Journal of Engineering and Development, Vol. 16, No.3, Sep. 2012 ISSN 1813- 7822

Using Multi wavelet Network for Identification

Ahmad Qasim A.
Computer Eng. Dept.
Univ. of Mousl

Asst. Prof.Thakwan M. Saleem
Computer Eng. Dept.
Univ. of Mousl

ABSTRACT

Multiwavelet neural network is a new class of artificial neural networks whose their activation functions are multiwavelet functions. The neurons in the hidden layers have been divided into two parts, each part contains wavelets of the same type. The multiwavenet algorithm consist of self-construction of the network and the minimization of the error. Here the constructed network is used for identification of a non-linear function. The network showed the excellent ability to identify the required function compared with the single wavelet network.

استخدام الشبكة متعددة المويجات لاغراض التطابق

الخلاصة:

في هذا البحث تم تطوير هيكلية المويجة الى الشبكة متعددة المويجات التي تعد صنفا جديدا للشبكات العصبية الاصطناعية، اذ تميزت الدوال الفعالة في الشبكة متعددة المويجة باحتوائها على مويجات متعددة من انواع مختلفة، حيث تم تقسيم الخلايا في الطبقة الخفية الى عدة اقسام، احتوى كل قسم على نفس النوع من المويجات، فضلا عن ذلك تم اجراء مقارنة بين اقسام الشبكة متعددة المويجة من ناحية نوعية المويجة المستخدمة في القسم والتكرارات ومعدل الخطأ0

1. INTRODUCTION:

Wavelet neural networks are a novel powerful class of neural networks that incorporate the most important advantages of multiresolution analysis introduced by Mallat in 1989 [6]. Zhang and Benveniste [1] found a link between wavelet decompositions and neural networks. They combine the good localisation properties of wavelets with the approximation abilities of neural networks. This kind of network uses wavelets as activation functions in the hidden layer and a type of backpropagation algorithm is used for its learning. the wavelet space is composed of the dilation and translation functions of mother wavelet, but multiwavelet space is composed of the dilation and translation functions of multiple mother wavelets that own some specific properties[7]. The idea of Multiwavenet is similar to the wavelet network but the activation function are multiwavelet functions.

2. Multiwavelet network algorithm

The Multiwavelet network approximates any desired signal y(t) by generalizing a linear combination of two sets of daughter wavelets $\psi_{1a_k,b_k}(t)$ and $\psi_{2a_m,b_m}(t)$ generated by dilation, a, and translation, b, from two mother wavelets $\psi_1(\tau)$ and $\psi_2(\tau)$, where $\tau = \dfrac{t-b}{a}$.

$$\psi_{1a_k,b_k}(t) = \psi_1\left(\frac{t-b_k}{a_k}\right) \quad (1)$$

$$\psi_{2a_m,b_m}(t) = \psi_2\left(\frac{t-b_m}{a_m}\right) \quad (2)$$

Where, a: Dilation factor, with a>0; b: Translation factor; t: Signal time interval.

The multiwavenet architecture is shown in Fig.(1). First the multiwavenet parameters dilation a 's , translations b 's , and weights w 's should be initialized, and the desired set of data, the input signal, where the desired output y(t), the number of wavelets K are given. The approximated signal of network $\hat{y}(t)$ can be represented by equation:

$$\hat{y}(t) = u(t)\left(\sum_{k=1}^{K} w_k \psi_{1a_k,b_k}(t) + \sum_{m=1}^{M} w_m \psi_{2a_m,b_m}(t)\right) \quad (3)$$

Where:

u(t): input signal.

K : number of wavelets in the first part of hidden layer.

M : number of wavelets in the second part of hidden layer.

w_k : weights of the first part.

w_m : weights of the second part.

$\psi_{1a_k,b_k}(t)$: set daughter wavelets in the first part of hidden layer.

$\psi_{2a_m,b_m}(t)$: set daughter wavelets in the second part of hidden layer.

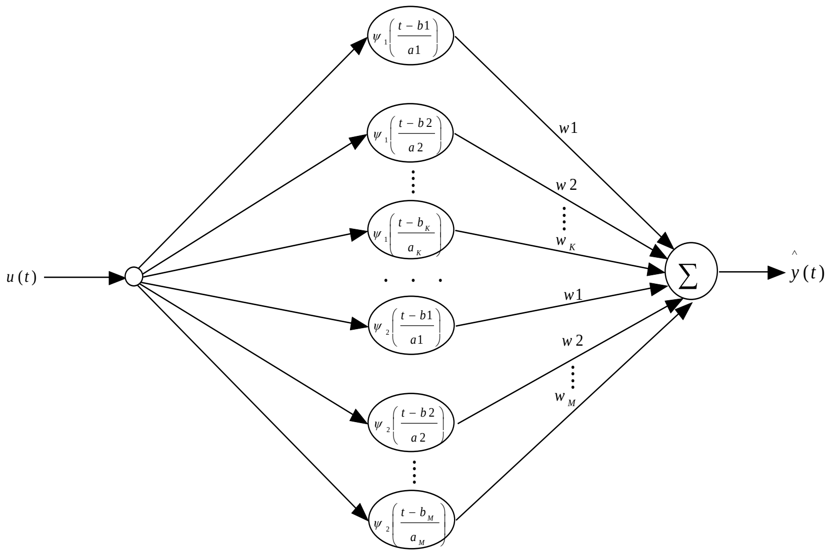

Fig.(1): structure of multiwavenet

The multiwavenet parameters *wk* , *wm* , *ak* , *am* , *bk* and , *bm* ,can be optimized in the LMS algorithm by minimizing a cost function or the energy function *E* over all time *t*, thus by denoting *e(t)* be a time-varying error function at time *t*[2][3].

$$e(t) = y(t) - \hat{y}(t) \quad (4)$$

the energy function will be defined by

$$E = \frac{1}{2}\sum_{t=1}^{T} e^{2}(t) \qquad (5)$$

where *T* is the length of the desired signal.

In the minimization of *E* the method of steepest descent is used here. This requires the computation of gradients $\frac{\partial E}{\partial w_k}, \frac{\partial E}{\partial w_m}, \frac{\partial E}{\partial b_k}, \frac{\partial E}{\partial b_m}, \frac{\partial E}{\partial a_k}$ and $\frac{\partial E}{\partial a_m}$ for the purpose of updating the incremental changes to each particular parameter *wk* , *wm* , *ak* , *am* , *bk* and , *bm* respectively. The gradient of *E* are given as follows:

$$\frac{\partial E}{\partial w_k} = -\sum_{t=1}^{T} e(t)\psi_1(\tau_k)u(t)$$
$$\frac{\partial E}{\partial w_m} = -\sum_{t=1}^{T} e(t)\psi_2(\tau_m)u(t) \qquad (6)$$

$$\frac{\partial E}{\partial b_k} = -\sum_{t=1}^{T} e(t)u(t)w_k \frac{\partial \psi_1(\tau_k)}{\partial b_k}$$
$$\frac{\partial E}{\partial b_m} = -\sum_{t=1}^{T} e(t)u(t)w_m \frac{\partial \psi_2(\tau_m)}{\partial b_m} \qquad (7)$$

$$\frac{\partial E}{\partial a_k} = -\sum_{t=1}^{T} e(t)u(t)w_k \tau_k \frac{\partial \psi_1(\tau_k)}{\partial b_k} = \tau_k \frac{\partial E}{\partial b_k}$$
$$\frac{\partial E}{\partial a_m} = -\sum_{t=1}^{T} e(t)u(t)w_m \tau_m \frac{\partial \psi_2(\tau_m)}{\partial b_m} = \tau_m \frac{\partial E}{\partial b_m} \qquad (8)$$

Where

$$\tau_k = \frac{t - b_k}{a_k}$$
$$\tau_m = \frac{t - b_m}{a_m} \qquad (9)$$

Derivatives of the various wavelet filter with respect to its translation $\frac{\partial \psi(\tau)}{\partial b}$ are given in [4]. Incremental changes of these coefficients are simply the negative of their gradients,

$$\Delta w_k = -\frac{\partial E}{\partial w_k}, \quad \Delta b_k = -\frac{\partial E}{\partial b_k}, \quad \Delta a_k = -\frac{\partial E}{\partial a_k}$$
$$\Delta w_m = -\frac{\partial E}{\partial w_m}, \quad \Delta b_m = -\frac{\partial E}{\partial b_m}, \quad \Delta a_m = -\frac{\partial E}{\partial a_m} \qquad (10)$$

Thus each coefficient of the network is updated in accordance with the rule given:

$$w_k(n+1) = w_k(n) + \eta_{w_k} \Delta w_k$$
$$w_m(n+1) = w_m(n) + \eta_{w_m} \Delta w_m \qquad (11)$$

$$a_k(n+1) = a_k(n) + \eta_{a_k} \Delta a_k$$
$$a_m(n+1) = a_m(n) + \eta_{a_m} \Delta a_m \qquad (12)$$

$$b_k(n+1) = b_k(n) + \eta_{b_k} \Delta b_k$$
$$b_m(n+1) = b_m(n) + \eta_{b_m} \Delta b_m \qquad (13)$$

The following nonlinear function with interval width 0.1 represents the desired output (target function)[5]:

$$y(t) = \begin{cases} -2.186x - 12.864 & \text{if } -10 \le x < -2 \\ 4.246x & \text{if } -2 \le x < 0 \\ 10e^{-0.05x-0.5}\sin((0.03x+0.7)x) & \text{if } 0 \le x \le 10 \end{cases} \quad (14)$$

3. Results

After training the network on the desired output, the mean squared error is 0.00023596 when 20 Morlet and 20 Slog1 are used, the following table shows the mean squared error for different type of wavelets.

Table(1) Mean square error for different types of wavelets combination

k=20, m=20 at iteration =100	
Wavelets types	MSE
Morlet+ polywog1	0.00067375
Morlet+slog1	0.00023596
Morlet+ slog2	0.00028724
slog1+ polywog1	0.0011
slog2+ slog1	0.0026
polywog1+ slog2	0.00065913

The following figures show how the network reaches the goal (desired output) when two wavelets are used in hidden layer (Morlet wavelet and slog1 wavelet).

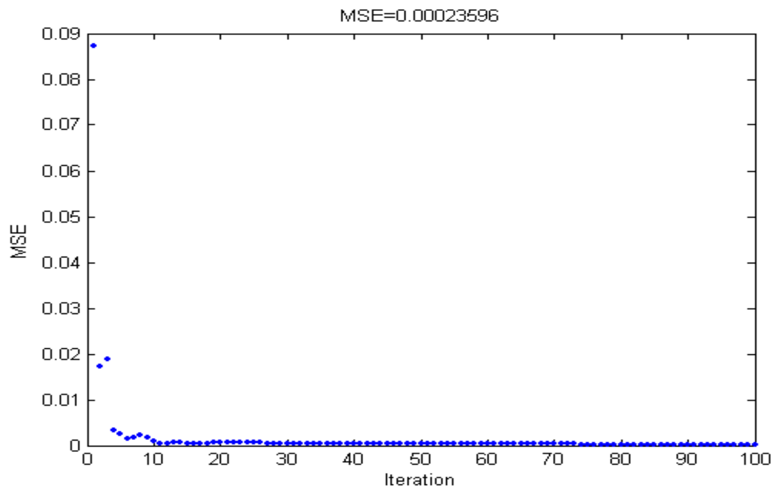

Fig.(2): MSE per learning iteration

From Fig.(2) at the beginning of the network operation, the Mean Square Error is the large, and then the error value begins to decrease after each iteration. After several iterations (here 10) we can see that the change of error is very small between each iteration compared to the large of error in the first set of iteration.

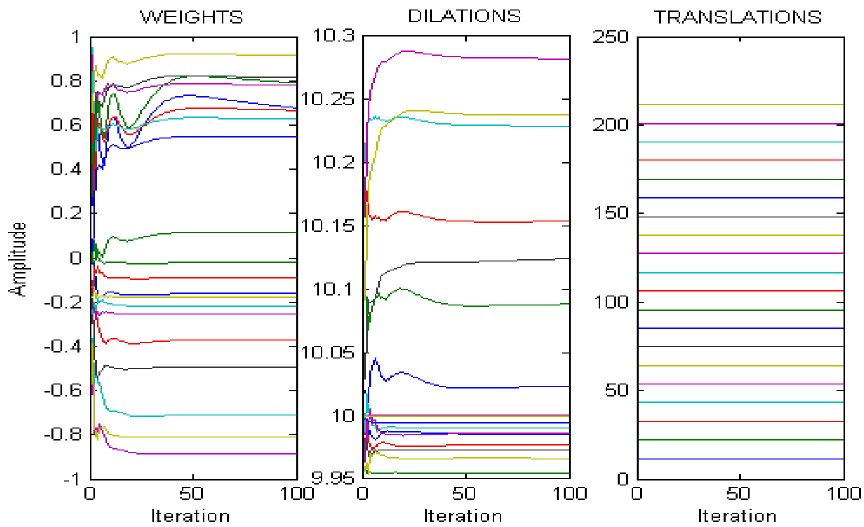

Fig.(3): MWN parameters updates for one type of wavelet for Morlet and Slog1 wavenet combination

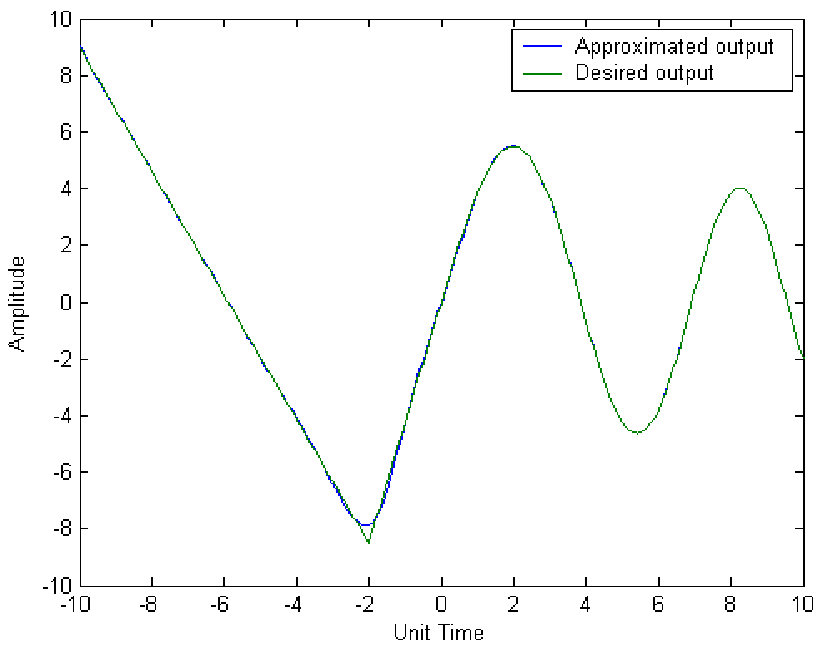

Fig.(4): Desired and approximated nonlinear output

To investigate the advantage of using multiwavenet over single wavelet network, a single Morlet wavelet network was designed, and the same designer output was used to be identified by the network which showed a MSE of 0.0011 . This showed the superiority of the multiwavenet MSE = 0.00023596 over the single wavelet network.

4. CONCLUTION

In all feedforward neural networks, the activation functions in multiwavelet neural network have best properties. In this paper we proposed multiwavenet with two different types of wavelets as an activation functions. The identification tasks discussed in this work showed that the success of the identification depends on the choice of the wavelet family used as activation function, as well as, the number of neurons, which should cover the whole extension of the signal efficiently (target). However the great amount of simulations realized showed that wavelets are quite efficient in the process of identification.

5. REFRENCES

[1] Q.zhang ,A.Benveniste , "**Wavelet Network**", IEEE Transactions on Neural Networks ,Vol.3,No.6,pp.889-898,Nov,1992.

[2] G.Lekutai, "**AdaptiveSelf - Tuning Neuro Wavelet Network Controllers**", Virginia Polytechnic Institute PHD thesis Blacksburg, Virginia, 1997.

[3] Chun-Ta Chen, Mahmood R. Azimi-Sadjadi and Chunhua Yuan , "**Signal Representation Using Adaptive Wave-Net**", IEEE , Department of Electrical Engineering, Colorado State University, Fort Collins, International Conference on Vol.4, pp.2225-2230, 9-12 June 1997.

[4] Y. Oussar, I. Rivals, L. Personnaz, G. Dreyfus. , " **Training wavelet networks for nonlinear dynamic input-output modeling**", 1996. Laboratory of Electronic Superior School of Physical and Chemistry Industrial 10, rue Vauquelin F - 75231 PARIS Cedex 05, FRANCE.

[5] Y. Oussar, G. Dreyfus, "**Initialization by Selection for Wavelet Network Training**", Laboratoire d'Électronique Vol. 34, pp. 131-143 (2000).

[6] S. G. Mallat, "**A theory for Multiresolution Signal Decomposition: The Wavelet Representation**", IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 11, No 7. July 1989.

[7] Liu Zhigang , **"Multiwavelet Neural Networks Construction Study"**, IEEE, Vol.2, pp.789-792, 14-15 July 2005.