

# EXTENDED FISHERFACE FOR FACE RECOGNITION FROM A SINGLE EXAMPLE IMAGE PER PERSON

*Shiguang Shan, Bo Cao, Wen Gao, Debin Zhao*

FRJDL, Institute of Computing Technology, Chinese Academy of Sciences, P.O.Box 2704, Beijing, China, 100080  
Department of Computer Science, Harbin Institute of Technology, P.O.Box 321, Harbin, China, 150001

## ABSTRACT

In this paper, we extend Fisherface for face recognition from one example image per person. Fisherface is one of the most successful face recognition methods. However, Fisherface requires several training images for each face, so it cannot be applied to the face recognition applications where only one example image per person is available for training. To tackle this problem, we extended the Fisherface method by proposing a method to derive multiple images of a face from one single image. Fisherface is then trained on these derived images. Experimental results on Bern face database and our 350 subjects database show that our method makes impressive performance improvement compared with the conventional Eigenfaces and template matching techniques.

## 1. INTRODUCTION

Face recognition technologies (FRTs) have a variety of potential applications in commerce and law enforcement, such as mug-shot database matching, identity authentication, access control, information security, and surveillance. Related research activities have significantly increased over the past few years [1,2].

As for early researches, geometric feature based methods and template matching methods used to be popular technologies, which were compared in 1992 by Brunelli and Poggio[1]. Their conclusion showed that template matching outperforms the former. Since the 1990s, appearance based methods have been dominant approaches, from which two FRT categories are derived: holistic appearance feature based and analytic local feature based. Popular methods belonging to the former paradigm include Eigenface [3], Fisherface [4], SVD and most NN based methods. Local Feature Analysis (LFA)[5] and Elastic Bunch Graph Matching (EBGM)[6] are typical instances of the latter category. In recent years, Eigenface [3], Fisherface[4,8], EBGM[6], Active Appearance Model (AAM)[7] based approaches have attracted much attention. FERET evaluation has provided extensive comparisons of these algorithms [9]. More recently, SVM has been applied to face recognition successfully [10]. Solutions to

pose and illumination variation problems include invariant feature based methods, 3D linear illumination subspace [4], linear object class [11], illumination and pose manifold [12], Shape-From-Shading [8], photometric alignment [13], Quotient Image [14], and illumination cone [15].

Among these popular methods, Fisherface is one of the most successful approaches. Compared with Eigenface, which extracts Most Expressive Features (MEFs), Fisherface is designed to extract features more suitable for classification purpose, so called Most Discriminating Features (MDFs). Therefore, Fisherface method is proved to outperform Eigenface method in most cases.

However, one of the limitations of Fisherface is that for each person it requires several images for training, which are not always available for some applications such as mug shot matching, suspect recognition etc. To apply Fisherface to these applications, we propose a technique to derive multiple samples from a single face image, by which Fisherface can be trained. To demonstrate the effectiveness of this method, experiments are conducted on the Bern face database and our own 350 subjects face database. The performance of our method is compared with two benchmarks, that is, Eigenface and template matching.

The remaining of the paper is organized as follows. In Section 2, Fisherface method is described briefly and its advantage is emphasized. Section 3 describes the method to derive multiple images of a face from a single image. Extensively comparative experiments are carried out in Section 4. A conclusion is drawn in the last section.

## 2. FISHERFACE METHOD

Fisher's Linear Discriminant (FLD) is a widely used method for feature extraction and dimensionality reduction in pattern recognition. FLD tries to find the "best" project direction in which training samples belonging to different classes are best separated. Mathematically, FLD selects the projection  $W_{fld}$  in such a way that the ratio of the determinant of the between-class scatter matrix of the projected samples and the within-class scatter matrix of the projected samples is maximized.

For a  $c$ -class problem, the between-class scatter matrix is defined as follows:

$$S_b = \sum_{i=1}^c \Pr(\Omega_i) (\mu_i - \mu)(\mu_i - \mu)^T$$

where  $\Pr(\Omega_i)$  is the prior class probability,  $\mu_i$  is the mean sample of class  $\Omega_i$  and  $\mu$  is the mean sample of all classes.

The within-class scatter matrix is defined as follows:

$$S_w = \sum_{i=1}^c (\Pr(\Omega_i) \times \frac{1}{N_i} \sum_{y_k \in \Omega_i} (y_k - \mu_i)(y_k - \mu_i)^T)$$

where  $N_i$  is the number of samples in class  $\Omega_i$ .

If  $S_w$  is non-singular, the project matrix  $W_{fld}$  can be chosen as follows:

$$W_{fld} = \arg \max_W \frac{|W^T S_b W|}{|W^T S_w W|}$$

Then,  $W_{fld}$  can be calculated by solving the generalized eigenvalue problem:

$$S_b W = S_w W \Lambda$$

In the face recognition applications, however, the within-class scatter matrix  $S_w$  is always singular. Because the rank of  $S_w \in R^{n \times n}$  is at most  $N-c$ , where  $N$  is the number of images in the training set, and  $N$  is typically much smaller than  $n$  which is the number of pixels in each image. To overcome this difficulty, PCA is first used to reduce the dimension of the images from  $n$  to  $N-c$  or less, then recalculated  $S_w$  will be non-singular and FLD can be used to find the project matrix  $W_{fld}$ , which is referred to as Fisherface in [4].

Principal Component Analysis (PCA), well known as Eigenface method in face recognition, is also a widely used feature extraction method. In PCA, the project matrix  $W_{pca}$  is chosen to maximize the determinant of the total scatter matrix of the projected samples. The total scatter matrix is defined as follows:

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

where  $N$  is the number of all samples and  $\mu$  is the mean sample.  $W_{pca}$  can then be calculated by solving the eigenvalue problem:

$$S W = W \Lambda$$

From the above definition of PCA, it can be concluded that the project matrix  $W_{pca}$  maximizes the distances between all the samples in the projected space. The

distances between training samples are due not only to differences between different classes but also to differences between the training samples of the same class. In a specific application, if the samples of the same class differ significantly, PCA may be unsuitable for classification purpose.

On the contrary, FLD minimizes the within-class scatter while maximizing the between-class scatter. Consequently, the samples of different classes in the projected space are separated and those of the same class are clustered. In the face recognition problem, it has been observed that "the variations between the images of the same face due to illumination and viewing direction are almost always larger than image variations due to change in face identity". Thus, Fisherface is more suitable for face recognition than Eigenface.

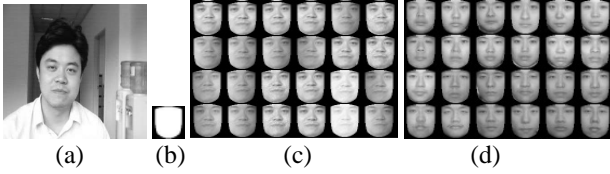
### 3. DERIVING MULTIPLE IMAGES FROM A SINGLE IMAGE

Though, as has discussed in above section, Fisherface does have many strengths, however, it is also well known that Fisherface cannot be applied to face recognition problems where only one training image per person are available, such as mug shot matching, suspect identification etc. Multiple training images per person are needed to compute the within-class scatter matrix  $S_w$ . To solve this problem, we propose a simple technique to derive multiple samples from a single example image. The technique is based on the following two intuitive propositions:

1. Proper geometric transforms, such as translation, rotation in image plane, scale variance etc., do not change the identity attribute of a face image visually.
2. Proper gray-level transforms, such as simulative directional lighting, man-made noise, etc., do not change the identity attribute of a face image visually.

In the proposed technique, the two kinds of transforms are combined to derive tens of training examples from a single example image, which are training sample images of the Fisherface method. Fig. 1(c) illustrates some normalized "virtual" examples derived from one face image as shown in Fig. 1(a) by using our technique.

In addition, to alleviate the influence of translation, rotation, lighting and scale variance, geometric and gray-levels normalization are adopted. As to geometric normalization, the locations of the two irises are first localized manually and then fixed at specific locations by affine transformation. A mask, as shown in Fig.1 (b), is covered over the face region to eliminate the alterable background and hairstyle. Finally all faces are warped to the size of 32x32 as shown in Fig.1 (d). Histogram equalization is conducted to normalize illumination, and all the face data are vectorized to unit length before they are fed into the training or testing procedure.



**Fig.1 Deriving multiple samples from a single image and normalization (a) a input face image (b) the mask (c) the derived multiple examples from the face in (a) (d) the normalized faces**

#### 4. EXPERIMENTAL RESULTS

To verify the effectiveness of the proposed approach, we also develop Eigenface method and template matching as benchmarks. Experiments are conducted on Bern database, and our own face database containing 350 subjects.

##### 4.1 Benchmarks And Performance Evaluation

Eigenface and template matching are de facto the standard benchmarks in the face recognition community. In our experiments Eigenface method is designed according to [3]. All faces are normalized as in Fig.1 (d). Template matching is also operated on these normalized faces.

Also, for Eigenface method, our experiments have shown that, recognition performance can be improved by using  $\cos(\theta)$  similarity measurement and by dropping some leading Eigenfaces. So, we have also reported the results of this improved Eigenface method to compare with our proposed method.

The performances of these approaches are evaluated and compared by using Cumulative Recognition Rate (CRR) defined as:

$$p(i) = \frac{N_i}{N}.$$

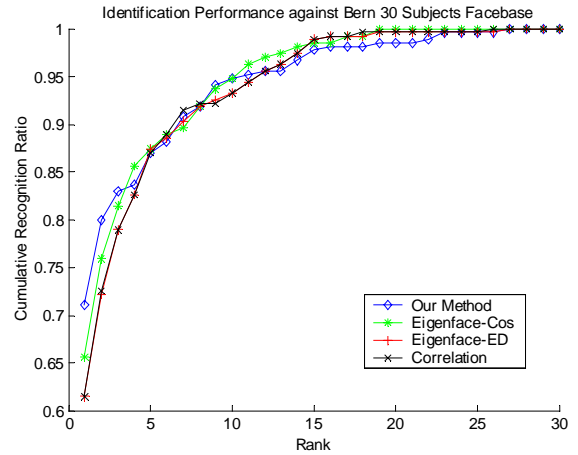
where  $N$  is the total number of face images tested,  $N_i$  is the number of test images recognized correctly in the top  $i$  candidates. For each algorithm, a CRR-Rank curve is plotted, whose horizontal axis is rank and vertical axis is the CRR.

##### 4.2 Experiments On Bern 30 Subjects Face Database

To verify the effectiveness of the proposed framework, comparative experiments are conducted on Bern 30 subjects multiple poses face database. The Bern database consist of 300 examples images of 30 subjects, for each person 10 gray-level images with variations of the head positions (1,2 looking right into the camera, 3,4 looking to the right, 5,6 looking to the left, 7,8 looking downwards, 9,10 looking upwards)\*.

In our experiments, the No.“1” examples (looking right into the camera) of each subject in the database are chosen as the example images to form the training set (30

examples totally). For Fisherface method, 17 novel face images derived from each normal face image as described in section 3 are used as the training sample images for each person, totally 510 training images are used. The performance curves of our method and benchmarks are shown in Fig. 2. It is clear that our method outperforms the other approaches. The Rank-1 (first-choice) recognition ratio of Fisherface method is 71.1%, while that of Eigenface method is 65.6% and that of correlation method is 61.5%.



**Fig.2 Performance comparisons on Bern face database**

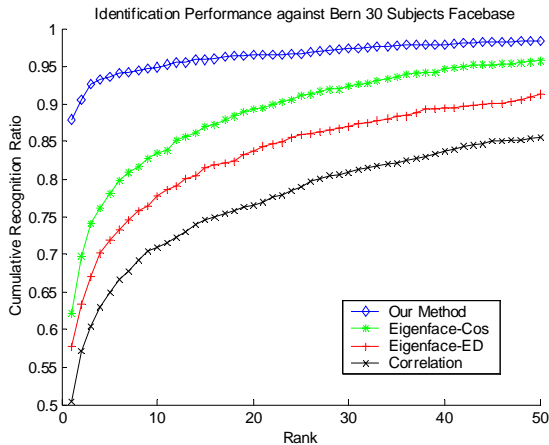
##### 4.3 Experiments On Our 350 Subjects Face Database

To further demonstrate the performance and scalability of Fisherface method on larger database, more detailed experiments are conducted on a 350 subjects face database. For totally 350 subjects, 1750 images are acquired, with 5 images per subject. All images are taken with a general USB camera. For each subject, 1 normal face (nearly frontal, neutral expression and ambient lighting condition) is chosen as the training example; therefore a training/gallery set containing 350 faces is constructed. All the remaining 1400 images (4 examples per subject) constitute the probe set, which contains face images with different expressions, lighting conditions and slight pose variance. Obvious difference can easily be seen between the images in the gallery set and the probe set. The algorithms tested on our database are as follows.

1. Our Fisherface method with cosine similarity, using the leading 90 Fisherfaces. [Denoted as “Our method” in Fig.3]
2. Eigenface method with cosine similarity, using the leading 300 Eigenfaces but the first 2 ones. [Denoted as “Eigenfaces-Cos” in Fig.3]
3. Eigenface method with Euclidean distance, using the leading 300 Eigenfaces but the first 2 ones. [Denoted as “Eigenfaces-ED” in Fig.3]
4. Template Matching on the normalized 32x32 template, [Denoted as Correlation in Fig.3]

\* Copyright 1995 University of Bern All Rights Reserved

Their performances are illustrated in Fig.3. Apparently the Fisherface method significantly outperforms all other algorithms. The Rank-1 (first-choice) recognition rate of our method is 87.9%, while that of the improved Eigenface method is 62.1% (Eigenfaces-Cos). Note that the dimensionality of eigenspace for Eigenface method is determined on the experiments, when the peak correction ratios are observed. 300 Eigenfaces are used in our experiments.



**Fig. 3 Performance comparisons on our 350 subjects face database.**

#### 4.4. Observations on Experimental Results

The experiments on the two face databases show that our Fisherface-based method outperforms the benchmarks significantly, especially on our 350 subjects face database.

Note that our Fisherface-based method does not perform in the Bern face databases as better as it does in our 350 subjects face databases (though it still outperforms the benchmarks). The reason is that the simple algorithm used in our paper to generate multiple training images from one image cannot compensate for the large pose variation in Bern face database. However, its excellent performance in our face database is courageous, which suggests more complicated “virtual” views generation approaches are valuable to be studied.

#### 5. CONCLUSIONS AND FUTURE WORKS

In this paper Fisherface is extended to face recognition from one training image by introducing a simple but effective technique to derive multiple samples from a single example image. Experiments on the Bern face databases and our 350 subjects databases demonstrate impressive performance improvement of our method over the conventional benchmarks, for face recognition training from one image while testing under some expression, illumination and slight pose variations.

Future researches will be focused on more complicated algorithms to generate more “virtual” views covering

image variation due to large illumination and pose variance.

#### 6. ACKNOWLEDGEMENT

This research is sponsored partly by Natural Science Foundation of China (No.69789301), National Hi-Tech Program of China (No.863-306-ZD03-01-2), 100 Talents Foundation of Chinese Academy of Sciences, and SiChuan Chengdu Yinchun Net. Co.

#### REFERENCES

- [1] R. Brunelli and T. Poggio, “Face Recognition: Features versus Template”, *TPAMI*, 15(10), pp1042-1052, 1993
- [2] R.Chellappa, C.L.Wilson ect. “Human and Machine Recognition of faces: A survey”, *Proc. of the IEEE*, 83(5), pp705-740, 1995.5
- [3] M.Turk and A.Pentland. “Eigenfaces for Recognition” *Journal of cognitive neuroscience*, 3(1), pp71-86, 1991.1
- [4] P.N.Belhumeur, J.P.Hespanha and D.J.Kriegman. “Eigenfaces vs Fisherfaces: recognition using class specific linear projection”. *TPAMI*, vol.20, No.7, 1997.7
- [5] P.Penev and J.Atick, “Local Feature Analysis: A General Statistical Theory for Object Representation,” *Network: Computation in Neural Systems*, vol.7, pp.477-500, 1996
- [6] L Wiskott, J.M.Fellous, N.Kruger and C.V.D.Malsburg, “Face Recognition by Elastic Bunch Graph Matching”, *IEEE Trans. On PAMI*, 19(7), pp775-779, 1997.7
- [7] T.F.Cootes, G.J.Edwards, C.J.Taylor, “Active Appearance Models”, *ECCV*, vol.2, pp484-498, 1998.
- [8] W.Zhao and R.Chellappa, “Robust Image-Based 3D Face Recognition”, *CAR-TR-932, N00014-95-1-0521, CS-TR-4091*, Center for Auto Research, UMD, 2000.1
- [9] P.J.Phillips, H.Moon, etc. “The FERET Evaluation Methodology for Face-Recognition Algorithms”, *IEEE TPAMI*, Vol.22, No.10, pp1090-1104, 2000
- [10]G.Guo, S.Z.Li and K.Chan, “Face Recognition by Support Vector Machines”, *FG’02*, pp196-201, Grenoble, 2000.3
- [11]T.Vetter and T.Poggio, “Linear Object Classes And Image Synthesis From A Single Example Image”, *IEEE Trans. On PAMI*, Vol.19, pp733-742, 1997
- [12]H.Murase, S.Nayar, Visual Learning and recognition of 3D object from appearance, *IJCV*, 14:5-24, 1995
- [13]A.Shashua, On Photometric Issues in 3D visual recognition from a single 2D Image, *International Journal of Computer Vision*, 21(1/2), 99-122, 1997
- [14]A.Shashua and T.Riklin-Raviv, “The Quotient Image: Class-Based Re-Rendering And Recognition With Varying Illuminations”, *IEEE Trans. on PAMI*, pp.129-139, 2001.2
- [15]A.S.Georgiades, P.N.Belhumeur and D.J.Kriegman, “From Few to Many: Illumination Cone Models for Face Recognition under Differing Pose And Lighting”, *IEEE TPAMI*, Vol.23, No.6, pp643-660, June 2001