

FUSION OF SVD AND LDA FOR FACE RECOGNITION

Yanwei Pang¹, Nenghai Yu¹, Rong Zhang¹, Jiawei Rong², Zhengkai Liu¹

¹Microsoft-USTC Intelligent Computing Research Center

University of Science and Technology of China, Hefei 230027, China

²Computer Science and Engineering Department, Fudan University, Shanghai, China

pyw@ustc.edu; {ynh, zrong, zhengkai}@ustc.edu.cn

ABSTRACT

A face recognition method based on the fusion of linear discriminant analysis (LDA) and singular value decomposition (SVD) is presented. In theory, fusion of different data or classifiers can achieve better performance when they are independent of each other or they can overcome shortcomings of each other. As one of the subspace methods, LDA-based method has a drawback that LDA is sensitive (variant) to translation, rotation and other geometric transforms. SVD-based method, as an algebraic feature extraction approach, has the merit of invariance to translation, rotation and mirror transforms. By combining these two methods, it is expected that better recognition performance can be obtained. Experiment results on ORL face database show the effectiveness of the proposed method.

1. INTRODUCTION

Face recognition has emerged as an active research area in the field of computer vision and pattern recognition due to its scientific challenges and its potential applications. It is a non-contact personal identification technique which has advantages over contact identifications such as fingerprints and iris texture.

There are two main approaches to face recognition [1]. The first approach is the feature-based matching approach using the relationship between facial features, such as eyes, mouth, and nose. The second approach is the template matching approach using the holistic features of a face image. Linear discriminant analysis or fisherface method is one of the most successful template matching approaches [2]. However, there are still many open problems, such as recognition under pose, expression, illumination and registration variations.

An area that has received significant attention in recent years is classifier combination and data fusion. For example, face recognition rate can be improved by the

fusion of visible images and thermal infrared imagery [3]. In [4], a hybrid recognition system integrating three classifiers, namely the profile approach, an HMM algorithm and the eigenface approach was developed. More recently, LDA and PCA were fused to get a better performance [5]. Further, face classifiers based on PCA, ICA and LDA were combined using the sum rule and RBF-based integration strategies [6].

However, PCA, ICA and LDA-based recognition methods are sensitive to translation, rotation and other geometric transforms. Thus no combined classifiers of PCA, ICA and LDA-based recognition methods can overcome this problem. This paper proposes face recognition method which is based on fusion of LDA and SVD. SVD-based method which uses singular values as the feature vectors, has the merit of invariance of the above geometric transforms [7]. However, recent research has shown that singular values do not contain adequate information for face recognition [8]. LDA has a powerful discriminate ability and is an efficient dimension reduction technique. Once they are combined, it is expected that better recognition performance can be achieved.

2. LDA AND SVD-BASED METHODS

2.1. LDA-based method

LDA is a well-known technique for dealing with the class separability problem. LDA can be used to determine the set of the most discriminant projection axes. After projecting all the samples onto these axes, the projected samples will form the maximum between-class scatter and the minimum within-class scatter in the projective feature space [9].

Given a set of N training face images $\{\mathbf{x}_i\}_{i=1}^N$, each of which is represented as a vector of length n ($=I_w \times I_h$), i.e., $\mathbf{x}_i \in \mathbf{R}^n$ belonging to one of C classes $\{Z_i\}_{i=1}^C$, where $(I_w \times I_h)$ is the image size and \mathbf{R}^n denotes a n -dimensional real space. There is a linear transformation matrix \mathbf{W} which

maps the original \mathbf{R}^n image space to feature space \mathbf{R}^m , where $m < n$. The new feature vectors $\mathbf{y}_i^{lda} \in \mathbf{R}^m$ are defined by the following transformation:

$$\mathbf{y}_i^{lda} = \mathbf{W}^T \mathbf{x}_i \quad i=1, \dots, N \quad (1)$$

LDA selects the optimal transformation \mathbf{W} in such a way that the ratio of the between-class scatter \mathbf{S}_b and the within-class scatter \mathbf{S}_w is maximized:

$$\mathbf{W}_{opt} = \arg \max \frac{|\mathbf{W}^T \mathbf{S}_b \mathbf{W}|}{|\mathbf{W}^T \mathbf{S}_w \mathbf{W}|} = [\mathbf{w}_1 \mathbf{w}_2 \dots \mathbf{w}_m]$$

where $\{\mathbf{w}_i \mid i=1, 2, \dots, m\}$ is the set of generalized eigenvectors of \mathbf{S}_b and \mathbf{S}_w corresponding to the m largest generalized eigenvalues $\{\lambda_i \mid i=1, 2, \dots, m\}$, i.e.,

$$\mathbf{S}_b \mathbf{w}_i = \lambda_i \mathbf{S}_w \mathbf{w}_i, \quad i=1, 2, \dots, m \quad (2)$$

\mathbf{S}_b and \mathbf{S}_w are defined respectively as

$$\mathbf{S}_b = \sum_{i=1}^C N_i (\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^T$$

$$\mathbf{S}_w = \sum_{k=1}^C \sum_{\mathbf{x}_i \in Z_k} (\mathbf{x}_i - \mathbf{u}_k)(\mathbf{x}_i - \mathbf{u}_k)^T$$

where \mathbf{u}_i is the mean image of class Z_i , and N_i is the number of samples in class Z_i .

However it is difficult to find the solution of equation (2) when the number of training images is less than the dimension of the images. A common way is applying PCA [10] before LDA is performed [2]. The principal components number t is selected such that $m+1 \leq C \leq t \leq N-C$. See [11] for details. For the experiments of this paper, t equals 160.

2.2. SVD-based method

SVD is an efficient algebraic feature extraction method which regards an image as a matrix while LDA treats an image as a high dimensional vector. Conventional SVD-based face recognition method uses the singular values as the feature vectors.

Theorem [7]. If $\mathbf{A} \in \mathbf{R}^{m \times n}$, then there exist orthogonal matrices

$$\mathbf{U} = [\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_m] \in \mathbf{R}^{m \times m} \quad \text{and} \quad \mathbf{V} = [\mathbf{v}_1 \mathbf{v}_2 \dots \mathbf{v}_n]^T \in \mathbf{R}^{n \times n}$$

such that

$$\mathbf{A} = \mathbf{U} \boldsymbol{\Sigma}_A \mathbf{V}^T \quad (3)$$

where $\boldsymbol{\Sigma}_A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$, $\lambda_1 \geq \dots \geq \lambda_p \geq 0$, $p = \min(m, n)$.

Suppose \mathbf{A}_i is matrix of a face image. After \mathbf{A}_i is decomposed by equation (3), the new feature vector $\mathbf{y}_i^{svd} \in \mathbf{R}^p$ is formed by arranging singular values:

$$\mathbf{y}_i^{svd} = [\lambda_1, \lambda_2, \dots, \lambda_p]^T \quad (4)$$

Literature [7] has proven that singular values are invariant to translation, rotation, mirror transforms. Further it is

stable when other variations occur such as noise or changing lighting.

3. FUSION OF SVD AND LDA

3.1. The motivation of the fusion

It is claimed that fusion methods are less important than the diversity of the team [12]. However, fusion (combination) is simple and economical to improve recognition performance. The goal of fusion is to get a higher accuracy than each individual classifier or feature vectors. The premise of fusion is that different classifiers or features can overcome the drawbacks of each other.

Feature vectors extracted by LDA have a powerful separability and LDA is a very efficient dimension reduction technique. However, LDA take each image as a point in high dimensional space where even a little amount of geometric transform such as translation, rotation degrades the recognition performance severely. By contrast, singular values are invariant to all the above geometric transforms [7]. Moreover, singular values are invariant to mirror transform [7]. This property is useful for face recognition since mirror images can provide more information. Note that a face is mirror symmetrical. In spite of those useful advantages, recent research has shown that singular values do not contain adequate information for face recognition [8]. Though above analysis, it is natural to consider to combine SVD-based method and LDA method. More specifically, we want to fuse the feature vectors extracted by SVD and LDA methods respectively. For the sake of convenience, we denote the proposed method as *SVD+LDA*.

3.2. The fusion strategy

Given two types of feature vector \mathbf{y}_i^{svd} (equation (4)) and \mathbf{y}_i^{lda} (equation (1)), the task is how to form the final distance measure. Two simple strategies are averaging and multiplying. Literature [12] has analyzed in theory when the averaging scheme is adopted and when the multiplying scheme is adopted. Here we employ the averaging strategy which will be explained as follows. The process of testing is :

- 1) Input a query face image \mathbf{x}_i , extract the SVD-based feature vectors \mathbf{y}_i^{svd} (equation (4)) and LDA-based feature vectors \mathbf{y}_i^{lda} (equation (1)).
- 2) Compute the distances of the query face image to all the training images:

$$\mathbf{d}_i^{lda} = (d_{i1}^{lda}, d_{i2}^{lda}, \dots, d_{iN}^{lda}) \quad \text{and}$$

$$\mathbf{d}_i^{svd} = (d_{i1}^{svd}, d_{i2}^{svd}, \dots, d_{iN}^{svd})$$

where N is number of training images.

Table 1: Filenames of different training subsets

Subset 1	1.pgm, 3.pgm, 6.pgm, 7.pgm, 10.pgm
Subset 2	1.pgm, 4.pgm, 5.pgm, 9.pgm, 10.pgm
Subset 3	3.pgm, 4.pgm, 5.pgm, 6.pgm, 8.pgm
Subset 4	2.pgm, 4.pgm, 5.pgm, 6.pgm, 10.pgm
Subset 5	3.pgm, 4.pgm, 8.pgm, 9.pgm, 10.pgm

If the SVD-based and LDA-based feature vectors of training images \mathbf{x}_j are \mathbf{y}_j^{svd} and \mathbf{y}_j^{lda} respectively, then

$$d_{ij}^{svd} = \|\mathbf{y}_i^{svd} - \mathbf{y}_j^{svd}\|$$

$$d_{ij}^{lda} = \|\mathbf{y}_i^{lda} - \mathbf{y}_j^{lda}\|.$$

- 3) Calculate the final distances which contain both the information of LDA and SVD:

$$\mathbf{d}_i = (d_{i1}, d_{i2}, \dots, d_{iN})$$

$$= \left(\frac{d_{i1}^{svd} + d_{i1}^{lda}}{2}, \frac{d_{i2}^{svd} + d_{i2}^{lda}}{2}, \dots, \frac{d_{iN}^{svd} + d_{iN}^{lda}}{2} \right)$$

$$\text{i.e. } d_{ij} = \frac{d_{ij}^{svd} + d_{ij}^{lda}}{2}.$$

- 4) Based on the distances $(d_{i1}, d_{i2}, \dots, d_{iN})$, the nearest neighbor classifier is employed to accomplish the task of classification. Categorize the face image \mathbf{x}_i into class label k where k is determined by

$$d_{ik} = \min_{1 \leq j \leq N} \{d_{ij}\} \quad 1 \leq k \leq C.$$

4. EXPERIMENTAL RESULTS

In this section, we describe our experiments on a well-known face database: the ORL (AT&T) face database. For the sake of convenience, we denote the proposed method as *SVD+LDA*. The *SVD+LDA* method are compared respectively with conventional PCA (eigenface) method, LDA (fisherface) method and SVD method. It is noted that no additional preprocessing step was adopted in our experiments.

4.1. The testing and training sets

The ORL database is made up of 10 different images of 40 people. The filenames of the ten images is "1.pgm", ..., "10.pgm". The size of each image is 112×92 . They are taken at different times, varying lighting slightly, facial expressions (open/closed eyes, smiling/nonsmiling), and facial details (glasses/no glasses). The persons are in upright, frontal position with tolerance for some side movement. Figure 1 shows ten images of a person.

Table 2: Recognition rates of different methods (%)

method	Training subset					Mean accuracy (%)
	1	2	3	4	5	
SVD	77.5	73.5	75	71	77.5	74.9
PCA	93	96	94.5	94.5	91.5	93.9
LDA	95	96.5	97	96	96	96.1
SVD+LDA	98	98.5	96.5	97.5	96.5	97.4

**Figure 1:** Example images of ORL database

In our experiments the cross-validation sampling technique was employed to determine face recognition rate. We randomly chose five images as training images and the remaining images as testing images. The test is repeated five times with the same numbers of testing images and training in every run. The final recognition rate (rank 1) is determined by averaging the results of every run. Since the test is repeated five times, there are five training sets and five corresponding testing sets. The five training sets are list in table 1 which shows which images are used in different training sets so that any reader can repeat our experiments and obtain the same recognition rates. For PCA method 50 principal components are used. For LDA method and SVD+LDA method the number of discriminate vectors is 39.

4.2. Results

Table 2 shows the recognition rates of different methods. The mean recognition rates of SVD-based, PCA-based, LDA-based and SVD+LDA-based methods are 74.9%, 93.9%, 96.1% and 97.4% respectively. The recognition rate of SVD+LDA is 1.3% higher than LDA, 3.5% than PCA, and 22.5% higher than SVD method. Moreover, for every training set and the corresponding testing set, the accuracy of SVD+LDA is never worse than individual SVD and individual LDA. So the goal of fusion is reached. It can be drawn from Table 2 that the proposed method can achieve better recognition rates.

5. CONCLUSION

We proposed a face recognition method which is based on the fusion SVD and LDA. LDA has a powerful discriminate ability and is an efficient dimension reduction technique. The performance of LDA degrades severely when translation, rotation mirror transforms occur. This drawback can be compensated partially by incorporating the information of singular values which is obtained by SVD. In other words, LDA provides rich information to SVD since singular values themselves do not contain adequate information for face recognition.

ACKNOWLEDGEMENT

This work has been fully supported by Nature Science Fund of Anhui Province, P.R. China, under grant No. 03042307.

REFERENCES

- [1] R. Chellappa, C.L. Wilson, S. Sirohey, "Human and machinerecognition of faces: a survey," Proc. IEEE, Vol. 83, No.5, pp. 705-740, 1995.
- [2] P. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," IEEE Trans. on PAMI, Vol. 19, No. 7, pp. 711-720, 1997.
- [3] Diego A. Socolinsky, Andrea Selinger, and Joshua D. Neuheisel, "Face recognition with visible and thermal infrared imagery", Computer Vision and Image Understanding, Vol. 91, pp. 72-114, 2003.
- [4] B. Achermann, H. Bunke, "Combination of face classifiers for person identification," Proceedings of the 13th IAPR international conference on Pattern recognition, vol. 3, pp. 416-420, 1996.
- [5] G.L. Marcialis and F. Roli, "Fusion of LDA and PCA for face verification," Proceedings of the Workshop on Biometric Authentication, M. Tistarel and J. Bigun Eds., Springer LNCS 2359, 2002.
- [6] X. Lu, Y. Wang, A.K. Jain, "Combining classifier for face recognition," International Conference on Multimedia and Expo, vol. 3, pp.16-19, 2003.
- [7] Z. Hong, Algebraic feature extraction of image for recognition, Pattern recognition, vol. 24, pp. 211-219, 1991.
- [8] Y. Tian, T. Tan, Y. Wang, Y. Fang, "Do singular values contain adequate information for face recognition?," Pattern recognition, vol. 36, pp. 649-655, 2003.
- [9] L.F. Chen, H.Y.M. Liao, M.T. Ko, J.C. Lin, G.J. Yu, "A new LDA-based face recognition system which can solve the small sample size problem," Pattern recognition, vol. 33, pp. 1713-1726, 2000.
- [10] M. Turk, A. Pentland, "Face recognition using eigenfaces," In: Proc. IEEE Conf. On Computer Vision and Pattern Recognition, pp. 586-591, 1991.
- [11] DL Swets and JJ Weng, "Using discriminat eigenfeatures for image retrieval," IEEE Trans. On PAMI, vol. 18, no.8, pp. 831-836, 1996.
- [12] David M.J. Tax, Martijn van Breukelen, Robert P.W. Duin, Josef Kittler, Combining multiple classifiers by averaging or by multiplying[J], Pattern recognition, vol. 33, pp. 1475-1485, 2000.