TWO-DIRECTIONAL TWO-DIMENSIONAL DISCRIMINANT LOCALITY PRESERVING PROJECTIONS FOR IMAGE RECOGNITION

Jiwen Lu and Yap-Peng Tan

School of Electrical and Electronic Engineering
Nanyang Technological University, Singapore

ABSTRACT

We propose in this paper an improved manifold learning method called two-directional two-dimensional discriminant locality preserving projections, $(2\mathrm{D})^2$-DLPP, for efficient image recognition. As the existing method of two-dimensional discriminant locality preserving projections (2D-DLPP) mainly relies upon the local structure information in the rows of images, we first derive an alternative 2D-DLPP algorithm that makes use of the information in the columns. Exploiting the local structure and discriminant information in both the rows and the columns, we develop the $(2\mathrm{D})^2$-DLPP method for efficient image feature extraction and dimensionality reduction. Experimental results on two benchmark image datasets show the effectiveness of the proposed method.

Index Terms— Locality preserving projections, two-directional two-dimensional analysis, image recognition.

1. INTRODUCTION

Locality preserving projections (LPP) has been shown to be efficient for image feature extraction and dimensionality reduction [1]. The aim of LPP is to seek an embedding that can best describe the essential manifold and preserve the local structure of images. As LPP is relatively insensitive to outliers, it has gained popularity in many applications of pattern recognition and computer vision, such as face recognition [1] and scene analysis [2]. However, a LPP-based image representation method needs to convert 2D images into 1D vectors, a step that compromises the structural information of images and usually leads to the problem of "curse of dimensionality".

To overcome this shortcoming, an improved LPP technique called two-dimensional locality preserving projections (2DLPP) [3, 4] has been recently proposed to directly project each image, rather than a lexicographically ordered vector, under a specific projection criterion. The effectiveness of 2DLPP is evidenced from experiments on several image databases [3, 4]. As 2DLPP is a unsupervised learning algorithm, it is suboptimal for image recognition and thus two-dimensional discriminant locality preserving projections (2D-DLPP) [5] has been more recently proposed to exploit discriminant information in 2DLPP, which has been successfully applied in facial expression recognition.

Like many other 2D dimensionality reduction methods, however, 2D-DLPP suffers from one major shortcoming: it needs many more coefficients for image representation when compared to LPP [3, 4] and DLPP [6]. Consider, for example, an image of size 128×128, the number of coefficients required by 2D-DLPP is $128 \times d$, where d is usually larger than 3 for satisfactory performance. Although this problem may be alleviated by applying PCA after 2D-DLPP, this additional dimensionality reduction may unduly compromise the image structure and the recognition performance.

Similar to some conventional two-dimensional subspace learning methods such as 2DPCA [13], 2DLDA [14] and 2DLPP [3, 4], 2D-DLPP also performs dimensionality reduction only in row direction. In other words, 2D-DLPP mainly relies on the local structure in the rows of the images. Inspired by the similar work done on principal component analysis (PCA) [7], Fisher's linear discriminant analysis (LDA) [8] and locality preserving projections [9], in this paper we first derive an alternative 2D-DLPP that exploits the local image structure in the other (the column) direction, and then develop the proposed $(2\mathrm{D})^2$-DLPP algorithm to perform DLPP in both the row and column directions, and evaluate its performance using two benchmark image datasets— ORL face database [10] and PolyU palmprint database [11, 12]—for face and palmprint recognition.

The remainder of this paper is organized as follows. Section 2 briefly reviews the existing 2D-DLPP algorithm. In Section 3, we derive the alternative 2D-DLPP method and propose the $(2\mathrm{D})^2$-DLPP method. In Section 4, we present the experimental results to show the effectiveness of the proposed $(2\mathrm{D})^2$-DLPP method. In Section 5, we conclude the paper by highlighting our contribution.

2. 2D-DLPP

Consider a training set consisting of N images X_i of $m \times n$ pixels, where $i = 1, 2, \cdots, N$. The 2D-DLPP minimizes an objective function defined as [5]:

$$J = \frac{\sum_{s=1}^{C} \sum_{i,j=1}^{N_s} (Y_i^s - Y_j^s)^T (Y_i^s - Y_j^s) S_{ij}^s}{\sum_{i,j=1}^{C} (M_i - M_j)^T (M_i - M_j) P_{ij}} \tag{1}$$

978-1-4244-2354-5/09/$25.00 ©2009 IEEE

where Y_i^s and Y_j^s denote the low-dimensional representation of X_i and X_j in the sth class, M_i and M_j are the mean samples of Y in the ith and jth classes, respectively, C is the number of classes, N_s denotes the number of training samples in the sth class, S_{ij}^s and P_{ij} are two affinity matrices, defined as

$$S_{ij}^s = \begin{cases} \exp\left(-\frac{\|X_i^s - X_i^s\|^2}{t_1}\right) & \text{if } L_{X_i} = L_{X_j} = \text{s} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

and

$$P_{ij} = \exp\left(-\frac{\|F_i - F_j\|^2}{t_2}\right) \quad (3)$$

where $F_i = \frac{1}{N_i}\sum_{k=1}^{N_i} X_k^i$ and $F_j = \frac{1}{N_j}\sum_{k=1}^{N_j} X_k^j$ are the mean samples of X in the ith and jth classes, L_{X_i} and L_{X_j} are the class labels of X_i and X_j, t_1 and t_2 are two empirically pre-specified parameters, respectively.

Let V be the transformation matrix and $Y_i = X_i V$, $i = 1, 2, \cdots, N$. By simple algebraic manipulations as shown in [5], one can reduce the numerator and denominator of (1) to $\frac{1}{2}V^T X^T L X V$ and $\frac{1}{2}V^T F^T H F V$, respectively, where $X^T = [X_1^T, X_2^T, \cdots, X_N^T]$ is an $mN \times n$ matrix obtained by arranging all the training images in a column form, $L = D - S$ is known as the Laplacian matrix with D being a diagonal matrix comprising elements $D_{ii} = \sum_j S_{ij}$, $F^T = [F_1^T, F_2^T, \cdots, F_C^T]$, $H = E - P$, and $E_{ii} = \sum_j P_{ji}$. Then, the projections of 2D-DLPP can be solved from the following generalized eigenvalue problem:

$$X^T L X v = \lambda F^T H F v \quad (4)$$

As matrices $X^T L X$ and $F^T H F$ are both symmetric and positive semi-definite, the eigenvalues obtained from (4) are no smaller than zero. Let $v_1, v_2, \cdots, v_d$ be the eigenvectors of (4) corresponding to the d smallest eigenvalues ordered according to $0 \le \lambda_1 \le \lambda_2 \le \cdots \le \lambda_d$. An $n \times d$ transformation matrix $V = [v_1, v_2, \cdots, v_d]$ can be obtained to project each $m \times n$ image X_i into an $m \times d$ feature matrix Y_i, as follows:

$$Y_i = X_i V, \quad i = 1, 2, \cdots, N \quad (5)$$

3. PROPOSED $(2D)^2$-DLPP

As 2D-DLPP mainly relies on the local structure in the rows of the images, we can easily derive an alternative 2D-DLPP that exploits the column directional information for feature extraction. Moreover, we consider perform feature extraction in both row and column directions, and thus introduce a new two-directional two-dimensional discriminant locality preserving projections ((2D)^2-DLPP) method, which performs dimensionality reduction and feature extraction both in the row and in column directions of image matrices.

Let $Z = VX_i$, and V be a $q \times m$ transformation matrix to be sought, where $X = [X_1, X_2, \cdots, X_N]$ is an $m \times Nn$ matrix obtained by arranging all the training images in a row form. Similarly, we can have an alternative of 2D-DLPP and its objective can be formulated as follows:

$$J(Z) = \frac{\sum_{s=1}^{C}\sum_{i,j=1}^{N_s}(Z_i^s - Z_j^s)(Z_i^s - Z_j^s)^T S_{ij}^s}{\sum_{i,j=1}^{C}(M_i - M_j)(M_i - M_j)^T W_{ij}} \quad (6)$$

where Z_i^s and Z_j^s denote the projected features of X_i and X_j in the sth class, respectively. After simple algebraic manipulations, we obtain the projections of this alternative 2D-DLPP through solving the following generalized eigenvalue problem:

$$XLX^T w = \eta F H F^T w \quad (7)$$

Let $w_1, w_2, \cdots, w_q$ be the eigenvectors of (7) corresponding to the q smallest eigenvalues ordered according to their values, i.e., $0 \le \eta_1 \le \eta_2 \le \cdots \le \eta_q$. An $q \times m$ transformation matrix $W = [w_1, w_2, \cdots, w_q]$ can thus be obtained to project each $m \times n$ image X_i into an $q \times n$ feature matrix Z_i, as follows:

$$Z_i = W X_i, \quad i = 1, 2, \cdots, N \quad (8)$$

To obtain an efficient DLPP-based image representation that exploits the local image structure and reduces the dimensions in both the row and column directions, we propose to seek transformation matrices V and W of size $n \times d$ and $q \times m$, respectively, that project each $m \times n$ image X_i into a $q \times d$ feature matrix, given as

$$T_i = W X_i V, \ i = 1, 2, \cdots, N \quad (9)$$

by optimizing the objective function

$$J(T) = \frac{\sum_{s=1}^{C}\sum_{i,j=1}^{N_s}(T_i^s - T_j^s)^T (T_i^s - T_j^s) S_{ij}^s}{\sum_{i,j=1}^{C}(Q_i - Q_j)^T (Q_i - Q_j) W_{ij}} \quad (10)$$

where Q_i and Q_j are the means of samples of T in the ith class and jth class, respectively,

To the best of our knowledge, there is no closed-form solution to (10). Hence, we apply a stepwise strategy to solve it. We propose to (i) obtain the $n \times d$ transforation matrix V by solving the generalized eigenvalue problem of (4) using the $m \times n$ training images X_i, $i = 1, 2, \cdots, N$; (ii) use V to project the training image X_i into $n \times d$ feature matrices Y_i; and (iii) obtain the $q \times m$ transformation matrix W by solving the generalized eigenvalue problem (7). Alternatively, we can obtain matrix W by solving (7) first and then matrix V by solving (4). Our empirical study has shown that similar performance in representation and recognition can be attained regardless of which transformation matrix is obtained first.

For efficient representation, the transformation matrix V and W can be used to project an $m \times n$ image X_k into a $q \times d$ feature matrix $T_k = W X_k V$. For recognition, we apply a nearest neighbor classifier on the distance between T_k and T_i (the feature matrix of training image X_i) defined as

$$d(T_k, T_i) = \|T_k - T_i\|_2 = (\sum_{x=1}^{q}\sum_{y=1}^{d}(T_k^{x,y} - T_i^{x,y}))^{1/2} \quad (11)$$

Fig. 1. Ten samples of one subject in ORL database.

where $T_k^{x,y}$ and $T_i^{x,y}$ denote the (x, y) element of matrices T_k and T_i, respectively.

4. EXPERIMENTAL RESULTS

We carried out several experiments on two benchmark image databases—ORL face database [10] and PolyU palmprint database [11, 12]—to evaluate the performance of the proposed (2D)^2-DLPP method in comparison with DLPP [6], 2D-DLPP [5], (2D)^{2}PCA [7], and (2D)^{2}LDA [8], and other conventional feature extraction methods such as PCA, LDA, 2DPCA [13] and 2DLDA [14]. The experiments were conducted on a PC with 3.4 GHz CPU and 1GB memory.

4.1. Results on ORL database

The ORL database contains 400 images from 40 subjects, and each subject has ten different images. The images of some of the subjects were acquired at different times. Furthermore, the images were taken with a tolerance for face tilting and rotation by up to 20° and a variation in image scaling by up to 10%. All images are in gray levels and normalized to a resolution of 112×92 pixels. Fig. 1 shows ten samples of one subject in the database.

For each subject, we randomly selected k images to construct the training set and the remaining images of each subject as the testing set. We performed two comparative experiments with $k = 2$ and 5, and empirically selected the parameter t by the cross validation strategy. Table 1 shows the performance of the evaluated feature extraction methods.

Table 1. Top recognition rate (%) and training time (s) with corresponding reduced dimension obtained by each method on ORL face database.

Method	Dim	$k = 2$		$k = 5$	
		Time	CRR	Time	CRR
PCA	30	0.188	78.8	0.766	90.5
LDA	39	0.125	75.9	0.609	92.0
DLPP	30	0.188	80.9	0.641	93.5
2DPCA	112×3	0.016	86.6	0.016	96.0
2DLDA	112×3	0.016	85.9	0.016	96.5
2DLPP	112×3	0.016	88.1	0.016	97.0
Alter. 2DPCA	5×92	0.016	84.7	0.016	94.5
Alter. 2DLDA	5×92	0.016	84.4	0.016	95.0
Alter. 2D-DLPP	5×92	0.1016	84.9	0.016	97.0
(2D)^{2}PCA	5×5	0.031	88.1	0.031	96.5
(2D)^{2}LDA	5×5	0.031	85.9	0.031	97.5
(2D)^2-DLPP	5×5	0.031	**88.4**	0.031	**98.5**

To evaluate the effects of face alignment on the proposed (2D)^2-DLPP method, we prepared the face images in two different ways: one is normalizing each image to align the two eyes at the same height, and the other is simply cropping from each image (without alignment) a subregion to include the main part of the face. Each processed image is of size 64×64 pixels and some samples are shown in Fig. 2. We then randomly selected 5 images of each subject to construct the training set and the remaining images as the testing set and applied DLPP, 2DLPP, alternative 2D-DLPP and (2D)^2-DLPP methods to perform face recognition.

Fig. 2. Sample images of the aligned images (the first five) and the cropped images (the second five).

Table 2. Top recognition rate (%) and training time (s) with corresponding reduced dimension obtained by each method on ORL face database with/without alignment.

Method	Aligned (64×64)			Cropped (64×64)		
	Dim	Time	CRR	Dim	Time	CRR
DLPP	30	0.266	95.5	30	0.266	87.0
2D-DLPP	64×3	0.016	98.0	64×3	0.016	89.0
Alter. 2D-DLPP	5×64	0.031	98.0	5×64	0.031	88.5
(2D)^2-DLPP	5×5	0.063	**98.5**	5×5	0.063	**94.0**

We can see from Table 2 that the proposed (2D)^2-DLPP always attains the highest correct recognition rate with the same and fewer coefficients (dimensions) among all the methods under comparison. Furthermore, similar to other 2D-based feature representation methods, (2D)^2-DLPP is also faster than 1D-DLPP method. From Table 2, we can easily see that without proper alignment, the recognition performance of DLPP, 2D-DLPP and the alternative 2D-DLPP methods reduces significantly, while that of the proposed (2D)^2-DLPP method can still maintain over 90% correct recognition rate. In other words, (2D)^2-DLPP appears to more robust than the other comparison method when the face samples are not perfectly aligned.

4.2. Results on PolyU palmprint database

The PolyU palmprint database [11, 12] contains the palmprints of 100 subjects with six samples from each subject. These palmprint images were collected in two sessions, and three samples were acquired in each session. Fig. 3 shows six cropped palmprint images of size 128×128 from one subject.

We randomly selected 4 palmprint images of each subject to construct the training set and the remaining 2 as the testing set. Table 3 shows the recognition performance of (2D)^2-DLPP versus other DLPP-based feature extraction methods. The superiority of the proposed (2D)^2-DLPP method is evidenced for its highest recognition accuracy despite using the

Fig. 3. Samples of the cropped palmprint images from one subject of PolyU database.

Table 3. Top recognition rate (%) and training time (s) with corresponding reduced dimension obtained by each method on PolyU palmprint database.

Method	Dim	Time	CCR)
DLPP	49	2.3125	92.0
2D-DLPP	128×8	0.0313	92.0
Alter. 2D-DLPP	8×128	0.0313	92.5
$(2D)^2$-DLPP	8×8	0.0625	**94.0**

same or fewer coefficients (dimensions) for feature representation.

To further reveal the relationship between the accuracy and dimension of the feature matrices, we conducted a series of experiments with different feature dimensions using the DLPP and the proposed $(2D)^2$-DLPP methods. It is easy to see from the results shown in Fig. 4 that the proposed $(2D)^2$-DLPP consistently achieves better recognition accuracy than the DLPP method under different feature dimensions.

5. CONCLUSIONS

We have proposed in this paper an efficient image representation and recognition method called $(2D)^2$-DLPP. The main difference between the proposed method and the existing 2D-DLPP method is that the latter only relies on the local structure in the row of the images, while our proposed method exploits the local structure in both the image rows and columns. As a result, the proposed method requires fewer coefficients for image representation and attains better recognition accuracy than the existing 2D-DLPP method and other dimensionality reduction methods. Experimental results on benchmark face and palmprint databases clearly show the efficacy of the proposed method.

6. REFERENCES

[1] X. F. He, S. C. Yan, Y. X. Hu, P. Niyogi, and H. J. Zhang, "Face recognition using Laplacianface," *IEEE Trans. on PAMI,* vol. 27, no. 3, pp. 328-340, 2005.

[2] L. Wang, and D. Suter, "Learning and matching of dynamic shape manifold for human action recognition," *IEEE Trans. on Image Processing,* vol. 29, no. 6, pp. 1646-1661, 2007.

[3] D. W. Hu, G. Y. Feng, and Z. T. Zhou, "Two-dimensional locality preserving projections (2DLPP) with its application to palmprint recognition," *Pattern Recognition,* vol. 40, no. 1, pp. 339-342, 2007.

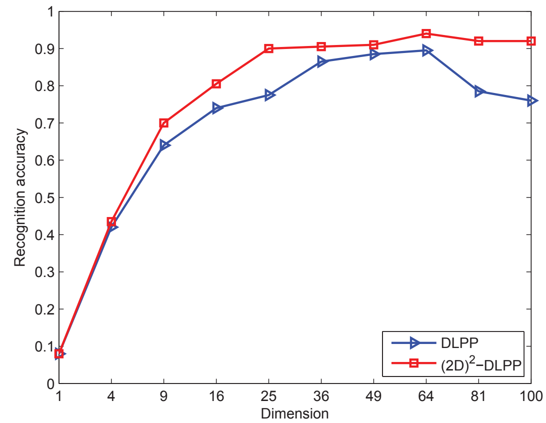

Fig. 4. Recognition accuracies of $(2D)^2$-DLPP and DLPP under different projection dimensions.

[4] B. Niu, Q. Yang, C.K.Shiu, and S. K. Pal, "Two dimensional Laplacianfaces method for face recognition," *Pattern Recognition,* vol. 41, pp. 3237-3243, 2008.

[5] R. Zhi, and Q. Ruan, "Facial expression recognition based on two-dimensional discriminant locality preserving projections," *Neurocomputing,* vol. 71, pp. 1730-1734, 2008.

[6] W. Yu, X. Teng, and C. Liu,"Face recognition using discriminant locality preserving projections," *Image and Vision Computing,* vol. 24, pp. 239-248, 2006.

[7] D. Zhang and Z. H. Zhou, $(2D)^2$PCA: Two-directional two-dimensional PCA for efficient face representation and recognition, *Neurocomputing,* vol. 69, pp. 224-231, 2005.

[8] P. Nagabhushan, D. S. Guru and B. H. Shekar, $(2D)^2$FLD: An efficient approach for appearance based object recognition, *Neurocomputing,* vol. 69, pp. 934-940, 2006.

[9] S. B. Chen, B. Luo, G. P. Hua, and R. H. Wang, "Bitateral two-dimensional discriminant locality preserving projections," *in Proc. IEEE ICASSP'2007,* pp. 601-604, 2007.

[10] The ORL (AT&T) face database. <Available at http://www.cam-orl.co.uk>.

[11] The PolyU palmprint database. <Available at http://www.comp.polyu.edu.hk/biometrics>.

[12] D. Zhang, W. K. Kong, J. You, and M. Wong, "Online palmprint identification," *IEEE Trans. on PAMI,* vol. 25, no. 9, pp. 1041-1050, 2003.

[13] 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. on PAMI,* vol. 26, no. 1, pp. 131-137, 2004.

[14] H. Xiong, M. N. Swamy, and M. O. Ahmad, "Two-dimensional FLD for face recognition," *Pattern Recognition,* vol. 38, no. 7, pp. 1121-1124, 2005.