Wavelet based Artificial Light Receptor – A feature extraction model for face recognition

Kabeer V.
Department of Computer Science, Farook College,
Calicut, Kerala, India.

ABSTRACT
This paper presents a novel biologically inspired and wavelet based model for extracting features of faces from face images. The biological knowledge about the distribution of light receptors, cones and rods, over the surface of the retina, and the way they are associated with the nerve ends for pattern vision forms the basis for the design of this model. A combination of classical wavelet decomposition and wavelet packet decomposition is used for simulating the functional model of cones and rods in pattern vision. The paper also describes the experiments performed for face recognition using the features extracted on the AT&T face database containing 400 face images of 40 different individuals. In the recognition stage we used k-Nearest Neighbour classifier. A feature vector of size 40 is formed for face images of each person and recognition accuracy is computed using k-NN classifier.

General Terms
Face Recognition, Pattern Recognition.

Keywords
Face recognition, Image analysis, Wavelet feature extraction, Pattern recognition, k-NN Classifier.

1. 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 has 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 [4, 5]. 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)[9,10], Independent Component Analysis(ICA)[11], Linear Discriminant Analysis (LDA)[12], a specific kind of genetic algorithm called Evolutionary Pursuit (EP)[13], Elastic Bunch Graph Matching (EBGM) where faces are represented as graphs, with nodes positioned at fiducial points [14], Kernel Methods which are a generalization of linear methods like KPCA, KLDA, KICA etc [15],The Trace transform, a generalization of the Radon transform[16], 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[17],Hidden Markov Models (HMM)[18] and Support Vector Machine (SVM)[19].

In the present study we propose an entirely new and biologically inspired model called Artificial Light Receptor Model for extracting the features of human faces from face images. The model is simulated using a combination of 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 coefficient matrices are subjected to simple statistical operations and the results are organized in such a fashion similar to the arrangements of cones and rods in the retina giving compact and meaningful feature vectors. Then, we use k-Nearest Neighbour classifier to classify the face feature vectors into person classes.

This paper is organised as follows. Section 2 gives an over view of wavelet transforms – here a brief account of wavelet transform, discrete wavelet transform, and classical wavelet transform and wavelet packet decomposition are given. In section 3 Wavelet based Artificial Light Receptor model for feature (WALR) extraction method is described. In Section 4 k-Nearest Neighbour algorithms are discussed. Section 5 presents the simulation experiment conducted using AT & T face database and reports the recognition results obtained using k-NN classifier. Finally, section 6 gives the conclusions and direction of future research.

2. WAVELET TRANSFORM – AN OVER VIEW

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 [20, 21] 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 [22]. 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 the 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 axis [22, 30].

3. WAVELET BASED ARTIFICIAL LIGHT RECEPTOR MODEL FOR FEATURE (WALRM) EXTRACTION

In the present work a combination of Classical Wavelet Decomposition and Wavelet Packet Decomposition methods is used to develop Wavelet based Artificial Light Receptor (WALRM) model for extracting feature vectors for representing a face image. 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 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 [28]. This association of rods and cones with the nerve end forms basis for the design of the feature extraction model in our approach.

The entire model for face recognition based on this method is shown in figure 1 and, the process is carried out in three stages. In the first stage, Wavelet cons are used to extract one component of the WALRM feature by making the face image to undergo CWD recursively to decompose it into 5th level of resolution (level 5 was found to be optimum experimentally as illustrated in Table 1); therefore, the approximation coefficient matrix at this level is sufficiently small representative of the original image and carries enough information content to describe face image characteristics coarsely.

This matrix can be considered equivalent to an image formed in retina at cones and each element in the matrix are sent to separate nerve ends as the case may be with cons in the human eyes. Let A_k represents this approximation matrix at decomposition level k, which can be written as:

$$A_k = \begin{bmatrix} A_{11} & A_{12} & \cdots\cdots & A_{1n} \\ A_{21} & A_{22} & \cdots\cdots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{m1} & A_{m2} & \cdots\cdots & A_{mn} \end{bmatrix}$$

Then, the cone component of WALR feature vector, V_{k1}, is given by,

$$V_{k1} = \bigcup_{i=1}^{m} \bigcup_{j=1}^{n} \left\{ A_{ij} \right\}$$

In the second stage, Wavelet rods are used to extract the other component of WALRM feature by decomposing each face image using WPD to its 5th level of resolution(level 5 was found to be optimum experimentally as illustrated in Table 1). Then find the best level of wavelet packet decomposition tree. The first coefficient matrix at the best level tree contain 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 rods component of the WALRM feature vector, $\mathbf{V}_{k2}$, is given by, $\mathbf{V}_{k2} = \{\mu_i\}$,for all ,, $i = 1, 2, 3 \ldots m$ (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 WALRM feature vector V.

$$\mathbf{V} = \bigcup_{i=1}^{2} \left\{ \mathbf{V}_i \right\}$$

We got a feature vector of size 40 for WALRM feature vector. The figure 2 shows a sample WALRM feature vector for the first person in AT & T face database.

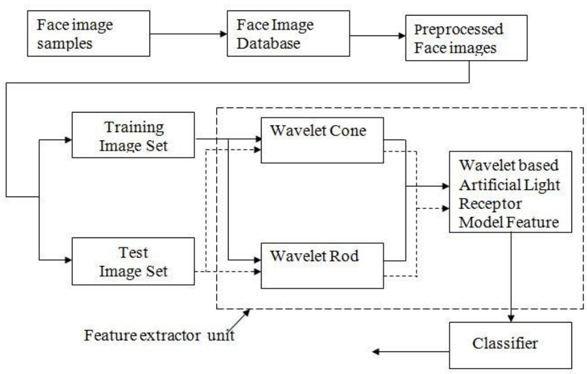

Fig 1. Face recognition system Model using Wavelet based Artificial Light Receptor feature extractor.

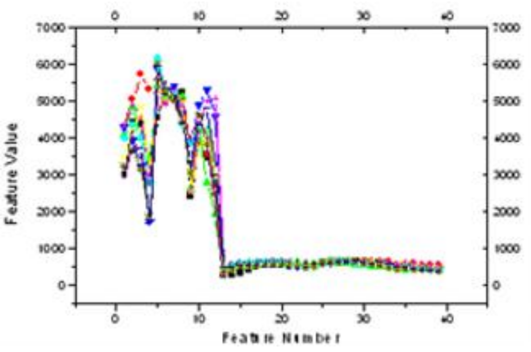

Fig. 2 Feature vector generated for ten face images of the first person in AT&T face database using Wavelet based Artificial Light Receptor model.

Table 1 Classification results on a representative subset of AT&T face database at different resolution levels using k-NN classifier

S.No.	Resolution Level	Feature Size	% Accuracy
1	1	2604	25
2	2	672	33
3	3	196	56
4	4	70	70.5
5	5	40	91.5
6	6	32	46
7	7	29	30
8	8	29	30

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.

4. CLASSIFICATION USING K-NN CLASSIFIER

Pattern classification by distance functions [23] is one of the earliest concepts in pattern recognition [24, 25]. Here the proximity of an unknown pattern to a class serves as a measure for its classification. A class can be characterized by single or multiple prototype pattern (s). The k-Nearest-Neighbor method is a well-known non-parametric classifier, where a posteriori probability is estimated from the frequency of nearest neighbours of the unknown pattern. It considers multiple prototypes while making a decision and uses a piecewise linear discriminant function.

Let us consider the case of m classes $\{C_i\}$ where i = 1: m and a set of N patterns $\{y_i\}$ where i= 1: N whose classification is priory known. Let $\boldsymbol{x}$ denote an arbitrary incoming pattern. The nearest neighbour classification approach classifies $\boldsymbol{x}$ in the pattern class of its nearest neighbour in the set $\{y_i\}$ where i= 1: N, i.e.

$$\text{if } \| x - y_j \| = \min \| x - y_i \| \quad 1 \le i \le N \quad \text{then } x \,\varepsilon\, C_j .$$

This scheme (can be termed the 1-NN rule since it employs only the classification of the nearest neighbour to x) can be modified by considering the k nearest neighbours to x and using a majority-rule type classifier. The following algorithm summarizes the classification process.

Algorithm: Minimum distance k -Nearest Neighbour Classifier.

Input: n – problem's dimension.

N – the number of pre-classified patterns

m – the number of pattern classes.

$(\mathbf{y}_i, c_i)$, $1 \le i \le N$ - N ordered pairs , where $\mathbf{y}_i$ is the ith pre-classified pattern and c_i its class number ($1 \le c_i \le m$ for all i).

k -the order of NN classifier (i.e. the k closest neighbours to the incoming patterns are considered).

$\boldsymbol{x}$- an incoming pattern.

Output: L – the number of class into which $\boldsymbol{x}$ is classified.

Step 1: Set S= $\{ (\boldsymbol{y}_i, \boldsymbol{c}_i) \}$ i = 1,… ,N

Step 2: Find (y_j, c_j) ε S which satisfies

$\| \boldsymbol{x} - \boldsymbol{y}_j \|^2 = \min \| \boldsymbol{x} - \mathbf{y}_i \|^2$, where $1 \le \imath \le m$.

Step 3: If k= 1 set $L = c_j$ and stop; else initialize an m – dimensional vector I:

$I(i') = 0,\ i' \ne c_j$; $I(c_j) = 1$ where $1 \le \imath \le m$ and set S=S - $\{ (y_j, c_j) \}$

Step 4: For i_0 = 1,2….,k-1 do steps 5-6

Step 5: Find $(\boldsymbol{y}_j, \boldsymbol{c}_j)$εS such that

$\| \boldsymbol{x} - \boldsymbol{y}_j \| = \min \| \boldsymbol{x} - \boldsymbol{y}_i \|$, where $1 \le \imath \le N$

Step 6: Set $I(c_j) = I(c_j) + 1$ and S = S-$\{ (\boldsymbol{y}_j, \boldsymbol{c}_j) \}$.

Step 7: Set $L = \max \{I(i')\}$, $1 \le i' \le$ m and stop.

The following section gives a detailed description about the experiments we conducted and the classification results obtained.

5. 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 fig.3.

Fig 3. Samples face images taken from the AT & T face database.

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.

As described in the previous section we used Wavelet based Artificial Light Receptor model feature vector and k-NN algorithm for classification. Dividing data set into training and test tests was part of the experimental studies. Figure 4 gives recognition accuracies for forty persons in the standard AT & T face image database. Overall recognition rate obtained is 86.5%. Table 2 gives comparative study of face recognition results of other methods in literatures with the present work. For methods not marked with * the results are as on FERET database. Table shows that the proposed method renders a better or comparable result.

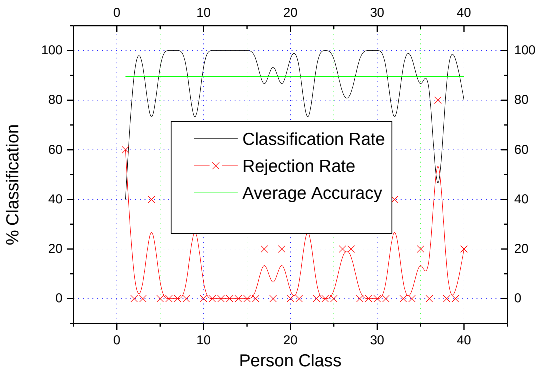

Fig 4. Classification accuracies – Person classes Vs %Recognition

Table 2 Some Popular Methods & Their Results

S.No.	Method	% Accuracy
1	WALRM*	86.50
2	PCA L1	69.41
3	LDA	77.53
4	ICA	84.50
5	EGBM	79.60
6	Kernel Space	81.50

6. CONCLUSION

This paper presented a robust approach to model a Wavelet based Artificial Light Receptor model for extracting face image feature vectors. A feature vector of size 40 is formed for face images of each person and recognition accuracy is computed using *k*-NN classifier. Overall recognition accuracy obtained for the AT & T face database is 86.5%. There is significant dimensionality reduction as we used a feature vector of 40-element size to represent a face image. More effective implementation of multiple classifier system is one of our future research directions and more research is needed to deal with building a fully functional system.

7. REFERENCES

[1] W. Zhao, R. Chellappa, P.J Philips and A. Rosenfeld, “Face Recognition: A Literature Survey”, ACM Computing Surveys, Vol.35 (4), 399-458, Dec.2003

[2] Ming-Hsuan Yang, David J Kriegman and Narendra Ahuja, “Detecting faces in Images: A Survey”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.24 (1), 34-58, Jan.2002.

[3] M.A. Grudin, “On internal representation of face recognition systems”, Pattern Recognition, Vol.33 (7), 1161-1177, 2000.

[4] J. Daugman, “Face detection: A survey”, Computer Vision and Image Understanding, Vol.83 (3), 236-274, Sept. 2001.

[5] Stan Z. Li, Anil 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, 1071-1082, 2004.

[7] J. Haddadnia , M. Ahmadi, K Faez, “An Efficient Method for Recognition of Human face using high order Pseudo Zernike Moment Invariant”, The 5th IEEE International Conference on Automatic Face and Gesture Recognition, Washington, DC,USA, 315-32, May 20-21,2000.

[8] J. Wang, T. Tan, ‘A new face detection method based on shape information’, Pattern Recognition Letter, Vol.21, pp.463-471, 2000.

[9] M. Turk, A. Pentland, Eigenfaces for Recognition, Journal of Cognitive Neurosicence, Vol. 3, No. 1, pp. 71-86, 1991.

[10] H. Moon, P.J. Phillips, Computational and Performance aspects of PCA-based Face Recognition Algorithms, Perception, Vol. 30, pp. 303-321, 2001.

[11] M.S. Bartlett, J.R. Movellan, T.J. Sejnowski, Face Recognition by Independent Component Analysis, IEEE Trans. on Neural Networks, Vol. 13, No. 6, pp. 1450-1464, November 2002.

[12] K. Etemad, R. Chellappa, Discriminant Analysis for Recognition of Human Face Images, Journal of the Optical Society of America A, Vol. 14, No. 8, pp. 1724-1733, August 1997.

[13] C. Liu, H. Wechsler, Evolutionary Pursuit and Its Application to Face Recognition, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 22, No. 6, pp. 570-582, June 2000.

[14] L. Wiskott, J.-M. Fellous, N. Krueuger, C. von der Malsburg, Face Recognition by Elastic Bunch Graph Matching, Chapter 11 in Intelligent Biometric Techniques in Fingerprint and Face Recognition, eds. L.C. Jain et al., CRC Press, pp. 355-396, 1999.

[15] M.-H. Yang, Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods, Proc. of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 20-21 May 2002, Washington D.C., USA, pp. 215-220 , 2002.

[16] Kadyrov, M. Petrou, The Trace Transform and Its Applications, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 8, pp. 811-828, August 2001.

[17] T.F. Cootes, K. Walker, C.J. Taylor, View-Based Active Appearance Models, Proc. of the IEEE International Conference on Automatic Face and Gesture Recognition, 26-30 March 2000, Grenoble, France, pp. 227-232, 2000.

[18] A.V. Nefian, M.H. Hayes, Maximum likelihood training of the embedded HMM for face detection and recognition, Proc. of the IEEE International Conference on Image Processing, ICIP 2000, Vol. 1, 10-13 September 2000, Vancouver, BC, Canada, pp. 33-36

[19] G. Guo, S.Z. Li, K. Chan, Face Recognition by Support Vector Machines, Proc. of the IEEE International Conference on Automatic Face and Gesture Recognition, 26-30 March 2000, Grenoble, France, pp. 196-201 , 2000.

[20] Mallat S, “Multifrequencies Channel Decompositionsof Images and Wavelets Models”, IEEE Transactions on Acoustics, Speech and Signal Processing, 37(12), 1989.

[21] Daubechies, “The Wavelet Transform, Time-Frequency Localization and Signal Analysis”, IEEE Transactions on Information Theory, 36(5), 961-1005, 1990.

[22] Christophe Garcia, Giorgos Zikos& Giorgos Tziritas, “A Wavelet-based Framework for Face Recognition”, ICS – Foundation for Research and Technology-Hellas, FORTH, 2000.

[23] J.T. Tou & R.C Gonzalez, Pattern Recognition Principles (Addison Wesley, London, 1974).

[24] Sankar K. Pal & Sushmitha Mitra, Neuro-fuzzy Pattern Recognition – Methods in Soft Computing (John Wiley & Sons, Inc., 1999).

[25] M. Friedman, A. Kandel, Introduction to Pattern Recognition – Statistical, Structural, Neural and Fuzzy Logic Approac, (World Scientific, 1999).

[26] V. Kabeer & N. K. Narayanan, Face Recognition using Wavelet based Feature parameter and K –NN Classifier, Proceedings Kerala Science Congress, Kannur, pp-887-888, January 2007.

[27] Jaideva C. Goswami & Andrew K. Chan, Fundamentals of Wavelets-Theory, Algorithms and Applications, Wiley Series, 1999.

[28] Rafael C Gonzales & Richard E. Woods, Digital Image Processing, PE, 2002.

[29] Chanda & D. Dutta Majumder, Digital Image Processing and Analysis,PHI, 2000.

[30] Raghuveer M.Rao & Ajith S. Bopardikar, Wavelet Transforms – Introduction to Theory and Applications, PE, 1998.

[31] Michael W. Frazier, An introduction to Wavelets through Linear Algebra, Springer, 1999.

[32] V. Kabeer & N. K. Narayanan, Face Recognition using Wavelet Hybrid Features and K–NN Classifier, Proceedings IEEE ACVIT-07, Aurangabad,, Maharashtra, India., pp 738-746, Nov. 2007.

[33] V. Kabeer & N K Narayanan, Wavelet based Artificial Light Receptor Model for Human Face Recognition, International Journal of Wavelets, Multiresolution and Information Processing (IJWMIP), World Scientific, No.5., 2009.