Journal of Engineering and Development, Vol. 18, No.6, November 2014, ISSN 1813- 7822

A Proposed Image Structure of Multiwavelet Network

Lecturer Dr. Wael Hussein Zayer
Electrical Engineering, computer and control, Amara Technical Institute
Wael_zayer@yahoo.com

Lecturer Dr. Ahmed Thamer Radi
Electrical Engineering, Power and Machine, Amara Technical Institute

Abstract:

The combination of wavelet theory and neural networks has lead to the development of wavelet networks. Wavelet networks are feed-forward neural networks using wavelets as activation function. Wavelet networks have been used in classification and identification problems with some success. The strength of wavelet networks lies in their capabilities of catching essential features in "frequency-rich" signals. In wavelet networks, both the position and the dilation of the wavelets are optimized besides the weights. Proposed multi wavelet network are used in identification problems of nonlinear systems. A multiwavelet network is constructed as an alternative to a neural network to approximate a nonlinear system.

Keywords: Wavelet transform, Multiwavelet Network, Neural network, Wavelet networks.

صورة مقترحة لهيكل شبكات متعدد الموجة

م.د. وائل حسين زاير
هندسة كهربائية / سيطره وحاسبات

م.د. احمد ثامر راضي
هندسة كهربائية / قدرة ومكائن

المعهد التقني-عمارة

الخلاصة:

كما معرف ان الدمج ما بين مفهوم نظرية الموجة و الشبكات العصبية يؤدي الى تكوين مفهوم جديد يسمى بالشبكات الموجية.شبكات الموجة هي شبكات عصبية بتغذية أمامية وباستخدام الموجة كدالة التفعيل فيها. تستخدم شبكات الموجة في تصنيف و تمثيل المشاكل بكل نجاح حيث تكمن أهمية الشبكات الموجة في إمكانيتها على اصطياد ترددات الاشارة وامكانية التعامل معها في مختلف المستويات وبكل دقة.

في هذا البحث تم اقتراح نموذج مطور للشبكات الموجة اسمه شبكات متعدد الموجة التي تحقق نتائج أفضل بكثير من شبكات الموجة وخصوصا مع الأنظمة اللاخطية وذلك لاستخدام اكثر من داله موجة واحدة في هذه المنظومة مما يؤدي الى التعامل مع تفاصيل ادق و الوصول اليها وبمختلف الترددات.

لذا يعتبر هذا النموذج المقترح هو امتداد للشبكات العصبية و للشبكات الموجة وبنتائج افضل بكثير في التمثيل و التصنيف للاشارة احادية البعد و متعددة الابعاد.

الكلمات المفتاحية: تحويل الموجة،شبكة متعدد الموجة،شبكة العصبية، شبكة الموجة.

1. Introduction:

Multiwavelet network is a network combining the ideas of the feed-forward neural networks and the wavelet decompositions. Zhang and Benveniste[1] provide an alternative to the feed-forward neural networks for approximating functions.

Multiwavelet networks use simple wavelets and multiwavelet network learning is performed by the standard back propagation type algorithm as the traditional neural network. Multiwavelet network can approximate any continuous functions on $[0,1]^n$ and have certain advantages such as the use of wavelet coefficients as the initial value for back propagation training and possible reduction of the network size while achieving the same level of approximation. In a feed-forward network, neurons take their inputs from the previous layer only and send the outputs to the next layer only. Since the signals go in one direction only, the network can compute a result very quickly. Basic neurons of a multiwavelet network are multidimensional wavelets and the neuron parameters are the dilation and translation coefficients. The output of a multiwavelet network is the linear combination of the values of several multidimensional wavelets.

Because of multiwavelet decomposition, wavelet networks provide universal approximation properties[2].

An explicit link exists between the network coefficients and appropriate transforms. This also derives from the properties of wavelet decomposition, translation, scale, and rotation may be used to design the constraints. Furthermore, the availability of direct and closed form formals for computing the continuous wavelet transform is helpful in initial guesses of the wavelet network coefficients [3].

2. Computation of Multiwavelet Network

By merging the Multiwavelet Transform theory with the basic concept of neural networks, a mapping network called Multiwavelet Network (MWN) or Multiwavenets is proposed as an alternative to feed forward neural networks and wavelet network for approximating arbitrary nonlinear functions.

It has been shown that Multiwavelet networks reveal the same approximation capabilities as traditional neural networks and wavelet networks, however they have better initialization characteristics due to the storing influence of wavelet theory in constructing the network.

The Multiwavelet network algorithms consist of two processes: the self construction of networks and the minimization of error. In the first process, the network structures applied to represent and determine using wavelet analysis. The network gradually recruits hidden units to effectively and sufficiently cover the time-frequency region occupied by a given target. Simultaneously, the network parameters are updated to preserve the network topology and take advantage of the later process.

In the second process, the approximations of instantaneous errors are minimized using an adaptation technique based on the Least Mean Square (LMS) algorithms. The parameter of the initialized network is updated using the steepest gradient-descent method of minimization.

Each hidden unit has a square window in the time-frequency plane. The optimization rule is only applied to the hidden units where the selected point falls into their windows. Therefore, the learning cost can be reduced[4,5] .

3. Proposal Method of Multiwavelet Network

The structure of Multiwavelet Network (MWN) is shown in **Figure (1)** where the input layer consist of two neurons since the input data consist of two dimension therefore the input layer can be consist from of neurons. If the input data are three dimension also the activation function is (one).The first hidden layer consists of six neurons each two of them take the minimum and then reach to the same neuron in the second hidden layer to select the output to the output layer.

The activation functions of the neurons in the first hidden layer must be different and two dimensional functions and the activation functions on the second hidden layer also must be different and one dimensional functions selected depending on the characteristics of the input date where this issue is very important in order to reach to the goal with minimum time and calculation. The output layer consists of one neuron and the activation function is the summation of the input to it.

The weights between the input and the hidden layer is equal to (one) but the weights between the second hidden layer and the output layer is modified according to the modification equations of it.

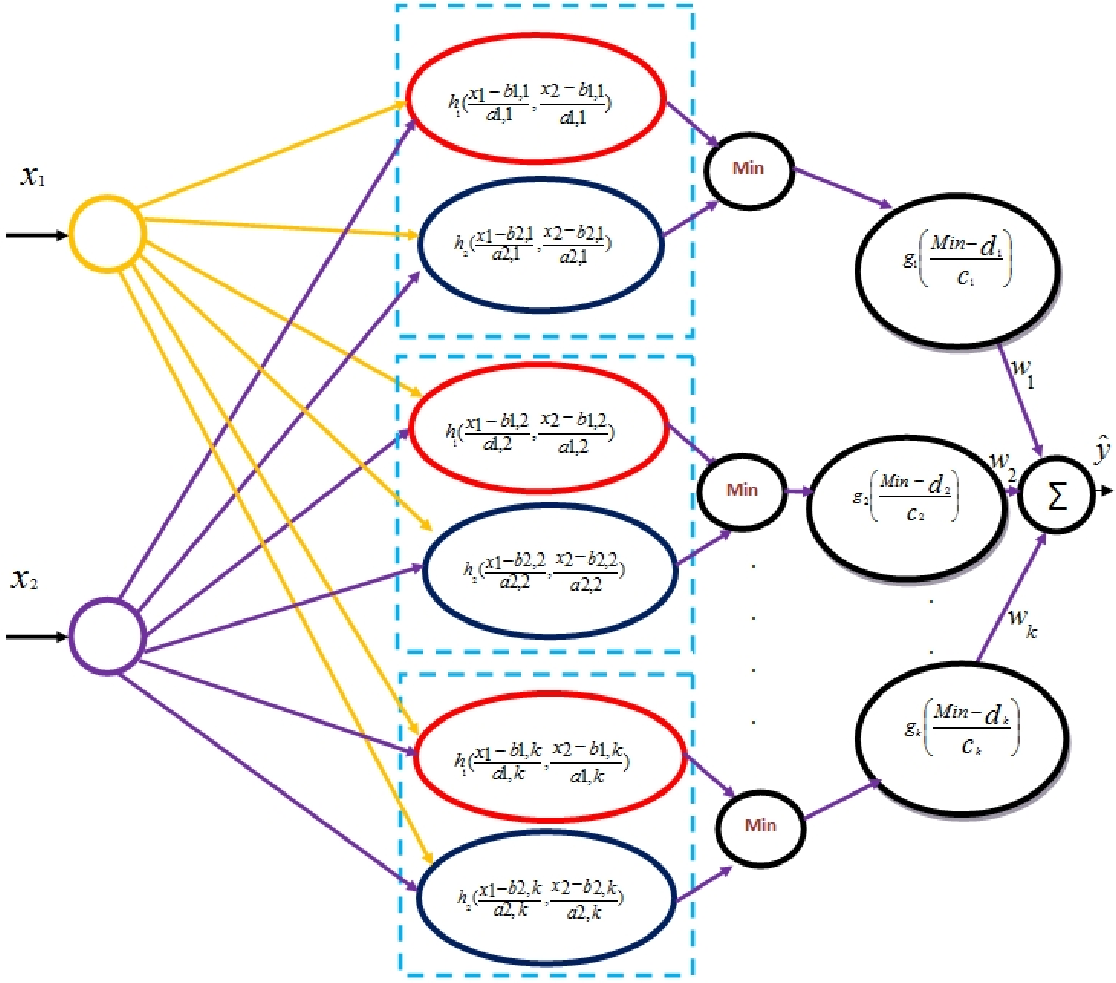


Fig .(1) The block diagram of the Proposed Algorithm of MWN.

4. Proposed Method for MWN single Output

The MWN architecture approximates any desired image y by generalizing a linear combination of two set of daughter wavelets $h_{1,a,b}(x1,x2)$ and $h_{2,a,b}(x1,x2)$, where the daughter wavelets $h_{1,a,b}(x1,x2)$ and $h_{2,a,b}(x1,x2)$ are generated by dilation, a, and translation, b, from two mother wavelets $h_1(t1,t2)$ *and* $h_2(t1,t2)$, where $t1=\dfrac{x1-b}{a}$, $t2=\dfrac{x2-b}{a}$.

The network architecture is shown in **Figure (1):**

$$h_{1,a,b}(x1,x2)=h_1\left(\frac{x1-b}{a},\frac{x2-b}{a}\right) \qquad \text{……………………. (1)}$$

$$h_{2,a,b}(x1,x2) = h_2\left(\frac{x1-b}{a}, \frac{x2-b}{a}\right) \qquad \text{……………………. (2)}$$

where

a : Dilation factor, with $a > 0$.

b: Translation factor.

$(x1, x2)$: are the two dimensions date.

A MWN is a 3-layers feed forward neural network. First the MWN parameters, dilation a's, translation b's, and weight w's should be initialized, and the desired sets of data, the input signal x is the data value, the desired output (target) y, the number of scaling functions p (p=2 in this work) and the number of wavelons$k$ are given(in our proposed method k=3).Assuming that the network output function satisfies the admissibility condition and the network sufficiently approximates the target. The approximated signal of the network $\hat{y}$ can be represented by equation:

$$\hat{y}(x_1, x_2) = \sum_{i=1}^{k} w_i g_{c_i, d_i}(Min) \qquad \text{……………………. (3)}$$

(x_1, x_2) are the two dimensions data.

w_i is the weight coefficients between hidden and output layers.

i=1,2,…, k. k is a number of wavelons in second hidden layer.

g_{c_i, d_i} is a daughter wavelets.

Similar to WN, after constructing the initial MWN and after calculating output signal of the network, the training of MWN starts. It is further trained by the gradient descent algorithms like least mean squares (LMS) to minimize the mean-squared error. During learning, the parameters of the network are optimized.

The MWN parameters $a_{j,i}$, and $b_{j,i}$ in the first hidden layer can be optimized in the LMS algorithm by minimizing a cost function or the energy function, E, over all function interval using equation (5).

$$E = \frac{1}{2}\sum_{t=1}^{T} e^2 \qquad \text{……………………. (4)}$$

$$E = \frac{1}{2}\sum_{t=1}^{T} (y - \hat{y})^2 \qquad \text{……………………. (5)}$$

The energy function is defined by equations (4) and (5), $y(x)$ is the desired output (target) and $\hat{y}(x)$ is the actual output signal of MWN.

To minimize E the method of steepest descent is used, which requires the gradients $\frac{\partial E}{\partial a_{j,i}}$, and $\frac{\partial E}{\partial b_{j,i}}$ for updating the incremental changes to each particular parameter $a_{j,i}$, and $b_{j,i}$, respectively. The gradients of E are:

$$\frac{\partial E}{\partial b_{j,i}} = -\sum_{t=1}^{T} E \times \frac{\partial h(t1,t2)}{\partial b_{j,i}} \qquad \text{………………….. (6)}$$

$$\frac{\partial E}{\partial a_{j,i}} = -\sum_{t=1}^{T} E \times (t1+t2) \frac{\partial h(t)}{\partial b_{j,i}} = (t1+t2) \frac{\partial E}{\partial b_{j,i}} \qquad \text{………………….. (7)}$$

$$t1 = \frac{x1 - b_{j,i}}{a_{j,i}}, t2 = \frac{x2 - b_{j,i}}{a_{j,i}} \qquad \text{…………………..(8)}$$

The incremental changes of each coefficient are simply the negative of their gradients as in equations (9) and (10). Thus each coefficients b, and a of the network is updated in accordance with the rule given in equations (11), (12).

$$\Delta b = -\frac{\partial E}{\partial b} \qquad \text{………………….. (9)}$$

$$\Delta a = -\frac{\partial E}{\partial a} \qquad \text{………………….. (10)}$$

$$b(t+1) = b(t) + m_b \Delta b \qquad \text{………………….. (11)}$$

$$a(t+1) = a(t) + m_a \Delta a \qquad \text{………………….. (12)}$$

Where ($m_b = m_a$ =0.1 or any value less than 1).

The training algorithm of the proposed multiwavelet network consists of the following six steps:

1) Initialize MWN parameters, dilation a's, translation b's, and weight w's, p=2, two mother wavelets filters $\left[h_1\left(\frac{x_1-b_i}{a_i},\frac{x_2-b_i}{a_i}\right),h_2\left(\frac{x_1-b_i}{a_i},\frac{x_2-b_i}{a_i}\right)\right]$, the desired sets of data, the input signal x, the desired output (target) y, and the number of wavelons is 3.
2) Set: the number of trainings, iter = 0, the incremental changes of each coefficient, $(\Delta a,\Delta b)$ =0, and the initial square error, $E_{iter}=0.1$ (or any value less than 1).
3) Calculate the approximated signal of the network $\hat{y}$ using equation (3).
4) Calculate the gradients of each coefficient using equations (6), (7) and calculate the coefficients incremental changes which are the negative of their gradients using equations (3).
5) Choose a constants $m_b, m_a = 0.1$ and calculate the new coefficients b_{iter+1}, and a_{iter+1} of the network in accordance with the rules given in equations (11), and (12).
6) Calculate the square error E_{iter+1} using equation (5).

If E_{iter+1} is small enough (as required or reach the desired value), then the training is good and the algorithm is stopped. Otherwise, set iter = iter + 1 and go to step (3).

In the second hidden layer

To minimize E the method of steepest descent is used, which requires the gradients $\frac{\partial E}{\partial w_i}$, $\frac{\partial E}{\partial c_i}$, and $\frac{\partial E}{\partial d_i}$ for updating the incremental changes to each particular parameter w_i, c_i, and d_i, respectively. The gradients of E are:

$$\frac{\partial E}{\partial w_i}=-\sum_{t=1}^{T}E\times g(t) \qquad \text{.......................} \ \mathbf{(13)}$$

$$\frac{\partial E}{\partial d_i}=-\sum_{t=1}^{T}E\times w_i\frac{\partial g(t)}{\partial d_i} \qquad \text{.......................} \ \mathbf{(14)}$$

$$\frac{\partial E}{\partial c_i}=-\sum_{t=1}^{T}E\times w_i t\frac{\partial g(t)}{\partial d_i}=t\frac{\partial E}{\partial d_i} \qquad \text{.......................} \ \mathbf{(15)}$$

Where

$$t = \frac{Min - d_i}{c_i} \qquad \text{......................(16)}$$

The derivatives of the various wavelet filter with respect to its translation, $\frac{\partial g(t)}{\partial d_i}$, are given [1].

Incremental changes of these coefficients are simply the negative of their gradients,

$$\Delta w = -\frac{\partial E}{\partial w}, \quad \Delta d = -\frac{\partial E}{\partial d}, \quad \Delta c = -\frac{\partial E}{\partial c} \qquad \text{...................... (17)}$$

Here these coefficients w_i, c_i and d_i, of the WN are updated in accordance with the following rules

$$\begin{aligned} w(n+1) &= w(n) - h_w \Delta w \\ d(n+1) &= d(n) - h_d \Delta d \\ c(n+1) &= c(n) - h_c \Delta c \end{aligned} \qquad \text{...................... (18)}$$

where h is the fixed learning rate parameter.

The incremental changes of each coefficient are simply the negative of their gradients as in equation (18). Thus each coefficient w, d, and c of the network is updated in accordance with the rule given in equations (13-15).

The training algorithm of the second hidden layer parameters are of the following six steps:

1) Initialize parameters: dilation c's, translation d's, and weight w's, p=1, two mother wavelets filters $\left[g\left(\frac{Min - d_i}{c_i}\right)\right]$, the desired sets of data, the input signal *x*, the desired output (target) *y*, and the number of wavelons is 3.
2) Set: the number of trainings, iter = 0, the incremental changes of each coefficient, $(\Delta w, \Delta c, \Delta d)$ =0, and the initial square error, $E_{iter} = 0.1$
3) Calculate the approximated signal of the network $\hat{y}$ using equation (3).

4) Calculate the gradients of each coefficient using equations (13), (14), (15) and calculate the coefficients incremental changes which are the negative of their gradients using equation (18).

5) Choose a constants $h_w, h_d, h_c = 0.1$ and calculate the new coefficients w_{iter+1}, d_{iter+1}, and c_{iter+1} of the network in accordance with the rules given in equation (18).

6) Calculate the square error E_{iter+1} using equation (5).

If E_{iter+1} is small enough(as required or reach the desired value),(E_{iter+1}=0.001), then the training is good and the algorithm is stopped. Otherwise, set iter = iter + 1 and go to (3).

Figure (2) shown the mechanism of the MWN training algorithm**.**

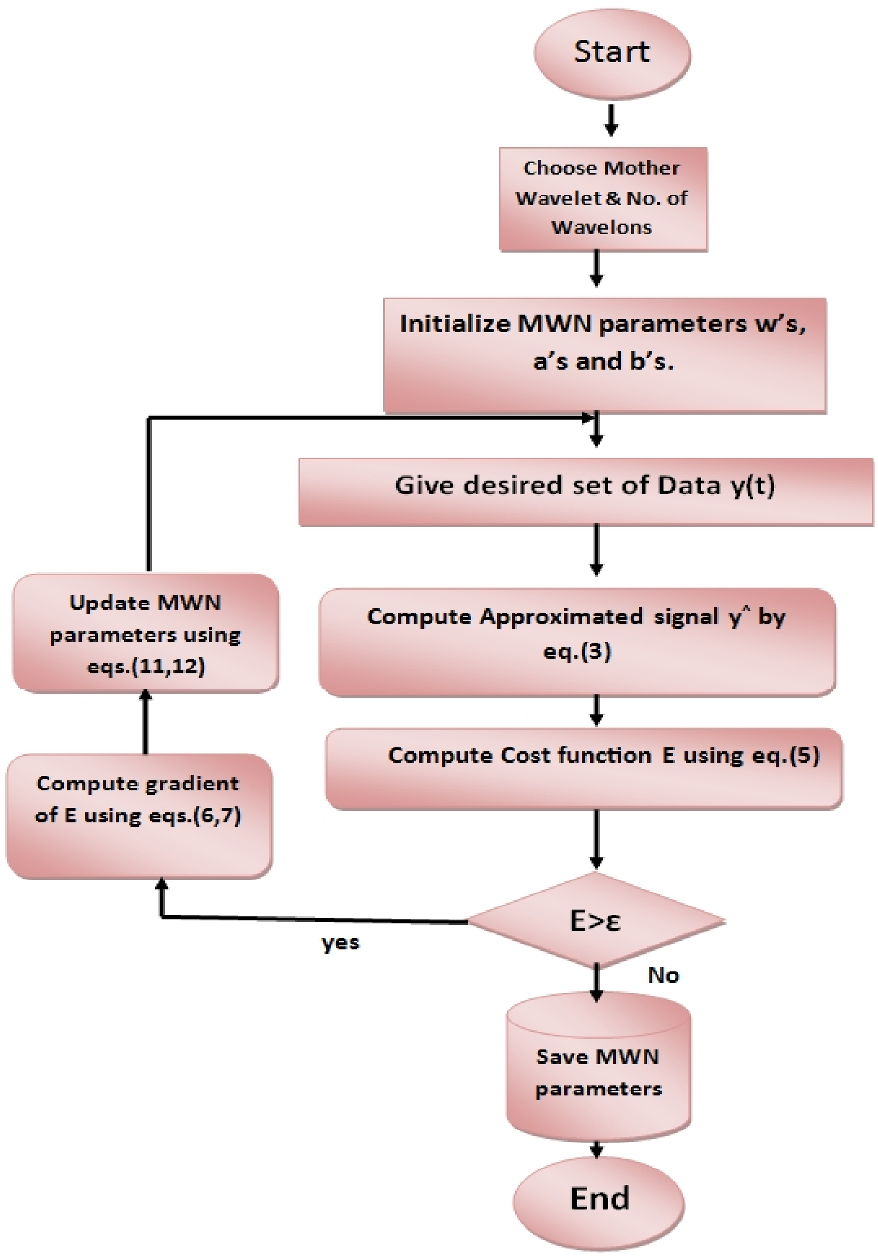


Fig .(2):The Mechanism of the MWN training Algorithm.

5. Conclusions:

Multiwavelet network is improved to wavelet network which is improved to neural network, therefore a comparison made among these three types. Since the structure of MWN is two dimensions therefore the comparison applied with input data two dimensions. The main points which make the difference are the time need to reach the aim and the number of neurons in the hidden layer. As the time of reach is less than and the number of neuronsare less the structure is better than the others.

Table (1) shows comparison among three types applied on two dimension image, from this table shown that the number of iteration is reduced to become approximately 10% of the wavelet network method and neural network method which mean less time need to identify.

Table .(1) Approximation results of two dimensions function.

Method	Number of units	Number of parameters	Number of iteration	Error
Multiwavelet Network	49	343	≈ 4000	0.001
Wavelet Network	49	442	≈ 40000	0.001
Neural Network	225	1126	≈ 40000	0.001

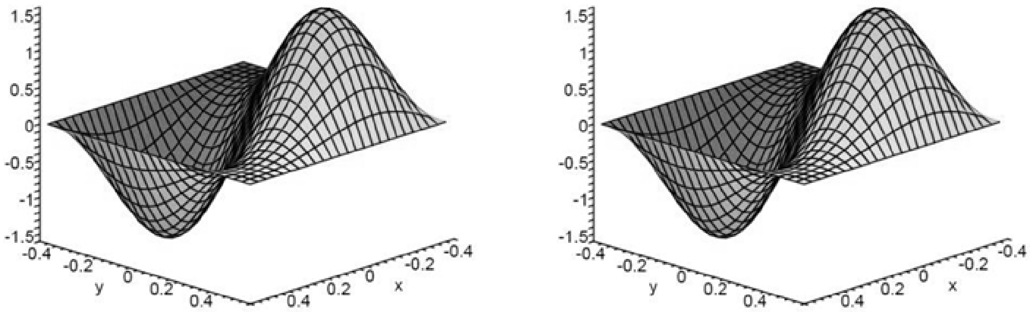
Simulation results demonstrate that the proposed MWN is quite effective in identification and control application.

In **Table (2)** the same units and parameters in **Table (1)** used with make the number of iteration is constant (1000) in order to find the error at this number of iteration,

Table .(2) Approximation results of two dimensions function with constant number of iteration.

Method	Number of units	Number of parameters	Number of iteration	Error
Multiwavelet Network	49	343	1000	0.034
Wavelet Network	49	442	1000	0.9511
Neural Network	225	1126	1000	0.9069

Figures (3-7) are examples of Matlab functions and images using MWN to approximate it.



(a) Original function **(b)approximate function**

Fig .(3) Approximation results of function=cos(πx)sin(2πy)-sin(2πx)cos(πy).

(a) Original Image(512*512) **(b) Approximate image.**

Fig .(4) Approximation results of Barbara image.

(a) Original Image(512*512) **(b) Approximate image.**

Fig .(5) Approximation results of Peppers image.


(a) Original Image(512*512) (b) Approximate image.

Fig .(6) Approximation results of Girl image.


(a) Original Image(512*512) (b) Approximate image.

Fig .(7) Approximation results of Lena image.

6. References

1. Zhang, Q., Benveniste, A., " Wavelet Networks", IEEE Transactions on Neural Networks, Vol. 3, No. 6, 1992.
2. Mahmoud, W., Al-Taai, H., "New Fast Method For Computing Multiwavelet coefficient From 1D up to 3D pro. Of 1^{st} Int. Conf on Digital Com,Comp.,App.,PP.412-422,Jordan,2007.
3. Wael H., "Person Identification Using Face and Fingerprint Authentication", Ph.D. thesis, University of Basrah, 2011.

4. **Mahmoud, W.,KattoushA., and NihadS.,"The performance of Multiwavelets based OFDM system under different channel conditions" International Journal of Digital Signal Processing, Elsevier Institute, Vol.20,No.2,PP 472-482,USA,2010.**
5. **Miller, B.E., Colgate, J. E.," Using A Wavelet Network to Characterize Real Environments for Haptic Display", e-mail :blemillerfgnwu.edu.2006.**