Selection of Eigenvectors for Face Recognition

Manisha Satone
Sinhgad College of Engineering
Pune,India

G.K.Kharate
University of Pune
Pune, India

Abstract—Face recognition has advantages over other biometric methods. Principal Component Analysis (PCA) has been widely used for the face recognition algorithm. PCA has limitations such as poor discriminatory power and large computational load. Due to these limitations of the existing PCA based approach, we used a method of applying PCA on wavelet subband of the face image and two methods are proposed to select best of the eigenvectors for recognition. The proposed methods select important eigenvectors using genetic algorithm and entropy of eigenvectors. Results show that compared to traditional method of selecting top eigenvectors, proposed method gives better results with less number of eigenvectors.

Keywords—face recognition; PCA; wavelet transform; genetic algorithm

I. INTRODUCTION

Many recent events, exposed defects in most sophisticated security systems. Therefore, it is necessary to improve security systems based on the body or behavioral characteristics, called biometrics. With the interest in the development of human and computer interface and biometric identification, human face recognition has become an active research area. Face recognition offers several advantages over other biometric methods. Nowadays, Principal Component Analysis (PCA) has been widely adopted for the face recognition algorithm. Yet still, PCA has limitations such as poor discriminatory power and large computational load [1].

In view of the limitations of the existing PCA-based approach, here we used a method of applying PCA on wavelet subband of the face image and two methods are proposed to select a best eigenvectors for recognition. In the proposed method, face image is decomposed into a number of subbands with different frequency components using the wavelet transform (WT). Out of the different frequency subbands, a mid-range frequency subband image is selected. The resolution of the selected subband is 16x16. The proposed method works on lower resolution, instead of 128 x 128 resolution of the original image. Working on lower resolution images, reduces the computational complexity. Experimental results show that applying PCA on WT sub-image with mid-range frequency components gives better recognition accuracy and discriminatory power than applying PCA on the entire original image [2][3]. In PCA, all the eigenvectors are not equally informative. This paper proposes two methods of eigenvector selection. In comparison with the traditional use of PCA, the proposed methods select the eigenvectors based on genetic algorithm and entropy of eigenvectors.

Paper is organized as follows: Section II reviews the wavelet. Section III reviews the background of PCA and eigenfaces. Section IV is about the eigenvector selection. Experiments and results are discussed in Sections V and conclusion in section VI.

II. WAVELET TRANSFORM

Multiresolution methods are powerful tools, which are widely used in feature extraction, image compression and denoising applications. Wavelet decomposition is a widely used multiresolution technique in image processing. Wavelet analysis has generated a great interest in both theoretical and applied mathematics, and the wavelet transform has proven to be an effective tool for data analysis, numerical analysis, and image processing. Wavelets are functions which give different frequency components of data. They have advantages over fourier methods in analyzing physical situations where the signal contains discontinuities and sharp spikes. The advantages of WT are good time and frequency localizations [4].

WT is chosen to be used in image frequency analysis and image decomposition because of the following reasons:

• The resolutions of the images can be reduced by decomposing an image using WT. Working on a lower resolution image reduces computational complexity.

• Wavelet transform provides local information in both space domain and frequency domain.

Wavelet transform can be performed for every scale and translation, resulting in Continuous Wavelet Transform (CWT), or only in multiples of scale and translation intervals, resulting in Discrete Wavelet Transform (DWT). CWT provides redundant information and requires a lot of computations, DWT is generally preferred. A two-dimensional wavelet transform is derived from two one-dimensional wavelet transform by taking tensor products. The implementation of WT is carried out by applying a one-dimensional transform to the rows of the original image data and the columns of the row transform data respectively.

Using wavelet transform, an image is decomposed into four subbands. The band *LL* is a coarser approximation to the original image. The band gives the edges along horizontal directions, whereas band *HL* record the edges along the vertical directions. The *HH* band records the diagonal edges present in the image. This is the first level decomposition. Further decomposition can be conducted on the *LL* subband. After applying a 3-level wavelet transform, an image is decomposed into subbands of different frequency components. Figure 1 shows 1-level and 3- level wavelet decomposition. In this paper, Daubechies wavelet D4 is adopted for image decomposition. G.C. Feng, P.C. Yuen and D.Q. Dai [5] has

shown that Daubechies wavelet D4 is better than other wavelets. Therefore, Daubechies wavelet D4 is adopted for image decomposition in our system.

Nastar et al. [2][3] have found the relationship between variations in facial appearance and their deformation spectrum. They found that facial expressions and small occlusion affect the intensity manifold locally. Under the frequency-based representation, only the high frequency spectrum was affected, whereas changes in illumination affected the intensity manifold globally, in which only the low frequency spectrum was affected. When there was a change in human face, all frequency components were affected. Based on Nastar et al.'s findings, we used subbands containing mid-range frequencies. Among subbands 1 to 4, subbands 2 to 4 were the mid-range frequency subbands. Further experimental results in [5] showed that applying PCA on subband 4 gave better recognition accuracy and class separability compared with applying PCA on the whole image, or applying PCA on subbands 1, 2 or 3. Hence we have chosen subband 4.

LL	LH
HL	HH

(a)

1 2	6	9
3 4		
5	7	
8		10
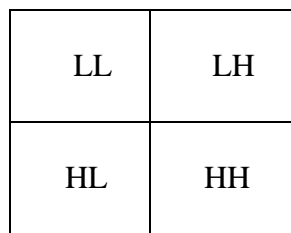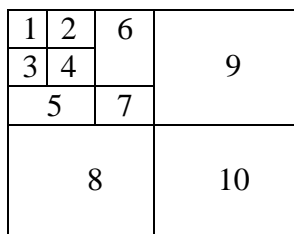
(b)

Fig.1. (a) 1-level wavelet decomposition and (b) 3-level wavelet decomposition

III. PRINCIPAL COMPONENT ANALYSIS

To linearly project an image in a low-dimensional space, PCA [6] is used where images are represented in eigenpace. This space is spanned by eigenvectors corresponding to the largest eigenvalues of the training images. After an image has been projected in the eigenspace, a feature vector containing the coefficients of the projection is used to represent the image. These features can be called as eigenfeatures [6].

Representing each image I (x, y) as a NxN vector Ii , first, the average face ψ is computed:

$$\psi = \frac{1}{R} \sum_{i=1}^{R} I_i$$

(1)

Where R is the number of faces in the training set.

Next, the difference $\emptyset$ of each face from the average face is computed:

$$\phi_i = I_i - \psi .$$

(2)

Then the covariance matrix is estimated by:

$$C = \frac{1}{R} \sum_{i=1}^{R} \phi_i \phi_i^T = AA^T$$

(3)

Where, $A=[\emptyset_1 \quad \emptyset_2 \quad \emptyset_3 \quad \emptyset_4]$

The eigenspace can then be defined by computing the eigenvectors μ_i of C.

Since C is very large ($N^2 X N^2$) , computing its eigenvectors will be very expensive. Instead, we can compute v_i, the eigenvectors of $A^T A$, an $R X R$ matrix.

Then, μ_i can be computed from v_i as follows:

$$\mu_i = \sum_{j=1}^{R} v_{ij}\phi_{j,} j = 1.....R$$

(4)

Given a new image, I , we subtract the mean ($\emptyset=I–\psi$) and compute the projection:

$$\phi = \sum_{i=1}^{R_k} w_i u_i$$

(5)

Where $w_i = u_i^T \phi$ are the coefficients of the projection.

Using projection coefficients images can be represented as linear combinations of the eigenvectors. The projection coefficients define a compact image representation. The eigenspace representation of images is very powerful and it has been used in various applications.

IV. THE IMPORTANCE OF EIGENVECTOR SELECTION

Feature selection is the first and one of the most important steps in pattern recognition. The aim of feature selection is driving an optimal subset of features from a given space leading to high classification performance.

In the eigenspace, all the eigenvectors are not equally informative. It has been found in several studies that different eigenvectors encode different kind of information [7][8].

Generally, in most of techniques the order of the eigenvalues determines the importance of eigenvectors, but as it is discussed in researches such as [9][10] ,this order is not always suitable to describe the data. For example, the first few eigenvectors seem to encode lighting while other eigenvectors seem to encode features such as glasses or moustaches [7]. Although many of the eigen-features are very important for face recognition, they might actually confuse the classifier in other applications. In this study, we consider GA and entropy of eigenvectors to select a good subset of eigen-features in order to improve face recognition performance. It reduces computation and increases recognition rate also.

A. Eigenvector Selection using GA

Selecting a set of eigenvectors for best classification performance is an optimization problem which can be solved by using the genetic algorithm (GA). GAs are inspired by the mechanisms of natural selection [11]. GAs operates iteratively on a population of structures, each of which represents a candidate solution to the problem, encoded as a string of symbols called chromosome. A randomly generated set of chromosomes forms the initial population from which the GA starts its search.

Each image is represented as a vector of eigen-features which are the coefficients of the linear expansion of the image in the eigenspace. In our encoding scheme, the chromosome is a bit string whose length is determined by the number of eigenvectors. Each eigenvector, computed using PCA, is associated with one bit in the string. If the ith bit is 1, then the ith eigenvector is selected, otherwise, that component is ignored. The goal of feature subset selection is to use less features to achieve the same or better performance. If we select only two eigenvectors then in chromosome only two bits are 1, all others are 0. Each chromosome thus represents a different eigen-feature subset.

The initial population is generated randomly. We will end up with a population where each individual contains for example two number of 1's only. The 1's are randomly scattered in the chromosome. Each chromosome is tested to see how good it is at solving the problem at hand and assign a *fitness score* accordingly.

The fitness score is a measure of how good that chromosome is at solving the problem at hand. The fitness of chromosome is measured in terms of recognition accuracy. Individuals with higher accuracy will outweigh individuals with lower accuracy. Mutation is the chance that a bit within a chromosome will be flipped. In proposed method only predefined numbers of bits are flipped. Chromosome with best fitness function is selected for face recognition.

Following are the steps to carryout genetic algorithm.

- Initialized the first population randomly (Only predefined number of bits are 1, all others are 0).

- Begin the evolution loop until the stoppings criteria (maximum number of trials) are reached.Here maximum numbers of trials are 200.

- For each chromosome in the population multiply the chromosome gene values to Eigenfaces.

- For each chromosome compute the fitness function.

- Apply the selection operator and select the fitted chromosomes population to generate the intermediate population.

- Apply mutation operators to form the next generation.

B. Feature Selection Using Entropy of Eigenvectors

Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image. Entropy is defined as

$$E = \sum_{i=1}^{n} P_i \log_2 P_i \qquad (6)$$

Where P_i is probability of occurrence of pixel value i.

In this paper method based on entropy ranking is used. To evaluate the importance and obtain the ranking list of eigenvectors each eigenvector is removed in turn and the corresponding entropy of the rest of the eigenvectors set is computed. If the removal of an eigenvector causes more disorder in the system than another, it shows more importance and higher rank of this eigenvector. For recognition only top ranking eigenvectors are selected.

V. EXPERIMENTS AND RESULTS

The experiment is performed on ORL, GTAV and YALE databases. Face database from AT&T (Olivetti) Research Laboratories, Cambridge (ORL database) [12] contains 40 individuals with each person having ten frontal images. So there are total 400 images in database. There are variations in facial expressions such as open or closed eyes, smiling or no smiling, and glasses or no glasses. All images are 8-bit grayscale of resolution   112 X 92 pixels. For testing of proposed methods we used 340 images (6 images per person) for training and 160 images   (4 images per person) for testing. Image features are extracted using D4 wavelet transform. Four level decomposition is carried out on images and LL subband A3 of each level is used for further processing.

For genetic algorithm base eigenvector selection, experiment is performed in two steps. First A4 subbands of randomly selected 20 images from 160 images are used for testing. Using genetic algorithm best eigenvectors are selected. In second step best eigenvectors selected in first step are used for recognition of all 160 test images. For eigenvector selection using entropy, calculations are done by considering all eigenvectors and then by removing each eigenvector in turn. Top eigenvectors which are causing more difference in entropy after removing are used for reducing dimensionality in PCA.

Same experiment is performed on GTAV [13] and YALE [14] database. In GTAV, 92 images are used as test images and 138 images are used for training. In YALE database 1900 images are used as test images and 570 images are used for training. Similarity measurement between the test image and the training images in the library was performed to determine whether the input test image was matched with any of the images in the library. Here Euclidean distance is used for similarity measurement. Given the test image representation X and the training image representation Y, the similarity measurement $d(X,Y)$ between the two images X and Y is defined by,

$$d(X,Y) = L_{P=2}(X,Y) = \|X - Y\| \qquad (7)$$

Table 1. shows the recognition rate for different number of eigenvectors for ORL database.Table 2 and Table 3 shows recognition rate for GTAV and YALE databases respectively.

TABLE I.   % RECOGNITION RATE FOR ORL DATABASE

No. of Eigenvectors	% Recognition Rate		
	Top Eigenvectors	*Genetic Algorithm*	*Entropy of Eigenvectors*
2	47.5	52.5	37.5
4	79.375	81.25	81.25
6	83.125	90.625	85.00
8	81.875	93.75	90.00
10	81.875	95.00	88.125
12	81.875	96.875	90.00
20	85.625	96.875	90.00

TABLE II.   % RECOGNITION RATE FOR GTAV DATABASE

No. of Eigenvectors	% Recognition Rate		
	Top Eigenvectors	*Genetic Algorithm*	*Entropy of Eigenvectors*
2	80.43	81.52	76.08
4	89.13	91.30	82.60
6	86.95	90.625	89.13
8	84.78	93.75	90.21
10	93.4	95.00	94.56
12	94.56	97.82	94.56
20	96.73	98.91	97.82

TABLE III.   % RECOGNITION RATE FOR YALE DATABASE

No. of Eigenvectors	% Recognition Rate		
	Top Eigenvectors	*Genetic Algorithm*	*Entropy of Eigenvectors*
2	37.00	49.00	35.00
4	76.00	79.00	79.00
6	78.00	81.50	80.00
8	78.42	81.50	80.00
10	79.00	85.00	82.00
12	80	86.00	83.00
20	80	86.00	83.00

VI.   CONCLUSION

Generally, in most of techniques in PCA the order of the eigenvalues determines the importance of eigenvectors, but this order is not always suitable to describe the data. In this paper two methods using genetic algorithm and entropy of eigenvectors are proposed to select best eigenvectors.

With less number of eigenvectors, proposed methods increase the recognition rate. Methods are implemented on A4 subband of wavelet transform which reduces computation.

REFERENCES

[1] B. Moghaddam, W Wahid and A pentland, Beyond eigenfaces: Probabilistic matching for face recognition, *Proceeding of face and gesture recognition*, pp. 30 –35, 1998.

[2] C. Nastar, "The image shape spectrum for image retrieval", *Technical report*, No. 3206, INRIA, June 1997.

[3] C. Nastar, B Moghaddam and A Pentland, "Flexible images: matching and recognition using learned deformations", *Computer Vision and Image Understanding*, Vol. 65, No. 2, pp. 179-191, 1997.

[4] I. Daubechies, "Ten Lectures on Wavelets", *CBMS-NSF series in Applied Mathematics*, Vol.61,SIAM Press, Philadelphia, 1992.

[5] G. C. Feng , P C Yuen and D Q Dai, "Human Face Recognition Using PCA on Wavelet Subband", *J. Electron. Imaging* 9, 226, 2000.

[6] M. Turk and A. Pentland. "Eigenfaces for recognition", *Journal of cognitive neuroscience*. vol 3, no.1 pp. 71-86, 1991.

[7] W. Yambor, B. Draper, and R. Beveridge, "Analyzing PCA-based Face Recognition algorithms: Eigenvector Selection and Distance Measures", *2nd Workshop on Empirical Evaluation in Computer Vision,*2000.

[8] D. Valentin et al, "Principal Component and Neural Network Analyses of Face Images: What can be Generalized in Gender Classification?", *Journal of Mathematical Psychology,* vol. 41, pp. 398-413, 1997.

[9] T. Xiang and S. Gong, "Spectral clustering with eigenvector selection," Pattern Recognition, vol. 41, pp. 1012-1029, 2008.

[10] F. Zhao, et al., "Spectral clustering with eigenvector selection based on entropy ranking," Neurocomputing, vol. 73, pp. 1704-1717, 2010.

[11] D. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison Wesley, 1989.

[12] ORL   database http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html.

[13] F.Tarrés, A. Rama, *GTAV Face Database*. "http://gps-tsc.upc.es/GTAV/ResearchAreas/UPCFaceDatabase/ GTAVFaceDatabase.htm"

[14] [Yale University Face Database, 2002. "http://cvc.yale.edu/projects/yale faces /yalefaces.html"