Face Recognition Based on PCA on Wavelet Subband of Average-Half-Face

M.P. Satone* and Dr. G.K. Kharate**

Abstract—Many recent events, such as terrorist attacks, exposed defects in most sophisticated security systems. Therefore, it is necessary to improve security data systems based on the body or behavioral characteristics, often called biometrics. Together with the growing interest in the development of human and computer interface and biometric identification, human face recognition has become an active research area. Face recognition appears to offer 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. This paper proposes a novel algorithm for face recognition using a mid band frequency component of partial information which is used for PCA representation. Because the human face has even symmetry, half of a face is sufficient for face recognition. This partial information saves storage and computation time. In comparison with the traditional use of PCA, the proposed method gives better recognition accuracy and discriminatory power. Furthermore, the proposed method reduces the computational load and storage significantly.

Keywords—Face Recognition, Principal Component Analysis, Subband, Wavelet Transform

1. INTRODUCTION

Biometric includes identification based on physiological characteristics such as face, fingerprints, finger geometry, hand geometry, hand veins, palm, iris, retina, ear, voice and behavioral traits such as gait, signature and keystroke dynamics [1]. Almost all biometric technologies require some voluntary action by the user, i.e. the user needs to place his hand on a hand rest for finger printing or hand geometry detection and has to stand in a fixed position in front of a camera for iris or retina identification. However, face recognition can be done passively without any explicit action or participation on the part of the user, since face images can be acquired from a distance by a camera. Hence the face recognition system is more appropriate for security and surveillance purposes.

Furthermore, data acquisition in general is fraught with problems for other biometric techniques that rely on hands and fingers. These can be rendered useless if the epidermal tissues are damaged in some way (i.e. bruised or cracked). Iris and retina identification require expensive equipment and are more sensitive to body motion. Voice recognition is susceptible to background noises in public places and auditory fluctuations on a phone line or tape recording. Sig-
natures can be modified or forged. That said, facial images can be easily obtained with a few inexpensive fixed cameras, cannot be modified or forged, and are not affected by background sound noise. Face recognition algorithms with appropriate preprocessing of the images may compensate for noise, slight variations in orientation, scale and illumination.

Face recognition by computer can be divided into two approaches [2, 3], namely, constituent-based and face-based. In the constituent-based approach, recognition is based on the relationship between human facial features such as eyes, mouth, nose, profile silhouettes and face boundary [4-7]. The success of this approach relies highly on the accuracy of the facial feature detection schemes. However, extracting facial features accurately is difficult. Every human face has similar facial features, yet a small deviation in the extraction may introduce a large classification error. The face-based approach [8-12] uses the face as a whole for recognition. Many face-based recognition algorithms have been developed and each has its strength. They all input a full face image into the algorithm. However, none of the methods currently exploit the inherent symmetry of the face for recognition. In this paper we have used the average-half-face as an input to face recognition algorithm. It increases the recognition rate and potentially decreases storage and computation time.

Analysis has been performed using the average-half-face for face recognition in paper [13]. The average-half-face is a transformation method that attempts to preserve the bilateral symmetry present in the face. We have compared the accuracy using both the average-half-face and the original full face image. Principal Component Analysis (PCA) [14-16] has been proven to be an effective face-based approach. Sirovich and Kirby [16] first proposed a method using Karhunen-Loeve (KL) transform to represent human faces. In their method, faces are represented by a linear combination of weighted eigenvectors known as eigenfaces. Turk and Pentland [17] developed a face recognition system using PCA. However, common PCA-based methods suffer from two limitations, namely, poor discriminatory power and large computational load. Swets and Weng [10] observed these drawbacks of the PCA approach and further improved the discriminability of PCA by adding linear discriminant analysis (LDA). In LDA, to get a precise result, a large number of samples for each class are required. On the other hand, O'Toole et al. [18] proposed a different approach for selecting the eigenfaces. They pointed out that the eigenvectors with large eigenvalues are not the best for distinguishing face images. They also demonstrated that although the low dimensional representation is not optimal for recognizing a human face, it gives good results in identifying the physical categories of the face, such as gender and race.

In view of the limitations of the existing PCA-based approach, we proposed a method of applying PCA on wavelet subband of the average-half-face. In the proposed method, an average half face 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 16 x16. Since the proposed method works on lower resolution, instead of 128 x 128 resolution of the original image, the proposed method reduces the computational complexity significantly when the training images are larger than 16 x 16. Moreover, experimental results demonstrated 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.

This paper is organized as follows: Section 2 reviews the average-half-face. Section 3 reviews the background of PCA and eigenfaces. The review of wavelet transform and selection for the proposed method is in Section 4. Experiments and results are discussed in Sections 5 and 6 re-
spectively, and Section 7 gives the conclusions.

2. AVERAGE-HALF-FACE

The face is roughly symmetrical [19]. Zhao and Chellappa [20] used the idea of facial symmetry to solve the problem of illumination in face recognition using Symmetric Shape-from-Shading. Ramanathan et al. [21] introduced the notion of ‘Half-faces’ to assist in computing a similarity measure between faces using images that have non-uniform illumination. In face recognition, the use of bilateral symmetry of the face has been limited to extracting facial profiles for recognition [15, 22]. The average-half-face [13] is inspired by the symmetry preserving singular value decomposition (SPSVD) [14].

An average-half-face is calculated automatically using following steps:

1. The face is centered and oriented in the image. This is done by finding the location of the tip of the nose.
2. The image is partitioned into two equal left and right halves.
3. The ordering of (or mirror) the columns of right half image is reversed.
4. The resulting mirror right half image is averaged with the left half image.

Fig. 1 displays the full face from an example image from the GTAV database [23], the right and left halves of the face after centering, and the average-half-face. The average-half-face can be seen as a preprocessing step to a face recognition algorithm. Less time is required to calculate the average-half-face. Feature selection as well as subspace computation can be performed on the set of average-half-faces just as it is done on a set of full faces. Therefore, the average-half-face can be applied to any face recognition algorithm that uses full frontal faces as an input. The average-half-face requires half the storage space than that for the full face, hence the computation time and storage are reduced.

![Fig. 1. (a) 2D full face image (b) Right half-face (c) Left half-face and (d) Average-half-face](image)

3. REVIEW OF PCA

Principal Component Analysis is based on second order statistics. It is a standard technique used in pattern recognition and signal processing for data dimensionality reduction and feature extraction. The pattern often contains redundant information. Mapping pattern to a feature vector can throw away this redundancy and preserve most of the essential information of the pattern.
These extracted features play a greater role in distinguishing patterns. Every test image can be transformed to a low-dimensional feature vector by projecting on the eigenface space. Eigenface was obtained from the training set. This low-dimensional feature vector can then be compared with the set of feature vectors obtained from the training set. Different distance measures can be used to classify the face. Some important details of the PCA are highlighted as follows [24]:

Let \( X = \{ X_n \in \mathbb{R}^d \mid n = 1, \ldots, N \} \) be an ensemble of vectors.

Each image is converted into a vector and \( X \) is formed by row concatenation of the image data, where \( d \) is the product of the width and height of an image. \( N \) is the total number of images and \( \mathbb{R}^d \) indicates dimension of matrix \( X \).

Let \( E(X) \) be the average vector in the ensemble,

\[
E(X) = \frac{1}{N} \sum_{n=1}^{N} X_n
\]  

(1)

After subtracting the average from each element of \( X \), we get a modified ensemble of vectors,

\[
\overline{X} = \{ \overline{X}_n \mid n = 1, \ldots, N \}
\]  

(2)

with

\[
\overline{X}_n = X_n - E(X)
\]  

(3)

The auto-covariance matrix \( M \) for the ensemble \( X \) is defined by,

\[
M = \text{Cov}(X) = E(\overline{X} \otimes \overline{X})
\]  

(4)

where \( M \) is a \( d \times d \) matrix, with elements,

\[
M(i,j) = \frac{1}{N} \sum_{n=1}^{N} \overline{X}_n(i) \overline{X}_n(j), 1 \leq i, d \leq j
\]  

(5)

Matrix \( M \) is positively definite or semi-definite and has real non-negative eigenvalues [24]. The eigenvectors of the matrix \( M \) form an orthonormal basis for \( \mathbb{R}^d \) which is referred to as the K-L basis. Let \( \{ Y_n, n = 1, \ldots, d' \} \) be the eigenvectors and let \( K \) be the \( d \times d \) matrix whose columns are the vectors \( Y_1, \ldots, Y_d \). The adjoint matrix of the matrix \( K \) is called the K-L transform, which maps the standard coordinates into K-L coordinates. The eigenvectors in \( K \) are sorted according to the eigenvalues in a descending order. Out of \( d \) eigenvectors only \( d' \) eigenvectors corresponding to top \( d' \) eigenvalues are selected. The PCA of the vector \( y \) related to the matrix \( X \) is obtained by projecting vector \( y \) onto the subspaces spanned by \( d' \) eigenvectors, where \( d' \) is smaller than \( d \). This projection has a vector containing \( d' \) coefficients \( a_1, \ldots, a_{d'} \). The vector \( y \) is represented by a linear combination of the eigenvectors with the weights \( a_1, \ldots, a_{d'} \).
Eigenface is the eigenvector obtained from PCA. For face recognition, each training image is transformed into a vector by row concatenation. From the set of training images the covariance matrix is constructed by a set of training images. The significant features, or eigenvectors associated with large eigenvalues, are called eigenfaces. The projection characterizes a face image by a weighted sum of eigenfaces. Recognition is performed by comparing the weight of each eigenface between reference and unknown faces.

PCA has been widely adopted in human face recognition. In spite of its popularity, PCA suffers from two major limitations: poor discriminatory power and large computational load. It is known that PCA gives a very good approximation in face image. However, in eigenspace, each class is closely packed. Moghaddam et al. [25] have plotted the largest three eigen coefficients of each class, where Moghaddam found that they overlapped each other. This shows that PCA has poor discriminatory power. The computational cost for finding eigenvectors is in cubic complexity $O(d^3)$. From matrix theory, we know that if the number of images on the ensemble $N$ is less than the dimension of covariance matrix $M$, then $M$ is singular and ranks greater than $N$. The eigenvectors can be determined by a linear combination of $N$ training images. The computational complexity is then reduced to $O(N^3)$, so the computational complexity can be expressed as $O(r^3)$ where $r=\min(N, d)$.

Pattern often contains redundant information. When we map the pattern to a feature vector, this redundancy can get rid of, and yet preserve, most of the intrinsic information content of the pattern. When right half and left half-face are averaged, all the features are averaged and concentrated in a small area rather than the full face. Therefore, the principle components of the average-half-face carries more information than that of the full face and accuracy increases. If only the right half-face is considered then the features of the left half-face are not at all used, and hence PCA carries less information than the average-half-face and accuracy is less. The same is the case if only the left half-face is considered for face recognition. We proposed a new method of applying PCA on the wavelet subband 4 of average-half-faces. The following sections elaborate our proposed method.

4. WAVELET TRANSFORM

Multiresolution methods provide powerful signal analysis tools, which are widely used in feature extraction, image compression and denoising applications. Wavelet decomposition is a widely used multiresolution technique in image processing. In the recent years, wavelet analysis has generated a great interest in both theoretical and applied mathematics, and the wavelet transform in particular 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 [28]. In the proposed system, WT is chosen to be used in image frequency analysis and image decomposition because of the following reasons:

- Decomposing an image using WT reduces the resolutions of the subband images. The computational complexity will be reduced by working on a lower resolution image.
• Wavelet decomposition provides local information in both space domain as well as in 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). Since CWT provides redundant information and requires a lot of computation, 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.

By using wavelet transform, an image is decomposed into four subbands as shown in Fig. 2(a). The band $LL$ is a coarser approximation to the original image. The bands $LH$ and $HL$ record the contours/edges along the horizontal and vertical directions respectively. 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 as shown in Fig. 2(b). In this paper, Daubechies wavelet D4 , the Daubechies wavelet transform is named after its inventor, the mathematician Ingrid Daubechies is adopted for image decomposition. In fact, in order to select a suitable wavelet, the recognition rates should be computed by applying different wavelets onto face image decomposition. It is shown by G.C. Feng, P.C. Yuen and D.Q. Dai [29] that Daubechies wavelet D4 is better than other wavelets. Therefore, Daubechies wavelet D4 is adopted for image decomposition in our system. The Daubechies D4 transform has four wavelet and scaling function coefficients. If the resolution of an image is 128x128, the subbands 1, 2, 3 and 4 are of size 16x16, the subbands 5, 6, 7 are of size 32x32 and the subbands 8, 9, 10 are of size 64x64. Fig. 3(a) shows an average-half-face image, while Fig. 3(b) shows the 3-level wavelet transform decomposition of the image in Fig. 3(a).

In choosing the WT subband, we used two criteria: First was to reduce the computational complexity. As discussed before, the computational complexity of the PCA-based method is in the cubic order of image resolution, $d$, or number of training images, $N$, depending on which value is smaller. To minimize the computational complexity, we chose to work on the subbands with lower resolution, i.e. 16 x 16. Accordingly, subband 1 to subband 4 were chosen. Applying PCA on a 16 x 16 subimage will bound the computational complexity to the cubic order of 256,

![Fig. 2. (a) 1-level wavelet decomposition and (b) 3-level wavelet decomposition](image-url)
M.P. Satone and Dr. G.K. Kharate

regardless of the size of the training image set. Nastar et al. [30, 31] have investigated 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 focused on subbands containing mid-range frequencies. Among subbands 1 to 4, subbands 2 to 4 were the mid-range frequency subbands. Further experimental results in [29] revealed 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 subband 4 was chosen in our proposed method.

5. EXPERIMENTS

5.1 Training Stage

The images were centered for both the average-half-face and full face recognition results. The system followed the face-based approach where it consisted of two stages, namely, training and recognition stages. The training stage first computed the average-half-face of a full face image, then it found the representational bases for images in the domain of interest (that is reference images), and converted them into training image representations. The training image representation of each image was stored in the library. The recognition stage translated the probe image into probe image representation using the average-half-face and the representational bases, and then matched it with the reference images stored in the library to identify the face image. First, a 3-level Wavelet Transform (WT), using the Daubechies wavelet D4 [26, 28], was applied to decompose the reference average-half-face images. Each reference average-half-face image was then decomposed into 10 subbands. In the second step, PCA was applied on the 16 x 16 subband 4 images. The collection of the subband 4 images would form a new ensemble. An autocorrelation matrix would be constructed for the ensemble as discussed. The output of this step would be a set of eigenvectors and eigenvalues. Eigenvalues were arranged in descending order. The third step was to select the $d'$ eigenvectors with the largest eigenvalues which were used as the representational bases. These $d'$ eigenvectors span the subspace $E = \text{span} \{e_1, \ldots, e_{d'}\}$. After that, the subband 4 images, $x_4$, of all reference images were represented by a linear combination of $M$ representational bases by projecting them into the $M$ eigen subspace. The resultant of the projection of $x_4$ into $E$ is called Training Image Representation. Finally, the training image rep-
representations and the representational bases were stored in the library. Fig. 4(a) shows one of the images from GTAV database and Fig. 4(b) shows the training image representation of the image in Fig. 4(a).

5.2 Recognition Stage

The recognition stage were divided into three steps. Firstly, when an unknown average half face image \( y \) was presented to the recognition stage, it was subtracted by the mean value of the reference half face images and a caricature image was obtained. Then a 3-level WT was applied to transform the caricature image same as in the training stage. In the second step, the 16 x 16 subband 4 subimage was represented as a probe image representation by projecting its subband 4 image into the subspace \( E = \text{span}\{e_1, \ldots, e_M\} \), which was obtained in the training stage. Therefore, we obtained the probe image representation. In the last step, similarity measurement between the probe image and the reference images in the library was performed to determine whether the input probe image was matched with any of the images in the library. Given the probe image representation \( A \) and the training image representation \( B \), the similarity measurement \( \rho(A,B) \) between the two images A and B is defined by,

\[
\rho(A,B) = \frac{\langle A, B \rangle}{\sqrt{\langle A, A \rangle \langle B, B \rangle}}
\]  \hspace{1cm} (6)

Where \( \langle I_1, I_2 \rangle \) denotes the inner product of the two vectors.
The similarity \( \rho \) will return a value between -1 and 1.
If \( \rho(A, B) = 1 \), then A and B are perfectly matched.
If \( \rho(A, B) = -1 \), then A and B are perfectly unmatched.

6. RESULTS

GTAV Face Database [23], ORL[32] and Yale [33] databases were used for testing. The experiment was performed for PCA on full face image, PCA on average-half-face image, PCA on wavelet subband of full face image, and PCA on wavelet subband of average-half-face image. The results were also compared by using the different numbers of training images. The well-known Daubechies wavelet D4 was adopted. Results are as shown in Tables 1 to 5.
**Table 1. Performance Comparison Using ORL Database**

<table>
<thead>
<tr>
<th>No. of Training Samples</th>
<th>Recognition Rate (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PCA/FF</td>
<td>PCA/AHF</td>
<td>PCA/WT/FF</td>
<td>PCA/WT/AHF</td>
</tr>
<tr>
<td>3</td>
<td>76</td>
<td>77.2</td>
<td>79.5</td>
<td>80.4</td>
</tr>
<tr>
<td>4</td>
<td>80</td>
<td>81.4</td>
<td>84</td>
<td>85.5</td>
</tr>
<tr>
<td>5</td>
<td>84</td>
<td>84.8</td>
<td>85.2</td>
<td>86.2</td>
</tr>
<tr>
<td>6</td>
<td>87</td>
<td>87.2</td>
<td>86.4</td>
<td>87.8</td>
</tr>
<tr>
<td>7</td>
<td>87.6</td>
<td>89</td>
<td>95.2</td>
<td>97.4</td>
</tr>
<tr>
<td>8</td>
<td>88</td>
<td>92</td>
<td>97.2</td>
<td>98.6</td>
</tr>
</tbody>
</table>

**Table 2. Performance Comparison Using GTAV Database**

<table>
<thead>
<tr>
<th>No. of Training Samples</th>
<th>Recognition Rate (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PCA/FF</td>
<td>PCA/AHF</td>
<td>PCA/WT/FF</td>
<td>PCA/WT/AHF</td>
</tr>
<tr>
<td>3</td>
<td>77</td>
<td>78.4</td>
<td>82.5</td>
<td>84.5</td>
</tr>
<tr>
<td>4</td>
<td>81</td>
<td>82.4</td>
<td>89.4</td>
<td>92.2</td>
</tr>
<tr>
<td>5</td>
<td>84.2</td>
<td>85</td>
<td>87</td>
<td>88</td>
</tr>
<tr>
<td>6</td>
<td>87.2</td>
<td>87.6</td>
<td>85.2</td>
<td>86.4</td>
</tr>
<tr>
<td>7</td>
<td>88.2</td>
<td>92</td>
<td>98.1</td>
<td>99.1</td>
</tr>
<tr>
<td>8</td>
<td>88</td>
<td>94.2</td>
<td>98.8</td>
<td>99.6</td>
</tr>
</tbody>
</table>

**Table 3. Performance Comparison Using Yale Database**

<table>
<thead>
<tr>
<th>No. of Training Samples</th>
<th>Recognition Rate (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PCA/FF</td>
<td>PCA/AHF</td>
<td>PCA/WT/FF</td>
<td>PCA/WT/AHF</td>
</tr>
<tr>
<td>3</td>
<td>70</td>
<td>82.4</td>
<td>76</td>
<td>90</td>
</tr>
<tr>
<td>4</td>
<td>72</td>
<td>82.7</td>
<td>76.4</td>
<td>90.5</td>
</tr>
<tr>
<td>5</td>
<td>73.2</td>
<td>84.4</td>
<td>78</td>
<td>91</td>
</tr>
<tr>
<td>6</td>
<td>74</td>
<td>85</td>
<td>78.4</td>
<td>92.5</td>
</tr>
<tr>
<td>7</td>
<td>76.4</td>
<td>85.4</td>
<td>79.2</td>
<td>93</td>
</tr>
<tr>
<td>8</td>
<td>77.8</td>
<td>86.7</td>
<td>80</td>
<td>93.5</td>
</tr>
</tbody>
</table>

**Table 4. Storage Required**

<table>
<thead>
<tr>
<th>Database</th>
<th>Full Face</th>
<th>Average-Half-Face</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORL</td>
<td>12 KB</td>
<td>6 KB</td>
</tr>
<tr>
<td>GTA V</td>
<td>32 KB</td>
<td>16 KB</td>
</tr>
<tr>
<td>Yale</td>
<td>72 KB</td>
<td>36 KB</td>
</tr>
</tbody>
</table>

**Table 5. Computation Time in milliseconds**

<table>
<thead>
<tr>
<th>Database</th>
<th>PCA/FF</th>
<th>PCA/AHF</th>
<th>PCA/WT/FF</th>
<th>PCA/WT/AHF</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORL</td>
<td>366</td>
<td>185</td>
<td>275</td>
<td>152</td>
</tr>
<tr>
<td>GTA V</td>
<td>2200</td>
<td>1120</td>
<td>1650</td>
<td>925</td>
</tr>
<tr>
<td>Yale</td>
<td>1002</td>
<td>531</td>
<td>725</td>
<td>400</td>
</tr>
</tbody>
</table>
7. CONCLUSION

This paper proposes a novel algorithm for face recognition based on average-half-face, wavelet subband and PCA application for human face recognition. Average-half-face was used to reduce storage and increase computational speed. The wavelet transform was adopted to decompose an image into different subbands with different frequency components. A mid-range frequency subband was selected for PCA representation. Experimental results showed that the proposed method gave better recognition accuracy, half of storage and less computation time than applying PCA on the whole original image. Nevertheless, there may have been other frequency bands, or even a combination of frequency bands, that would have given a better performance. We would like to further our studies in the future by finding the optimal frequency band(s) for face recognition. We would also like to apply the average-half-face and wavelet to facial feature extraction and to solve the problems of illumination, facial expression, aging and occlusions.

REFERENCES


M.P. Satone

M.P. Satone received her ME in Electronics from RTMU, Nanded in 1998. She is working as an Associate Professor at KKWIEER, Nashik, and is now undertaking a doctorate course. Her research interests are in image and Video Processing, Face Recognition, Microprocessor and Computer Network.
Dr. G.K. Kharate

Dr. G.K. Kharate received his Ph.D. in Electronics and Telecommunication Engineering from Pune University in 2007. He is a Principal at Matoshri College of Engineering and Research, Nashik and also serves as a Dean of Pune University. His research interests are in the areas of Digital Signal Processing, Communication Engineering, Image Processing and VLSI Design.