Face Recognition Using Multi-feature and Radial Basis Function Network

Su Hongtao[(a,b)], David Dagan Feng[(a,c)], Zhao Rong-chun[(b)]

[a]Center of Multimedia Signal processing, Dept. of EIE, Hong Kong Polytechnic University. Hong Kong

[b]Dept. of Computer Science and Engineering, Northwestern Polytechnical University. Xi'an, China

[c]School of Information Technologies, University of Sydney, Australia

Email: su_hongtao@yahoo.com.cn

Abstract

In this paper, a face recognition algorithm using multi feature and Radial basis Function Network (RBFN) is proposed. The algorithm consists of three steps. In the first step, a coarse classification is performed using Fourier frequency spectrum feature, and only the first k gallery images with minimum Euclidean distance to the probe image are retained. In the second step, the Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) features of frequency spectrum are extracted, which will be taken as the input of the RBFN in the third step. In the last step, the classification is carried out by using RBFN. The proposed approach has been tested on ORL face database and Shimon database. The experimental results have demonstrated that the performance of this algorithm is much superior to the other algorithms on the same database.

Keywords: Face Recognition, Classification, Radius Base Function Network, Principal Component Analysis, Linear Discriminant Analysis, Frequency Domain, Fourier Frequency Spectrum

1 Introduction

Face recognition has drawn considerable interest and attention from many researchers for the last two decades because of its potential applications, such as in the areas of surveillance, secure trading terminals, Closed Circuit Television (CCTV) control and user authentication. A number of face recognition methods have been proposed [2][3] and some related face recognition systems have been developed. In general, face recognition methods can be divided into two categories: constituent based and face based. The constituent based face recognition approaches are based on the relationship between individual facial features, such as eyes, mouth, nose and face boundary, etc. To detect facial feature, deformable template could be used as a common method. Yuille et al. [5] proposed a method of eye detection using deformable template in 1992, Xie et al [6] modified the eye template and energy function for eye detection. Okada et al. [7] proposed to use Gabor wavelet to extract facial features. Lam and Yan [8] presented a method of using eye corners to guide the template matching process. After the facial features are extracted, the classification of face can be performed using the distance of these features or elastic graph matching techniques [9], Constituent based method can provide flexibility in dealing with non-rigid facial features, such as eyes and mouth, but it is obvious that the performance of this method highly rely on the accuracy of the facial feature detection, moreover, the computation of template matching is very time consuming. Whereas face based approach uses raw pixel information or a processed form of which as a representation of face. Principal Component Analysis (PCA) is a typical and successful face based techniques. Turk and Pentland developed a face recognition system using PCA in 1991 [10]. To overcome the difficulty caused by illumination variation, Belhumeur et al. proposed Fisherface technique based on Linear Discriminant Analysis (LDA) in 1997 [11]. Comparing to constituent based method, face based method consider face as a global feature, and then the detection of local feature is not required. Furthermore, the computational complexity of this method is not high. However, due to the variation of facial expression, orientation and illumination direction, single feature is usually not enough to represent human face, so that the performance of this approach is quite limited. To overcome this problem, we propose a new algorithm now, which uses multi feature based on frequency domain to represent face and uses RBFN as classifier. The performance of this new algorithm will be tested and validated using ORL face database and Shimon database.

A typical face recognition system includes the following steps: (1) extract human face area from images, i.e. detect and locate face; (2) find a suitable representation of the face region; (3) classify the representations. A comprehensive survey on the techniques used in step (1) can be found in [1]. In this paper we will focus our attention to the techniques used in steps 2 and 3.

In order to better understand the proposed algorithm, some background materials on Principal Component Analysis and Linear Discriminant Analysis will be provided in appendix, the details of the proposed algorithm will be given in section

2, the experiments and results will be presented in section 3. Finally, section 4 will conclude this paper.

2 Algorithm

2.1 PCA and LDA feature in frequency domain

Because the PCA and LDA features of images reflect different properties of images, i.e. PCA features represent the Most Expressive Features and LDA features represent the Most Discriminant Features. If both of them are used into classification, it will have better classification performance than using single feature. To further improve the performance of classification, we adopt the PCA and LDA feature based on frequency domain of image rather than which based on raw image data. Because the Fourier Frequency spectrum is a kind of feature with good classification property, the PCA and LDA feature based on frequency domain should have better classification property than which based on raw image data (spatial domain). This point will be proved by experiments in Section 3. We take the Fourier spectrum of images as the raw data, and then perform PCA and LDA feature extraction base on these data.

2.2 RBFN Classifier

The RBFN classifier is used in our algorithm. In training phase, The PCA feature of Fourier spectrum (FS) of all gallery images is used as input of the first RBFN (we assume it RBFN1), The LDA feature of Fourier spectrum of all gallery images is used as input of the second RBFN (we assume it RBFN2); the TARGET vector is the membership of the input feature corresponding to each class, if there are m classes, and the input vector belong to the k class, then the TARGET vector should be (t_1, $t_{2,\ldots,}t_k,t_{k+1},\ldots t_n$), here

$$\begin{aligned} t_i &= 0, \quad (\mathrm{i} \neq \mathrm{k}) \\ \mathrm{t_i} &= 1, \quad (\mathrm{i} = \mathrm{k}) \end{aligned} \tag{1}$$

Figure 3 shows the implementation architecture of the training and test of RBFN classifier.

2.3 Algorithm architecture

The proposed algorithm consists of three steps. The first step is the coarse classification. Firstly calculate the Fourier spectrum of each gallery image and probe image, then compute the Euclidean distance of the Fourier spectrum of the probe image and each gallery image, and the first k gallery images with minimum distance will be retained as the candidates of the third steps. The second step is feature extraction. Unlike to the general PCA and LDA feature extraction, we extract the PCA and LDA features of Fourier spectrum of image data. The third step is classification using RBFN. Figure 1 describes its implementation architecture. The whole algorithm is given in Figure 2.

The ultimate classification is decided by the OUTPUT vector of RBFN1 and RBFN2, let O_1 be output of RBFN1, O_2 be output of RBFN2, and then the ultimate vector used for classification is

$$O = w_1 \cdot O_1 + w_2 \cdot O_2 \tag{2}$$

where Weight Vector $W = (w_1, w_2)$,and w_1 is the RBFN correct recognition rate only using PCA feature of Fourier spectrum data, while w_2 is the RBFN correct recognition rate only using LDA feature of Fourier spectrum data, and they can be estimated by a recognition test in a small set. If $O=(o_1,o_2,..o_m,o_{m+1},..o_n)$, and assume that $o_m = \max\{o_i \mid i = 1,2,...n\}$, while the k candidate images retained in step 1 belong to $C_{pre} = (c_1, c_2,...c_j)\ (j \leq k)$, if $m \in C_{pre}$, then the test image is classified to the m class, otherwise the test image is classified to the p class, and $o_p = \max\{o_{c_i} \mid i = 1,2,...j\}$.

3 Experiment

3.1 Face databases

Two face databases are used in experiments, Olivetti research laboratory (ORL)face database and Shimon database.

3.1.1 The ORL face database

The database consists of 400 images acquired from 40 persons with variations in facial expression (e.g. open / close eyes, smiling / non-smiling), and facial details (e.g. wearing glasses / not wearing glasses). All images were taken under a dark background, and the subjects were in an upright frontal position, with tilting and rotation tolerance up to 20 degree, and tolerance of up to about 10% scaly. All image are grey scale with a 92*112 pixels resolution. Figure 3 shows two individual samples in ORL database.

3.1.2 Shimon Edelman's face database

There are 16 well-controlled images of each of 11 faces in the database. All faces are of males without distinctive features such as glasses, beards, or mustaches. All images were taken by the same camera under tightly controlled conditions of illumination and viewpoint. The frontal view and neural expression face images under 16 different illumination directions are considered (see Figure 4).

3.2 Experiment Results

3.2.1 Comparison of PCA and LDA feature based on spatial domain and frequency domain

To validate the effectiveness of the PCA and LDA feature of Fourier frequency spectrum, a comparative experiment was carried out on ORL database. The training image number of each class is from 3 to 7, and the rest images are taken as test image. The experiment result was given in Table I, obviously, the recognition performance by using PCA or LDA feature based on frequency domain is superior to which of using corresponding feature based on spatial domain.

3.2.2 Comparison of the performance of the proposed algorithm and others algorithms

Two experiments were carried out on ORL database and Shimon database respectively. The results were given in Table II and III. The training image number of each class

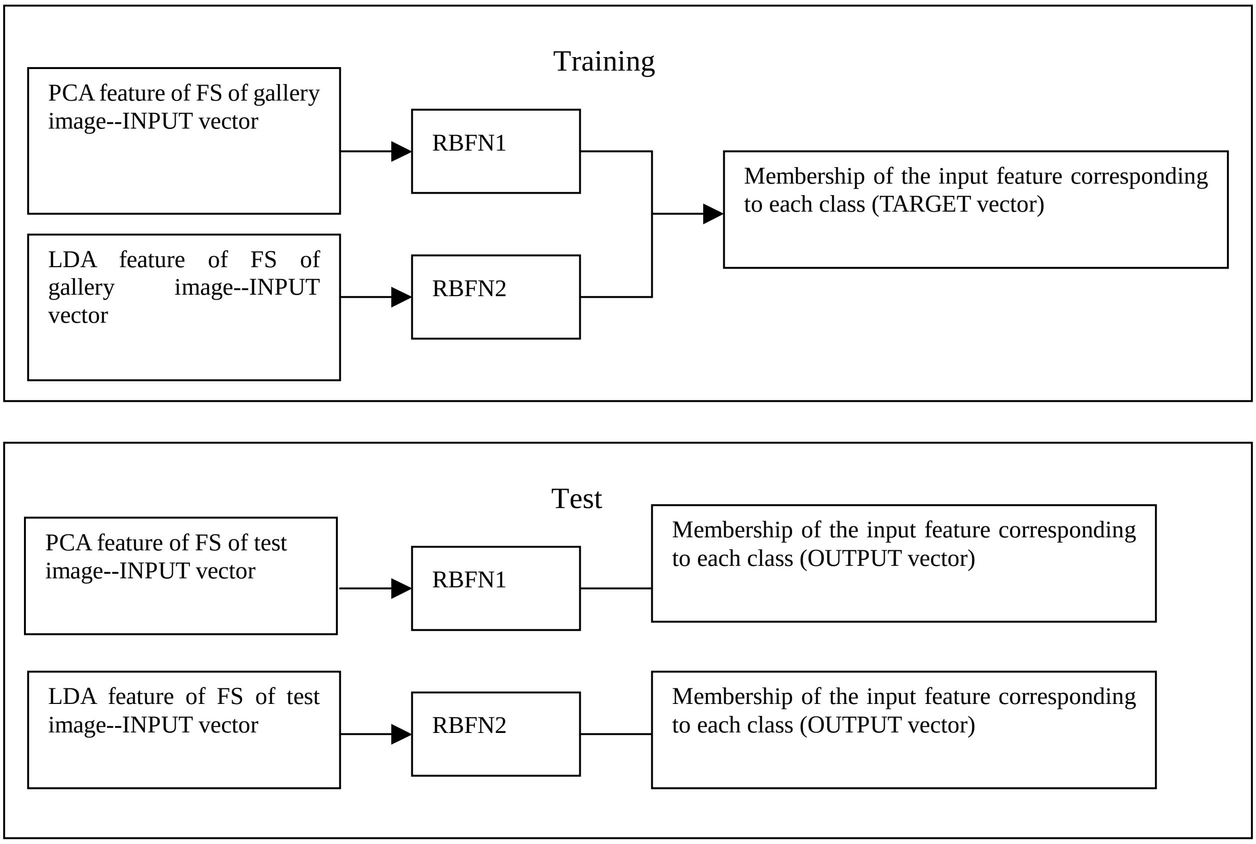

Figure 1 RBFN Classifier

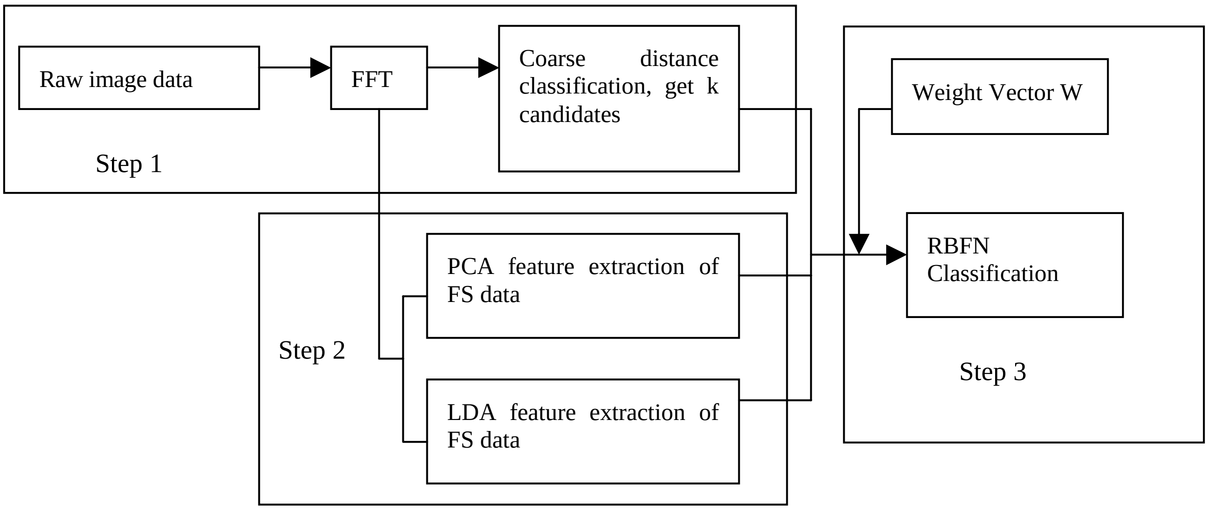

Figure 2 Whole Algorithm

Figure 3 Examples of two individual’s face images in ORL database

Figure 4 An example of same face in Shimon face database

Method \ Training image number	3	4	5	6	7
PCA on spatial domain	80.3%	83.3%	85.0%	90.0%	95.8%
PCA on frequency domain	81.1%	84.6%	87.5%	93.8%	96.7%
LDA on spatial domain	84.3%	88.3%	89.0%	91.3%	93.3%
LDA on frequency domain	87.1%	91.3%	93.5%	96.3%	96.7%

Table I: The correct recognition rate of PCA and LDA feature based on spatial domain and frequency domain using ORL database

Method \ Training image number	3	4	5	6	7
PCA+RBFN	88.9%	92.1%	94.5%	97.5%	99.2%
LDA+RBFN	87.5%	90.8%	94.0%	96.9%	99.2%
FS+RBFN	86.8%	91.7%	92.0%	97.5%	99.2%
Proposed method	91.1%	94.2%	99.5%	99.4%	100%

Table II:Result on ORL face database

Method \ Training image number	6	7	8	9	10
PCA+RBFN	87.3%	92.9%	95.5%	97.4%	98.5%
LDA+RBFN	89.1%	91.9%	96.6%	97.4%	100%
FS+RBFN	89.1%	89.9%	95.5%	98.7%	100%
Proposed method	90.9%	96.0%	100%	98.7%	100%

Table III: Result on Shimon face database

of two experiments is from 3 to 7 and from 6 to 10 respectively, the rest images are test images. In these two experiments, The FS feature is the first 120 coefficients in low frequency. For both ORL database and Shimon database, the proposed algorithm achieved better performance than other algorithms that using single feature and RBF neural network.

4 Conclusion

An algorithm using multi-feature and RBFN was proposed in this paper. The algorithm consists of three steps. In the first step, a coarse classification is performed using Fourier transform coefficient feature, and only first k gallery images with minimum Euclidean distance to the probe image are retained; in the second step, the PCA and LDA features of frequency spectrum are extracted, they will be taken as the input of the RBFN in the next step; In the last stage, the recognition is carried out by using RBFN. Experiments run on different face databases validate this algorithm. The results demonstrated that the proposed algorithm worked very well on the face database with different expression, scale and small-scale variations of pose, and illumination.

To overcome the problem caused by great non-uniform illumination variation, some modification of the proposed algorithm are required, and it will be our future work.

Acknowledgments

This research was supported by the UGC and ARC Grants.

5 References

1 M.H. Yang, D. Kriegman, N. Ahuja (2002): Detecting faces in images: a survey. PAMI 24(1): 34-58.

2 R. Chellappa, C.L. Wilson, S. Sirohey (1995): Human and machines recognition of faces: a survey. Proc. IEEE 83(5): 705-740.

3 A.Samal and P.A.Iyengar (1992): Automatic recognition and analysis of human faces and facial expressions: A survey. Pattern Recognition 25(1): 65-77.

4 W. Y. Zhao, R. Chellappa, A. Rosenfeld, and P. J. Phillips (2000): Face recognition: A literature survey. UMD CfAR Technical Report CAR-TR-948.

5 A.L. Yuille, P.W. Hallinan, D.S. Cohen (1992): Feature extraction from faces using deformable templates, Int. J Computer Vision, 8 (2): 99-111.

6 X. Xie, R. Sudhakar, H. Zhuang (1994): On improving eye feature extraction using deformable templates. Pattern Recognition 27 (6): 791-799.

7 H. Okada, J. Ste!ens, T. Maurer, H. Hong, E. Elagin, H. Neven, C. Malsburg (1998): The Bochum/USC face recognition system, Face Recognition: From Theory to Applications, Springer, Berlin: 186-205.

8 K.M. Lam, H. Yan (1996): Location and extracting the eye in human face images. Pattern Recognition 29: 771-779.

9 L. Wiskott, J.M. Fellous, N. Kruger, C. Malsburg (1997): Face recognition by elastic bunch graph matching. IEEE Trans. PAMI.19: 775-780.

10 M.Turk and A.Pentland (1991): Face recognition using eigenfaces. Proc.IEEE Conf. On Computer vision and Pattern Recognition: 586-591.

11 P. Belhumeur, J. Hespanha, D. Kriegman (1997): Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. IEEE Transaction on PAMI. 19(7): 711-720.

12 Y.Adini, Y.Moses, and S.Ullman (1997): Face Recognition: The Problem of Compensating for Changes in Illumination Direction. IEEE Transaction on PAMI 19(7): 721-732.

13 T.Huntberger, J.Rose, S.Ramaka (1998): Fuzy-Face: A Hybrid Wavelet/Fuzzy Self-Organizing Feature Map System for Face Processing. Journal of Biological Systems Vol.6: 281-298.

14 J. Wu, Z. Chen, Z.-H. Zhou (2001): Projection Incorporated Subspace Method for Face Recognition. Proceedings of the 6th International Conference for Young Computer Scientists, vol.2: 962-965.

15 Ralph Gross, Iain Matthews, Simon Baker (2002): Eigen Light-Fields and Face Recognition Across Pose. Proc. of the 5th Conf. on Face and Gesture Recognition.

16 Georghiades, A., Belhumeur, P., Kriegman, D. (2000): From few to many: Generative models for recognition under variable pose and illumination. In: Proc. of the 4th Conf. on Face and Gesture Recognition.

17 K.Etemad and R.Chellapa (1996): Face recognition using discriminant eigenvectors. In: Proc. of ICASSP.

18 P.Hallinan (1994): A low-dimensional representation of human faces for arbitrary lighting conditions. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, 1994: 995-999.

19 Piater, J. H., &Grupen, R. A. (1999): Toward learning visual discrimination strategies. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition: 410--415.

20 Javad Haddadnia, Karim Faez (2002): A Neural based Human Face Recognition System Using an Efficient Feature Extraction Method with Pseudo Zernike Moment. Journal of Circuits, System, and Computers, 11(3): 283-304.

21 A. Jonathan Howell, Hilary Buxton (1998): Learning identity with radial basis function networks. Neurocomputing: 15-34.

22 T. Kondo, H. Yan (1999): Automatic human face detection and recognition under non-uniform illumination. Pattern Recognition 32(10): 1707-1718.

23 J. H. Lai, P. C. Yuen, G. C. Feng (2001): Face recognition using holistic Fourier invariant features. Pattern Recognition 34: 95-109.

24 Sing-Tze Bow (2002): Pattern Recognition and Image Preprocessing. Marcel Dekker, Inc. New York.

25 S. Ranganath and K. Arun (1997): Face recognition using transform features and neural networks. Pattern Recognition 30: 1615-1622.

Appendix

A.1 Principal Component Analysis

Principal Components Analysis (PCA), also known as Karhunen-Loeve methods, is a technique now commonly used for dimensionality reduction in computer vision, particularly in face recognition. A method called Eigenface based on PCA was used in face recognition in [10], which chooses a dimensionality reducing linear projection that maximizes the scatter of all projected samples. More formally, suppose a set of N sample images $\{x_1,x_2,\ldots,x_N\}$ taking values in an n-dimensional image space, and assume that each image belongs to one of the c classes $\{X_1\ X_2,\ldots,X_c\}$. In order to reduce dimensionality, we need to find a linear transformation which maps the original n-dimensional image space into an m-dimensional feature space, where m < n. The new feature vectors $y_k \in \Re^m$ are defined by the following linear transformation:

$$y_k = W^T X_k \qquad k = 1,2,\ldots\ N \tag{1}$$

where $W \in \Re^{n\times m}$ is a matrix with orthonormal columns. Let S_T be the total scatter matrix,

$$S_T = \sum_{k=1}^{N}(x_k - \mu)(x_k - \mu)^T \tag{2}$$

where N is the number of sample images, μ is the mean image of all samples. After applying the linear transformation W^T, the scatter of the transformed feature vectors $\{y_1,y_2,\ldots,y_N\}$ is $W^T S_T W$. In PCA, the projection *W is* chosen to maximize the determinant of the total scatter matrix of the projected samples, i. e.

$$W_{opt} = \arg\max_W |W^T S^T W| = (w_1\ w_2\ldots\ w_m) \tag{3}$$

where $\{\ w_i \mid i = 1,2,\ldots m\ \}$ is the set of n-dimensional eigenvectors of S corresponding to the m largest eigenvalues. Since these eigenvectors have the same dimension as the original images, they are referred to as Eigenpictures or Eigenfaces and form a basis set for face images. The Eigenfaces span a subspace S that is called the face space. In the training phase, every training face image is projected onto space S and its projection is represented by a m-dimension vector. In the performance phase, an unknown image X is given and projected onto space S. The training image corresponds to the nearest neighbor projection is considered as the best match of X, and its class label is assigned to X.

A.2 Linear Discriminant Analysis

Linear Discriminant analysis (LDA), also known as Fisher Discriminant method, is a very useful statistical tool. It takes into account the different variables of an object and works out which group the object most likely belongs to. Whether or not two or more groups are significantly different from each other with respect to the mean of particular variables.

Let a training set of N face images represent c different subjects. The face images in the training set are two-dimensional intensity arrays, represented as vectors of dimension n. Different instances of the same person are defined to belong to the same class while faces of different subjects should belong to different classes. The basic LDA-based face recognition algorithm is described as following.

A.2.1 Feature extraction

The between-class scatter matrix S_B and within-class scatter matrix S_W are defined by Equation (4) and (5).

$$S_B = \sum_{i=1}^{c} P_i(\mu_i - \mu)(\mu_i - \mu)^T \tag{4}$$

$$S_W = \sum_{i=1}^{c} P_i \sum_{j=1}^{N_i}(x_j^{(i)} - \mu_i)(x_j^{(i)} - \mu_i)^T \tag{5}$$

where $x_j^{(i)}$ is the j th sample vector with n dimension belonging to class i ; μ_i is the mean vector of class i ; μ is the overall mean of sample vectors:

$$\mu_i = \frac{1}{N_i}\sum_{j=1}^{N_i} x_j^{(i)} \tag{6}$$

$$\mu = \frac{1}{\sum_{i=1}^{c} N_i}\sum_{i=1}^{c}\sum_{j=1}^{N_i} x_j^{(i)} \tag{7}$$

S_B represents the scatter of the mean vector μ_i of class i around the overall mean vector μ . S_W represents the average scatter of sample vector of class i . If S_W is non-singular, the LDA selects a matrix $W^{opt} \in \Re^{n\times k}$ with orthonormal columns which maximizes the ratio of the determinant of the between class scatter matrix of the projected vector samples to the determinant of the within-class scatter matrix of the projected samples, i.e.

$$W^{opt} = \arg\max_W \frac{|W^T S_B W|}{|W^T S_W W|} = [w_1, w_2,\ldots, w_k] \tag{8}$$

where $\{w_i \mid i = 1,2,\ldots k\}$ is the set of generalized eigenvectors of S_B and S_W corresponding to the set of decreasing eigenvalues $\{\lambda_i \mid i = 1,2,\ldots k\}$,i.e.

$$S_B w_i = \lambda_i S_W w_i \tag{9}$$

From [10], the upper bound of k is c-1. The matrix W^{opt} describes the Optimal Linear Discriminant Transform or the Foley-Sammon Transform. However, the matrix W^{opt} cannot be calculated directly from Equation (9) because S_W is singular generally. In order to deal with this case, many solutions have been presented in the literatures. The method called Fisherfaces was introduced in [10], where the problem of S_W being singular is avoided by projecting the image set onto a lower dimensional space. The feature space can be reduced to to N-c by using PCA, and then, the dimension can be further reduced to c-1 by applying the standard linear discriminant defined in Equation (9). More formally W^{opt} is given by

$$W^{opt} = W_{fld} W_{pca} \tag{10}$$

where

$$W_{pca} = \underset{W}{\arg\max} |W^T C W| \tag{11}$$

and

$$W_{fld} = \arg\max_{W} \frac{|W^T W_{pca}^T S_B W_{pca} W|}{|W^T W_{pca}^T S_W W_{pca} W|} \tag{12}$$

where C is the covariance matrix of the set of training images .The columns of W^{opt} are orthogonal vectors that are called Fishfaces. Unlike the Eigenface, the Fisherfaces do not correspond to face-like patterns. All sample face images in the training set are projected on the vectors corresponding to the columns of the W_{fld} and the features set is extracted for each sample face image, which can be used directly for classification.

A.2.2 Classification

Like the classification of PCA method, in LDA method, firstly a test face image is projected onto the Fisherface subspace achieved in feature extraction phase, secondly the recognition can be performed by using Euclidian distance in the feature space. In order to improve classification performance, a weighted mean absolute distance (or square distance) is presented in [17], where the weights are obtained on the base of the reliability of the decision axis.