Face Identification Using
Back-Propagation Adaptive Multiwavenet

Prof. Dr. Waleed Ameen Mahmoud
University of Baghdad/ College of Engineering
Electrical Engineering Department

Ali Ibrahim Abbas
University of Baghdad/ College of Engineering
Electrical Engineering Department

Asst. Prof. Nuha Abdul Sahib Alwan
University of Baghdad/ College of Engineering
Computer Engineering Department

Abstract
Face Identification is an important research topic in the field of computer vision and pattern recognition and has become a very active research area in recent decades. Recently multiwavelet-based neural networks (multiwavenets) have been used for function approximation and recognition, but to our best knowledge it has not been used for face Identification. This paper presents a novel approach for the Identification of human faces using Back-Propagation Adaptive Multiwavenet. The proposed multiwavenet has a structure similar to a multi-layer perceptron (MLP) neural network with three layers, but the activation function of hidden layer is replaced with multiscaling functions. In experiments performed on the ORL face database it achieved a recognition rate of 97.75% in the presence of facial expression, lighting and pose variations. Results are compared with its wavelet-based counterpart where it obtained a recognition rate of 10.4%. The proposed multiwavenet demonstrated very good recognition rate in the presence of variations in facial expression, lighting and pose and outperformed its wavelet-based counterpart.

Keywords: Face Identification, multiwavelet neural network, Back-Propagation Adaptive Multiwavenet

.

.

(Back‑Propagation Adaptive Multiwavenet)

(MLP)   .

ORL   %97.75   .

.%10.4   .

.

Prof. Dr. Waleed Ameen Mahmoud
Ali Ibrahim Abbas
Asst. Prof. Nuha Abdul Sahib Alwan

FACE IDENTIFICATION USING BACK-PROPAGATION
ADAPTIVE MULTIWAVENET

[

1. Introduction

Over the last decade, face recognition has become a popular area of research in computer vision. A general statement of the face recognition problem can be formulated as follows: Given still or video images of a scene, identify or verify one or more persons in the scene using a stored database of faces. A survey of face recognition techniques has been given by (Zhao W., et al., 2003).

In general, face recognition techniques can be divided into two groups based on the face representation they use:

- Appearance-based, which uses holistic texture features and is applied to either whole-face or specific regions in a face image;
- Feature-based, which uses geometric facial features (mouth, eyes, brows, cheeks etc.) and geometric relationships between them.

Among various solutions to the problem (Turk M., et al., 1991) the most successful seems to be appearance-based approaches, which generally operate directly on images or appearances of face objects and process the image as two-dimensional patterns. Most effort in the literature have focused mainly on developing feature extraction methods (Zhao W., et al., 2003), (Zhao W., et al., 1997), (Deniz O., et al., 2003) and employing powerful classifiers such as probabilistic (Moghaddam B., 2002), Hidden Markov models (HMMs) (Othman H., et al., 2003) neural networks (NNs) (Er M.J., et al., 2003), (Er M.J., et al., 2005) and support vector machine (SVM) (Lee K., et al., 2002).

The use of neural networks for face recognition has been addressed in (Pang S., et al., 2005), (Zhang B., et al., 2004), (Fan X., et al., 2005), (Lu X., et al., 2003). Recently, (Li G., et al., 2006) suggested the use of a non-convergent chaotic neural network to recognize human faces. (Lu K., et al., 2006) suggested a semi-supervised learning method that uses support vector machines for face recognition. (Zhou W., et al., 2006) suggested using a radial basis function neural network that is integrated with a non-negative matrix factorization to recognize faces. (Huang L.L., et al., 2006) proposed using two neural networks whose outputs are combined to make a final decision on classifying a face. (Park C., et al., 2006) used a momentum back propagation neural network for face and speech verification.

Back Propagation Wavenet (BPW) is an artificial neural network (ANN) that is integrated with wavelet techniques and has been used successfully in many fields. Instead of conventional nonlinear sigmoid transfer functions, the transfer function of the nodes in a wavelet neural network is wavelet bases. Because wavelet bases are localized in time and frequency, the ability of a BPW in mapping complicated nonlinear functions is enhanced considerably (Yang X., et al., 2009).

(Shen Y., et al., 2004) used BPW for object recognition. (Ensafi A.A., et al., 2007) applied BPW for the determination of sulfide and thiocyanate in real samples such as tap, waste and river waters with satisfactory results. to improve the detection rate for anomaly state and reduce the false positive rate for normal state in the network anomaly detection. (Liu L., et al., 2009) proposed a novel method based on BPW trained by Modified Quantum-behaved Particle Swarm Optimization (MQPSO) algorithm. (Long-yun X., et al., 2008) used BPW for gear faults diagnosis. (Zhao Y.Z., et al., 2009) used BPW for learning the wear out pattern of the milling machine cutters to predict their remaining useful life. (Bin Z., et al., 2010) applied BPW to solve the problem of tunnel surrounding rock deformation prediction.

As an extension of wavelets, a multiwavelet can preserve all the advantages the wavelet has. Furthermore, it can simultaneously have several properties very useful in practical applications such as orthogonality, regularity, symmetry, and compact support, which is impossible for a scalar wavelet. Therefore, the networks using the dilations and translations of a multiscaling function as node functions have better performances which are worth investigating (Jiao L.C., et al. 2001). Based on the considerations above, a model of multiwavelet-based neural network for face identification is used in this paper.

Section 2 provides an introduction of the Multiwavenet. Section 3 presents a face identification system based on a back-propagation multiwavenet classifier. Section 4 presents experimental results. Conclusions are presented in Section 5.

2. Multiwavenet

Suppose a multiplicity-r multiscaling function $\phi(t) = \left[\phi^1(t), \phi^2(t), \ldots, \phi^r(t) \right]^T$ satisfies a dilation equation

$$\phi(t) = \sum_{k \in Z} p_k \phi(2t - k) \qquad (1)$$

and the dilations and translations of $\phi^l(t)$s, denoted by $\phi^l_{j,k}(t) = 2^{j/2}\phi^l_{j,k}(2^j t - k)$, $l = 1, \ldots, r$ $k \in Z$, span the scale space V_j $j \in Z$.

We are interested in the case of orthogonal multiscaling functions. In that case, $\{\phi^l_{j,k}(t) \mid l = 1, \cdots, r, \quad k \in Z\}$ form an orthonormal basis of V_j, and the associated multiwavelet $\psi(t) = [\psi^1(t), \psi^2(t), \ldots, \psi^r(t)]^T$ make $\psi^l_{j,k}(t) = 2^{j/2}\psi^l_{j,k}(2^j t - k)$, $l = 1, \ldots, r$, $k \in Z$, form an orthonormal basis of the orthogonal complementary subspace W_j of V_j in V_{j+1}. From the theory of multiresolution analysis, we know that $\bigcup_{j \in z} V_j = L^2(R)$ and hence, for any $f \in L^2(R)$, there exists a natural number J_0, such that $\|f - f_J\|_2 < \varepsilon$, $J > J_0$ where $\|\bullet\|_2$ is L^2 norm, ε is an arbitrary positive number, and $f_J \in V_J$

$$f_J = \sum_{l=1}^{r}\sum_{k \in z}\langle f, \phi^l_{J,k}\rangle \phi^l_{J,k}(t). \tag{2}$$

From the viewpoint of neural networks, f_J can be learned by a neural network, which is called multiwavelet neural network (multiwavenet) because of its connection with the multiwavelet theory though the node functions used are the associated multiscaling functions (Jiao L.C., et al. 2001).

3. The Proposed Face Identification System

The general structure of the proposed face identification system is shown in **Fig. 1**.

Function of each block is clarified below:

- **Preprocessing:** Locating faces in image and extracting the face part only. This is done manually.
- **Normalization:** Histogram equalization of images to reduce the effect of lightning source amplitude's variations.
- **Feature Extraction:** Down sample the image to a low resolution using a bicubic interpolation method and put the result in a one-dimensional feature vector.
- **Classifier:** Multiwavenet classifier is used. After the training is done, the parameters of the multiwavenet classifier, namely

weights, translations and dilations are stored and are used in the identification stage when the classifier works in simulation mode.

The processing in both training and identification stages is similar except for the classifier, which works in training mode in the training stage and in simulation mode in the identification stage.

In this section, a classifier for the face identification system is proposed. Each output of the multiwavenet classifier corresponds to a class. The position of the output with the highest value determines the class to which the input belongs. For example if there are five classes, five outputs are needed. During the training phase, each data in the training set is presented to the classifier to the classifier along with its desired output, where desired output for an input belonging to class c is an all zero vector with a one in the cth position. When a new input is presented to the classifier, the position of the highest value output determines the class.

The architecture of the proposed Back Propagation Adaptive Multiwavenet (BPAMW) is basically the same as the back propagation neural network, except that the sigmoid function of hidden layer node of the back propagation neural network is replaced with two or more scaling functions of a multiwavelet system. For each distinct individual in the training set, an output is needed and for each output a set of weights is required. A competitive layer is added after the output layer. Function of this layer is to produce the final classification result by choosing the output with the highest value among all outputs as the winner and returning its position as the final classification. The architecture of the BPAMW classifier is shown in **Fig. 2**.

The kth output of the BPAMW classifier before the competitive layer is

$$y_k(U_s) = \sum_{i=1}^{M}\sum_{L=1}^{r} w_{iLk}\varphi_L\left(\frac{z_{is} - t_i}{\lambda_i}\right) \tag{3}$$
$$k = 1\cdots K, \quad s = 1\cdots P$$

$$z_{is} = \sum_{j=1}^{N} v_{ji}u_{js}, \tag{4}$$

Prof. Dr. Waleed Ameen Mahmoud
Ali Ibrahim Abbas
Asst. Prof. Nuha Abdul Sahib Alwan

FACE IDENTIFICATION USING BACK-PROPAGATION
ADAPTIVE MULTIWAVENET

where M is the number of multiwavelons, r is the multiplicity of the multiscaling function and each multiwavelon has r wavelons, t_i and λ_i are the translation and dilation of ith multiwavelon's scaling functions respectively, φ_L is the L'th scaling function. $U_s = \{u_{1s}, u_{2s}, ..., u_{Ns}\}$ is the sth input vector of the total P input vectors in the training set, N is the number of elements of each input vector (input dimension), z_{is} is the inner product between the input vector U_s and the ith input weight vector $V_i = \{v_{1i}, v_{2i}, ..., v_{Ni}\}$ (weights between input nodes and ith multiwavelon), w_{iLk} is the weight between Lth wavelon of ith multiwavelon and the kth output, $y_k(U_s)$ is the kth output of the network and there are K nodes in the output layer, K should be equal to the number of individuals in the training database.

(Plonka G., et al. 1998) presented an efficient method for creating multi-scaling functions with given approximation order, regularity, symmetry and short support. In second example of their paper, multi-scaling functions with approximation order 4, compact support (0,2), and symmetry was constructed which will be used in this work. Unlike GHM (Geronimo-Hardin-Massopust) multi-scaling functions, these functions have a closed form expression:

$$\phi_1(u) = \begin{cases} \left(-2u^3 + 3u^2\right) & u \in [0,1) \\ (2-u)^2(2u-1) & u \in [1,2] \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

$$\phi_2(u) = \begin{cases} -u^2(3u-3) & u \in [0,1) \\ -(2-u)^2(3u-3) & u \in [1,2] \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

To calculate partial derivatives, the derivative of ϕ_L is also required. Each ϕ_L is composed of two polynomial functions, to obtain the derivative of ϕ_L each polynomial is differentiated as an independent function and results are combined.

$$\frac{d\phi_1(u)}{du} = \begin{cases} -6u(u-1) & u \in [0,1) \\ 6(u-1)(-2+u) & u \in [1,2] \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

$$\frac{d\phi_2(u)}{du} = \begin{cases} -3u(3u-2) & u \in [0,1) \\ -3(3u-4)(-2+u) & u \in [1,2] \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

The output of the competitive layer is

$$c = \arg\max_{k \in \{1 \cdots K\}} (y_k) \quad (9)$$

where y_k is the kth output and K is the total number of outputs.

Gradient descent method is used for training of the network parameters. The objective function to be minimized is

$$C_k = \frac{1}{2P} \sum_{s=1}^{P} (y_k(U_s) - f_k(U_s))^2, \quad k = 1 \ldots K \quad (10)$$

where P is the number of training pairs, $y_k(U_s)$ is the kth output of the BPAMW classifier before the competitive layer and $f_k(U_s)$ is the kth desired output. The competitive layer is not included in the training, because it has a fixed function that does not need training and only operates when the classifier is in identification stage where it determines the final classification result.

Batch training mode is used where all training pairs $\{U_s, f_k(U_s)\}$, $s = 1, ..., P$ should be processed before parameters could be updated. Parameters are modified in the opposite direction of the gradient of C_k. To speed up the convergence rate, momentum term is included in parameter's update. Let

$$\tau_{is} = \frac{z_{is} - t_i}{\lambda_i}, \quad (11)$$

$$\phi_L(\tau_{is}) = \phi_L(\frac{z_{is} - t_i}{\lambda_i}), \text{ and} \quad (12)$$

$$e_{sk} = y_k(U_s) - f_k(U_s). \quad (13)$$

Partial derivatives are expressed as follows:

$$\frac{\partial C}{\partial w_{iLk}} = \sum_{s=1}^{P} e_{sk}\phi_L(\tau_{is}) \qquad (14)$$

$$\frac{\partial C}{\partial v_{ji}} = \sum_{s=1}^{P}\sum_{k=1}^{K}\sum_{L=1}^{r} e_{sk}\, w_{iLk}\frac{\partial \phi_L(\tau_{is})}{\partial \tau_{is}} u_{js}\lambda_i^{-1} \qquad (15)$$

$$\frac{\partial C}{\partial t_i} = -\sum_{s=1}^{P}\sum_{k=1}^{K}\sum_{L=1}^{r} e_{sk}\, w_{iLk}\frac{\partial \phi_L(\tau_{is})}{\partial \tau_{is}}\lambda_i^{-1} \qquad (16)$$

$$\frac{\partial C}{\partial \lambda_i} = \sum_{s=1}^{P}\sum_{k=1}^{K}\sum_{L=1}^{r} e_{sk}\, w_{iLk}\frac{\partial \phi_L(\tau_{is})}{\partial \tau_{is}}\tau_{is}\lambda_i^{-1}. \qquad (17)$$

Parameters can be updated as follows:
h = iteration number

$$w_{iLk}^{h+1} = w_{iLk}^{h} - \eta\frac{\partial C}{\partial w_{iLk}} + \alpha\Delta w_{iLk}^{h} \qquad (18)$$

$$v_{ji}^{h+1} = v_{ji}^{h} - \eta\frac{\partial C}{\partial v_{ji}} + \alpha\Delta v_{ji}^{h} \qquad (19)$$

$$t_i^{h+1} = t_i^{h} - \eta\frac{\partial C}{\partial t_i} + \alpha\Delta t_i^{h} \qquad (20)$$

$$\lambda_i^{h+1} = \lambda_i^{h} - \eta\frac{\partial C}{\partial \lambda_i} + \alpha\Delta\lambda_i^{h}. \qquad (21)$$

where $\Delta x^h = x^h - x^{h-1}$.

Parameter initialization has a significant impact on the convergence rate of the BPAMW. A heuristic method for parameter initialization is proposed here. The following equations express the parameter initialization.

$v_{ji} = rand() * 4 - 2$ (22)
for $j = \{1,\ldots,N\}$, $i = \{1,\ldots,M\}$

$w_{iL} = 0$ (23)
for $i = \{1,\ldots,M\}$, $L = \{1,\ldots,r\}$

$$U_{max} = \max_{s\in\{1,\cdots,P\}}(U_s) \qquad (24)$$

$$U_{min} = \min_{s\in\{1,\cdots,P\}}(U_s) \qquad (25)$$

$$z_{i\,max} = V_i^T \cdot U_{max} \quad \text{(inner product)} \qquad (26)$$

$$z_{i\,min} = V_i^T \cdot U_{min} \quad \text{(inner product)} \qquad (27)$$

$$\lambda_i = \frac{z_{i\,max} - z_{i\,min}}{2} \qquad (28)$$

$$t_i = \frac{z_{i\,max} + z_{i\,min}}{2} \qquad (29)$$

where *rand*() is a function that generates random number in the range (0, 1).

The BPAMW classifier's training is needed to be done only once and after that, it can recognize new and unseen images of the persons in the training set. When operating in the identification mode, the classifier uses the stored parameters from training stage to calculate the outputs. The index of the output with the highest value is the class number of the input that is calculated by the competitive layer.

4. Experimental Results

AT&T "The Database of Faces" (formerly "The ORL Database of Faces") database contains 400 grayscale images of 40 persons. Each person has 10 images, each having a resolution of 92 x 112 and 256 gray levels. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement). The first three images from each individual of the database are selected for the test set and the rest of the face images are included in the training set. Therefore, a total of 280 face images are used for training and another 120 face images are used for testing. A sample of the face images are shown in **Fig. 3**.

Images were resized to size 6*7 and put in a one-dimensional vector, then pixel values were normalized to range (0,1) by dividing each element of the vector by the maximum value of the vector.

For the BPAMW classifier; number of multiwavelons was set to 40, the iteration number

823

Prof. Dr. Waleed Ameen Mahmoud
Ali Ibrahim Abbas
Asst. Prof. Nuha Abdul Sahib Alwan

FACE IDENTIFICATION USING BACK-PROPAGATION
ADAPTIVE MULTIWAVENET

[

was set to 5 and the learning rate was 0.01. The momentum term coefficient was 0.9. Results are summarized in **Table 1**. **Fig. 4** shows the training performance of the network. An average recognition rate of 97.75% was obtained over ten trials. Each trial involves training and testing. Since training uses the *rand()* function for parameter initialization, the final stored parameters upon convergence may differ slightly. This is why averaging the recognition rate over ten trials is performed to achieve stochastic ensemble averaging.

Table 1 Result of the BPAMW Classifier over ORL Database

Average Training MSE	Average Testing MSE	Average Recognition Rate
0.0034	0.0034	97.75

To compare the performance of the proposed BPAMW classifier with the BPW classifier, the same experiments were performed with the same parameters. Results are summarized in **Table 2**.

Fig. 5 shows the training performance of the network. An average independent recognition rate of 10.4% were obtained over ten trials.

Table 2 Result of the BPW Classifier over ORL Database

Average Training MSE	Average Testing MSE	Average Recognition Rate
0.2490	0.9846	10.4167

The proposed BPAMW classifier has a very good recognition rate of 97.75% in a very few iterations, which indicates that it has a very good recognition rate when there are variations in the facial expression and limited pose variations. Also the results show that BPAMW classifier is superior to BPW classifier because it achieved a higher recognition rate in the same number of iterations.

5. Conclusions

In this paper a classifier has been proposed for identification of human faces. The classifier is based on multiwavelet neural network. As an extension of wavelets, a multiwavelet can preserve all the advantages the wavelet has. Furthermore, it can simultaneously have several properties very useful in practical applications such as orthogonality, regularity, symmetry, and compact support, which is impossible for a scalar wavelet. Therefore, the network using the multiwavelet as activation functions has a better performance than the same network with wavelets as activation functions.

Experiments that were performed on AT&T "The Database of Faces" (formerly "The ORL Database of Faces") showed that BPAMW classifier has a very good recognition rate of 97.75% in the presence of facial expression, lighting and pose variations.

References

Bin Z., De-yuan H., and Tao L., "Research on the Application of Wavelet Neural Network in the Surrounding Rock Displacement Prediction," 2nd International Asia Conference on Informatics in Control, Automation and Robotics (CAR 2010), Wuhan, China, pp. 266-269, 6-7 March 2010.

Deniz O., Castrillfion M., and Hernfiandez M., "Face recognition using independent component analysis and support vector machines", Pattern Recognition letters, vol. 24, pp. 2153-2157, 2003.

Ensafi A.A., Khayamian T., and Tabaraki R., "Simultaneous Kinetic Determination of Thiocyanate and Sulfide Using Eigenvalue Ranking and Correlation Ranking in Principal Component-Wavelet Neural Network," Talanta, Vol. 71, Issue 5, pp. 2021-2028, March 2007.

Er M.J., Chen W., and Wu S., "High speed face recognition based on discrete cosine transform and RBF neural network", IEEE Trans. On Neural Network, vol. 16, no.3, pp. 679-691, 2005.

Er M.J., Wu S., Lu J., and Toh L.H., "face recognition with radial basis function (RBF) neural networks", IEEE Transactions on Neural Networks, vol. 13, no. 3, pp. 697-710, 2003.

Fan X., and Verma B., "A Comparative Experimental Analysis of Separate and Combined Facial Features for GA-ANN based Technique," Sixth International Conference on Computational

Intelligence and Multimedia Applications, pp. 279-284, Aug. 2005.

Huang L.L., and Shimizu A., "Combining Classifiers for Robust Face Detection," In Lecture Notes in Computer Science, 3972, Springer-Verlag, pp. 116-121, 2006.

Jiao L.C., Pan J., and Fang Y.W., "Multiwavelet neural networks and its approximation properties," IEEE Transactions on Neural Networks, Vol. 12, Issue 5, pp. 1060-1066, 2001.

Lee K., Chung Y., and Byun H., "SVM based face verification with feature set of small size", Electronic letters, vol. 38, no.15, pp. 787-789, 2002.

Li G., Zhang J., Wang Y., and Freeman W.J., "Face Recognition Using a Neural Network Simulating Olfactory Systems," In Lecture Notes in Computer Science, 3972, Springer-Verlag, pp. 93-97, 2006.

Liu L., and Liu Y., "MQPSO based on wavelet neural network for network anomaly detection," 5th International Conference on Wireless Communications, Networking and Mobile Computing (WiCom'09), Beijing, China, pp. 1-5, 24-26 Sep. 2009.

Long-yun X., Zhi-yuan R., and Rui-cheng F., "Gear Faults Diagnosis Based on Wavelet Neural Networks," Proceedings of IEEE International Conference on Mechatronics and Automation, Takamatsu, Japan, pp. 452-455, 5-8 Aug. 2008.

Lu K., He X., and Zhao J., "Semi-supervised Support Vector Learning for Face Recognition," In Lecture Notes in Computer Science, 3972, Springer-Verlag, pp. 104-109, 2006.

Lu X., Wang Y., and Jain A.K., "Combining Classifiers for Face Recognition," Proceedings of International Conference on Multimedia and Expo (ICME '03), Vol. 3, pp. 13-16, 2003.

Moghaddam B., "Principal manifolds and probabilistic subspaces for visual recognition", IEEE Transactions on pattern Anal. Machine Intelligence, vol. 24, no.6, pp. 780-788, 2002.

Othman H., and Aboulnasr T., "A separable low complexity 2D HMM with application to face recognition", IEEE Trans. Pattern. Anal. Machie Inell., vol. 25, no.10, pp. 1229-1238, 2003.

Pang S., Kim D., and Bang S.Y., "Face Membership Authentication Using SVM Classification Tree Generated by Membership-Based LLE Data Partition". In IEEE Transactions on Neural Networks, Vol. 16, Issue 2, pp. 436-446, 2005.

Park C., Ki M., Namkung J., and Paik J.K., "Multimodal Priority Verification of Face and Speech Using Momentum Back-Propagation Neural Network," In Lecture Notes in Computer Science, 3972, Springer-Verlag, pp. 140-149, 2006.

Plonka G., and Strela V., "Construction of Multi-Scaling Functions with Approximation and Symmetry," SIAM Journal on Mathematical Analysis, Vol. 29, Issue 2, pp. 481-510, March 1998.

Shen Y., Wang Q., and Yu S., "A Target Recognition of Wavelet Neural Network Based on Relative Moment Features," Proceedings of the 5th World Congress on Intelligent Control and Automation, Hangzhou, P.R. China, Vol. 5, pp. 4089–4091, 15-19 June 2004.

Turk M., and Pentland A., "Eigen faces for face recognition", Journal cognitive neuroscience, vol. 3, no.1, 1991.

Yang X., Kumehara H., and Zhang W., "Back Propagation Wavelet Neural Network Based Prediction of Drill Wear from Thrust Force and Cutting Torque Signals," Canadian Center of Science and Education (CCSE), Computer and Information Science Journal, Vol. 2, Issue 3, pp. 75-86, Aug. 2009.

Zhang B., Zhang H., and Ge S., "Face Recognition by Applying Wavelet Subband Representation and Kernel Associative Memory," In IEEE Transactions on Neural Networks, Vol. 15, pp. 166-177, 2004.

Zhao W., Chellappa R., and Krishnaswamy A,, "Discriminant analysis of principal component for face recognition", IEEE Transactions on Pattern Anal. Machine Intelligence, vol. 8, 1997.

Prof. Dr. Waleed Ameen Mahmoud
Ali Ibrahim Abbas
Asst. Prof. Nuha Abdul Sahib Alwan
[

FACE IDENTIFICATION USING BACK-PROPAGATION
ADAPTIVE MULTIWAVENET

Zhao W., Chellappa R., Phillips J., and Rosenfeld A., "Face recognition in still and video images: A literature survey", ACM Comput Surv vol. 35, pp. 399–458, 2003.

Zhao Y.Z., and Li X., "A Dual-Momentum Hybrid Wavelet Neural Net (DM-HWNN): Its Performance Evaluation and Application," 7th IEEE International Conference on Industrial Informatics (INDIN 2009), Cardiff, Wales, UK, pp. 325-330, 23-26 June 2009.

Zhou W., Pu X., and Zheng Z., "Parts-Based Holistic Face Recognition with RBF Neural Networks," In Lecture Notes in Computer Science, 3972, Springer-Verlag, pp. 110-115, 2006.

List of Symbols:

|	
$C(\cdot)$	Objective Function to be Minimized in a Neural Network
$f(\cdot)$	Desired Output
J	Resolution Level of Multiscaling Function
K	Number of Output Nodes
L	Index of the Scaling Functions of a Multiscaling Function
M	Number of Wavelons or Multiwavelons
N	Number of Input Nodes (Input Dimension)
P	Number of Training Patterns
U	Input Vector
V	Input Weight Matrix
Z	The Set of All Integers
C	Output of the Competitive Layer
E	Difference Between Desired Output and Actual Output (Error)
H	Iteration Number (Epoch)
I	Index of the Wavelon or Multiwavelon
K	Index of Output Nodes
R	Multiplicity of Multiwavelet
s	Index of the Training Samples
T	Translation
U	Scalar Input
V	Weight Between the Input Layer and the Hidden Layer
W	Output Weight
Y	Output
z	Inner Product Between the Input Vector and the Input Weight Vector
Δ	Difference Operator
$\sum$	Summation Operator
$\Phi(\cdot)$	Multiscaling Function
$\Psi(\cdot)$	Multiwavelet Function
α	Momentum Coefficient
ε	Training Error Threshold
η	Learning Rate
λ	Dilation
τ	Input to the Wavelet or Scaling Function
$\phi(\cdot)$	Father Scaling Function
$\psi(\cdot)$	Mother Wavelet Function
$\psi'(\cdot)$	Derivative of Wavelet Function

Training
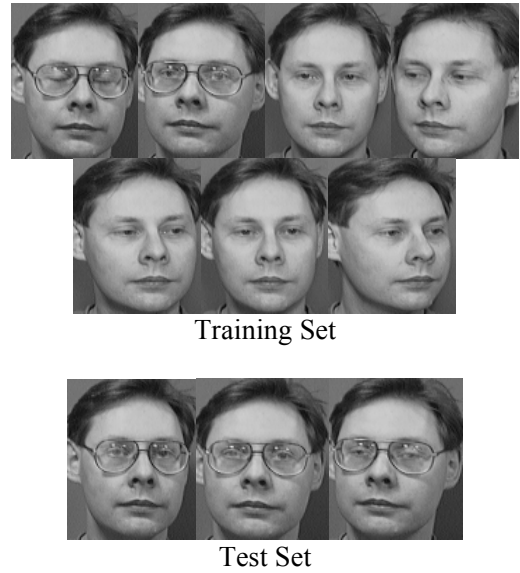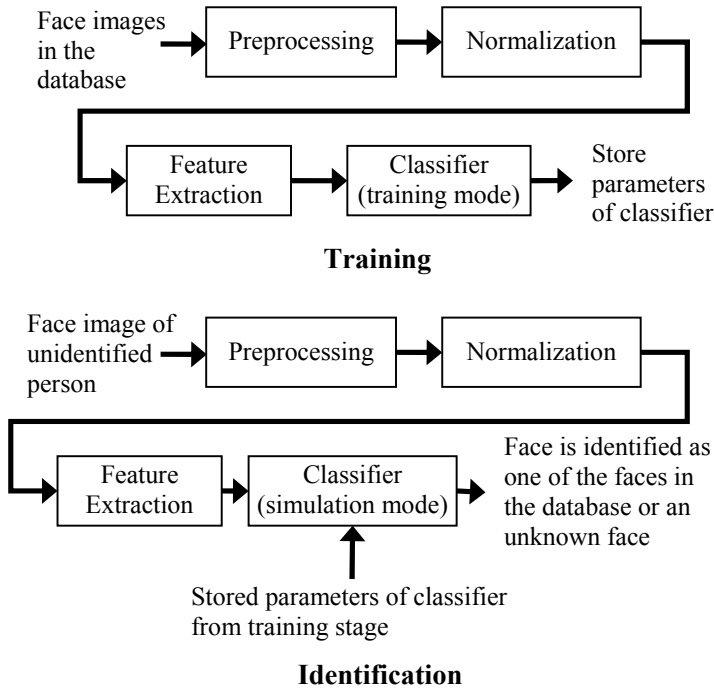
Identification

Fig. 1 The Proposed Face Identification System Block Diagram

Training Set

Test Set

Fig. 3 Sample Images from the ORL Face Database

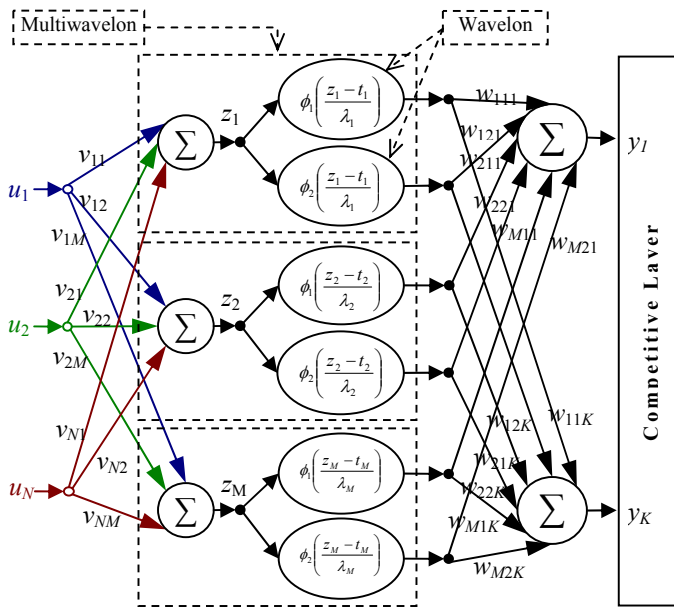**Fig. 2** The proposed BPAMW Classifier
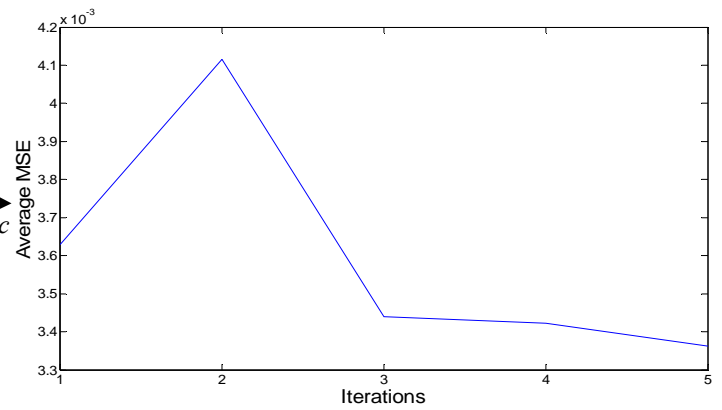
Fig. 4 BPAMW Classifier Performance over ORL Database: Average MSE vs. Iterations

Prof. Dr. Waleed Ameen Mahmoud
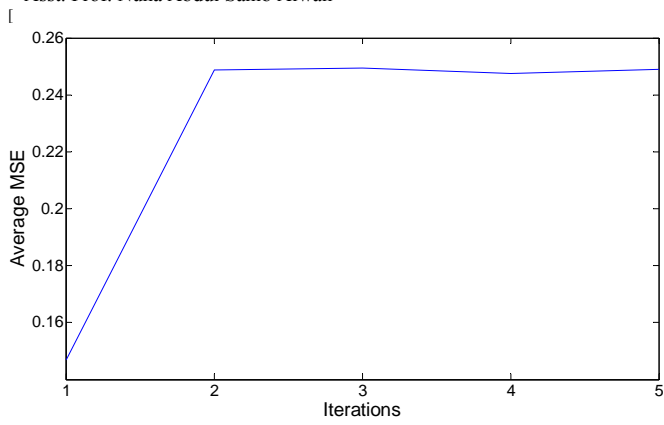Ali Ibrahim Abbas
Asst. Prof. Nuha Abdul Sahib Alwan

[

Fig. 5 BPW Classifier Performance over ORL
Database: Average MSE vs. Iterations