Journal of Theoretical and Applied Information Technology
10th January 2013. Vol. 47 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

TWO-DIMENSIONAL MAXIMUM MARGIN PROJECTION FOR FACE RECOGNITION

[1]YOUMAO BAI, [2]ZIQIANG WANG, [1]XU QIAN

[1] College of Mechanical Electronic and Information Engineering, China University of Mining and Technology (Beijing), Beijing China

[2]School of Information Science and Engineering, Henan University of Technology, Zhengzhou, China

ABSTRACT

To effectively cope with the high dimensionality problem in face recognition, a novel two-dimensional maximum margin projection (2DMMP) algorithm for face recognition is proposed in this paper. Specially, 2DMMP is based on the maximum margin projection (MMP) and fully considers the intrinsic tensor structure of face image. By utilizing both local manifold structure and discriminative information, as well as characterizing the correlations among pixels with the tensor structure, 2DMMP can effectively project the high-dimensional face image space into lower-dimensional feature space for face recognition. Experimental results on three face databases show that the proposed 2DMMP algorithm outperforms other related algorithms in terms of recognition accuracy.

Keywords: *Face Recognition, Maximum Margin Projection (MMP), Tensor Structure, Manifold Learning*

1. INTRODUCTION

Face recognition has received increasing attention from researchers in the past few decades due to its wide applications, such as identity authentication, surveillance, image retrieval and human-computer interaction [1]. In general, a face image of size of $n_1 \times n_2$ is represented as a vector in the image space. Through the image space is of high dimensionality, the face image space is often a low dimensional feature space which is embedded in the ambient space. Therefore, it is often essential to conduct dimensionality reduction to acquire an efficient and discriminative representation before formally conducting classification. Dimensionality reduction could effectively avoid the "curse of dimensionality", improve performance and computational efficiency of face recognition algorithms, and alleviate storage requirement [2]. Thus, dimension reduction is an important data preprocessing step for face recognition applications. Principal component analysis (PCA) and linear discriminant analysis (LDA) [3] are two of the well-known dimensionality reduction techniques for face recognition.

PCA also known as Karhunen-Loeve transformation, is a classical dimensionality reduction technique widely used in the areas of data mining and pattern recognition. PCA projects the data points into a lower dimensional subspace, and aims to find a set of mutually orthogonal bases that capture the global information of the data points in terms of variance. However, PCA is unsupervised, it does not fully use the class label information of the given face images. LDA also called Fisher's linear discriminant, is a supervised dimensionality reduction method. LDA aims to find an optimal transformation that maps the data into a lower-dimensional space (while preserving the class label information) that minimizes the within-class scatter and simultaneously maximizes the between-class scatter, thus achieving maximum discrimination. Both PCA and LDA have been extensively used in face recognition and created the popular Eigenfaces and Fisherfaces [3], respectively. However, both PCA and LDA effectively find only the Euclidean structure, they fail to discover the underlying face image manifold structure [4].

In recent years, various researchers have shown that face space is usually a sub-manifold of very low dimensionality which is embedded in the ambient space of very high dimensional image. In order to discover the intrinsic manifold structure, many manifold-related algorithms have been proposed for dimensionality reduction, such as isometric feature mapping(ISOMAP) [5] , locally linear embedding(LLE) [6], Laplacian eigenmap (LE) [7], and locality preserving projections (LPP) [4], four of the most popular manifold learning algorithms. However, the former three algorithms all suffer from the out of sample problem, i.e., we cannot obtain the low-dimensional representation of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

samples not in the training set [8]. Although LPP adopts an linearization procedure to construct explicit maps over new testing set, all these algorithms are imperfect for supervised learning tasks(such as face recognition) since they are unsupervised and only consider the intra-class geometry. In order to overcome the above shortcomings, a novel manifold learning algorithm called maximum margin projection (MMP) [9], has been recently proposed for dimensionality reduction. By jointly considering the local manifold structure and discriminative information together for dimensionality reduction, MMP and its extensions have been successfully applied to various tasks. However, it still has the following problems that are not properly addressed till now: it unfold input image data into vectors before dimensionality reduction. In fact, face images are intrinsically in the form of second or higher order tensors. As a result, such vectorization process ignores the underlying data structure and often leads to the curse of dimensionality and the small sample size problems. Therefore, it is often helpful to process the face image in their original form and order. Several research groups have shown that the image as tensor representation can lead to good classification performance for different applications [10-12]. Nevertheless, how to conduct MMP for dimensionality reduction by encoding a face image as a two-order tensor structure in the context of face recognition is still a research area where few people have tried to explore. In this paper, we propose a new two-dimensional maximum margin projection (2DMMP) for face recognition and demonstrate that the proposed algorithm alleviate the above problems when using the vector representation.

The rest of the paper is organized as follows. In Section 2, we provide a brief review of the MMP algorithm. The two-dimensional MMP algorithm for face recognition is developed in Section 3. The experimental results are reported in Section 4. Finally, the conclusions are presented in Section 5.

2. BRIEF REVIEW OF MMP

MPP is a recently proposed manifold learning algorithm for dimensionality reduction [9]. It is based on two local adjacency graphs and explicitly considers the local manifold structure and the class relationship between the face images.

Given a set of face images $x_1, x_2, \ldots, x_n \in R^p$, Let $X = [x_1, x_2, \ldots, x_n]$. Let S_b and S_w be weight matrices of between-class graph G_b and within-class graph G_w, respectively. The optimal projection matrix of MMP can be obtained by solving the following maximization problem:

$$a_{opt} = \arg\max_a a^T X\left(\beta L_b + (1-\beta) S_w\right) X^T a \quad (1)$$

with the constraint

$$a^T X D_w X^T a = 1 \quad (2)$$

where $\beta \in [0,1]$ is an positive constant which control the tradeoff between L_b and S_w, it is empirically set to be 0.5 in our experiments. $L_b = D_b - S_b$ is the Laplacian matrix of G_b, D_b is a diagonal matrix whose entries on diagonal are column sum of S_b, i.e., $D_{b,ii} = \sum_j S_{b,ij}$, D_w is a also diagonal matrix whose entries on diagonal are column sum of S_w, i.e., $D_{w,ii} = \sum_j S_{w,ij}$. The definitions of two weight matrices S_b and S_w are as follows:

$$S_{w,ij} = \begin{cases} \gamma, & \text{if } x_i \text{ and } x_j \text{ share the same label} \\ 1, & \text{if } x_i \text{ or } x_j \text{ is unlabeled} \\ & \text{but } x_i \in N_w(x_j) \text{ or } x_j \in N_w(x_i) \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

$$S_{b,ij} = \begin{cases} 1, & \text{if } x_i \in N_b(x_j) \text{ or } x_j \in N_b(x_i) \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

$$N_b(x_i) = \left\{ x_i^j \,\middle|\, l(x_i^j) \neq l(x_i), j = 1, \ldots, k \right\} \quad (5)$$

$$N_w(x_i) = N(x_i) - N_b(x_i) \quad (6)$$

where $N(x_i)$ and $l(x_i)$ denote the set of k nearest neighbors and the label of x_i, respectively. Specially, $N_b(x_i)$ contains the neighbors having different labels, and $N_w(x_i)$ contains the rest of the neighbors.

The objective function of MMP incurs a heavy penalty if neighboring points x_i and x_j in the within-class graph is mapped far apart. Meanwhile, it incurs a heavy penalty if neighboring points x_i and x_j in the between-class graph is mapped close together. Finally, the optimal projection matrix a that maximizes (1) is given by the maximum

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

eigenvalue solution to the following generalized eigenvalue problem:

$$X\left(\beta L_b + (1-\beta) S_w\right) X^T a = \lambda X D_w X^T a \qquad (7)$$

Note that since the number of images is less than the dimensionality of images, it implies that both $X(\beta L_b + (1-\beta)S_w)X^T$ and XD_wX^T are singular. To overcome this problem, one can first project the face image into the PCA subspace by throwing away those zero singular values. After the eigenvector are obtained by solving (7), then for the new testing image x , its lower-dimensional embedding is as follows:

$$x \rightarrow y = A^T x \qquad (8)$$

where y is a lower feature representation of the face image x , and $A = [a_1, a_2, \cdots, a_l]$ is the projection matrix of MMP.

From the above computing process of MMP, we can observe the following drawbacks of MMP: It suffers from the singular problem, which stems from MMP unfold face image into vectors before dimensionality reduction. Such kind of vectorization largely increases the computational costs of dimensionality reduction and seriously destroys the intrinsic two-order tensor structure of face image. In order to overcome the above problem of MMP, we propose the tensor extension of MMP for face recognition in the following section.

3. TENSOR EXTENSION OF MMP FOR FACE RECOGNITION

Given a set of face images $X_1, X_2, \cdots, X_n$ in the two-order tensor $\mathbb{R}^{n_1 \times n_2}$, tensor MMP aims to find two transformation matrices $U \in \mathbb{R}^{n_1 \times l_1}$ and $V \in \mathbb{R}^{n_2 \times l_2}$ that project each face images X_i into a lower-dimensional feature representation $Y_i \in \mathbb{R}^{l_1} \times \mathbb{R}^{l_2}$ by using $Y_i = U^T X_i V$ such that Y_i represents X_i in terms of local manifold structure and discriminative information, where $i = 1, \cdots n$, $l_1 < n_1$, and $l_2 < n_2$.

In order to make the connected points of the between-class graph G_b stay as distant as possible while the connected points of the within-class graph G_w stay as close together as possible, two-dimensional (tensor) MMP aims to find two transformation matrices by solving the following optimal objective function:

$$\begin{aligned} &\min \sum_{ij} \left\| Y_i - Y_j \right\|^2 W_{w,ij} \\ &= \min \sum_{ij} \left\| U^T X_i V - U^T X_j V \right\|^2 W_{w,ij} \end{aligned} \qquad (9)$$

$$\begin{aligned} &\max \sum_{ij} \left\| Y_i - Y_j \right\|^2 W_{b,ij} \\ &= \max \sum_{ij} \left\| U^T X_i V - U^T X_j V \right\|^2 W_{b,ij} \end{aligned} \qquad (10)$$

with the constraint

$$Y^T D_w Y = 1 \Rightarrow \left(U^T X V\right)^T D_w U^T X V = 1 \qquad (11)$$

According to the matrix theory $\|A\|^2 = Tr\left(AA^T\right)$, (9) and (11) can be rewritten as follows:

$$\begin{aligned} &\sum_{ij} \left\| Y_i - Y_j \right\|^2 W_{w,ij} T \\ &= \sum_{ij} Tr\left[\left(U^T X_i V - U^T X_j V\right)\left(U^T X_i V - U^T X_j V\right)^T \right] W_{w,ij} \\ &= 2Tr\left[\sum_i U^T X_i V D_{w,ii} V^T X_i^T U - \sum_{ij} U^T X_i V W_{w,ij} V^T X_j^T U \right] \\ &= 2Tr\left[\sum_i U^T X_i V D_{w,ii} V^T X_i^T U - \sum_{ij} U^T X_i V W_{w,ij} V^T X_j^T U \right] \end{aligned} \qquad (12)$$

$$\begin{aligned} &Tr\left(Y^T D_w Y\right) \\ &= Tr\left[\left(U^T X V\right)^T D_w U^T X V \right] \\ &= Tr\left[\left(U^T X V\right)^T D_w U^T X V \right] \\ &= Tr\left[U^T \left(X^T V D_w V^T X \right) U \right] \end{aligned} \qquad (13)$$

By combining (12) and (13), the minimization problem of (9) can be transformed into the following maximization problem:

$$\begin{aligned} &Tr\left(\sum_{ij} U^T X_i V W_{w,ij} V^T X_j^T U \right) \\ &= Tr\left[U^T \left(\sum_{ij} X_i V W_{w,ij} V^T X_j^T \right) U \right] \end{aligned} \qquad (14)$$

In addition, by using $\|A\|^2 = Tr\left(AA^T\right)$, (10) can be rewritten as follows:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$\begin{aligned}
&\sum_{ij}\left\|Y_i - Y_j\right\|^2 W_{b,ij} \\
&= \sum_{ij}\left\|U^T X_i V - U^T X_j V\right\|^2 W_{b,ij} \\
&= \sum_{ij} Tr\left[\left(U^T X_i V - U^T X_j V\right)\left(U^T X_i V - U^T X_j V\right)^T\right] W_{b,ij} \\
&= 2Tr\left[\sum_i U^T X_i V D_{b,ii} V^T X_i^T U - \sum_{ij} U^T X_i V W_{b,ij} V^T X_j^T U\right] \\
&= 2Tr\left[U^T\left(\sum_i X_i V D_{b,ii} V^T X_i^T - \sum_{ij} X_i V W_{b,ij} V^T X_j^T\right) U\right]
\end{aligned} \tag{14}$$

In order to concisely describe in the following, we define

$$Q_{vw} = \sum_{ij} X_i V W_{w,ij} V^T X_j^T \tag{15}$$

$$Q_{vb} = \left(\sum_i X_i V D_{b,ii} V^T X_i^T - \sum_{ij} X_i V W_{b,ij} V^T X_j^T\right) \tag{16}$$

$$Q_w = \left(X^T V D_w V^T X\right) \tag{17}$$

Then, the optimal objective function of tensor MMP can be rewritten as the following maximization problem:

$$\max_{U,V} \frac{Tr\left[U^T\left(\beta Q_{vb} + (1-\beta) Q_{vw}\right) U\right]}{Tr\left[U^T Q_w U\right]} \tag{18}$$

Similarly, by using $\|A\|^2 = Tr\left(A^T A\right)$, the optimal objective function of tensor MMP can be also rewritten as the following maximization problem:

$$\max_{U,V} \frac{Tr\left[V^T\left(\beta Q_{ub} + (1-\beta) Q_{uw}\right) V\right]}{Tr\left[V^T P_w V\right]} \tag{19}$$

where

$$Q_{uw} = \sum_{ij} X_j^T U W_{w,ij} U^T X_i \tag{20}$$

$$Q_{ub} = \left(\sum_i X_i^T U D_{b,ii} U^T X_i - \sum_{ij} X_i^T U W_{b,ij} U^T X_j\right) \tag{21}$$

$$P_w = \left(X U D_w U^T X^T\right) \tag{22}$$

Finally, we obtain the optimal problems of tensor MMP:

$$\max_{U,V} \frac{Tr\left[U^T\left(\beta Q_{vb} + (1-\beta) Q_{vw}\right) U\right]}{Tr\left[U^T Q_w U\right]} \tag{23}$$

$$\max_{U,V} \frac{Tr\left[V^T\left(\beta Q_{ub} + (1-\beta) Q_{uw}\right) V\right]}{Tr\left[V^T P_w V\right]} \tag{24}$$

The above two maximization problems (23) and (24) depend on each other, and hence can not be solved independently. In the following, we propose a simple iteration method to solve these two optimization problems. In this algorithm, we first initialize the projection matrices U and V, then each projection matrix U (or V) can be iteratively solved by fixing the other projection matrix V (or U) in alternation. Specially, we first fix V and use the maximization problems (23), then U can be computed by solving the following generalized eigenvector problem:

$$\left(\beta Q_{vb} + (1-\beta) Q_{vw}\right) U = \lambda Q_w U \tag{25}$$

Let the column vector $U_1, U_2, \cdots, U_d$ be the solution to (25) according to their eigenvalues $\lambda_1 > \lambda_2 \cdots > \lambda_d$, thus the projection matrix $U = \left(U_1, U_2, \cdots, U_d\right)$.

Once U is obtained, by using the maximization problems (24), V can be updated by solving the following generalized eigenvector problem:

$$\left(\beta Q_{ub} + (1-\beta) Q_{uw}\right) V = \lambda P_w V \tag{26}$$

Let the column vector $V_1, V_2, \cdots, V_d$ be the solution to (26) according to their eigenvalues $\lambda_1 > \lambda_2 \cdots > \lambda_d$, thus the projection matrix $V = \left(V_1, V_2, \cdots, V_d\right)$.

Thus, the optimal projection matrices U and V can be obtained by iteratively computing the generalized eigenvector problems of (25) and (26). In our experiments, matrix V is initially set to the identity matrix.

Once the final optimal projection matrices U and V of tensor MMP is obtained, for a new testing X, its lower-dimensional embedding can computed according to

$$X \rightarrow Y = U^T X V \tag{27}$$

where Y is a lower feature representation of the face image X.

After the transformation by two-dimensional (tensor) MMP, a feature matrix is obtained for each image. Then, the face recognition becomes a pattern classification task, and different pattern classifier can be applied for face recognition. In this paper, we apply the nearest-neighbor classifier for its simplicity, and the Euclidean metric is used as our distance measure.

Journal of Theoretical and Applied Information Technology
10th January 2013. Vol. 47 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. EXPERIMENTAL RESULTS

In this section, we investigate the performance of our proposed two-dimensional MMP (2DMMP) algorithm for face recognition. The system performance is compared with the two-dimensional PCA (2DPCA) [10], two-dimensional LDA (2DLDA) [11], two-dimensional LPP (2DLPP) [12] and the original MMP [9] algorithms, where 2DPCA, 2DLDA, and 2DLPP are three of the most tensor dimensionality reduction algorithms in face recognition. We use the same graph structures in the MMP and 2DMMP algorithms. The settings of other algorithms are identical to the description in the corresponding papers.

In the following experiments, three face databases were tested for face recognition: the Yale database, the Olivetti Research Laboratory (ORL), and the PIE (pose, illumination, and expression) database from CMU. In all the experiments, preprocessing to locate the faces was applied. Original images were normalized (in scale and orientation) such that the two eyes were aligned at the same position. Then, the facial areas were cropped into the final images for recognition. In addition, in order to reduce the influence of some extreme illumination, histogram equalization is also done as preprocessing. Some sample images after preprocessing of the three databases are shown in Figure 1 to Figure 3, respectively.

Figure 1: Face Image Examples From The Yale Database

Figure 2: Face Image Examples From The Orl Database

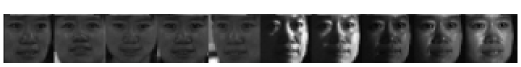

Figure 3: Face Image Examples From The Cmu Pie Database

In short, to perform face recognition, we first obtain the face subspaces by different dimensionality reduction algorithms. Then, facial images are projected into the face subspaces. Finally, the nearest-neighbor classifier is applied to recognize different facial images in the reduced feature spaces, where the Euclidean metric is used as the distance measure.

The Yale face database (http://cvc.yale.edu/projects/yalefaces/yalefaces.html) was constructed at the Yale Center for Computational Vision and Control. It contains 165 gray scale images of 15 individuals. The images demonstrate variations in lighting condition, facial expression (normal, happy, sad, sleepy, surprised, and wink). We randomly select five images of each individual to construct the training set and the rest images of the database to form the testing set. Thus, the numbers of the training samples and testing samples are 75 and 90, respectively. For each evaluation, 10 rounds of experiments are repeated with random selection of the training data, and the average result is reported as final recognition accuracy. In general, the performance of all dimensionality reduction algorithms varies with the number of reduced dimensions. We show the best recognition accuracies and the corresponding optimal dimensionality obtained by 2DPCA, 2DLDA, 2DLPP, MMP, and 2DMMP algorithms in Table 1. Figure 4 shows the plots of recognition accuracy versus reduced dimensionality on the Yale face database. It can be found that our proposed 2DMMP outperforms other algorithms in terms of recognition accuracy.

Table 1: Recognition Accuracy Comparisons On The Yale Database

Algorithm	Accuracy	Dimensionality
2DPCA	79.3%	20×20
2DLDA	84.6%	15×15
2DLPP	88.5%	15×15
MMP	92.4%	50
2DMMP	97.7%	10×10

Table 2: Recognition Accuracy Comparisons On The ORL Database

Algorithm	Accuracy	Dimensionality
2DPCA	88.2%	15×15
2DLDA	93.8%	10×10
2DLPP	95.7%	10×10
MMP	96.9%	80
2DMMP	98.3%	10×10

Journal of Theoretical and Applied Information Technology
10th January 2013. Vol. 47 No.1

ISSN: **1992-8645** www.jatit.org E-ISSN: **1817-3195**

The ORL database (http://www.uk.research.att.com/facedatabase.html) contains 400 images grouped into 40 distinct subjects with ten different images for each. These images were captured at different times, and for some subjects, the images may vary in facial expressions and facial details. All the images were taken against a dark homogeneous background with the tolerance for some side movement of about 20° . In this experiment, a random subset with five images per individual was chosen to form training set, and the rest of the database was considered to be the testing set. Thus, the numbers of the training samples and testing samples are 200 and 200, respectively. Likewise, we average the results over 10 random splits of the database as final recognition accuracy. The experimental protocols are set as the same with those applied in the Yale database. Table 2 lists the recognition results. Figure 5 shows the plots of recognition accuracy versus reduced dimensionality on the ORL face database, so we can observe that our proposed 2DMMP algorithms has the best performance.

Table 3: Recognition Accuracy Comparisons On The CMU PIE Database

Algorithm	Accuracy	Dimensionality
2DPCA	84.6%	20×20
2DLDA	95.1%	15×15
2DLPP	95.8%	15×15
MMP	96.2%	100
2DMMP	97.9%	15×15

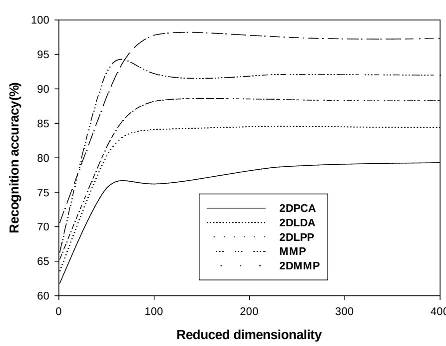

Figure 4: Recognition Accuracy Versus Reduced Dimensionality On The Yale Database

The CMU PIE face database contains 68 subjects with 41368 face images as a whole [13]. The face images were captured by 13 synchronized cameras and 21 flashes, under varying pose, illumination, and expression. We used 170 face images for each individual in our experiment, 85 for training and the other 85 for testing. Thus, the numbers of the training samples and testing samples are 5780 and 5780, respectively. Likewise, we average the results over 10 random splits of the database as final recognition accuracy. Table 3 shows the recognition results. Figure 6 shows the plots of recognition accuracy versus reduced dimensionality on the CMU PIE face database. As can be seen, our proposed 2DMMP performs much better than other algorithms.

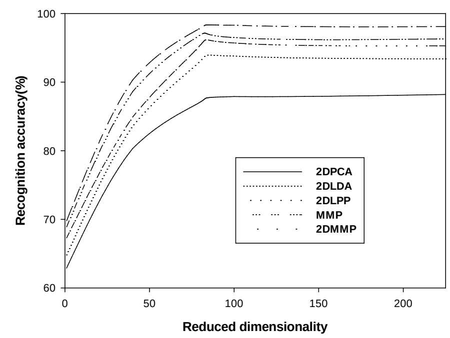

Figure 5: Recognition Accuracy Versus Reduced Dimensionality On The Orl Database

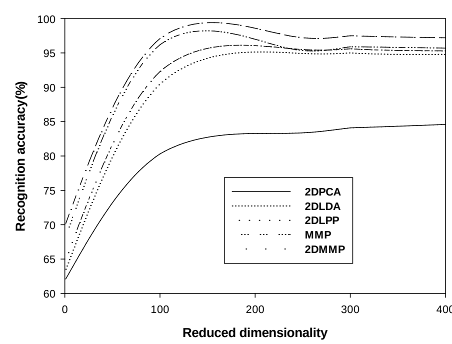

Figure 6: Recognition Accuracy Versus Reduced Dimensionality On The Cmu Pie Database

In summary, three experiments on three face databases have been systematically performed. These experimental results reveal the following observations:

1) Our proposed 2DMMP algorithm consistently outperforms 2DPCA, 2DLDA, 2DLPP, and MMP algorithms, which shows that the tensor extension can effectively improve the recognition performance of 2DMMP algorithm.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2) The performance of 2DPCA is the worst among the compared algorithms. The possible explanation is that 2DPCA is unsupervised and neglects the valuable discriminative information.

3) 2DLPP performs better than 2DPCA and 2DLD.The main reason could be attributed to the fact that both 2DPCA and 2DLDA see only the Euclidean structure, while failing to discover the underlying manifold structure.

4) Although MMP performs much better than LPP by simultaneously using local manifold structure and discriminative information, it still performs worse than our proposed 2DMMP algorithm, which demonstrates the importance of utilizing both local manifold structure and discriminative information, as well as characterizing the correlations among pixels with the tensor structure. Meanwhile, it further proves that the face image-as-tensor representation can lead to better recognition performance for face recognition.

5. CONCLUSION

In this paper, we have proposed a novel two-dimensional maximum margin projection (2DMMP) algorithm for face recognition. It combines the discriminative manifold preserving power of maximum margin projection (MMP) and tensor structure representation of face image to provide an effective method for face recognition. The experimental results on three face databases show that the proposed 2DMMP algorithm outperforms other related algorithms in terms of recognition accuracy.

ACKNOWLEDGMENTS

This work is supported by NSFC (Grant No. 70701013), the National Science Foundation for Post-doctoral Scientists of China (Grant No. 2011M500035), and the Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No.20110023110002).

REFERENCES:

[1] W.Zhao, R. Chellappa, P. Phillips, and A. Rosenfeld, "Face recognition: a literature survey", *ACM Computing Surveys*, Vol.35, No.4, 2003, pp.399-458.

[2] H.Wang, S.Chen, Z.Hu, and W.Zheng, "Locality-preserved maximum information projection", *IEEE Transactions on Neural Networks*, Vol.19, No.4, 2008, pp.571-585.

[3] P.Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.19, No.7, 1997, pp.711-720.

[4] X.He, S.Yan, Y.Hu, P.Niyogi, and H.-J. Zhang, "Face recognition using Laplacianfaces", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.27, No.3, 2005, pp.328-340.

[5] J.Tenenbaum, V.Silva, and J. Langford, "A global geometric framework for nonlinear dimensionality reduction", *Science*, Vol.290, No.5500, 2000, pp.2319-2323.

[6] S.Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding", *Science*, Vol.290, No.5500, 2000, pp.2323-2326.

[7] M.Belkin and P.Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation", *Neural Computation*, Vol.15, No.6, 2003, pp.1373-1396.

[8] D.Cai, X.He, J.Han, and H.J.Zhang, "Orthogonal Laplacianfaces for face recognition", *IEEE Transactions on Image Processing*, Vol.15, No.11, 2006, pp.3608-3614.

[9] X.He, D.Cai, and J.Han, "Learning a maximum margin subspace for image retrieval", *IEEE Transactions on Knowledge and Data Engineering*, Vol.20, No.2, 2008, pp.189-201.

[10]Yang, D.Zhang, A.F.Frangi, and J.Y.Yang, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.26, No.4, 2004, pp.131-137.

[11] J.Ye, R.Janardan, and Q.Li, "Two-dimensional linear discriminant analysis", *Neural Information Processing Systems*, 2005, pp. 1569-1576.

[12] X.He, D.Cai, and P.Niyogi, "Tensor subspace analysis", *Advances in Neural Information Processing Systems*, 2005, pp.1-8.

[13] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illuminlation, and expression database", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.25, No.12, 2003, pp.1615-1618.