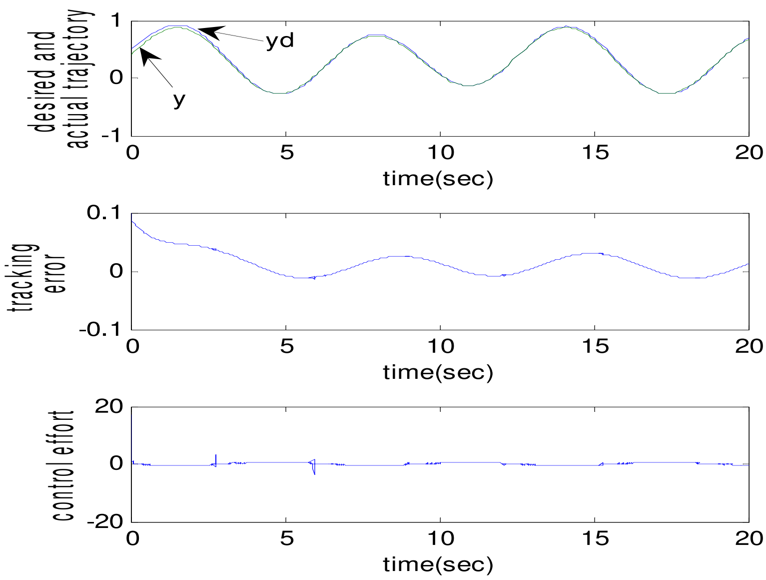

Figure 2: System output and tracking error.

8 Conclusion

A reinforcement learning WNN based adaptive tracking control strategy is proposed for a class of systems with unknown system dynamics. Adaptive wavelet networks are used for approximating the unknown system dynamics. Adaptation laws are developed for online tuning of the wavelet parameters. The stability of the overall system is guaranteed by using the Lyapunov functional. The theoretical analysis is validated by the simulation results.

References

[1] Wen, J. W. and Liu, F. Robust Model Predictive Control for Fuzzy Systems Subject to Actuator Saturation. In: *Sixth International Conference on Fuzzy Systems and Knowledge Discovery* (2009) 400–404.

[2] Chen, J. and Liu, F. Robust fuzzy control for nonlinear system with disk pole constraints. *Control and Decision* (2007) 983–988.

[3] He, P. and Jagannathan, S. Reinforcement learning neural-network-based controller for nonlinear discrete-time systems with input constraints. *IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics* **37**(2) (2007) 425–436.

[4] Lin, W.S., Chang, L.H. and Yang, P.C. Adaptive critic anti-slip control of wheeled autonomous robot. *IET Control Theory Applications* **1**(1) 2007.

[5] Crespo, L.G. Optimal performance, robustness and reliability base designs of systems with structured uncertainty. In: *Proceeding of American Control Conference* (2003) 4219–4224.

[6] Zadeh, N., Jamali, N. and Hajiloo, A. Frequency-based reliability Pareto optimum design of proportional-integral-derivative controllers for systems with probabilistic uncertainty. *Journal of Systems and Control Engineering (IMECHE)* **22**(18) (2007) 1061–1066.

[7] Peters, J. and Schaal, S. Policy gradient methods for robotics. In: *Proc. of the IEEE International Conference on Intelligent Robotics Systems (IROS)* (2006) 2219–2225.

[8] Sharma, M., Kulkarni, A. and Puntambekar, S. Wavelet based adaptive tracking control for uncertain nonlinear systems with input constraints. In: *Proc. of the IEEE International Conference on Advances in Recent Technoiogies in Communication and Computing (ARTCOM'09)* (2009) 694–698.

[9] Zhang, Q. and Benveniste, A. Wavelet networks. *IEEE Transactions on Neural Networks* **3**(6) (1992) 889–898.

[10] Zhang, J., Walter, G.G., Miao, Y. and Lee, W. Wavelet neural networks for function learning. *IEEE Transactions on Signal Processing* **43**(6) (1995) 1485–1497.

[11] Jaddu, H. and Hiraishi, K. A Wavelet Approach for Solving linear quadratic optimal control problems. *SICE-ICASE International Joint Conference* (2006) 18–21.

[12] Minu, K. K., Lineesh, M. C. and John, C.J. Wavelet Neural Networks for Nonlinear Time Series Analysis. *Applied Mathematical Sciences* **4**(2010) 2485–2495.

[13] Lin, X. and Balakrishnan, S. N. Convergence analysis of adaptive critic based optimal control. In: *Proc. of the American Control Conference* (2000) 1929–1933.

[14] Murray, J. J., Cox, C., Lendaris, G.G. and Saeks, R. Adaptive dynamic programming. *IEEE Transactions on System, Man, Cybernatics* **32**(2) (2002) 140–153.

[15] Prokhorov, D. V. and Wunsch, D.C. Adaptive critic design. *IEEE Transactions on Neural Networks* **8**(5) (1997) 997–1007.

[16] Hasselt, H. V. and Wiering, D.C. Reinforcement learning in continous action space. *IEEE symposium on Approximate Dynamic Programming and Reinforcement Learning* (2007) 272–279.

[17] Talmoudi, S., Abderrahim, K., Abdennour, R.B. and Ksouri, M. Multimodal approach using neural networks for complex systems modelling and identification. *Nonlinear Dynamics and Systems Theory* **8**(3) (2008) 299–316.