Making Discriminative Common Vectors Applicable to Face Recognition with One Training Image per Person

Lei Zhu
Institute of Biomedical Engineering & Instrument
Hangzhou Dianzi University
Hangzhou 310018, China
zhulei@hdu.edu.cn

Yongying Jiang
Institute of mechanical Engineering
Wenzhou University
Wenzhou 325035, China
yyjiang2003@126.com

Lihua Li
Institute of Biomedical Engineering & Instrument
Hangzhou Dianzi University
Hangzhou 310018, China
lilh@hdu.edu.cn

***Abstract*—Though discriminant common vector (DCV) method has obtained some success in face recognition task, it fails when only one training image per person is available. In this paper, we propose an approach to make DCV method applicable when each person has one training image. Our approach is based on the assumption that human faces share similar intrapersonal variation. The intrapersonal variation of the training set can be estimated from the collected generic face set. The proposed method was compared with PCA, E(PC)2A and SVD perturbation algorithm, and experimental results on the subset of FERET face database show the promising performance of the proposed method.**

***Keywords*—Face recognition, Discriminative common vectors, One training image per person**

I. INTRODUCTION

In the past few decades, face recognition has been one of the hottest research areas in computer vision and pattern recognition[1]. Subspace methods[2] are probably the most popular and widely applied techniques in face recognition. Two of the most popular techniques of subspace methods are Eigenface[3] and Fisherface[4]. Eigenface, which is based on principal components analysis (PCA), produces an expressive subspace for face representation. Fisherface uses linear discriminant analysis (LDA) and extracts discrimination information by maximizing the between-class scatter matrix, while minimizing the within-class scatter matrix in the projective subspace. LDA is widely believed to have better performance than PCA when enough training samples are available[5]. But when the dimension of the sample space is larger than the number of samples in the training set, the LDA can not be implemented. This is called a "small sample size" (SSS) problem. Recently a pattern recognition method called discriminant common vector (DCV)[6, 7] was proposed for the small sample size case. DCV yields an optimal solution based on a modified criterion and obtains a promising performance in pattern recognition field.

Unfortunately, in some specific case such as law enforcement, there may be only one image per person available to the system. Under this condition, most traditional subspace methods such as Fisherface and DCV failed, because in order to get within-class variation, they require at least two samples per person. Recently, a lot of methods have been proposed to solve this problem[8]. In[9] and [10], face recognition methods are proposed to solve the problem of single training image using $E(PC)^2A$ and SVD perturbation, respectively.

In this paper, we propose a simple but effective approach to making DCV method applicable in the case where each person has only one training sample.

II. METHODOLOGY

A. Review of DCV

The idea of DCV was originally introduced for isolated word recognition problem. Recently, some face recognition methods based on DCV have been proposed. These approaches eliminate the unwanted information of the face images, such as environmental effects, personal differences, and extract the common properties of classes in the training set.

Let the training set $\boldsymbol{X}$ be composed of a set of vectors $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N\}$, where $\boldsymbol{x}_i \subset R^D$. Each original face image data $\boldsymbol{x}_i$ belongs to one of C object classes $\{X_1, X_2, \cdots, X_C\}$.

The objective function of DCV can be written as:

$$J(\boldsymbol{A}_{opt}) = \underset{|\boldsymbol{A}^T\boldsymbol{S}_W\boldsymbol{A}|=0}{\arg\max} |\boldsymbol{A}^T\boldsymbol{S}_B\boldsymbol{A}| = \underset{|\boldsymbol{A}^T\boldsymbol{S}_W\boldsymbol{A}|=0}{\arg\max} |\boldsymbol{A}^T\boldsymbol{S}_T\boldsymbol{A}| \quad (1)$$

where $\boldsymbol{S}_W$ denotes the within class scatter matrix, and $\boldsymbol{S}_T$ denotes the total scatter matrix as follows.

$$\boldsymbol{S}_W = \sum_{i=1}^{C}\sum_{j=1}^{N_i}(x_j^i - \boldsymbol{\mu}_i)(\boldsymbol{x}_j^i - \boldsymbol{\mu}_i)^T \quad (2)$$

$$\boldsymbol{S}_T = \sum_{i=1}^{C}\sum_{j=1}^{N_i}(\boldsymbol{x}_j^i - \boldsymbol{\mu})(\boldsymbol{x}_j^i - \boldsymbol{\mu})^T \quad (3)$$

where $\boldsymbol{\mu}$ denotes the mean of all samples in the training set, and $\boldsymbol{\mu}_i$ denotes the mean of samples in the i th class.

Consider the matrices $\boldsymbol{U} = [\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_d]$ and $\overline{\boldsymbol{U}} = [\boldsymbol{\alpha}_{d+1}, \cdots, \boldsymbol{\alpha}_D]$. $\{\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_d\}$ and $\{\boldsymbol{\alpha}_{d+1}, \cdots, \boldsymbol{\alpha}_D\}$ denote the sets of orthonormal eigenvectors of nonzero eigenvalues and zero eigenvalues of $\boldsymbol{S}_W$, respectively. So face image $\boldsymbol{x}_i$ has a unique decomposition:

$$\boldsymbol{x}_j^i = \boldsymbol{y}_j^i + \boldsymbol{z}_j^i \quad (4)$$

where $\boldsymbol{y}_j^i = \boldsymbol{U}\boldsymbol{U}^T\boldsymbol{x}_j^i$, $\boldsymbol{z}_j^i = \overline{\boldsymbol{U}}\overline{\boldsymbol{U}}^T\boldsymbol{x}_j^i$.

It has been proved that with the above decomposition, the same unique vector for all samples of the same class can be obtained as:

$$\boldsymbol{x}_{com}^i = \boldsymbol{x}_j^i - \boldsymbol{U}\boldsymbol{U}^T\boldsymbol{x}_j^i = \boldsymbol{x}_j^i - \overline{\boldsymbol{U}}\overline{\boldsymbol{U}}^T\boldsymbol{x}_j^i, i = 1, \cdots, C, \; j = 1, \cdots, N_i \quad (5)$$

where $\boldsymbol{x}_{com}^i$ denotes the common vector of i th class samples.

B. *Proposed method*

As mentioned above, DCV needs to calculate the within class scatter matrix $\boldsymbol{S}_W$. When we have only one training sample per person, it is impossible to calculate the intrapersonal variation. Wang et al[11] pointed out that human faces share similar intrapersonal variation. Under this assumption, the within class scatter matrix $\boldsymbol{S}_W$ of training samples can be estimated from faces of others.

The summary procedure of our proposed DCV face recognition method is shown as follows:

Step1 Estimate the intra person variation and compute the vectors that span the range space of $\boldsymbol{S}_W{}'$.

We collect a set of generic face images with multiple samples per person and use them to computer the intra person variation, denoted as $\boldsymbol{S}_W{}'$. Then we compute the eigenvectors $\boldsymbol{U} = [\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_d]$ corresponding to the nonzero eigenvalues of $S_W{}'$, where d is the rank of $\boldsymbol{S}_W{}'$.

Step 2 Compute the common vector for each subject.

Project any sample from each class onto the null space of $\boldsymbol{S}_W{}'$ and get the common vectors in equation (5).

Step 3 Compute the projection matrix $\boldsymbol{A}$.

Consider the matrix

$$\boldsymbol{S}_{com} = \sum_{i=1}^{C}(\boldsymbol{x}_{com}^i - \boldsymbol{\mu}_{com})(\boldsymbol{x}_{com}^i - \boldsymbol{\mu}_{com})^T, \; i = 1, ..., C \quad (6)$$

where $\boldsymbol{\mu}_{com}$ denotes the mean of all common vectors.

The projection matrix $\boldsymbol{A} = [\boldsymbol{a}_1, ..., \boldsymbol{a}_{C-1}]$ is composed of C-1 eigenvectors associated with the nonzero eigenvalues.

Step 4 Obtain the feature vector and recognition

The feature vector of each person can be obtained by

$$\boldsymbol{V}_i = \boldsymbol{A}^T\boldsymbol{x}_i, \; i = 1, ..., C \quad (7)$$

To recognition a test sample, the feature vector is found by

$$\boldsymbol{V}_{Test} = \boldsymbol{A}^T\boldsymbol{x}_{Test} \quad (8)$$

Recognition is performed by nearest neighbor classification.

III. EXPERIMENTS AND RESULTS

The proposed DCV face recognition method with one training sample per person was tested using the subset of the FERET face database. The database contains 200 gray-level frontal view face images from 100 persons. Each person has two images (ba and bj) which cover a wide range of variations in light conditions, facial expressions and so on. We focused on the first 100 persons. Their ba images were used for training, while bj images were used for testing. The generic set were made up of the rest 100 persons' images (ba and bj). All face images were cropped into 64*64 from original frames based on the location of the two eyes. Figure 1 shows some sample images.

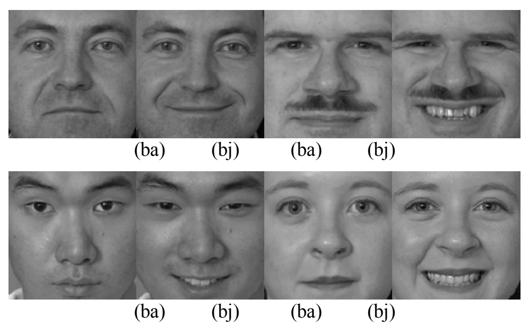

Figure 1. Some images in the experimental database

TABLE I. RECOGNITION PERFORMANCE RESULTS ON THE FERET FACE DATABASE

Methods	Feature dimensions	Recognition rate (%)
PCA	69	73.0
$E(PC)^2A$	57	72.0
SVD perturbation	56	68.0
Proposed method	99	84.0

We compared the proposed method with PCA, $E(PC)^2A$. and SVD perturbation algorithm. For PCA, the eigenvectors accounting for 98% of the total energy (original information) was calculated. $E(PC)^2A$ and SVD perturbation algorithm applies PCA to extract features on an enlarged or modified training set for subsequent recognition process. Nearest neighborhood (NN) algorithm in L_2 norm sense was used to construct the classifiers. The results of the performance of the four algorithms mentioned above are shown in Table 1.

From Table I, it can be seen that the accuracy is 84.0% for our method, 73.0% for PCA, 68.0% for $E(PC)^2A$ and 68.0% for SVD perturbation method. PCA method outperforms $E(PC)^2A$ method and SVD perturbation method in our experiments. The proposed method achieves an improvement of 11% in accuracy compared to the PCA algorithm. The results show that the proposed method is more accurate than the other three methods.

To study the effect of the generic sets, we chose images from two different databases to estimate the intra person variations.

1) fa, fb images in FERET database

Like ba and bj images in FERET database, fa and fb images are front facial images which cover a wide range of variations in illumination, facial expressions and so on.

2) Yale database

Yale face database contains 165 grayscale face images from 15 persons. We chose 2 images per person to form generic set. We manually cropped the face portion of the images. All cropped images were aligned at the centers of eyes and normalized with resolution 64*64.The training set and test set are same as the first experiment. The training set, test set and generic set do not pairwise overlapped. The experimental results are listed in Figure 2.

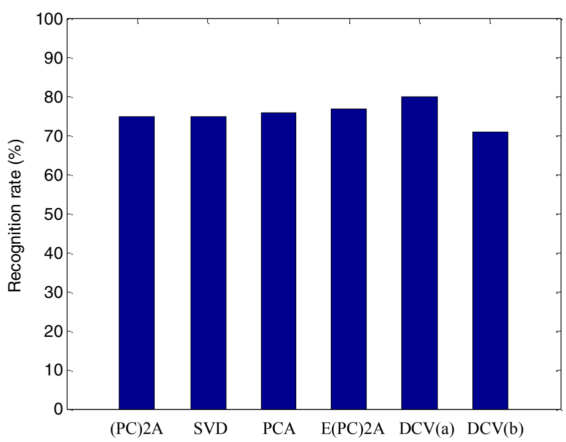

Figure 2. Experiments results with different generic set

In figure 2, DCV(a) denotes the generic set is collected with fa, fb images in FERET database, while DCV(b) means the generic set is formed with images from Yale database. We can see the result of DCV(a) is the highest (80%), and DCV(b) is only 71%. The reason may be FERET and Yale are different databases. The intra person variations in one database may not be well estimated with images from another database.

It should also be noted that, we would expect the recognition rate of these face recognition methods mentioned above to improve if a more sophisticated classifier was used.

IV. CONCLUSION

This paper proposed a method making DCV applicable when only one training sample per person is available. The experimentations on FERET face database indicate that the proposed method is not only feasible but also better than PCA, $E(PC)^2A$ and SVD perturbation algorithm in case of one training sample scenario. The proposed face recognition algorithm is based on DCV which is a linear feature extraction method. Our future work is to extend the proposed method to nonlinear form by kernel trick[12].

ACKNOWLEDGMENT

This paper is partially supporter by National Natural Science Foundation of China (60775016, 60705016), National Distinguished Young Research Scientist Award from the National Science Foundation of China (60788101), and a grant from the Science and Technology Program of Zhejiang Province of China (2006C14026). And the authors would like to thank the anonymous reviews for their constructive advice.

REFERENCES

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *Acm Computing Surveys,* vol. 35, pp. 399-459, Dec 2003.

[2] G. Shakhnarovich and B. Moghaddam, *Face Recognition in Subspaces*. New York: Springer-Verlag, 2004.

[3] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Cognitive Neuroscience,* vol. 3, pp. 71-86, 1991.

[4] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection " *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 19, pp. 711-720, 1997.

[5] L. Juwei, K. N. Plataniotis, and A. N. Venetsanopoulos, "Face recognition using LDA-based algorithms," *Neural Networks, IEEE Transactions on,* vol. 14, pp. 195-200, 2003.

[6] Y. He, L. Zhao, and C. Zou, "Face recognition using common faces method," *Pattern Recognition,* vol. 39, pp. 2218-2222, 2006.

[7] H. Cevikalp, M. Neamtu, M. Wilkes, and A. Barkana, "Discriminative common vectors for face recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 27, pp. 4-13, 2005.

[8] X. Tan, S. Chen, Z.-H. Zhou, and F. Zhang, "Face recognition from a single image per person: A survey," *Pattern Recognition,* vol. 39, pp. 1725-1745, 2006.

[9] S. Chen, D. Zhang, and Z.-H. Zhou, "Enhanced (PC)2A for face recognition with one training image per person," *Pattern Recognition Letters,* vol. 25, pp. 1173-1181, 2004.

[10] D. Zhang, S. Chen, and Z.-H. Zhou, "A new face recognition method based on SVD perturbation for single example image per person," *Applied Mathematics and Computation,* vol. 163, pp. 895-907, 2005.

[11] W. Xiaogang and T. Xiaoou, "Unified subspace analysis for face recognition," 2003, pp. 679-686 vol.1.

[12] V. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer, 1995.