Directional two-dimensional neighborhood preserving projection for face recognition

Li Yiying [1] Tian Qichuan [2]
[1] Key Laboratory on Integrated Services Networks,
XIDIAN University
Xi'an,China
lyymangguo126@126.com

Gao Quanxue [1] Xu jing [1]
[2] School of Electronic Information Engineering,
Taiyuan University of Science and technology,
Taiyuan, China
qxgao@xidian.edu.cn

Abstract—**This paper presents a novel manifold learning method, namely Directional two-dimensional neighborhood preserving embedding (Dir-2DNPE), for feature extraction. In contrast to standard NPE, Dir-2DNPE directly seeks the optimal projective vectors from the directional images without image-to-vector transformation. Moreover, Dir-2DNPE can well reserve the spatial correlations between variations of rows and those of columns of images. Experiments on the ORL and Yale databases show the effectiveness of the proposed method.**

Keywords-Neighborhood preserving embedding (NPE), Directional-image, 2-Dimensional NPE, Dir-2DNPE, face recognition

I. INTRODUCTION

Dimensionality reduction (DR) is one of the fundamental problems in computer vision, machine learning and biometric recognition. The aim of DR is to capture the meaningful low-dimensional structures embedded in high-dimensional data and obtain useful representation of data for subsequent analysis, such as classification, visualization, and clustering. Up to now, it assembles numerous methods, all striving to present high-dimensional data in low-dimensional space, in a way that faithfully captures the meaningful structures and unexpected relationships embedded in images [1]. In the absence of prior knowledge, such representation must be learned or discovered automatically. Automatic methods which discover hidden structure from the statistical regularities of large data sets can be studied in the general framework of unsupervised learning [2].

Principal component analysis (PCA) is one of the classical unsupervised dimensionality reduction methods, and is designed to operate when data points are linearly or almost linearly separable. Thus, PCA can fail to discover the meaningful structures of nonlinear manifolds. To address this problem, a large volume of manifold learning methods have been proposed, such as Isomap [3], Locally Linear Embedding (LLE) [4], Laplacian Eigenmap (LE) [5], Locality Preserving Projection (LPP) [6], and Neighborhood Preserving Embedding (NPE) [7], among which NPE is one of the classical methods. Different from PCA, which aims at preserving the global Euclidean structure of data, NPE can well preserve the local intrinsic structure, i.e. similarity, of data, and achieves impressive results in real-world applications. Based on this content, many NPE-based methods have been developed to improve the recognition accuracy. Among all these methods, the 2D face images must be transformed into 1D vectors column-by-column or row-by-row. However, concatenating the entries in 2D matrices into 1D vectors often leads to a high-dimensional data space and the small sample size (SSS), where it is difficult to calculate the local scatter matrix accurately due to the SSS. Furthermore, computing the eigenvectors of a large size local scatter matrix is very time-consuming [8,12].

Recently, a new technique called 2-dimensional neighborhood preserving embedding (2DNPE), which is a special case of tensor NPE [9], was proposed to solve the above problems. The main idea behind 2DNPE is to find the optimal projective vectors in the row direction of images without the image-to-vector transformation. That is, it calculates the so-called local scatter matrix, which can well measure the local intrinsic structure of nonlinear data, and total scatter matrix of data from the rows of images, and then computes eigenvectors of generalized eigen-function which is composed of local scatter matrix and total scatter matrix. Since the size of the scatter matrices is equal to the number of columns of images, which is quite small compared with the size of scatter matrices in NPE, so 2DNPE evaluates the local and total scatter matrices more accurately and computes the corresponding optimal projection more efficiency than NPE.

However, the projective vectors of 2DNPE only reflect variations between rows of images, while the omitted variations between columns of images are usually also useful for recognition [10-12]. In that case, 2DNPE can hardly obtain improved accuracy. In this paper, a novel method called Directional two-dimensional neighborhood preserving embedding (Dir-2DNPE) is proposed. In contrast to 2DNPE, Dir-2DNPE seeks the optimal projective vectors from the directional face images and therefore the correlations between variations of rows and those of columns of images can be kept. Experimental results on a subset of ORL and Yale databases show that Dir-2DNPE is much more accurate than both NPE and 2DNPE.

II. DIRECTIONAL NEIGHBORHOOD PRESERVING EMBEDDING

Our motivation for developing the Dir-2DNPE method originates from an essential observation on the recently proposed 2DNPE and Dia-PCA[12]. That is, 2DNPE can be

978-0-7695-4202-7/10 $26.00 © 2010 IEEE
DOI 10.1109/CASoN.2010.87

approximately seen as the row-based NPE, thus 2DNPE only reflects the information between rows, which implies some structure information (such as information between columns of images) cannot be uncovered by it. We attempt to solve that problem by transforming the original face images into corresponding directional face images, namely row directional images. Because the rows (columns) in the row directional face images simultaneously integrate the information of neighbors of rows/columns in original images, it can reflect both information between rows and those between columns. Through the entanglement of row and column information, it is expected that Dir-2DNPE may capture some useful structure information for improving recognition accuracy.

Given an arbitrary image matrix $A \in R^{m \times n}$ (without loss of generality, suppose $m \geq n$), the row direction image can be defined as:

Definition 1. (Row directional image) Re-sample and re-arrange the matrix A along the column direction to generate the row-directional image B, as show in Figure 1. Where parameter both l and h control the resulted directional image. For the convenient expression, both l and h are 2 in Figure 1.

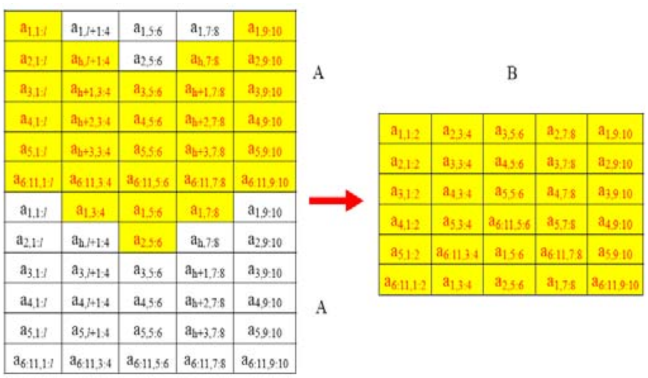


Figure 1. Row directional image formation

Given N training face images $A_i \in R^{m \times n}$ $(i = 1,2,\cdots,N)$. For each image, the corresponding rearranging face images $B_i \in R^{m \times n}$, i.e. row-directional images, can be obtained according to the definition 1.

Denote by $X^T = [B_1^T \, B_2^T \cdots B_N^T]$ the training image matrix, based on the row directional images and tensor-NPE, the objective function of Dir-2DNPE can be expressed as

$$\alpha^* = \arg \min_{\alpha^T \alpha = 1} \frac{\alpha^T X^T \left(M \otimes I_m\right) X\alpha}{\alpha^T X^T \left(I_N \otimes I_m\right) X\alpha} \qquad (1)$$

Where α denotes the projection map; $\otimes$ the Kronecker product, and T the transpose operator. $M = (I - W)^T (I - W)$, I denotes the identity matrix. The elements of matrix W can be calculated from

$$\min \sum_i \left\| A_i - \sum_j W_{ij} A_j \right\|_F^2 \qquad (2)$$

With constraints $\sum_j W_{ij} = 1, \; \forall i$ and if A_i is among k nearest neighbors of A_j or A_j is among k nearest neighbors of A_i, then $W_{ij} = 0$. Where $\|\bullet\|_F$ is Frobenius norm.

According to the matrix theory, the optimal projection α^* is the eigenvector of the following generalized eigen-equation corresponding to minimization nonzero eigen-value

$$X^T (M \otimes I_m) X\alpha = \lambda X^T (I_N \otimes I_m) X\alpha \qquad (3)$$

Where λ is the eigenvalue corresponding to eigenvector α.

Let $\mathrm{P} = [\alpha_1 \alpha_2 \cdots \alpha_d]$ denotes the projective matrix, projecting training faces A_i $(i = 1,2,\cdots N)$ onto P, yielding m by d feature matrices

$$C_i = A_i \mathrm{P} \qquad (4)$$

Given a test face image A, first use Eq. (3) to get the feature matrix $C = A\mathrm{P}$,then a nearest neighbor classifier can be used for classification. Here the distance between C and C_i can be defined as

$$d(C, C_i) = \|C - C_i\| = \sqrt{\sum_{k=1}^{m} \sum_{j}^{d} \left(C^{(k,j)} - C_i^{(k,j)}\right)^2} \qquad (5)$$

Where $C^{(k,j)}$ denote the k^{th} row j^{th} column of C, $C_i^{(k,j)}$ denote the k^{th} row j^{th} column of C_i .If $d(C, C^*) = \min_i d(C, C_i)$, and C^* belongs to class τ_p ,then A is assigned to the class τ_p, i.e. $A \in \tau_p$.

III. Experiments

In this section, Dir-2DNPE was evaluated on the two face database (ORL and Yale) and compared with the classical PCA, NPE, 2DPCA, and 2DNPE.

The ORL (http://images.ee.umist.ac.uk/danny/database. html) database contains images from 40 individuals, each providing 10 sample images. For each subject, the images were taken at different times. The facial expressions (open or closed eyes, smiling or non-smiling) and occlusion (glasses or no glasses) also vary. The images were taken with a tolerance for tilting and rotation up to 20 degrees. There is also some variation in the scale of up to 10 percent. All images were normalized to a resolution of 112×92. In experiment, the first 5 images per-person was selected as training, and the remaining images for testing. Thus the training images and testing images are both 200. Table 1 presents the comparisons of the five methods on top recognition accuracy. Fig. 2 shows the recognition accuracy of five methods versus the number of projected vectors.

The Yale database (http://cvc.yale.edu/projects/yalefaces/ yalefaces. html) contains 165 images of 15 individuals (each person providing 11 different images) under various facial expressions and lighting conditions. Each image is manually cropped and resized to 32×32 pixels [6]. The images from the first 3, 6, 9 images per-person were selected as training images; the remaining images for testing images. Table 2 lists the five methods and the associated number of features. Fig. 3 plots the recognition accuracy of five methods versus the number of projected vectors.

From the table 1, table 2, Figure 2, and Figure 3, it is easy to have that the top recognition accuracy of the proposed method is superior to that of PCA and 2DPCA when the parameter h is suitable selected. The reason may be that global structure preserved by PCA and 2DPCA can not well reflect the intrinsic structure of nonlinear data. While the proposed method uses an adjacency graph to group data and can well preserve the local intrinsic structure of data; Dir-2DNPE is also superior to NPE and 2DNPE, the main reason may be that NPE can not well preserve the spatial structure among pixels which is important for classification. 2DNPE can well preserve the relationship among pixels of rows, but it omits the relationships among pixels of column of images which is also useful for data classification. Differently, the proposed method employs the directional row-directional image to estimate the scatter matrices, thus the information among pixels of rows and columns of images can be kept in estimating the optimal projection. Note that, parameter l also plays an important role for classification [10,11]. The better classification accuracy can be achieved when l is 2, so in all experiments, we set l =2.

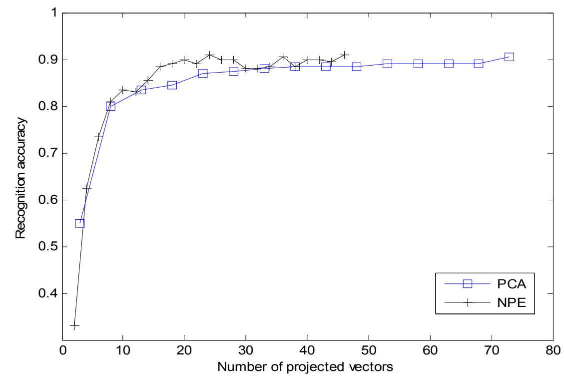
TABLE I. THE RECOGNITION ACCURACY (%) AND CORRESPONDING NUMBER OF FEATURES OF DIFFERENT SCHEMES ON THE ORL DATABASE.

Method	PCA	NPE	2DPCA	2DNPE	Dir-2DNPE		
					h=5	*h=23*	*h=40*
Accuracy	90.50	91.00	93.00	92.50	92.50	93.50	93.00
Dimension	73	24	112*7	112*3	112*7	112*12	112*12

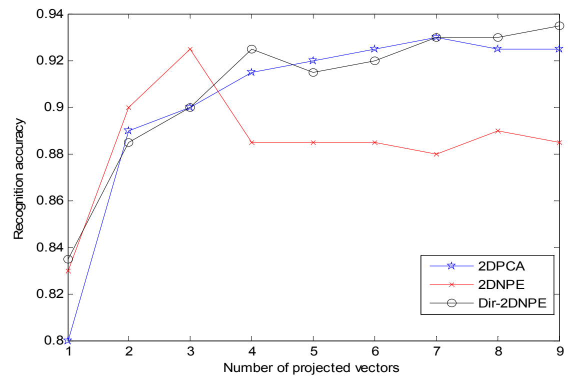
TABLE II. THE RECOGNITION ACCURACIES (%) OF DIFFERENT SCHEMES ON THE YALE DATABASE. THE VALUES IN PARENTHESES ARE THE CORRESPONDING NUMBER OF FEATURES.

Training number	PCA	NPE	2DPCA	2DNPE	Dir-2DNPE		
					h=4	*h=8*	*h=16*
3	59.17 (24)	60.00 (38)	61.67 (4)	61.67 (2)	62.50 (3)	67.50 (3)	63.33 (6)
6	66.67 (13)	68.00 (21)	68.00 (1)	73.33 (2)	74.67 (4)	74.67 (2)	73.33 (3)
9	86.67 (46)	93.33 (21)	86.67 (3)	93.33 (4)	93.33 (2)	93.33 (4)	93.33 (5)

Figure 2. The recognition accuracy of five methods on the ORL database show in figure (a) and (b) as follows:

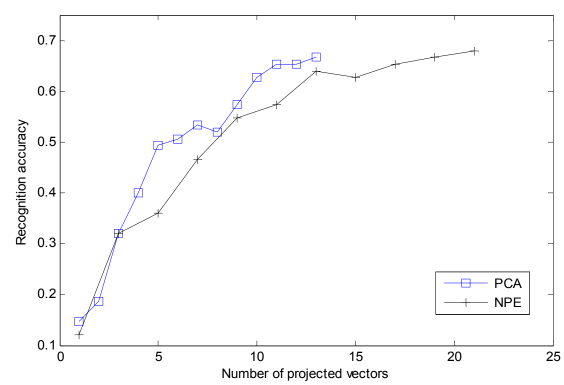


(a) vector-based methods

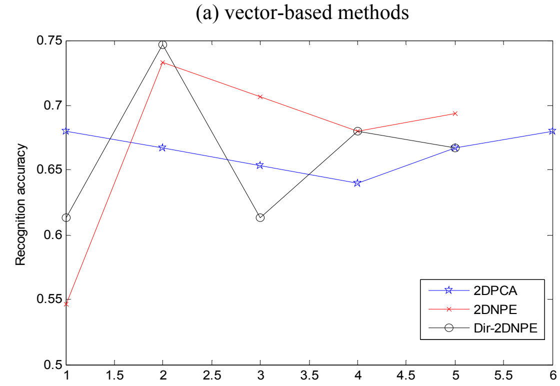


(b) matrix-based methods

Figure 3. The recognition accuracy of five methods on the Yale database show in figure (a) and (b) as follows:



(a) vector-based methods



(b) matrix-based methods

IV. CONCLUSION

A novel face recognition method called directional two-dimensional neighborhood preserving embedding (Dir-2DNPE) is proposed in this paper. The essential idea of the proposed method is to generate the directional face image from the original training images, from which the optimal projective vectors are sought, therefore the correlations between variations of rows and those of columns of images can be reserved. Experimental results on ORL and Yale databases show that Dir-2DNPE is much more accurate than PCA, 2DPCA, NPE, and 2DNPE.

ACKNOWLEDGMENT

We would like to thank all the anonymous reviewers for their constructive advices. This project is supported by National Natural Science Foundation of China under Grants no. 60802075, Key Laboratory on Integrated Services Networks of China under Grants 2010, 111 Project of China (B08038).

REFERENCES

[1] Y. Koren, and L. Carmel, "Robust linear dimensionality," IEEE Transactions on Visualization and Computer Graphics, 10, pp. 459-470, 2004.

[2] K. S. Lawrence, and T. R. Sam, "Think Globally, Fit Locally: Unsupervised Learning of Low Dimensional Manifolds," Journal of Machine Learning Research, 4, pp. 119-155, 2003.

[3] J. B. Tenenbaum, V. D. Silva, and J.C. Langford, "A global geometric framework for nonlinear dimensionality reduction," Science, 290(5500), pp.2319-2323, 2000.

[4] S. T. Roweis and L. K. Saul, " Nonlinear Dimensionality reduction by locally linear embedding," Science, 290(5500), pp.2323-2326, 2000.

[5] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," Neural Computation, 15, pp.1373-1396, 2003.

[6] X. He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang, "Face Recognition Using Laplacianfaces," IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(3), pp.328-340, 2005.

[7] X. He, D. Cai, S. Yan, and H. Zhang, "Neighborhood Preserving Embedding," Proc. ICCV, 2005.

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

[9] G. Dai and D. Y. Yeung, "Tensor embedding methods," Proceedings of the National Conference on Artificial Intelligence, pp.330-335, 2006.

[10] Q. Gao, L. Zhang and D. Zhang, "Directional independent component analysis with tensor representation," In proc. Of IEEE conference on Computer Vision and Pattern Recognition (CVPR), 2008.

[11] Q. Gao, L. Zhang, D. Zhang, and H. Xu, "Independent components extraction from image matrix," Pattern Recognition Letters, 2010.

[12] D. Zhang, Z. H. Zhou, and S. Chen, "Diagonal principal component analysis for face recognition," Pattern Recognition, 39, pp.140-142, 2006.