International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:4, No:8, 2010

A new biologically inspired pattern recognition approach for face recognition

V. Kabeer and N.K Narayanan*
School of Information Science and Technology, Kannur University,
Mangattuparamba Kannur, Kerala 670 567, India.
vkabeer@gmail.com, *csirc@rediffmail.com

***Abstract*—This paper reports a new pattern recognition approach for face recognition. The biological model of light receptors - *cones and rods* in human eyes and the way they are associated with pattern vision in human vision forms the basis of this approach. The functional model is simulated using CWD and WPD. The paper also discusses the experiments performed for face recognition using the features extracted from images in the AT & T face database. Artificial Neural Network and k-Nearest Neighbour classifier algorithms are employed for the recognition purpose. A feature vector is formed for each of the face images in the database and recognition accuracies are computed and compared using the classifiers. Simulation results show that the proposed method outperforms traditional way of feature extraction methods prevailing for pattern recognition in terms of recognition accuracy for face images with pose and illumination variations.**

***Index Terms*—Face recognition, Image analysis, Wavelet feature extraction, Pattern recognition, Classifier algorithms**

I. INTRODUCTION

FACE recognition is a task that humans perform routinely and effortlessly in daily life. Machine simulation of human vision has been a subject of intensive research for scientists and engineers for the last three decades. However automatic face recognition is yet to achieve a completely reliable performance. There are several challenges involved in automatic face recognition. Large variation in facial appearance, head size, orientation, changes in illumination and poses, occlusion, presence or absence of structural components etc are some of them to list. The interest devoted to this work is not only by the exciting challenges associated, but also the huge benefits that a Face-recognition system, designed in the context of a commercial application, could bring. Moreover, wide availability of powerful and low-cost desktop and embedded computing systems have also contributed to enormous interest in automatic processing of digital images and videos in a number of applications - Entertainment, Smart cards, Information security, Low enforcement and Surveillance are some of them [1], [2], [3], [4].

Face recognition lies at the core of the discipline of pattern recognition where the objective is to recognize an image of face from a set of face images. A complete face recognition system generally consists of three stages. The first stage involves detecting and localizing the face in arbitrary images [5], [6], [7], [8]. The second stage requires extraction of pertinent feature from the localized image obtained in the first stage. Finally, the third stage involves classification of facial images based on derived feature vector obtained in the previous stage. In order to design high accuracy recognition system, the choice of feature extraction method is very crucial. Two main approaches to feature extraction have been extensively used in conventional techniques [5], [6]. The first one is based on extracting structural facial features that are local structures of face images, for example, the shapes of the eyes, nose and mouth. The structure-based approaches deals with local information rather than global information, and, therefore are not affected by irrelevant information in an image. However, because of the explicit model of facial features, the structure-based approaches are sensitive to unpredictability of face appearance and environmental conditions. The second method is statistical-based approach that extracts features from the entire image and, therefore uses global information rather than local information.

There have been a lot of popular attempts towards automated face recognition which kept the research in the area active and vibrant. Some of them are Eigenfaces-PCA based approach [9], [8], [10] Independent Component Analysis(ICA)[11], Linear Discriminant Analysis (LDA) [12], [10], [13], [14] a specific kind of genetic algorithm called Evolutionary Pursuit (EP)[15], Elastic Bunch Graph Matching (EBGM) where faces are represented as graphs with nodes positioned at fiducial points [16], Kernel Methods which are a generalization of linear methods [1] like KPCA, KLDA, KICA etc., Trace transform, a generalization of the Radon transform [17], Active Appearance Model (AAM) is an integrated statistical model which combines a model of shape variation with a model of the appearance variations in a shape-normalized frame [18], Hidden Markov Models (HMM) [19], and Support Vector Machine (SVM)[20].

This paper presents a new pattern recognition approach for face recognition. The biological model of light receptors - *cones and rods* in human eyes and the way they are associated with pattern vision in human vision forms the basis of this approach. The functional model is simulated using *Continuous Wavelet Decomposition* (CWD) and *Wavelet Packet Decomposition* (WPD). The paper also discusses the experiments performed for face recognition using the features extracted from images in the AT & T face database. Artificial Neural Network and k-Nearest Neighbour classifier algorithms are employed for the recognition purpose. A feature vector is formed for each of the face images in the database and recognition accuracies are computed and compared using the classifiers. Simulation results reveals that the proposed method outperforms traditional way of feature extraction methods prevailing for pattern recognition in terms of recognition accuracy for face images with pose and illumination variations. Moreover, ANN classifier based approach yields better result than kNN classifier.

This paper is organised as follows. Section II gives an overview of Wavelet Transform. Section III describes Wavelet based Artificial Light Receptor Model(WALRM) for extraction of Artificial Light Receptor Feature Vector (*ALR Feature Vector*). Section IV presents the simulation experiment conducted using AT & T face database and reports the recognition results obtained using all the three classifiers. Finally, section V gives the conclusions and direction for future research.

II. Wavelet Transform - An overview

In the last decade, wavelets have become very popular, and new interest is rising on this topic. The main reason is that a complete framework has been recently built [21], [22] in particular for what concerns the construction of wavelet bases and efficient algorithms for its computation. The main characteristic of wavelets (if compared to other transformations) is the possibility to provide a multi-resolution analysis of the image in the form of coefficient matrices. Strong arguments for the use of multi-resolution decomposition can be found in psycho-visual research, which offers evidence that the human visual system processes the images in a multi-scale way [23]. Moreover, wavelets provide a spatial and a frequential decomposition of an image at the same time.

Wavelets are also very flexible: several bases exist, and one can choose the basis which is more suitable for a given application. This is still an open problem, and up to now only experimental considerations rule the choice of a wavelet form. However, the choice of an appropriate basis can be very helpful.

The wavelet transform is a transformation to basis functions that are localized in scale and in time as well (where the Fourier transform is only localized in frequency, never giving any information about where in space or time the frequency happens). The frequency (similar in that sense to Fourier-related transforms) is derived from the scale. Wavelets, as basis functions, are scaled and convolved with the functions that are analyzed all along the time.

A. Wavelet transform

In mathematics and signal processing, the continuous wavelet transform (CWT) of a function f is a wavelet transform defined by

$$\gamma(\tau, s) = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{|s|}}\psi^*(\frac{t-\tau}{s}) \qquad (1)$$

Where τ represents translation, s represents scale and ψ is the "mother" wavelet. ψ^* is the complex conjugate. The original function f can be reconstructed with the inverse transform

$$f(t) = \frac{1}{C_\psi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\gamma(\tau, s)\frac{1}{\sqrt{|s|}}\psi(\frac{t-\tau}{s})d\tau\frac{ds}{s^2} \qquad (2)$$

where

$$C_\psi = \int_{-\infty}^{+\infty}\frac{\hat{\psi}|\zeta|^2}{|\zeta|}d\zeta \qquad (3)$$

$\hat{\psi}$ is called the admissibility constant and is the Fourier transform of ψ. For a successful inverse transform, the admissibility constant has to satisfy the admissibility condition: $0 < C_\psi < +\infty$ It is possible to show that the admissibility condition implies that $\hat{\psi}(0) = 0$, so that a wavelet must integrate to zero.

The function ψ serves as the prototype for the *daughter* wavelets the signal is convolved with. For this reason, it is called the *mother* wavelet. The daughter wavelets are the scaled and shifted copies of the mother wavelet:

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{|s|}}\psi(\frac{t-\tau}{s}) \qquad (4)$$

B. Discrete Wavelet Transform

In numerical analysis and functional analysis, the discrete wavelet transform (DWT) refers to wavelet transforms for which the wavelets are discretely sampled. The first DWT was invented by the Hungarian mathematician Alfrd Haar. For an input represented by a list of 2^n numbers, the Haar wavelet transform may be considered to simply pair up input values, storing the difference and passing the sum. This process is repeated recursively, pairing up the sums to provide the next scale: finally resulting in 2^{n-1} differences and one final sum. Figure 1 shows this process. The process of starting with

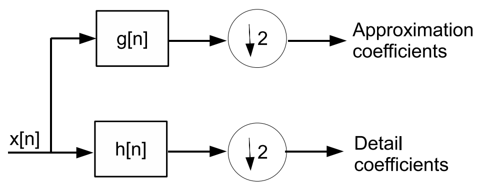

Fig. 1. Decomposition of a signal x[n] using Discrete Wavelet Transform.

a sequence of the approximation coefficients at some level of resolution and then generating approximation and detail coefficients at coarser levels through decimation is referred to as a decomposition or analysis of the sequence the reverse process is called reconstruction or synthesis.

C. Classical Wavelet Decomposition (CWD)

In classical wavelet decomposition, the image is split into an approximation and details images. The approximation is then split itself into a second-level approximation and details.

For n-level decomposition, the signal is decomposed in the following way [23]:

$$\begin{aligned} A_n &= [H_x * [H_y * A_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2} \\ D_{n1} &= [H_x * [G_y * A_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2} \\ D_{n2} &= [G_x * [H_y * A_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2} \\ D_{n3} &= [G_x * [G_y * A_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2} \end{aligned}$$

where * denotes the convolution operator, $\downarrow 2, 1(\downarrow 1, 2)$ sub-sampling along the rows (columns) and $A_0 = I(x, y)$ is the original image. A_n is obtained by low-pass filtering and is the approximation image at scale *n*. The details images D_{ni} are obtained by band-pass filtering in a specific direction and thus contain directional detail information at scale *n*. The original image *I* is thus represented by a set of sub-images at several scales $\{A_n, D_{ni}\}$.

D. Wavelet Packets Decomposition (WPD)

Wavelet Packet Decomposition (WPD), sometimes known as just Wavelet Packets (WP) is a wavelet transform where the signal is passed though more number of filters than the DWT. In the DWT, each level is calculated by passing the previous approximation coefficients though a high and low pass filters. However in the WPD, both the detail and approximation coefficients are decomposed. Figure 2 shows the process of WPD.

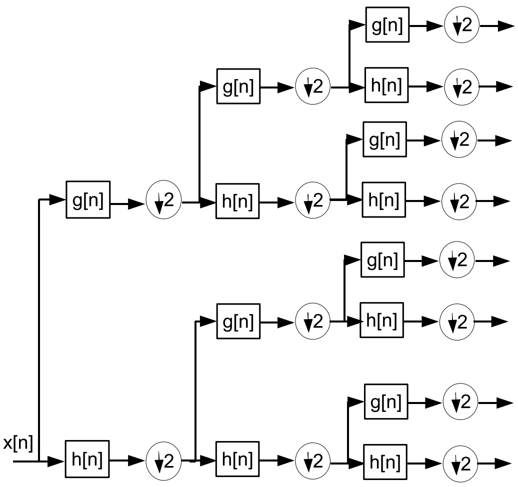

Fig. 2. Decomposition of a signal x[n] using Wavelet Packet Decomposition.

III. Wavelet based Artificial Light Receptor Model(WALRM) for Feature Extraction

Pattern vision is afforded by the distribution of light receptors over the surface of the retina. There are two classes of receptors called *cones* and *rods*. The *cones* in each eye number between 6 and 7 million, and are located primarily in the central portion of the retina. These *cones* help human to resolve fine details they see around largely because each one is connected to its own nerve end. On the other hand, *rods* are very huge in number (*nearly 120 million*) when compared to *cones* and several *rods* are connected to a single nerve end, which in turn reduces the amount of detail carried by these receptors [24]. Figure 3shows this arrangement of *rods* and *cones* in retina and biological signal processing structure from retina to brain. This association of *rods* and *cones* with the nerve end forms basis for the design of the feature extraction model in our approach. The model is simulated using a combination of

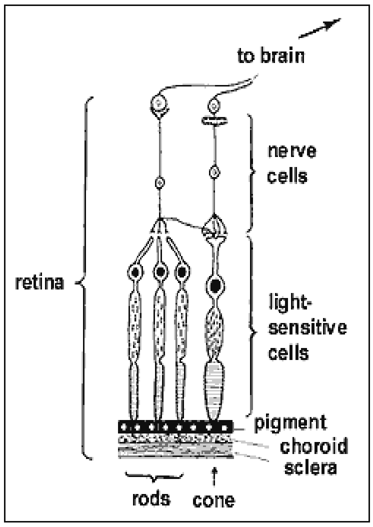

Fig. 3. Rods and Cones in Retina and Biological Image Signal Passing Structure.

Classical Wavelet Decomposition (CWD) and *Wavelet Packet Decomposition* (WPD). Each face image is described by a subset of band filtered images containing wavelet coefficients. The elements from these coefficients matrices are subjected to simple statistical operations and the results are organized in such a fashion similar to the arrangements of *rods* and *cones* in retina giving compact and meaningful feature vectors. Figure 4 shows the block diagram for the entire recognition system using Wavelet based Artificial Light Receptor Feature Extraction Model.

A. Feature Extraction Process

The feature extraction process consists of three stages. In the first stage one component of *ALR Feature Vector* is derived by subjecting the face image to undergo CWD recursively to decompose it into fifth level of resolution (fifth level has been found to be optimum experimentally as illustrated by Table 1). Therefore, the approximation matrix at this level of resolution is a significantly small representative of the original image and carries enough information content to describe face image characteristics coarsely. This matrix can be considered as analogous to an image formed in retina at *cone* area. We call this functional unit in our model as *Wavelet Cones*. And, each element in the matrix is sent to separate units (nerve ends) as the case may be with human visual system.

International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:4, No:8, 2010

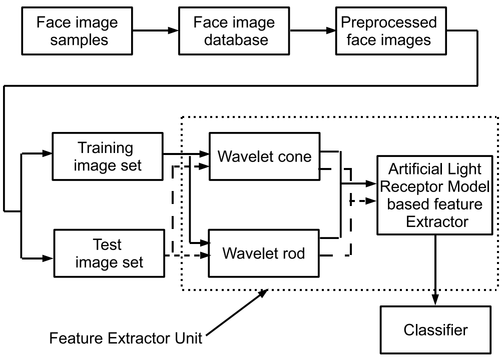

Fig. 4. Face Recognition System using Wavelet based Artificial Light Receptor Model for feature extraction.

Let $\mathbf{A}_k$ represents this approximation matrix at decomposition level k, which can be written as:

$$\mathbf{A}_k = \begin{pmatrix} A_{11} & A_{12} & ... & A_{1n} \\ A_{21} & A_{22} & ... & A_{2n} \\ ... & ... & ... & ... \\ A_{m1} & A_{m2} & ... & A_{mn} \end{pmatrix}$$

Then, the *cone* component of *ALR Feature Vector*, $\mathbf{V}_{k1}$, is given by,

$$\mathbf{V}_{k1} = \bigcup_{i=1}^{m} \bigcup_{j=1}^{n} \{A_{i,j}\} \tag{5}$$

In the second stage, *wavelet rods* are used to extract the other component of *ALR Feature Vector*. This is materialized by means of *Wavelet Packet Decomposition*(WPD) of the face images at their best level of wavelet packet decomposition tree. The first coefficient matrix at the best level tree contains enough information to represent the given input face image without loss of much face image features. Let μ represent mean of one row in this coefficient matrix then the *rod* component of the *ALR Feature Vector*, $\mathbf{V}_{k2}$, is given by,

$$\mathbf{V}_{k2} = \{\mu_i\}, \forall i = 1, 2, \ldots, m \tag{6}$$

where m is the number of rows in the best level coefficient matrix.
In the 3rd stage we combined $\mathbf{V}_{k1}$ and $\mathbf{V}_{k2}$ to form the final *ALR Feature Vector* $\mathbf{V}$ as

$$\mathbf{V} = \bigcup_{i=1}^{2} \{\mathbf{V}_{ki}\} \tag{7}$$

As the *wavelet cone* feature component is of size 12 and *wavelet rod* component is 28 the estimated *ALR Feature Vector* dimension is constraint to forty.

Feature vectors of representative samples had been generated from the database at different decomposition levels, and these feature vectors were subjected to classification using *kNN* classifier - comparatively faster classification algorithm with lesser accuracy. Table 1 gives the classification results on this representative subset using *kNN* classifier.

Analysis of Table 1 shows that feature vector generated at resolution level 5 is better than feature vectors at other resolution levels. This analysis lead us to decide the features at resolution level 5 is optimal for recognition.

TABLE I
CLASSIFICATION RESULTS ON A REPRESENTATIVE SUBSET OF AT & T FACE DATABASE AT DIFFERENT RESOLUTION LEVELS USING *k-NN*.

Resolution Level	Feature size	%Accuracy
1	2604	25
2	672	33
3	196	56
4	70	70.5
5	40	81.5
6	32	46
7	29	30
8	29	30

Figure 5 shows graph of *ALR Feature Vector* obtained from face database plotted for ten samples of first person in AT&T face database along with the mean curve. The graphs obtained for different samples of same person are found to be quite similar while the graphs for different persons are highly distinguishable.

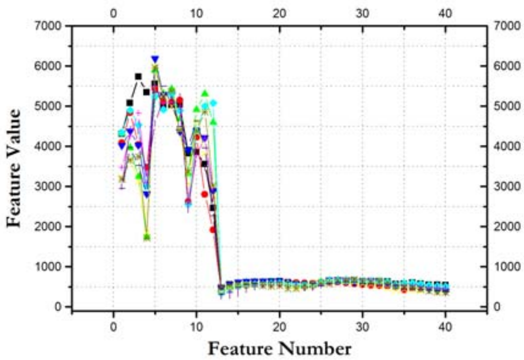

Fig. 5. Feature vector generated for ten face images of the first person in AT & T face database using.

IV. EXPERIMENTS AND RESULTS

All the experiments were carried out using the AT & T face database, which contains face images of 40 distinct persons. Each person has ten different images, taken at different times. Face images of five individuals (in five rows) in the AT & T face database are shown in figure 6. There are variations in facial expressions such as open/closed eyes, smiling/ non-smiling, and facial details such as glasses/no glasses. All the images were taken against a dark homogeneous background with the subjects in an up-right, frontal position, with tolerance for some side movements. There are also some variations in scale. For the classification purpose two typical classifiers namely, k-Nearest-Neighbor(*kNN*) and Artificial Neural Networks (*ANN*) are used.

The k-Nearest-Neighbor [25], [26] method is a well-known non-parametric classifier, where a posteriori probability is

Fig. 6. Sample face images taken from the AT & T face database.

estimated from the frequency of nearest neighbors of the unknown pattern. It considers multiple prototypes while making a decision and uses a piecewise linear discriminant function. On using *kNN* algorithm a top most overall accuracy of 89.5% has been observed for the value of $k = 5$. To fix up the value of k to be 5 we repeated the experiments for various values of k between 1 and 10. At $k = 5$ the aforesaid accuracy has been observed. Artificial Neural Networks (ANN) are massively

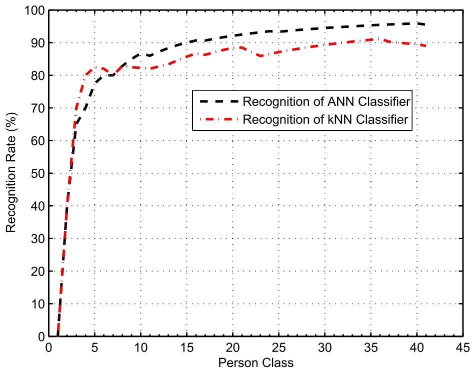

Fig. 7. Classification accuracies - Person classes Vs % Accuracy.

parallel adaptive networks of simple non-linear computing elements called neurons which are intended to abstract and model some functionalities of human nervous system in an attempt to partially capture some of its computational strengths [27], [28]. In other words, a set of processing units when assembled in a closely interconnected network, offers a surprisingly rich structure exhibiting some features of the biological neural network [29]. Artificial neural networks are useful only when the processing units are organized in a suitable manner to accomplish a given pattern recognition task [27], [29].

We used multilayer feed forward network for classification. There are three layers in the architecture designed. We used 40 neurons in the first layer (input layer), 16 neurons in second layer (hidden layer), and 40 neurons in the output layer. For training we used back propagation learning algorithm. During the experiment we made nodes in the middle layer to vary between one and thirty and we analyzed classification accuracy in all the cases with various network parameters. And, we found that a middle layer with sixteen nodes derive better classification than the other cases. Dividing data set into training and test sets was part of the experimental studies. There has been a recognition accuracy of 95.5% for *ALR Feature Vector* with *ANN*. Figure 7 shows the comparison graph of results obtained for both kNN and ANN classifiers.

V. Conclusion

This paper presented a robust biologically inspired Artificial Light Receptor Model for extracting face image feature vectors called *ALR Feature Vector*. The model is simulated using a combination of CWD and WPD. A feature vector of size 40 is formed for face images of each person and recognition accuracy is computed using kNN and ANN classifiers. The best overall recognition accuracy obtained for the AT & T face database is 95.5% with a combination of *ALR Feature Vector* and ANN classifier. There is significant dimensionality reduction as we used a feature vector of 40-element size to represent a face image. Suitability of the model developed with other kinds of images are to be studied in future researches.

References

[1] D. J. K. Ming-Hsuan Yang and N. Ahuja, "Detecting faces in images: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 34–58, 2002.
[2] J. Daugman, "Face detection: A survey," *Computer Vision and Image Understanding*, vol. 83, pp. 236–274, 2001.
[3] M. A. Grudin, "On internal representation of face recognition systems," *Pattern Recognition*, vol. 33, pp. 1161–1177, 2000.
[4] P. P. W. Zhao, R. Chellappa and A. Rosenfeld, "Face recognition: A literature survey," *ACM Computing Surveys*, vol. 35, pp. 399–458, 2003.
[5] S. Z. Li and A. K. Jain, *Handbook of Face Recognition*. Springer, 2004.
[6] J. Haddadnia and M. Ahmadi, "N-feature neural network human face recognition," *Image and Vision Computing*, vol. 22, p. 1071, 2004.
[7] A. M. J. Haddadnia and K. Faez, "An efficient method for recognition of human face using high order pseudo zernike moment invariant," in *5th IEEE International Conference on Automatic Face and Gesture Recognition*, Washington, DC,USA, 2000, pp. 315–32.
[8] J. Wang and T. Tan, "A new face detection method based on shape information," *Pattern Recognition Letter*, vol. 21, pp. 463–471, 2000.
[9] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, pp. 71–86, 1991.
[10] B. C. L. Shaokang Chen and T. Shan, "Robust adapted principal component analysis for face recognition," *International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI)*, vol. 23, pp. 491–520, 2009.
[11] J. M. M.S. Bartlett and T. Sejnowski, "Face recognition by independent component analysis," *IEEE Trans. on Neural Networks*, vol. 13, pp. 1450–1464, 2002.
[12] K. Etemad and R. Chellappa, "Discriminant analysis for recognition of human face images," *Journal of the Optical Society of America*, vol. 14, pp. 1724–1733, 1997.
[13] X. L. Dong Xu, Dacheng Tao and S. Yan, "Face recognition - a generalized marginal fisher analysis approach," *International Journal of Image and Graphics (IJIG)*, vol. 7, pp. 583–591, 2007.
[14] W.-S. Chen, "Wavelet-face based subspace lda method to solve small sample size problem in face recognition," *International Journal of Wavelets,Multiresolution and Information Processing (IJWMIP)*, vol. 7, pp. 199–214, 2009.
[15] C. Liu and H. Wechsler, "Evolutionary pursuit and its application to face recognition," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 570–582, 2000.
[16] N. L. Wiskott, J.-M. Fellous and C. Malsburg, *Face Recognition by Elastic Bunch Graph Matching Intelligent Biometric Techniques in Fingerprint and Face Recognition*, L. J. et al, Ed. CRC Press, 1999.
[17] A.Kadyrov and M. Petrou, "The trace transform and its applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, pp. 811–828, 2001.

[18] K. W. T.F. Cootes and C. Taylor, "View-based active appearance models," *IEEE International Conference on Automatic Face and Gesture Recognition*, vol. 17, pp. 227–232, 2000.
[19] A. Nefian and M. Hayes, "Maximum likelihood training of the embedded hmm for face detection and recognition," *IEEE International Conference on Image Processing*, pp. 33–36, 2000.
[20] S. L. G. Guo and K. Chan, "Face recognition by support vector machines," *IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 196–201, 2000.
[21] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Transactions on Information Theory*, vol. 36, pp. 961–1005, 1990.
[22] S. Mallat, "Multifrequencies channel decompositionsof images and wavelets models," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, 1989.
[23] G. Z. Christophe Garcia and G. Tziritas, "A wavelet-based framework for face recognition," *ICS - Foundation for Research and Technology-Hellas*, vol. fourth, 2000.
[24] R. C. Gonzales and R. E. Woods, *Digital Image Processing*. Pearson Edition, 2002.
[25] J. Tou and R. Gonzalez, *Pattern Recognition Principle*. Addison Wesley, 1975.
[26] B. Chanda and D. D. Majumder, *Digital Image Processing and Analysis*. PHI, 2000.
[27] S. Haykin, *Neural Networks a comprehensive foundation*. Pearson Education, 2001.
[28] T. Kohonen, "An introduction to neural computing," *Neural Networks*, vol. 1, pp. 3–16, 1988.
[29] B. Yegnanarayana, *Artificial neural networks*. PHI, 2005.