Incremental Face-Specific Subspace for Online-Learning Face Recognition

Wenchao Zhang[1], Shiguang Shan[2], Wen Gao[2], Jianyu Wang[1] and Debin Zhao[1,2]

[1](Department of Computer Science, Harbin Institute of Technology, Harbin 150001, P.R.China)
[2](ICT-ISVISION Joint R&D Laboratory for Face Recognition, CAS, Beijing 100080, P.R.China)
[1]{wczhang, jywang, dbzhao}@jdl.ac.cn, [2]{sgshan, wgao}@ict.ac.cn

ABSTRACT

A practical face recognition system is expected to have the ability to learn online to adapt to different variations of the imaging conditions in order to achieve better recognition performance, especially when batch training is impossible. This is commonly achieved by updating the face model incrementally for each face. Based on our previous Face-Specific Subspace (FSS) face recognition method, in this paper, an incremental subspace updating method is further applied to FSS in order to make it have the ability to learn online, which is named Incremental Face Specific Subspace (IFSS). Since in the FSS face recognition method, each individual face is represented as one specific face subspace, therefore, the face model can be updated even only single live example face image is available as well as its corresponding class label. Experiments on the Harvard face database show that better recognition performance can be achieved when more example images are incrementally fed into the IFSS system. Furthermore, Experiments also show that IFSS has comparable performance with batch training FSS for the same training sets, but less computational resource is needed and the training examples need not be stored.

1. INTRODUCTION

Face recognition technology (FRT) has numerous commercial and law enforcement applications. These applications range from static matching of controlled format photographs such as passports, credit cards to real-time matching of surveillance video images presenting different constraints in terms of processing requirements. Related research activities have significantly increased over the past few years, many algorithms are proposed to solve the face recognition problem [1][2].

However, the performance of almost all current face recognition systems, both academic and commercial systems, is heavily subject to the variance in the imaging conditions. It has been discovered by the FERET testing that pose and illumination variations are two bottlenecks for a practical face recognition system [2]. Obviously, one of the intuitive solutions to these problems is to let the system "see" more examples captured under different imaging conditions. However, when new examples are fed into the recognition system one by one, it is obviously inefficient and impractical to batch-training the face recognition system from all the previous examples and the new input ones.

Therefore, a practical face recognition system is expected to have the ability to learn online to adapt to different variations of the imaging conditions in order to achieve better recognition performance, especially when batch training is impossible. This is commonly achieved by updating the face model incrementally for each face. By doing so, less spatial-temporal computational resource is needed, because those previous examples need not be stored and relearned.

Among the typical face recognition methods, *Eigenface* is one of the most successful approaches in face recognition area [5][7][9][10][11][12][13][14]. The method projects the input faces onto a dimensionally reduced space where the recognition is carried out. And in the Eigenface method, the common eigenspace is generally pre-computed by batch training from a "large" training set. Although the batch-training algorithm is in some sense the best case, it is computationally expensive [3]. So, some eigenspace-updating algorithms for image analysis are proposed in the vision literature [3][4][5][6].

One of the demerits of Eigenspace-based methods is that the most discriminating features of a specific face are not accurately represented. Thus, Face-Specific Subspace (FSS) is proposed to represent each face by one individual face subspace [8]. In FSS, the reconstruction error, that is, distance from the face-specific subspace, is exploited as the "similarity" measurement for recognition. Since individual subspace is modeled for each face, FSS affords a good basis for online updating each face model and further improve the performance of the recognition system, even if only single example for some face is available at any moment.

The rest of this paper is organized as follows: In Section 2, the FSS-based face recognition method is briefly described. In Section 3, the method we exploit to learn face subspace incrementally is introduced. Experimental results are shown in Section 4 and final conclusions are drawn in Section 5.

2. FACE-SPECIFIC SUBSPACE (FSS)

2.1.Eigenface method

An eigenspace model, Ω , constructed over N face images $\mathbf{y} \in \Re^m$ and a mean face, a set of eigenvectors, the eigenvalue associated with each eigenvector. The mean face is given by:

$$\Psi = \frac{1}{N}\sum_{i=1}^{N}\mathbf{y}^i \qquad (1)$$

Eigenspace models can be computed using eigenvalue decomposition (EVD) of the covariance matrix $\sum$ of a set of N face images,

$$\Sigma = \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{y}^i - \Psi\right)\ \left(\mathbf{y}^i - \Psi\right)^T \qquad (2)$$

The eigenvector (eigenface) are columns of the matrix U , and the spread of the face images by the corresponding eigenvalue in the diagonal matrix Λ . The eigenvectors and their eigenvalues are solutions to the eigenproblem:

$$\sum U = U\Lambda \qquad (3)$$

In eigenface[7][13], the projection W_{opt} is chosen to maximize the determinant of the total scatter matrix of the projected faces, i.e.,

$$W_{opt} = \arg\max\left|U^T \Sigma U\right| = \left[w_1\ w_2\ \cdots\ w_n\right] \qquad (4)$$

Where $\{w_i \mid i = 1,2,\cdots,n\}$ is the set of m -dimensional eigenfaces of Σ corresponding to the n largest eigenvalues.

2.2. Learning Face-Specific Subspace (FSS)

[8] proposes to represent each face by using one individual face subspace, named Face-Specific Subspace (FSS), which is learnt from the training images of the face. Formally, the FSS can be learnt as the following procedure:
Let the class set of the faces to be identified is:

$$C = \{\Omega_1, \Omega_2, \cdots, \Omega_p\} \qquad (5)$$

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:

$$U_k{}^T \Sigma_k U_k = \Lambda_k \qquad (6)$$

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 $U_k = \left[\mu_1^{(k)}, \mu_2^{(k)}, \cdots, \mu_{d_k}^{(k)}\right]$ is the matrix formed by the eigenvectors of Σ_k , where $\mu_1^{(k)}, \mu_2^{(k)}, \cdots, \mu_{d_k}^{(k)}$ 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:

$$U_k = \left(\mu_1^{(k)}, \mu_2^{(k)}, \cdots, \mu_{d_k}^{(k)}\right) \qquad (7)$$

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

$$\Re_k = \left(U_k, \Psi_k, \Lambda_k, d_k\right) \qquad (8)$$

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

2.2. 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:

$$W^{(k)} = U_k{}^T \Phi^{(k)}, \text{ where } \Phi^{(k)} = \Gamma - \Psi_k \qquad (9)$$

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

$$\Phi_r^{(k)} = U_k W^{(k)} \qquad (10)$$

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

$$\varepsilon^{(k)} = \left\|\Phi^{(k)} - \Phi_r^{(k)}\right\| \qquad (11)$$

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:

$$\Gamma \in \Omega_m \text{ if } \varepsilon^{(m)} = \min_{1 \le k \le p}\{\varepsilon^{(k)}\} \qquad (12)$$

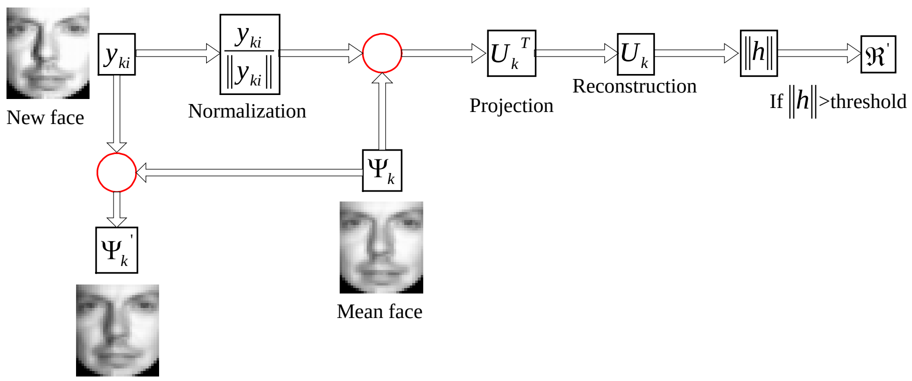

Fig.2. Framework of the Incremental FSS

3.INCREMENTAL FACE-SPECIFIC SUBSPACE (IFSS)

In FSS model $\Re_k = (U_k, \Psi_k, \Lambda_k, d_k)$, we modify it to $\Re_k = (U_k, \Psi_k, \Lambda_k, d_k, N_k)$, and N_k is the number of observations in the k^{th} face class. When $N_k -> N_k + 1$, estimate the FSS model

$$\Re_k{}' = (U_k{}', \Psi_k{}', \Lambda_k{}', d_k{}', N_k + 1) \tag{13}$$

directly, from the $\Re_k$ model we can obtain the update model $\Re_k{}'$ and needn't to obtain the new covariance matrix $\Sigma_k{}'$ of the k^{th} face. Fig.2 shows the procedure of incremental FSS. ($\mathbf{y}_{ki} \in R^{N_k \times m}$, with N_k the number of observation in the k^{th} face class, and m is the image vector dimension.) When a new face image $\mathbf{y}_{ki}$ is added, it is projected into the FSS model to obtain a p ($p << m$) dimensional vector $\mathbf{q}$, using the U_k as a basis. This $\mathbf{q}$ can be reconstructed into the m dimensional space, but with the loss represented by the residue vector $\mathbf{h}$.

$$\mathbf{q} = U_k{}^T(\mathbf{y}_{ki} - \Psi_k) \tag{14}$$

$$\mathbf{h} = (\mathbf{y}_{ki} - \Psi_k) - U_k \mathbf{q} \tag{15}$$

if the new face lies exactly within the k^{th} face class, the residue $\|\mathbf{h}\|$ is zero, then only $\Psi_k -> \Psi_k{}'$ in the FSS model. We could match the residue with a selected residue, if the residue larger than the selected residue we update FSS.

$$\Sigma_k{}' U_k{}' = U_k{}' \Lambda_k{}' \tag{16}$$

$$\Sigma_k{}' = \frac{N}{N+1}\Sigma_k + \frac{N}{(N+1)^2}\mathbf{y}_{ki}{}'(\mathbf{y}_{ki}{}')^T \tag{17}$$

And the new eigenvectors $U_k{}'$ must be a rotation of the current eigenvectors plus some new orthogonal unit vector, the rotation matrix represents with R .

Setting

$$U_k{}' = \left[U_k, \frac{\mathbf{h}}{\|\mathbf{h}\|}\right] R \tag{18}$$

Substitution of Equations 17 and 18 into 16, and following transformation we can obtain the result,

$$\left(\frac{N}{N+1}\begin{bmatrix} \Lambda_k & 0 \\ 0 & 0 \end{bmatrix} + \frac{N}{(N+1)^2}\begin{pmatrix} \mathbf{q}\mathbf{q}^T & \mathbf{h}^T\mathbf{y}_{ki}{}'\mathbf{q} \\ \mathbf{h}^T\mathbf{y}_{ki}{}'\mathbf{q}^T & (\mathbf{h}^T\mathbf{y}_{ki}{}')^2 \end{pmatrix} \right) R = R\Lambda_k{}' \tag{19}$$

We will show the experimental result in section 4.

4. EXPERIMENTS

To verify the effectiveness of the proposed approach, we also develop Eigenface method and FSS as benchmarks. Experiments are conducted on Harvard Face Database.

4.1. Benchmark designs

The Eigenface method is the standard benchmark in the face recognition community. Its performance can reflect the difficulty of the given face recognition task to a certain extent.

We design the Eigenface method according to [7]. The training set used to learn common face subspace is the same as the gallery set constituted by all the training images of all the faces.

4.2. Experiments On Harvard Face Database

The Harvard face image database is used in our experiment. In the database, five subsets (subset1 through subset5) contain the cropped, masked, single light source, frontal-view images of 10 individuals from the Harvard database. For each face, there are 6, 9, 13, 17, 21 face images in the five subsets for each person. These sets contain images from each individual in Fig3 but of increasing extremity of illumination with the number of the set.

Fig.3 Face Images of the 10 individuals in Harvard

Experiment 1:

To verify the online learning ability of the IFSS, comparisons are made between the Eigenface method, the FSS and the IFSS. In this experiment, for Eigenface and FSS, the training set is just the subset 1. FSS is trained for each face by the 6 images in his/her subset1. However, IFSS is trained for each face by the 6 images in his/her subset1 as FSS, and then added the nine images in his/her subset2 one by one by using the proposed IFSS method according to the section 3. The images in subset3 through subset5 are tested. The experiment results are shown in Table-1.

Table 1 Experimental results on Harvard Face Database

Method tested	**Train Set**	**Subset for Testing (%)**		
		Subset3	**Subset4**	**Subset5**
Eigenface	Subset1	53.8	30.0	16.9
FSS	Subset1	82.3	64.1	35.3
IFSS	Subset 1+2	90.0	71.8	53.2

Note: In experimental of IFSS, the residue error is 0.0006

From Table.1, It is obvious that IFSS has achieved a better performance than FSS and Eigenface after incrementally learn 9 additive examples for each face. This observation is consistent with the intuition. For this illumination face database case, more examples under different lighting conditions have significantly improved the performance of the recognition system against more complex lighting conditions.

Experiment 2:

We also conduct experiments to compare the online-learning IFSS method with the batch-learning Eigenface and FSS. FSS and Eigenface are both batch-trained from the subset 1 and subset 2 and tested on the subset 3, 4 and 5. IFSS is first trained from subset 1 and add the 9 examples in subset 2 one by one for each face. The experimental results are shown in Table.2. From Table.2, IFSS has comparable performance with the batch FSS and much better performance than the batch Eigenface. The advantages of the IFSS against the batch FSS include (1) IFSS does not need to store the previous learned examples, therefore it need less storage; (2) IFSS can update each face model individually even if only one additive example is available.

Table 2 Experimental results on Harvard Face Database

Method tested	**Train Set**	**Subset for Testing (%)**		
		Subset 3	**Subset 4**	**Subset 5**
Eigenface	Subset 1+2	89.2	66.5	38.3
FSS	Subset 1+2	92.3	74.1	53.7
IFSS	Subset 1+2	90.0	71.8	53.2

5. CONCLUSIONS

Based on the Face-Specific Subspace (FSS) face recognition method, in this paper, IFSS is proposed to make FSS have the ability to learn online. Since in the FSS face recognition method, each individual face is represented as one specific face subspace, therefore, each face model can be updated even if only single example face image is available along with its corresponding class label. Our experiments on the Harvard illumination face database show that better recognition performance can be achieved when more example images are incrementally fed into the IFSS system. Experiments also show that IFSS has comparable performance with the batch FSS, but less computational resource is needed and the training examples need not be stored. Another advantage of IFSS is that it can update each face model individually even if only one additive example is available.

6. ACKNOWLEDGEMENTS

This research is partially sponsored by Natural Science Foundation of China under contract No.60332010, National Hi-Tech Program of China (No. 2001AA114190 and 2002AA118010), and ISVISION Technologies Co., Ltd.

7. REFERENCES

[1] R.Chellappa, C.L.Wilson ect. "Human and Machine Recognition of faces: A survey", *Proc. of the IEEE*, 83(5), pp705-740, 1995.5

[2] W. Zhao, R. Chellappa, A. Rosenfeld, and J. Phillips, "Face Recognition: A Literature Survey," *Technical Report*, CS-TR4167, Univ. of Maryland, 2000

[3] S.Chandrasekaran, B.S.Manjunath, Y.F.Wang, J.Winkler, H.Zhang. An eigenspace update algorithm for image analysis. *Graphical Models and Image Processing*, 59(5):321-332, 1997.

[4] S.Chaudhuri, S.Sharma, S.Chatterjee. Recursive estimation of motion parameters. *Computer Vision and Image Understanding*, 64(3): pp434-442, 1996

[5] H.Murakami, B.V.K.V Kumar. Efficient calculation of primary images from a set of images. *IEEE Trans Pattern Analysis and Machine Intelligence*, 4(5): pp511-515, 1982

[6] P.M.Hall, D.Marshall, R.Martin, Incremental Eigenanalysis for Classification, *British Machine Vision Conf.*, pp.286-295, 1998

[7] M.Turk, and A.Pentland, "Eigenfaces for Recognition," *J. Cognitive Neuroscience*, vol.3, no. 1, pp. 71-86, 1991.

[8] Shiguang Shan, Wen Gao, Debin Zhao, Face Identification Based On Face-Specific Subspace, *International Journal of Image and System Technology, Special issue on face processing, analysis and synthesis*, 13(1), pp23-32, (2003), Publisher: John Wiley & Sons Inc., 605 Third Avenue, New York, NY 10158, USA

[9] H.Murase and S.K.Nayar, "Visual learning of object modules from appearance," *Proc. Image Understanding Workshop* 1993, (San Diego, CA), pp.547-555, 1993

[10] H.Murase and M.Lindenbaum, "Partial eigenvalue decomposition of large images using the spatial temporal adaptive method," *IEEE Trans-IP*, Vol.4 (5), pp.620-629, 1995

[11] E.Oja, Subspace methods of pattern recognition, John Wiley, 1983

[12] Pentland, B.Moghaddam, and T.Starner, "View-based and modular eigenspaces for face recognition," Proc. *IEEE Conf. On Computer Vision and Pattern Recognition*, 1994, (Seattle, Washington), pp.84-91, June 1994

[13] M.Turk and A.Pentland, "Face Recognition Using Eigenfaces," Proc. *IEEE Conf. On Computer Vision and Pattern Recognition*, 1991,pp.586-591

[14] Y. Zhang, J. Weng, "Convergence analysis of complementary candid incremental principle component analysis," *Technical Report*, MSU-CSE-01-23, Computer Science and Engineering Department of Michigan State University, August, 2001