Constructing Fuzzy Wavelet Network Modeling

Wang Jun[1], Peng Hong[2]

[1]School of E1ectrical Information, Xihua University, China, 610039
[2]School of Mathematics & Computer Science, Xihua University, China, 610039

jiayu@mail.sc.cninfo.net

Abstract

Using B-spline wavelet as membership function, a fuzzy wavelet network (FWN) is proposed for approximating nonlinear function in this paper. The approximation capability is proved. Simulation examples are given for static system and dynamic system respectively. Compared with the other methods, the presented FWN not only reserves the multi-resolution capability of WNN, but also has the advantages of the approximation accuracy and good generalization performance.
Keyword: Fuzzy neural network, wavelet network, system identification

I. Introduction

During the past few years, the nonlinear system modeling has been extensively studied. Neural network and fuzzy system are universal approximates and have been shown to be parsimonious. The back-propagation algorithm implements a nonlinear gradient optimization procedure; it can be trapped at a local minimum and converges slowly. Wavelet transform reveals properties of the function in localized regions of the joint time-frequency space and of multi-resolution analysis.

Since the first proposal by Zhang and Benveniste [1] in 1992, wavelet neural network (WNN) has been widely developed in theory research. WNN can avoid local minimum. Multi-layer perception (MLP) and the radial basis function (RBF) are the most often used neural network. The idea of using wavelets in MLP and RBF neural networks design was proposed in [1] and [2]. In [4], Yang et al introduced multi-resolution wavelet network. From the viewpoint of frequency domain, the three wavelet networks have different characteristic: MLP wavelet neural network (MLP-WNN) lose low frequency information and RBF wavelet neural network (RBF-WNN) lose high frequency information. The multi-resolution wavelet network includes most extensive information.

Recently, combination with fuzzy technique and wavelet transformation was proposed in some papers[5][7][8][12][13][14].The method proposed by [14] and [7] gave the idea multi-resolution fuzzy model, but they had no deeply researcher and had not better experimental for the model. Another fuzzy wavelet network (FWN) was proposed in [5] which combined WNN into the consequents of the rules in TS fuzzy model. As a result, the model accuracy and the generalizing capability can be improved. But the model structure became more complicate.

In this paper, taking the advantages of the property of locality of the scaling function, we discuss another new FWN with simpler structure that could overcome the disadvantage of ANN in low convergence rate and the shortage of the wavelet neural in needing too many training data. It also provides an efficient way to determine the wavelet coefficient.

This paper is organized as follows. Section 2 discusses wavelet multi-resolution theory. The self-constructing fuzzy neural network based wavelet basis function is described in Section 3. An approximation property of FWN is proven in Section 4. The simulation example is provided to verify the performance of the FWN in Section 5. Finally, a brief conclusion is drawn in Section 6.

II. Wavelet Theory with Multi-resolution Analysis

Then exists a function $\varphi(x) \in L^2(R)$ which is called scaling function. Defining the following functions:

$$\varphi_{j,k}(x) = 2^{-j/2}\varphi(2^{-j}x - n) \qquad (j,k \in Z) \tag{1}$$

basis $\{\varphi_{j,k}(x) | k \in Z\}$ is standard and orthonormal. j and k is dilation and translation of the scaling function, respectively. V_j with standard orthonormal basis is called the scaling subspace of $L^2(R)$, i. e., a nested chain of closed subspace.

$$\ldots \in V_{-1} \in V_0 \in V_1 \in V_2 \in \ldots \tag{2}$$

such that

$$\bigcap_m V_j = \{0\}, \qquad clos\left\{\bigcup_j V_j\right\} = L^2(R) \tag{3}$$

where V_j is the subspace spanned by $\{\varphi_{j,k}\}_{k=0}^{k=+\infty}$. In order to form an orthonormal supplementary space of V_j, i.e.

$$V_{j-1} = V_j \oplus W_j, \qquad V_j \perp W_j \qquad (j \in Z) \tag{4}$$

$$L^2(R) = \bigoplus_{j=-\infty}^{j=+\infty} W_j \tag{5}$$

It is easy to show the properties of MRA.

$$V_0 = V_1 \oplus W_1 = V_2 \oplus W_2 \oplus W_1 = V_3 \oplus W_3 \oplus W_2 \oplus W_1 = \ldots \tag{6}$$

and

$$V_j = span\left\{2^{-\frac{j}{2}}\varphi(2^{-j}x - k)\right\}, \qquad j \in Z \tag{7}$$

$$W_j = span\left\{2^{-\frac{j}{2}}\psi(2^{-j}x - k)\right\}, \qquad j \in Z \tag{8}$$

There existed a function $\psi(x) \in W_0$ such that $\{\psi(x-k), k \in Z\}$ is made to be the standard orthonormal basis of W_0. Traditionally $\psi(x)$ is called wavelet function. Function $f(x) \in L^2(R)$ can be decomposed in the following two methods:

$$f(x) = \sum_{j=-\infty}^{J} \sum_{k=-\infty}^{+\infty} d_{j,k}\varphi_{j,k}(x) + \sum_{k=-\infty}^{+\infty} c_{J,k}\psi_{J,k}(x) \tag{9}$$

When $J \to \infty$, (8) can be expressed as.

$$f(x) = \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} d_{j,k}\psi_{j,k}(x) \tag{10}$$

For any $\varepsilon > 0$, there existed a $k \in Z$, such that

$$\left\| f(x) - \sum_k c_{J,k}\phi_{J,k}(x) \right\| < \varepsilon \tag{11}$$

which is big enough a linear combination of the scale function can be used to approximate a $L^2(R)$ function. (9), (11) and (12) are just foundation that the WNN works.

III. The Structure of the FWN

In Takagi-Sugeno model, a set of fuzzy rules can be described by

R_i: if x_1 is A_1, and …, and x_p is A_p, then y is b_i

Where A_i is the fuzzy set characterized by the Gaussian type or triangular type membership function. x_i is ith input. y is output variable, p is the number of input variables. Instead, if wavelet function with symmetry, positive and singular value, i.e., spline wavelet is selected as A_i, membership function with fine resolution can capture the local behavior of the function accuracy. So, it leads to different fuzzy sub-model with variable dilation and translation factor. In this paper, we construct FWN using B-spline wavelet. Assuming when dilation factor is j, corresponding fuzzy sub-system is expressed by i. kth fuzzy rule can be denoted

$$R_i^{(k)} : \textit{if } x_1 \textit{ is } A_{j_i, t_i^{k1}} \textit{, and} \ldots \textit{, and } x_p \textit{ is } A_{j_i, t_i^{kp}} \textit{, then } y \textit{ is } b_i^k$$

where $A_{j_i, t_i^{k1}}$ corresponding to B-spline wavelet with j scale factor and t_i^{kl} translation factor of input x_i is expressed as $\mu_{j_i, t_i^{kl}}$,$l=1,\ldots,p$; b_i^k is the output of local model for rule $R_i^{(k)}$。 A typical fuzzy subsystem can be obtained:

$$\hat{y}_i = \frac{\sum_{k=1}^{N_i} b_i^k \prod_{l=1}^{p} \mu_{j_i, t_i^{kl}}(x_l)}{\sum_{i=1}^{c}\sum_{k=1}^{N_i}\prod_{l=1}^{p} \mu_{j_i, t_i^{kl}}(x_l)} = \sum_{k=1}^{N_i} b_i^k \hat{\mu}_i^k(x) \tag{12}$$

The whole the output of FWN system is

$$\hat{y} = \sum_{i=1}^{c} \hat{y}_i = \sum_{i=1}^{c}\sum_{k=1}^{N_i} b_i^k \hat{\mu}_i^k(x) \tag{13}$$

let

$$\hat{\mu}_i^k(x) = \frac{\prod_{l=1}^{p} \mu_{j_i, t_i^{kl}}(x_l)}{\sum_{i=1}^{c}\sum_{k=1}^{N_i}\prod_{l=1}^{p} \mu_{j_i, t_i^{kl}}(x_l)} \tag{14}$$

where $0 \le \hat{\mu}_i \le 1$, $\sum_{i=1}^{c} \hat{\mu}_i^k = 1$; Number of fuzzy subsystem is c. Number of fuzzy rule is $N_1,\ldots,N_c$ in every subsystem, respectively. According to fuzzy rule and (13) and (14), FWN system can be described as Fig.1 by a multilayer forward network.

We change c rule into a rule which member function is dilation function $\varphi_i(x)$ and (c-1) rule which member function $\psi_i(x)$ is wavelet function in (13). (13) can be described as follows:

$$\hat{y} = \sum_{i=1}^{c} \hat{y}_i = \sum_{k=1}^{N_1} b_1^k \hat{\varphi}_1^k(x) + \sum_{i=2}^{c}\sum_{k=1}^{N_i} b_i^k \hat{\psi}_i^k(x) \tag{15}$$

Where

$$\varphi_{j_i, t_i^{kl}}(x_l) = 2^{j_i/2} \varphi(2^{j_i} x_l - t_i^{kl}) \tag{16}$$

$$\psi_{j_i, t_i^{kl}}(x_l) = 2^{j_i/2} \psi(2^{j_i} x_l - t_i^{kl}) \tag{17}$$

Then (9) similar to (15). When k change from $-\infty$ to ∞, $\hat{y}$ is entirely identical with approximate function y. (15) resemble with multi-resolution wavelet proposed by Yang[4] which is called MRFWN (multi-resolution fuzzy wavelet network).

If first item is omitted in (15),it resemble with MLP wavelet proposed by Zhang[1] which structure resemble to MLP and be called MLP-FWN (multilayer perceptron fuzzy wavelet network). In finite input space and using finite neural cell, $\hat{y}$ can approximate nonlinear function y with

$$\hat{y}=\sum_{i=1}^{c}\frac{\sum_{k=1}^{N_i} b_i^k \prod_{l=1}^{p}\psi_{j_i,t_i^{kl}}(x_l)}{\sum_{i=1}^{c}\sum_{k=1}^{N_i}\prod_{l=1}^{p}\psi_{j_i,t_i^{kl}}(x_l)}=\sum_{i=1}^{c}\sum_{k=1}^{N_i} b_i^k \hat{\psi}_i^k(x)+g_0 \qquad (18)$$

where $g_0 \in R$. $\psi_i(.)$ is three order spline wavelet.

If second item is omitted in (15),it resemble with RBF wavelet proposed by Zhang[2] which structure resemble to RBF and be called RBF-FWN（Radial basis function fuzzy wavelet network). Then the approximated function of FWN is

$$\hat{y}=\frac{\sum_{k=1}^{N_1} b_1^k \prod_{l=1}^{p}\varphi_{j_i,t_i^{kl}}(x_l)}{\sum_{k=1}^{N_1}\prod_{l=1}^{p}\varphi_{j_i,t_i^{kl}}(x_l)}=\sum_{k=1}^{N_1} b_1^k \hat{\varphi}^k(x) \qquad (19)$$

According to multi resolution analysis of wavelet theory, when the dilation factor is finite, MLP-FWN will lose low frequency signal and RBF-FWN will lose high frequency signal. From the view of frequency domain, MRFWN include most extensive frequency signal. MRFWN can be decomposed into two parts that are low frequency part corresponding to V_j space with slow signal and high frequency part corresponding to $W_1 \oplus W_2 \cdots \oplus W_J$ space with accurately signal. The existing space of MRNN is $V_J \oplus W_J \oplus W_{J-1} \oplus \cdots \oplus W_1$. So, MRNN can approximate nonlinear function from coarse to fine with multi-resolution character.

In conclusion, in the case of finite neural cell and multi-resolution, above wavelet network correspond to different frequency domain. When selecting network structure, we should consider frequency character of signal.

Ⅳ. FWN Model Based on B-Spline Wavelet

Conclusion:Based on B-spline wavelet, MRNN-WNN, RBF-WNN and MLP-WNN are identical in the structure.

Proof: The *j* order B-spline dilation function $\varphi_i(x)$ and wavelet function $\psi_i(x)$ is [7]

$$\varphi_{j,k}(x)=\sum_{i=0}^{m}2^{-m+1}\binom{m}{i}2^{j/2}N_m(2^{j+1}x-2k-j) \qquad (20)$$

$$\psi_{j,k}(x)=\sum_{n=0}^{3m-2}h_n N_m(2^{j+1}x-2k-n) \qquad (21)$$

If using the dilation function and wavelet of three order B-spline wavelet, substituting (20) and (21) into (15), we can obtain:

$$f(x)=\sum_{j=-\infty}^{J}\sum_{k=-\infty}^{+\infty}d_{j,k}\sum_{n=0}^{3m-2}h_n N_m(2^{j+1}x-2k-n)++\sum_{k=-\infty}^{+\infty}c_{J,k}\sum_{i=0}^{m}2^{-m+1}\binom{m}{i}2^{J/2}N_m(2^{J+1}x-2k-i) \qquad (22)$$

Letting $2k+n=k_1, 2k+i=k_2, j_1=j+1$, above can be simplified as followed:

$$\begin{aligned} f(x) &= \sum_{j_1=-\infty}^{J+1}\sum_{k_1=-\infty}^{+\infty} D_{(j_1-1),k_1} N_m(2^{j_1}x-k_1) + \sum_{k_2=-\infty}^{+\infty} C_{J,k_2} N_m(2^{J+1}x-k_2) \\ &= \sum_{j_1=-\infty}^{J}\sum_{k_1=-\infty}^{+\infty} D_{(j_1-1),k_1} N_m(2^{j_1}x-k_1) + [\sum_{k_1=-\infty}^{+\infty} D_{J,k_1} N_m(2^{J+!}x-k_1) + \sum_{k_2=-\infty}^{+\infty} C_{J,k_2} N_m(2^{J+1}x-k_2)] \\ &= \sum_{j_1=-\infty}^{J}\sum_{k_1=-\infty}^{+\infty} D_{(j_1-1),k_1} N_m(2^{j_1}x-k_1) + \sum_{k_3=-\infty}^{+\infty} D_{J,k_3} N_m(2^{J+!}x-k_3) \end{aligned} \quad (23)$$

First part is composed by wavelet space $W_1 \oplus W_2 \cdots \oplus W_J$ under small dilation factor to compensate to fast signal. Second part is composed by wavelet space V_J under big dilation factor to compensate to slow signal. The two parts is expressed as wavelet dilation function such that become a part:

$$f(x) = \sum_{j_3=-\infty}^{J+1}\sum_{k_4=-\infty}^{+\infty} D_{(j_3-1),k_4} N_m(2^{j_3}x-k_4) \quad (24)$$

(24) is identical with MLP-FWN in (18) and RBF-FWN in (19). So FWN based on B-spline wavelet can approximate any nonlinear function.

Ⅴ. Simulation

The performance index is defined as

$$I_D = \frac{J_{i-1} - J_i}{J_{i-1}} \quad (25)$$

where J_i is the performance index at ith iteration. We take the measure in [1] and [5]

$$J_i = \sqrt{\frac{\sum_{l=1}^{c} (\hat{y}_l - y_l^d)^2}{\sum_{l=1}^{c} (y_l^d - \bar{y})^2}} \quad \text{with} \quad \bar{y} = \frac{1}{c}\sum_{l=1}^{c} y_l^d \quad (26)$$

where $\hat{y}_l$ is the estimated output and y_l^d is the desired output. For our FWN, initializing wavelets and fuzzy rules is done according to the method of [5]. Secondly, purifying the wavelets by orthogonal least-square algorithm[1]. After setting up the network initialization, we used the gradient method to adjust all translation factor, dilation factor and weight coefficients of the wavelet network simultaneously. Details of the gradient method can be found in many reports[9]. In this section, two examples are given to demonstrate the validity of the presented FWN.

A. Nonlinear static system identification

In order to compare with [1] and [5], we select the following function

$$f(x) = \begin{cases} -2.18x - 12.86 & -10 \le x < -2 \\ 4.246x & -2 \le x < 0 \\ 10e^{-0.05x-0.5} & 0 \le x \le 10 \end{cases} \quad (27)$$

In this example, we sampled 200 points distributed uniformly over [-10,10] as training data. The number of rules used in our FWN is 10. Fig.2 illustrates the simulation result. The performance of FWN is shown in the Table 1. Obviously, the performance of our FWN is superior to that of [1] and [5] with low unknown parameters and good accurate.

B. Nonlinear dynamic system identification

The nonlinear dynamic system to be identified is define as follows:

$$y(k) = 0.3y(k-1) + 0.6y(k-2) + f(u(k)) + e(k) \quad (28)$$

Table1 Comparison FWN and WNN

Method	Number of unknown parameters used for training	Performance index
Our FWN	20	0.022
FWN [5]	28	0.021
WNN [1]	22	0.05057

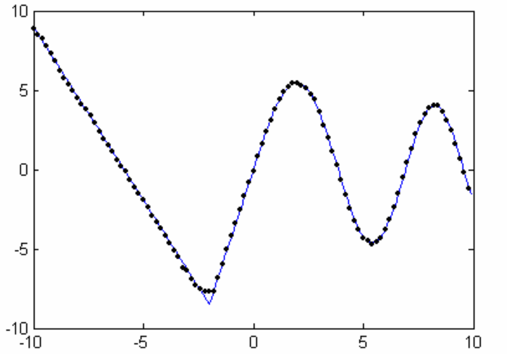

Fig. 2 The comparison of outputs between original function and the final from FWN(dotted line)

where $f(u(k)) = u^3 + 0.3u^2 - 0.4u$,e(k) is white noise signal, average is 0, variance is 0.2. In order to verify the predicting capability of model, input is selected as

$$u(k) = \begin{cases} \sin(2\pi k/125) & k \le 500 \\ 0.2\sin(2\pi k/40) + 0.8\sin(2\pi k/25) & k > 500 \end{cases} \quad (29)$$

We collected data on-line for 1000 time steps to form input-output sample pairs for constructing the FWN. The performance index is 0.08. Due to compact support of wavelet membership function, the study algorithm has local learning. During online dynamic identification, only finite scale variables and wavelet nodes are used. Fig. 3 shows that our FWN has good generalization capability and fast convergence.

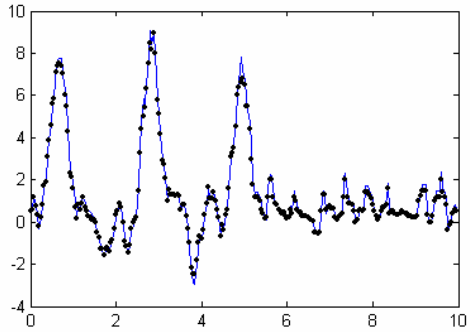

Fig. 3 The identification of dynamic system using FWN(dotted line)

Ⅵ. Conclusion

In this paper, we construct a new FWN based B-spline wavelet for function approximation. The presented FWN integrates fuzzy concepts with the WNN such that the contribution degree of different sub-WNNs at different resolution levels can be controlled flexibly. Compared with other algorithms for selecting wavelets, the presented FWN not only reserves the multi-resolution capability of WNN, but also has the advantages of the approximation accuracy and good generalization performance. However, it need to deeply research for high dimension system identification.

Acknowledgement: This work is supported by the importance project fund of the science and technology department of Sichuan province, China (No. 0229957) and the doctor fund of Ministry of Education of the People's Republic of China (No. 20040613013)

References

[1] Zhang Q. and Benveniste, "A Wavelet networks," IEEE Trans. neural networks, Vol. 3, No. 6, 1992, pp. 889-898.
[2] Zhang Qinghua, "A Using Wavelet Network for Nonparametric Estimation," IEEE Trans. Neural Networks, Vol. 8, No.2, 1997, pp. 227-236.
[3] Zhang Jun, Walter G G, Miao Y, et al, "Wavelet Neural Networks for Function Learning," IEEE Trans. Signal Processing, Vol. 43, No. 6, 1995, pp. 1483-1497.
[4] Zi-Jian Yang, Setsuo Sagara and Teruo Tsuji, "System Impulse Response Identification Using a Multiresolution Neural Network," Automatica, Vol. 33, No. 7,1997, pp.1345-1350.
[5] Daniel J., HO W. C., et. al., "Fuzzy wavelet networks for function learning. IEEE Trans. on Fuzzy Systems, Vol. 9, No. 1, 2001, pp. 200-211.
[6] Wang Haiqing, Song Zhihuan, et. al., "Nonlinear system identifaction using orthogonal wavelet networks," Control theory and applications, Vol. 2, 2001, pp.200-204.
[7] Wang Hui, "multi-resolution based fuzzy system and its application in locomotive adhesion control," Doctor degree dissertation, Southwest Jiaotong University, Vol. 5, 2004, pp.110-114.
[8] L.C. Jiao, F. Liu and L.Wang, "FWNN for interval estimation with interval learning algorithm," Proceedings of the 1998 IEEE International Symposium on Circuits and Systems, 1998, 3, pp. 119-122.
[9] Lixin wang, *Adaptive Fuzzy System and Control-Design and Stability analysis*, Beijing: Guofang Industry Publish, 1995.
[10] J.S.R.Jang, C.T.Sun, E.Mizutani, *Neoro-fuzzy and Soft Computing*, Prrentice Hall, 1997.
[11] B. J. Park, W. Pedrycz and S.K.Oh, "Identification of fuzzy models with the aid of evolutionary data granulation," IEE Proc. Control Theory Appl. Vol. 148, 2001, pp.406-418.
[12] Lin C. K. and Wang S. D., "Fuzzy modeling using wavelet transform," Electronics Letters,Vol.32, 1996, pp. 2255-2256.
[13] Lin C. K. and Wang S. D., "Constructing a Fuzzy Model from Wavelet Transforms," Soft Computing in Intelligent Systems and Information Processing of the 1996 Asian, 1996, pp. 394-399.
[14] M. Thuillard, "Wavelets in Soft Computing," World Scientific Series in Robotics and Intelligent Systems, Vol.25, World Scientific, Singapore, 2000, pp. 104-106.

Wang Jun received the B.S. degree and the M.S. degree from Chongqing University, Chongqing, China in 1988 and 1991, respectively. She is currently working toward the Ph.D degree in the school of electrical engineering in South-West Jiaotong university, Chengdu, China. Since 1991, she has been with Xihua university where she is a professor. Her current research interests are the application of neural network, fuzzy control and wavelet analysis.

Peng Hong received the B.S. degree and the M.S. degree from Sichuan Normal University, Chengdu, China in 1987 and 1990, respectively. Since 1990, he has been with Xihua university where he is a professor. His current research interests are support vector machine and its application.