Face recognition based on DCT and 2DLDA

Hongtao Yin, Ping Fu, and Jiaqing Qiao
Department of Automatic Test and Control, Harbin Institute of Technology, China
yinhongtao@gmail.com

Abstract

A face recognition method based on the discrete cosine transform (DCT) and two dimensional linear discriminant analysis (2DLDA) is presented. First, in this paper, the dimensionality of the original face image is reduced by using the DCT and the upper-left corner of the DCT matrix is selected to be the features of face image. Next, the proper feature are abstracted from the truncated DCT coefficient matrix by 2DLDA. The proposed algorithms are compared with both the DCT-based algorithm and the DCT+LDA algorithm which are proposed for face recognition. Experimental results on the ORL face database show that the proposed approach is feasible and has higher recognition performance than the other two algorithms.

1. Introduction

Face recognition is one of the most active research areas in computer vision and pattern recognition with practical applications that include forensic identification, access control and human computer interface. Many face recognition algorithms have been developed, such as eigenface[1,2], linear discriminant analysis(LDA)[3,4] and Independent Component Representations[5], etc. However, the computational requirements of these approaches are greatly related to the dimensionality of the original data and the number of training samples. When the face database becomes larger, the time for training and the memory requirement will significantly increase. As a consequence, it is impractical to apply the PCA in systems with a large database. The discrete cosine transform (DCT) has been employed in face recognition [6,7]. The DCT has several advantages over the PCA. First, the DCT is data independent. Second, the DCT can be implemented using a fast algorithm.

Yang et al [8] present a 2DPCA method which significantly reduces the feature extraction time than PCA. Fisher linear discriminant analysis (FLDA) has been successfully applied to face recognition area in the past few years. Nevertheless, FLDA usually encounters the small sample size (S3) problem in which the within-class scatter matrix becomes singular and thus the traditional FLDA algorithm fails to use. To address this problem, Based on the idea of 2DPCA, Yang et al [9] presented the two-dimensional linear discriminant analysis (2DLDA), which employed the image projection technique, and has been developed for image feature extraction. The 2DLDA becomes an interesting technique in face recognition, since it can extract discriminative feature faster than the one-dimensional discrimination analysis.

In this paper, we propose a method to apply 2DLDA for face recognition in DCT domain. The effectiveness of the proposed method is verified using the ORL (Olivetti Research Laboratory) database.

2. Discrete Cosine Transform

The DCT has been widely applied to solve numerous problems among the digital signal processing community. In particular, many data compression techniques employ the DCT, which has been found to be asymptotically equivalent to the optimal Karhunen-Loeve Transform (KLT) for signal decorrelation. Another merit of the DCT is it can be implemented efficiently using the Fast Fourier Transform (FFT). DCT has been widely used in image coding and face recognition. Given an input $M \times N$ image $f(x,y)$, its DCT, $C(u,v)$ is obtained by the following equation：

$$C(u,v) = a(u)a(v) \times \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\cos\frac{(2x-1)u\pi}{2M}\cos\frac{(2y-1)v\pi}{2N} \tag{1}$$

where $u = 0,1,\cdots,M-1$, $v = 0,1,\cdots,N-1$, and $a(u)$, $a(v)$ are defined by:

$$a(u) = \begin{cases} \sqrt{1/M} & u = 0 \\ \sqrt{2/M} & u = 1,2,\cdots,M-1 \end{cases} \tag{2}$$

$$a(v) = \begin{cases} \sqrt{1/N} & v = 0 \\ \sqrt{2/N} & v = 1,2,\cdots,N-1 \end{cases} \tag{3}$$

0-7695-2882-1/07 $25.00 ©2007 IEEE

where, x and y are spatial coordinates in the sample domain while u, v are coordinates in the transform domain.

For an $M \times N$ face image, we have an $M \times N$ DCT coefficient matrix covering all the spatial frequency components of the image. Figure 1. shows a $M \times N$ face image and its DCT coefficients.

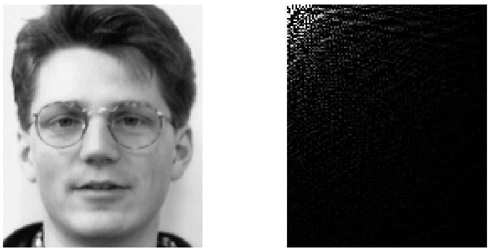

Figure. 1 Face image and its DCT coefficients

The DCT coefficients with large magnitude are mainly located in the upper-left corner of the DCT matrix. It can be observed that a large amount of information about the original image is stored in the upper-left corner of the DCT matrix. They are the low spatial frequency DCT components in the image.

3. Outline of 2DLDA

As opposed to conventional FLDA, 2DLDA is based on 2D matrices rather than 1D vector. The initial idea of 2DLDA is to perform the uncorrelated image matrix-based linear discriminant analysis (IMLDA)[10] twice: the first one is in horizontal direction and the second is in vertical direction. IMLDA can eliminate the correlations between image columns and compress the discriminant information optimally into a few of columns in horizontal direction. However, it disregards the correlations between image rows and the data compression in vertical direction. So, its compression rate is far lower than LDA and more coefficients are needed for the representation of images. This must lead to a slow classification speed and large storage requirements for large-scaled databases. After the two sequential IMLDA transforms, the discriminant information is compacted into a small matrix.

3.1 The IMLDA in horizontal direction

Suppose there are c known pattern classes. S is the total number of training samples, and S_i is the number of training samples in class i. In class i, the jth training image is denoted by an $m \times n$ matrix $A_j^{(i)}$. The mean image of training samples in class i is denoted by $\overline{A}^{(i)}$ and the mean image of all training sample is $\overline{A}$.

Based on the given training image samples (image matrices), the image between-class scatter matrix and image within-class scatter matrix can be constructed by

$$G_b = \frac{1}{S}\sum_{i=1}^{c} S_i \left(\overline{A}_i - \overline{A}\right)^T \left(\overline{A}_i - \overline{A}\right) \tag{3}$$

$$G_w = \frac{1}{S}\sum_{i=1}^{c}\sum_{j=1}^{S_i} \left(A_j^{(i)} - \overline{A}^{(i)}\right)^T \left(A_j^{(i)} - \overline{A}^{(i)}\right) \tag{4}$$

The generalized Fisher criterion can be defined by

$$J(\Phi) = \frac{\Phi^T G_b \Phi}{\Phi^T G_w \Phi} \tag{5}$$

It is easy to find a set of optimal discriminating vectors $\Phi_d = [\varphi_1, \varphi_2, \cdots, \varphi_d]$ by maximizing the Fisher criterion $J(\Phi)$. where $\Phi_d = [\varphi_1, \varphi_2, \cdots, \varphi_d]$ is the set of generalized eigenvectors of G_b and G_w corresponding to the d largest generalized eigenvalues , i.e., $G_b \varphi_i = \lambda_i G_w \varphi_i$, where $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_d$. The obtained eigenvectors Φ_d are used for image feature extraction. Let

$$B = A\Phi_d \tag{6}$$

where $\Phi_d = [\varphi_1, \varphi_2, \cdots, \varphi_d]$, the resulting feature matrix B is used to represent image A. The transform is called the uncorrelated image matrix-based linear discriminant analysis (IMLDA).

3.2 The IMLDA in vertical direction

After the first IMLDA transform in horizontal direction, we get the feature matrix B of sample A using Eq.(4). Constructing the image between-class and within-class scatter matrices $\mathbf{H}_b$ and $\mathbf{H}_w$ based on $\mathbf{B}^T$, we have

$$H_b = \frac{1}{S}\sum_{i=1}^{c} S_i \left(\overline{B}_i - \overline{B}\right)\left(\overline{B}_i - \overline{B}\right)^T \tag{7}$$

$$H_w = \frac{1}{S}\sum_{i=1}^{c}\sum_{j=1}^{S_i} \left(B_j^{(i)} - \overline{B}^{(i)}\right)\left(B_j^{(i)} - \overline{B}^{(i)}\right)^T \tag{8}$$

where $B_j^{(i)} = A_j^{(i)}\Phi_d$, $\overline{B}^{(i)} = \overline{A}^{(i)}\Phi_d$,and $\overline{B} = \overline{A}\Phi_d$. It is easy to find a set of optimal discriminating vectors $\Omega_e = [\ \omega_1\ ,\ \omega_2\ ,\ \cdots\ ,\ \omega_e\ ]$ by maximizing the Fisher criterion denotes as

$$J(\Omega) = \frac{\Omega^T H_b \Omega}{\Omega^T H_w \Omega} \tag{9}$$

We get the IMLDA feature matrix of B^T by

$$C^T = B^T \Omega_e \tag{10}$$

Thus

$$C = \Omega_e^T B = \Omega_e^T A \Phi_d \qquad (11)$$

The resulting feature matrix C is an $e \times d$ matrix, which is arranged a feature vector for classification. The process of 2DLDA is illustrated in Figure 2.

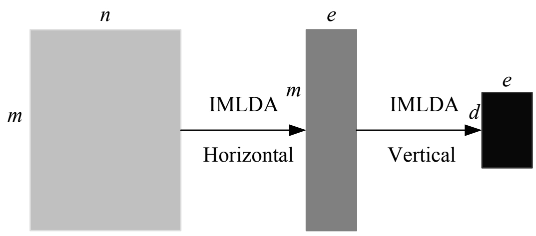

Figure 2. Illustration of 2DLDA transformation

As shown in Figure 2, the first 2DLDA transform $B = A\Phi_d$ compresses the 2D-data matrix ($m \times n$) in horizontal direction, making the image information pack into a small number of columns matrix ($m \times d$). While the second 2DPCA transform $C^T = B^T \Omega_e$ performs the compression of 2D-data in vertical direction, eliminating the correlations between columns of image **B** and making its image information further compact into a small number of rows matrix ($e \times d$). Ultimately, the image information is packed from matrix ($m \times n$) into the feature matrix ($e \times d$).

4. Proposed method

In our algorithm, to obtain the feature vector representing a face, first, its DCT is computed, and only a subset of the obtained coefficients is retained. The size of this subset is chosen such that it can sufficiently represent a face, but it can in fact be quite small. We select the low-to-mid frequency subset in the upper-left corner of its DCT coefficients. It can be expected that the DCT coefficients exhibit the expected behavior in which a relatively large amount of information about the original image is stored in a fairly small number of coefficients.

Second, IMLDA are applied to these selected DCT coefficient matrixes in horizontal and vertical directions respectively. After acquiring feature vectors with 2DLDA, the next step is to classify an unknown face image. In this paper, a nearest neighbor classifier is used to decide the class of an unknown face image. The block diagram of our method is shown in figure 3.

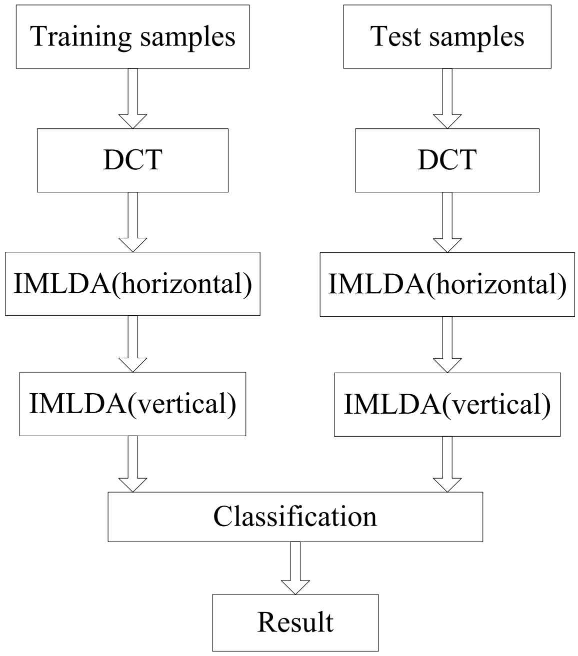

Figure 3. Block diagram of face recognition based on DCT and 2DLDA

5. Experimental results

To test the performance of the proposed method, some experiments are performed on a face database, which contains a set of face images taken at the Olivetti Research Laboratory (ORL) in Cambridge University, U.K. There are 400 images of 40 individuals in this database. For some subjects, the images were taken at different times, which contain quite a high degree of variability in lighting, facial expression (open/closed eyes, smiling/non-smiling etc), pose (upright, frontal position etc), and facial details (glasses/no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position, with tolerance for some tilting and rotation of up to 20°. The variation in scale is up to about 10%.

In the experiments, we randomly select 3,4,5 images from each subject to construct the training data set, the remaining images being used as the test images. Each experiment is repeated 20 times. The average recognition rates on the test sets are, respectively, summarized in Table 1.

Table 1. The recognition rates (%) on the ORL database

	10×10	15×15	20×20	25×25
3	92.66	92.64	92.98	92.75
4	95.48	95.67	95.23	95.37
5	96.92	96.65	97.15	97.30

We also compare our results with two face recognition algorithms, DCT[6] and DCT+LDA[7], that are based on DCT. The experiments are performed on ORL face database. We randomly select 5 face images from each subject to construct the training data set, the remaining images are used as the test images. Each experiment is repeated 20 times. The average recognition rates and testing times on the test sets are, respectively, summarized in Table 2 and Table 3.

Table 2. Comparison the most recognition rates (%) on the ORL database

	3	4	5
DCT	89.80	93.56	94.97
DCT+LDA	91.93	94.04	96.65
DCT+2DLDA	93.00	95.73	97.30

Table 3. Comparison of different methods in terms of training time and testing time

	Training times(s)	Recognition times(s)
DCT	3.8750	0.0216
DCT+LDA	8.2650	0.0434
DCT+2DLDA	3.9060	0.0199

From table 2 and table 3, we can see that our method is better than DCT and DCT+LDA methods both recognition rate and recognition time.

6. Conclusions

In this paper, we have proposed a face recognition algorithm based on DCT and 2DLDA. It implements 2DLDA in DCT domain instead of space domain. Experimental results on the ORL face database show that the proposed method is not only faster but also better in recognition performance than DCT and DCT+LDA algorithms.

7. References

[1] M. Kirby and L. Sirovich, “Application of the KL Procedure for the Characterization of Human Faces”. *IEEE Trans. Pattern Analysis and Machine Intelligence*. 12(1), 1990, pp. 103-108.

[2] M Turk and A Pentland, “Eigenfaces for recognition”. *Journal of Cognitive Neuroscience*, 3(1), 1991, pp. 71-86.

[3] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriengman, “Eigenfaces vs.Fisherfaces: recognition using class specific linear projection”. *IEEE Trans. Pattern Anal. Machine Intell.*, 19 (7), 1997, pp. 711–720.

[4] H. Yu and J. Yang, “A direct LDA algorithm for high-dimensional data --- with application to face recognition”. *Pattern Recognition*, 34(10), 2001, pp. 2067-2070.

[5] M. S. Bartlett, T. J. Sejnowski, “Independent Components of Face Images: a Representation for Face Recognition”. *Proc. 4th Annual J. Symp. Neural Computation*, 1997: 3~10.

[6] Z. M. Hafed and M. D. Levine, “Face Recognition Using the Discrete Cosine Transform”. *International Journal of Computer Vision*, 43(3), 2001, pp. 167~188.

[7] ZhangYan-Kun and Liu Chong-Qin, “A Novel Face Recognition Method Based on Linear Discriminant Analysis”. *J. Infrared Millim.Wav*. 22(5), 2003, pp. 327-330.

[8] J. Yang, D. Zhang, A.F. Frangi, J. Yang. “Two-dimensional PCA: a new approach to appearance based face representation and recognition”. *IEEE Trans. Pattern Anal. Machine Intell*. 26 (1), 2004, pp. 131 – 137.

[9] J. Yang, D. Zhang, x. Yong, and J.-Y. Yang, “Two-dimensional discriminant transform for face recognition”. *Pattern Recognition*, 38 (7), 2005, pp. 1125-1129.

[10] J. Yang, J.-y. Yang, A.F. Frangi, and D. Zhang, “Uncorrelated projection discriminant analysis and its application to face image feature extraction, Internat”. *J. Pattern Recogn. Artif. Intell.*, 17 (8), 2003, pp. 1325–1347.