PERGAMON

Pattern Recognition 35 (2002) 1997–1999

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Rapid and Brief Communication

From image vector to matrix: a straightforward image projection technique—IMPCA vs. PCA

Jian Yang*, Jing-yu Yang

Department of Computer Science, Nanjing University of Science and Technology, Nanjing 210094, People's Republic of China

Received 29 November 2001; accepted 27 December 2001

Abstract

The conventional principal component analysis (PCA) and Fisher linear discriminant analysis (FLD) are both based on vectors. Rather, in this paper, a novel PCA technique directly based on original image matrices is developed for image feature extraction. Experimental results on ORL face database show that the proposed IMPCA are more powerful and efficient than conventional PCA and FLD. © 2002 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Image principal component analysis (IMPCA); Principal component analysis (PCA); Linear discriminant analysis (FLD); Image feature extraction

1. Introduction

The conventional principal component analysis (PCA) and Fisher linear discriminant analysis (FLD) are both based on vectors. That is to say, if we use them to deal with the image recognition problem, the first step is to transform original image matrices into same dimensional vectors, and then rely on these vectors to evaluate the covariance matrix and to determine the projector. Two typical examples, the famous Eigenfaces [1] and Fisherfaces [2] both follow this strategy. The drawback of this strategy is obvious. For instance, considering an image of 100×100 resolution, its corresponding vector is 10 000-dimensional. To perform PCA or FLD on basis of such high-dimensional image vectors is a very time-consuming process.

In this paper, a straightforward image projection technique, termed *image principal component analysis* (IMPCA), is proposed to overcome the weakness of the conventional PCA as applied in image recognition. Our main idea is to construct the *image total covariance matrix* directly based on original image matrices, and then to utilize it as a generation matrix to perform principal component analysis. Since the scale of image total covariance matrix of IMPCA is generally much smaller than that of PCA, much computational time will be saved. We will outspread our idea in the following section.

2. Image principal component analysis (IMPCA)

2.1. Idea and fundamentals

Let X denote an n-dimensional column vector, our idea is to project the image A, an $m \times n$ matrix, onto X by the following linear transformation [3]

$$Y = AX. \tag{1}$$

Thus, we get an m-dimensional projected vector Y, which is called projected feature vector of image A. How to determine a good projection vector X? In fact, the total scatter of the projected samples can be introduced to measure the discriminatory power of the projection vector X. From this point of view, we can adopt the following criterion

$$J_p(X) = tr(TS_x), \tag{2}$$

* Corresponding author. Tel.: +86-25-431-6840; fax: +86-25-431-5510.

E-mail addresses: tuqingh@mail.njust.edu.cn (Jian Yang), yangjy@mail.njust.edu.cn (Jing-yu Yang).

PII: S0031-3203(02)00040-7

where, TS_x denotes the total covariance matrix of projected feature vectors of training images, and $tr(TS_x)$ denotes the trace of TS_x. Now, the problem is how to evaluate the covariance matrix TS_x.

Suppose there are L known pattern classes, and M denotes the total number of samples of all classes. The jth training image is denoted by an $m \times n$ matrix A_j $(j = 1, 2, \ldots, M)$, and the mean image of all training sample is denoted by $\bar{A}$.

After the projection of training image onto X, we get the projected feature vector

$$Y_j = A_j X, \quad j = 1, 2, \ldots, M. \tag{3}$$

Suppose the mean vector of all projected feature vectors is denoted by $\bar{Y}$, it is easy to get $\bar{Y} = \bar{A}X$. Then, TS_x can be evaluated by

$$TS_x = \frac{1}{M}\sum_{j=1}^{M}(Y_j - \bar{Y})(Y_j - \bar{Y})^{\mathrm{T}}$$

$$= \frac{1}{M}\sum_{j=1}^{M}[(A_j - \bar{A})X][(A_j - \bar{A})X]^{\mathrm{T}}. \tag{4}$$

So

$$tr(TS_x) = X^{\mathrm{T}}\left(\frac{1}{M}\sum_{j=1}^{M}(A_j - \bar{A})^{\mathrm{T}}(A_j - \bar{A})\right)X. \tag{5}$$

Now, let us define the matrix below

$$G_t = \frac{1}{M}\sum_{j=1}^{M}(A_j - \bar{A})^{\mathrm{T}}(A_j - \bar{A}). \tag{6}$$

The matrix G_t is called *image total covariance* (*scatter*) *matrix*. And, it is easy to verify that G_t is an $n \times n$ nonnegative definite matrix by its definition.

Accordingly, the criterion in Eq. (2) can be expressed in this form

$$J_P(X) = X^{\mathrm{T}} G_t X. \tag{7}$$

2.2. Image principal component analysis

The criterion in Eq. (7) is called *generalized total scatter criterion*. In fact, in the special case of image matrix being row vectors, it is not hard to prove that the criterion is the common total scatter criterion. The vector X maximizing the criterion is called the optimal projection axe, and its physical meaning is obvious, i.e., after projection of image matrix onto X, the total scatter of projected samples is maximized.

It is evident that the optimal projection axe is the eigenvector corresponding to the maximal eigenvalue of G_t. Generally, in many cases, one optimal projection axe is not enough, we usually select a set of projection axes subject to the orthonormal constraints and maximizing the criterion in Eq. (7). In fact, the optimal projection axes $X_1, \ldots, X_d$ of IMPCA can be selected as the orthonormal eigenvectors of G_t associated with the first d largest eigenvalues.

Fig. 1. Five images of one person in ORL face database.

2.3. Feature extraction method

The obtained optimal projection vectors $X_1, \ldots, X_d$ of IMPCA are used for feature extraction. Let

$$Y_k = AX_k, \quad k = 1, 2, \ldots, d. \tag{8}$$

Then, we get a family of image projected feature vectors $Y_1, \ldots, Y_d$, which are used to form an $N = md$ dimensional resulting projected feature vector of image A as follows:

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_d \end{bmatrix} = \begin{bmatrix} AX_1 \\ AX_2 \\ \vdots \\ AX_d \end{bmatrix}. \tag{9}$$

3. Experimental results

The proposed method is tested on the ORL database that contains a set of face images taken at the Olivetti Research Laboratory in Cambridge, UK. There are 10 different images for 40 individuals. For some persons, the images were taken at different times. And the facial expression (open/closed eyes, smiling/non-smiling) and facial details (glasses/no glasses) are variable. The images were taken against a dark homogeneous background and the persons are in upright, frontal position with tolerance for some tilting and rotation of up to 20°. Moreover, there is some variation in scale of up to about 10%. All images are grayscale and normalized with a resolution of 92×112. The five images of one person in ORL are shown in Fig. 1.

In this experiment, we use the first five images of each person for training and the remaining five for testing. Thus the total amount of training samples and testing samples are both 200. The proposed IMPCA is used for feature extraction. Here, since the size of image total covariance matrix G_t is 92×92, it is very easy to work out its eigenvectors. And the number of selected eigenvectors (projection vectors) varies from 1 to 10. Note that if the projection vector number is k, the dimension of corresponding projected feature vector is $112 \times k$. Finally, a minimum distance classifier and a nearest neighbor classifier are, respectively, employed to classify in the projected feature space. The recognition rates are shown in Table 1. The Eigenfaces [1] and Fisherfaces [2] are used for feature extraction as well, and their optimal performance lies in Table 2. Moreover, the CPU

Table 1
Recognition rates (%) based on IMPCA projected features as the number of projection vectors varying from 1 to 10

Projection vector number	1	2	3	4	5	6	7	8	9	10
Minimum distance	73.0	83.0	86.5	88.5	88.5	88.5	90.0	90.5	91.0	91.0
Nearest neighbor	85.0	92.0	93.5	94.5	94.5	95.0	95.0	95.5	93.5	94.0

Table 2
Comparison of the maximal recognition rates using the three methods

Recognition rate	Eigenfaces	Fisherfaces	IMPCA
Minimum distance	89.5% (46)	88.5% (39)	91.0%
Nearest neighbor	93.5% (37)	88.5% (39)	95.5%

Note: The value in parentheses denotes the number of axes as the maximal recognition rate is achieved.

Table 3
The CPU time consumed for feature extraction and classification using the three methods

Time (s)	Feature extraction time	Classification time	Total time
Eigenfaces (37)	371.79	5.16	376.95
Fisherfaces (39)	378.10	5.27	383.37
IMPCA (112×8)	27.14	25.04	52.18

time consumed for feature extraction and classification using the above methods under a nearest neighbor classifier is exhibited in Table 3.

Tables 1 and 2 indicate that our proposed IMPCA outperforms Eigenfaces and Fisherfaces. What's more, in Table 3, it is evident that the time consumed for feature extraction using IMPCA is much less than that of the other two methods. So our methods are more preferable for image feature extraction.

Acknowledgements

We wish to thank National Science Foundation of China under Grant No.60072034 for supporting.

References

[1] M. Turk, A. Pentland, Face recognition using Eigenfaces, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1991, pp. 586–591.
[2] P.N. Belhumeur, J.P. Hespanha, D.J. Kriengman, Eigenfaces vs. Fisherfaces: recognition using class specific linear projection, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 711–720.
[3] K. Liu, Y.-Q. Cheng, J.-Y. Yang, et al., Algebraic feature extraction for image recognition based on an optimal discriminant criterion, Pattern Recognition 26 (6) (1993) 903–911.

About the Author—JIAN YANG was born in Jiangsu, China, on 3 June 1973. He received his M.S. degree in Applied Mathematics from Changsha Railway University in 1998. Now, he is a teacher in the Department of Applied Mathematics of Nanjing University of Science and Technology (NUST). At the same time, he is working for his Ph.D. degree in Pattern Recognition and Intelligence System. His current interests include face recognition and detection, handwritten character recognition and data fusion.

About the Author—JING-YU YANG received the B.S. Degree in Computer Science from NUST, Nanjing, China. From 1982 to 1984 he was a visiting scientist at the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign. From 1993 to 1994 he was a visiting professor at the Department of Computer Science, Missuria University. And in 1998, he acted as a visiting professor at Concordia University in Canada. He is currently a professor and Chairman in the department of Computer Science at NUST. He is the author of over 100 scientific papers in computer vision, pattern recognition, and artificial intelligence. He has won more than 20 provincial awards and national awards. His current research interests are in the areas of pattern recognition, robot vision, image processing, data fusion, and artificial intelligence.