FACE IDENTIFICATION FROM A SINGLE EXAMPLE IMAGE BASED ON FACE-SPECIFIC SUBSPACE (FSS)

Shiguang Shan[1], *Wen Gao*[1,2], *Debin Zhao*[2]

[1]FRJDL, Institute of Computing Technology, CAS, P.O.Box 2704, Beijing, China, 100080
[2]Department of Computer Science, Harbin Institute of Technology, Harbin, China, 150001

ABSTRACT

In this paper, a novel face representation, named Face-Specific Subspaces (FSS), is proposed first. Then the corresponding face identification strategy is provided. This method is motivated, but essentially different, from Eigenface. In the proposed method, each face holds one private face subspace, while in Eigenface all faces share one common face subspace. To enable the proposed approach to tackle single example problem, a technique to derive multi-samples from single example is further developed. Extensive experiments on Yale face database Bern face database, and our 350 subjects database show that our method makes impressive performance improvement when compared with conventional Eigenface and template matching, which intensively indicates the robustness of our approach against appearance variances due to expression, illumination and pose.

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 researches, from which two FRT categories were 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 cones [15].

This paper extends the Eigenface method by proposing the idea of representing each face by using a face subspace. Corresponding similarity measure and minimum reconstruction error rule are presented. Extensive experiments demonstrate the effectiveness of our approach.

2. FSS BASED FACE IDENTIFICATION

2.1 Eigenface methods

As is well known in the face recognition community, Eigenface is essentially based on the idea that face images can be regarded as points in the high dimensional image space. They are believed to approximately form a subspace, so called "face subspace", which can be expanded by some leading Eigenfaces. Fig.1 visually illustrates the idea to describe a face image as the linear combination of some leading Eigenfaces.

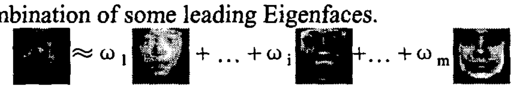


Fig.1 One face image is represented as the linear combination of *m* leading Eigenfaces

A recognized nature of Eigenface method is the Distance From Face Subspace (DFFS), i.e., reconstruction error, which can be used to detect the occurrence of faces. It is just this point that inspires us deriving the idea that, reconstruction error can be employed to detect the occurrence of a specific face, if a subspace is learnt from the face examples of the specific face.

2.2 Learning Face-Specific Subspace (FSS)

In this section, we propose to represent each face by using one face subspace, named Face-Specific Subspace (FSS), learnt from the training images of the face. Formally, the FSSs can be learnt as follows:

Let the class set of the faces to be identified is:

$$C=\{\Omega_1,\Omega_2,\cdots,\Omega_p\},$$

where p is the number of faces to be recognized. Then for the k^{th} face class $\Omega_k, k = 1,2,\cdots,p$ in C, eigen-decomposition is conducted as:

$$U_k^T \Sigma_k U_k = \Lambda_k,$$

where Σ_k is the covariance matrix of the k^{th} face, Λ_k is the diagonal matrix whose diagonal elements are the decreasingly ordered eigenvalues $\lambda_1^k, \lambda_2^k, \cdots, \lambda_{d_k}^k$ of Σ_k, and $U_k = [\mu_1^{(k)}, \mu_2^{(k)}, ..., \mu_{d_k}^{(k)}]$ is the matrix formed by the eigenvectors of Σ_k, where $\mu_1^{(k)}, \mu_2^{(k)}, ..., \mu_{d_k}^{(k)}$ are eigenvectors corresponding to eigenvalues $\lambda_1^k, \lambda_2^k, \cdots, \lambda_{d_k}^k$ respectively. So the following bases matrix spans the k-th Face-Specific Subspace:

$$U_k = (\mu_1^{(k)}, \mu_2^{(k)}, ..., \mu_{d_k}^{(k)}).$$

To sum up, the k^{th} face is represented as a *4*-tuple, that is the k^{th} FSS, by:

$$\Re_k = (U_k, \Psi_k, \Lambda_k, d_k),$$

where Ψ_k is the mean of the k^{th} face, and d_k is the dimension of the FSS.

2.3 Identify Faces Based on FSS

After FSS for each face is learnt, similar to DFSS in Eigenface method, the similarity of any image to a face can be measured by using the Distance From FSS (DFFSS): less DFFSS means more probability that the image belongs to the corresponding face. It can be formulated as follows:

Let Γ be any input image. It can be projected to the k^{th} FSS by:

$$W^{(k)} = U_k^T \Phi^{(k)}, \text{ where } \Phi^{(k)} = \Gamma - \Psi_k.$$

Then $\Phi^{(k)}$ can be reconstructed by:

$$\Phi_r^{(k)} = U_k W^{(k)}.$$

So, Γ's distance from k^{th} FSS (DFFSS) is computed as:

$$\varepsilon^{(k)} = \left\| \Phi^{(k)} - \Phi_r^{(k)} \right\|.$$

The DFFSS reflects the quantity of the k^{th} face pattern "hiding" in the input image Γ, or in other words, the power of the k^{th} FSS to reconstruct the input pattern Γ. So it can be regarded as the similarity of the input pattern Γ to the face corresponding to the k^{th} FSS. Therefore, the following minimal distance classifier can be naturally formulated:

$$\Gamma \in \Omega_m \text{ if } \varepsilon^{(m)} = \min_{1 \le k \le p} \{\varepsilon^{(k)}\}.$$

To demonstrate the rationality of the above recognition strategy intuitively, further reconstruction experiments are conducted on FSS for different input patterns. To get comparable visual effects, the reconstruction is carried out by the following formula:

$$\Gamma' = \left\| \Gamma - \Psi^{(k)} \right\| \cdot (\Phi_r^{(k)} + \Psi^{(k)}).$$

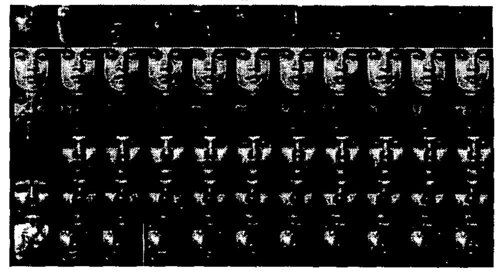


Fig.2 Eigenfaces of No.196 face and its ability to reconstruct different patterns

Fig. 2 illustrates the power of one specific FSS to reconstruct various input patterns. The first line shows the leading 11 Eigenfaces of No.196 FSS trained from one face image of subject No.196. The first columns from line 2 to 6 are the input patterns. The subsequent pictures in each line illustrate the reconstructed patterns by using the leading 10~19 Eigenfaces of No.196 FSS. The input face image in line 2 belongs to subject No.196, and we can see that the reconstructed faces are quite similar in appearance to the input face. The input faces in line 3~5 are non-No.196 faces, and it is clear that the reconstructed faces are quite different from the corresponding input face but still very similar to the No.196's face. The last line illustrates the case when one non-face pattern is fed into the reconstructing procedure, where much more difference between the input pattern and reconstructed ones can be seen clearly. Obviously, FSS has the favorable nature to reconstruct its own face patterns perfectly, however it is not the case for face patterns of other subjects. This strongly suggests that the FSS based face representation has excellent class discriminating power.

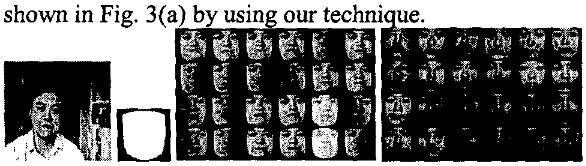
3. FACE RECOGNITION FROM SINGLE EXAMPLE IMAGE

As we know, to learn a face subspace, multiple training example images are required. For FSS, it means more than one example per face is needed to train his/her FSS. But for some face recognition applications, such as mug shot matching, suspect identification etc., only few (even single) face images are available for each subject involved, therefore, FSS based method cannot be applied to them directly. To solve this problem, we further propose a simple technique to derive multiple samples from 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 changes 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 single example image, which are then fed into the FSS learning procedure. Fig. 3(c) illustrates some normalized "virtual" example images derived from one face image as shown in Fig. 3(a) by using our technique.



(a) (b) (c) (d)

Fig.3 Deriving multiple samples from single image and normalization (a) input face image; (b) mask; (c) derived multiple examples from face in (a); (d) normalized faces

In addition, to alleviate the influence of translation, ration, 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 placed at fixed locations by affine transformation. A mask, as shown in Fig.3 (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.3 (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.

4. EXPERIMENTAL RESULTS

To verify the effectiveness of the proposed approaches, we also develop Eigenface method and template matching as benchmarks. Extensive experiments are conducted on Yale database, Bern database, and our own face database containing 350 different 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.3 (d). Furthermore, some improvements are taken. Template matching is also operated on these normalized faces.

Their performances are evaluated and compared by using Cumulative Recognition Rate (CRR). For each algorithm, a CRR-Rank curve is plotted, whose horizontal axis is rank and vertical axis is the probability of cumulative correct identification.

4.2 Experiments On Yale 15 Subjects Face Database

The Yale face database contains 165 images of 15 subjects, with 11 images per subject, among which there is a normal image with neutral expression, taken under ambient lighting conditions, while the left 10 images cover different cases including faces with/without glasses, images with various expressions, images illuminated by center-light, left-light and right-light. Refer to [4] for details.

In our experiment, all the 15 normal face images (one for each subject) are chosen to form the training set and gallery set, and all the other images (150 images) constitute the probe set for all the algorithms tested. For our FSS based method, 15 FSSs are learnt respectively from the 15 normal face images. The performance curves of different methods are plotted in Fig. 4. It is clear that the proposed method extraordinary outperforms the other approaches. The Rank-1 (first-choice) recognition ratio of our method is 95.33%, while that of Eigenface is 74.67%.

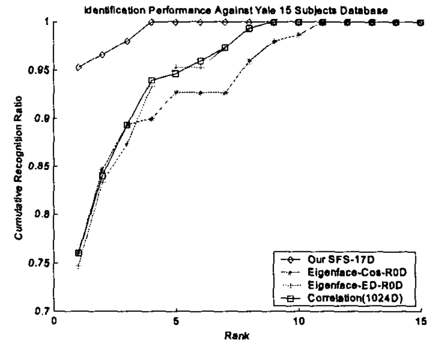


Fig.4 Performance comparisons on Yale database.

4.3 Experiments On Bern 30 Subject Multi-Pose Face Database

To verify the effectiveness of the proposed framework on multi-pose face recognition problem, comparative experiments are conducted on Bern 30 subjects multiple poses face database. The Bern database consist of 300 examples images of 30 subject, for each person 10 gray-level images with slight variations of the head positions (1,2 right into the camera, 3,4 looking to the right, 5,6 looking to the left, 7,8 downwards, 9,10 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). Performance curves are plotted in Fig.5. Again, our FSS based method outperforms significantly other algorithms, since the first-choice recognition rate of our method is 77.78%, while Eigenface method is 65.56%.

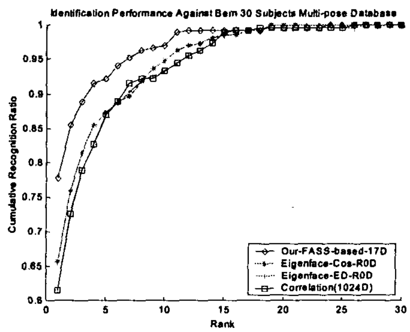


Fig. 5 Performance comparison on Bern face database

4.4 Experiments On Our 350 Subjects Face Database

To further demonstrate the performance and scalability of our FSS based method on larger database, more detailed experiments are conducted on a 350 subjects face database. For the 350 subjects, 1750 images are acquired, with 5 images per subject. All images are taken with a general USB camera. For each subject, one 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 cover 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.

Several algorithms are tested on our database. Their performance curves are plotted in Fig 6. Apparently the proposed method significantly outperforms all other algorithms. The Rank-1 recognition rate of our method is 88.36%, while that of the improved Eigenface is 61.57%.

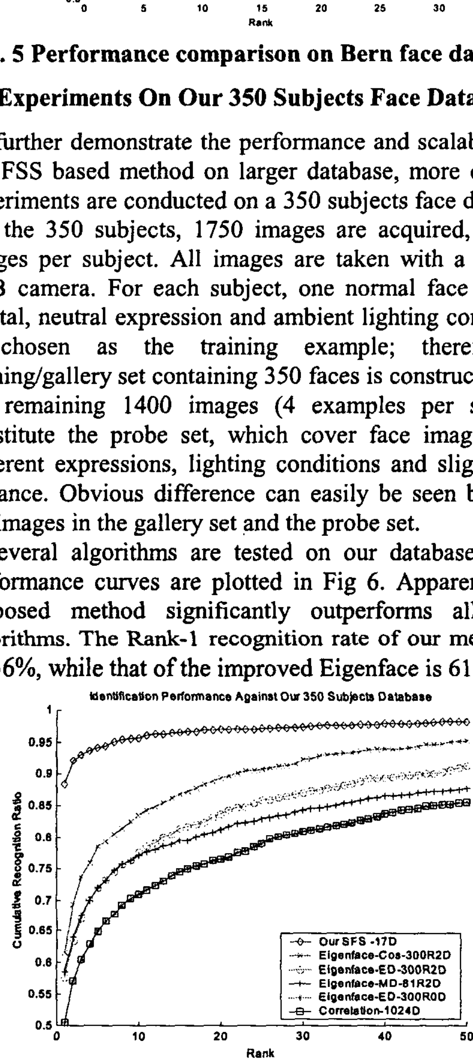


Fig.6 Performance comparison on our 350 subjects face database.

5. CONCLUSIONS AND FUTURE WORKS

We propose in this paper to represent each face by using Face-Specific Subspace (FSS) and present the FSS based face recognition method. Aiming at face recognition from single example image, a technique to derive multiple samples from single example image is further proposed. Extensive experiments demonstrate the excellent performance of our new method over conventional benchmarks under expression, illumination, and slight pose variance.

Though the proposed method requires more storage for each face and more time to recognize one face, however its computational complexity is linear and the recognition can be thoroughly conducted in parallel.

Future efforts will be devoted to more deliberate virtual views synthesis algorithms to derive multiple samples from single example view.

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 Recogniton 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.Georghiades, 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*