Two-stage Tensor Locality-Preserving Projection Face Recognition*

Ying Liu and Dimitris A. Pados†
Dept. of Electrical Engineering
The State University of New York at Buffalo
Buffalo, NY 14260, USA
Email: {yl72, pados}@buffalo.edu

Chia-Hung Yeh
Dept. of Electrical Engineering
National Sun Yat-sen University
Kaohsiung 80424, Taiwan
Email: yeh@mail.ee.nsysu.edu.tw

***Abstract*—Locality-preserving projection (LPP) is an efficient dimensionality reduction approach that preserves local relationships within data sets and uncovers essential manifold structures. In this paper, we develop a two-stage tensor locality-preserving projection for face recognition, in which first-stage tensor LPP is performed in the original tensor space of face images and second-stage tensor LPP is performed in the reduced-dimension tensor subspace of the first-stage projection. For classification, we seek a non-negative sparse representation in the final low-dimensional tensor subspace and determine the class of an unknown face image by minimum sparse representation error. Experimental studies demonstrate that our proposed two-stage tensor LPP scheme along with the non-negative sparse representation classifier effectively exploits the locality structure of face images and outperforms existing state-of-the-art face recognition schemes.**

***Index Terms*—Classification, face recognition, locality-preserving projection, non-negative sparse representation, tensor subspace.**

I. Introduction

Face recognition from captured face images under varying conditions (expression, illuminance, and occlusions) has always been a challenging task. The challenges become accentuated in the era of big data. Many face recognition procedures were developed over the past few decades. One of the most successful and well-studied techniques for face recognition is arguably the appearance-based method [1], [2]. In appearance-based methods, we usually represent an image of size $m \times n$ pixels by a vector in a D-dimensional space, $D = mn$. The resulting D-dimensional space is frequently too large to allow fast and effective face recognition. A common way to resolve this problem is to use dimensionality reduction techniques. Two of the most popular techniques for this purpose are principal component analysis (PCA) [1] and linear discriminant analysis (LDA) [3]. PCA is an unsupervised scheme that seeks a linear transform matrix that maximizes the data variance in the projection subspace. LDA is a supervised scheme that aims at minimizing the within-class variances as well as maximizing the between-class distances in the projection subspace. While PCA and LDA are methods that preserve the global structure of the underlying high dimensional face images, the locality-preserving projection (LPP) method [4], [5] preserves also local relationships of the data and uncovers their essential manifold structure. LPP computes a score to measure the similarity between each pair of face images; the higher the score, the closer the two face images. LPP then seeks a linear transformation that minimizes the sum of the pairwise image Euclidean distances in the projection domain. The distances are weighted by their corresponding similarity scores. In short, LPP aims at pulling similar face images close together in the projection space.

All of the face recognition theory and practice described above considers an $m \times n$ image as a high dimensional vector in $\mathbb{R}^{mn}$, while an image represented on the plane is intrinsically a matrix. In this direction, 2-dimensional extensions of PCA and LPP, i.e. 2DPCA [6] and 2DLPP [7], were developed. To better protect the geometry of the images, tensor subspace analysis (TSA) [8] was proposed in which an image is considered as the second-order tensor in $\mathbb{R}^m \otimes \mathbb{R}^n$ where $\mathbb{R}^m$ and $\mathbb{R}^n$ are two vector spaces. Examples of TSA algorithms are the bilateral 2D-PCA (B2DPCA) [9], tensor LDA [10], and tensor LPP [11] procedures, which essentially seek two projection matrices, row-wise and column-wise, respectively, that optimize objective functions similar to their one-dimensional (1D) counterparts in PCA [1], LDA [3], and LPP [4], [5].

After the projection matrix or matrices are trained with a set of training face images, proper classifiers can work in the projection domain for face recognition. A commonly used classifier is the nearest-neighbor classifier (NNC) that finds in the projection space the training sample(s) closest to the testing sample in Euclidean distance. The tested sample is then declared to have the same class label as the "closest" training sample(s).

In this paper, we utilize the advantages of TSA and propose a two-stage tensor LPP method for face image dimensionality reduction, which can more effectively capture the local structure of face images than single-stage tensor LPP. Then, instead of classic nearest-neighbor classification (NNC), we propose a non-negative sparse representation classifier (NNSRC) that performs classification in the obtained tensor subspace. The proposed classifier is an improved version of sparse represen-

*This work was supported in part by the National Science Foundation under Grant ECCS-1462341.
†Corresponding author.

tation classification [12] in which we impose a non-negative constraint on the sparse coefficients that avoids information loss when negative sparse coefficients may occur.

The remainder of this paper is organized as follows. In Section II, we briefly review the background of vector-space and tensor-space LPP. In Section III, the proposed two-stage tensor LPP procedure is presented together with the new non-negative sparse representation classifier. Experimental results and performance analysis are presented in Section IV. Finally, a few conclusions are drawn in Section V.

II. Background on Locality-Preserving Projection

In locality-preserving projection (LPP) [4], given N training data samples $\mathbf{x}_i \in \mathbb{R}^D$ with known class labels ℓ_i, $i = 1, ..., N$, the similarity between two samples $\mathbf{x}_i$ and $\mathbf{x}_j$ is measured by

$$S_{ij} = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|_2^2/t) \tag{1}$$

with $t > 0$ is a tuning parameter. The objective of LPP is to find an optimal embedding of $\mathbf{x}_i$ in a lower dimensional subspace, say $\mathbf{y}_i \in \mathbb{R}^d$, $d < D$, that minimizes $\frac{1}{2}\sum_{i,j=1}^{N} \|\mathbf{y}_i - \mathbf{y}_j\|_2^2 S_{ij}$, which is the sum of the pairwise sample distances in the lower dimensional subspace weighted by the pairwise similarity S_{ij}.

If a linear transform matrix $\mathbf{W} \in \mathbb{R}^{D\times d}$ is used to explicitly perform the embedding by $\mathbf{y}_i = \mathbf{W}^{\mathrm{T}}\mathbf{x}_i$, then $\mathbf{W}$ can be optimized by

$$\min_{\mathbf{W}\in\mathbb{R}^{D\times d}} \frac{1}{2}\sum_{i,j=1}^{N} \|\mathbf{W}^{\mathrm{T}}(\mathbf{x}_i - \mathbf{x}_j)\|_2^2 S_{ij}. \tag{2}$$

The solution for the columns of $\mathbf{W}$ in (2), $\mathbf{w}_1$, $\mathbf{w}_2$, ..., $\mathbf{w}_d$, is given by the d eigenvectors corresponding to the d minimum eigenvalues of the following generalized eigenvalue problem,

$$\mathbf{XLX}^{\mathrm{T}}\mathbf{w}_i = \lambda_i \mathbf{XDX}^{\mathrm{T}}\mathbf{w}_i, \quad i = 1, ..., d, \tag{3}$$

where $\mathbf{X} \triangleq [\mathbf{x}_1, ..., \mathbf{x}_N]$, $\mathbf{D} \in \mathbb{R}^{N\times N}$ is a diagonal matrix with diagonal elements $D_{ii} = \sum_{j=1}^{N} S_{ij}$, $i = 1, ..., N$, $\mathbf{L} = \mathbf{D} - \mathbf{S}$ is the Laplacian matrix, and λ_i is the eigenvalue associated with the eigenvector $\mathbf{w}_i$.

To classify an unknown test sample $\mathbf{x}$, $\mathbf{x}$ is first transformed to the d-dimensional vector $\mathbf{y} = \mathbf{W}^{\mathrm{T}}\mathbf{x}$. Then, if NNC is adopted, the Euclidean distances between $\mathbf{y}$ and all embedded training samples $\mathbf{y}_i$, $i = 1, ..., N$ are compared and classification is performed by

$$\widehat{i} = \arg\min_{1\le i\le N} \|\mathbf{y} - \mathbf{y}_i\|_2, \tag{4}$$

$$\widehat{\ell} = \ell_i, \tag{5}$$

which means that the class label $\widehat{\ell}$ assigned to $\mathbf{x}$ is the same as that of the training sample $\mathbf{x}_i$ that minimizes $\|\mathbf{y} - \mathbf{y}_i\|_2$.

To generalize now this approach to powerful tensor geometry [11], assume $\mathbf{X}_i \in \mathbb{R}^{m\times n}$ is a training face image, $i = 1, ..., N$. Tensor LPP refers to seeking two projection matrices $\mathbf{U} \in \mathbb{R}^{m\times d}$ and $\mathbf{V} \in \mathbb{R}^{n\times d}$ that minimize the sum of pairwise distances in the reduced-dimension tensor subspace as follows:

$$\min_{\mathbf{U}\in\mathbb{R}^{m\times d},\mathbf{V}\in\mathbb{R}^{n\times d}} \frac{1}{2}\sum_{i,j=1}^{N} \|\mathbf{U}^T\mathbf{X}_i\mathbf{V} - \mathbf{U}^T\mathbf{X}_j\mathbf{V}\|^2 S_{ij}. \tag{6}$$

The optimal solutions for $\mathbf{U}$ and $\mathbf{V}$ in (6) cannot be obtained simultaneously. Nevertheless, a sub-optimal solution can be obtained by iteratively solving for one variable $\mathbf{U}$ ($\mathbf{V}$) while the other one $\mathbf{V}$ ($\mathbf{U}$) is fixed. When $\mathbf{U}$ is fixed, (6) is reduced to

$$\min_{\mathbf{V}\in\mathbb{R}^{n\times d}} \mathrm{tr}\{\mathbf{V}^T(\mathbf{D}_u - \mathbf{S}_u)\mathbf{V}\} \tag{7}$$

where

$$\mathbf{D}_u \triangleq \sum_{i=1}^{N} D_{ii}\mathbf{X}_i^T\mathbf{U}\mathbf{U}^T\mathbf{X}_i \tag{8}$$

and

$$\mathbf{S}_u \triangleq \sum_{i,j=1}^{N} S_{ij}\mathbf{X}_i^T\mathbf{U}\mathbf{U}^T\mathbf{X}_j. \tag{9}$$

When $\mathbf{V}$ is fixed, (6) is reduced to

$$\min_{\mathbf{U}\in\mathbb{R}^{m\times d}} \mathrm{tr}\{\mathbf{U}^T(\mathbf{D}_v - \mathbf{S}_v)\mathbf{U}\} \tag{10}$$

where

$$\mathbf{D}_v \triangleq \sum_{i=1}^{N} D_{ii}\mathbf{X}_i\mathbf{V}\mathbf{V}^T\mathbf{X}_i^T \tag{11}$$

and

$$\mathbf{S}_v \triangleq \sum_{i,j=1}^{N} S_{ij}\mathbf{X}_i\mathbf{V}\mathbf{V}^T\mathbf{X}_j^T. \tag{12}$$

Tensor LPP also aims at maximizing the global variance in the tensor subspace

$$\begin{aligned} \mathrm{var}(\mathbf{Y}) &= \sum_{i=1}^{N} \|\mathbf{U}^T\mathbf{X}_i\mathbf{V}\|^2 D_{ii} \\ &= \mathrm{tr}\{\mathbf{V}^T\mathbf{D}_u\mathbf{V}\} \\ &= \mathrm{tr}\{\mathbf{U}^T\mathbf{D}_v\mathbf{U}\}. \end{aligned} \tag{13}$$

Therefore, the final optimization procedure is to iteratively solve the following two problems: (i) when $\mathbf{V}$ is fixed, solve

$$\min_{\mathbf{U}\in\mathbb{R}^{m\times d}} \frac{\mathrm{tr}\{\mathbf{U}^T(\mathbf{D}_v - \mathbf{S}_v)\mathbf{U}\}}{\mathrm{tr}\{\mathbf{U}^T\mathbf{D}_v\mathbf{U}\}}, \tag{14}$$

and (ii) when $\mathbf{U}$ is fixed, solve

$$\min_{\mathbf{V}\in\mathbb{R}^{n\times d}} \frac{\mathrm{tr}\{\mathbf{V}^T(\mathbf{D}_u - \mathbf{S}_u)\mathbf{V}\}}{\mathrm{tr}\{\mathbf{V}^T\mathbf{D}_u\mathbf{V}\}}. \tag{15}$$

In fact, problems (14) and (15) are equivalent to

$$\max_{\mathbf{U}\in\mathbb{R}^{m\times d}} \frac{\mathrm{tr}\{\mathbf{U}^T\mathbf{S}_v\mathbf{U}\}}{\mathrm{tr}\{\mathbf{U}^T\mathbf{D}_v\mathbf{U}\}}, \tag{16}$$

and

$$\max_{\mathbf{V}\in\mathbb{R}^{n\times d}} \frac{\mathrm{tr}\{\mathbf{V}^T\mathbf{S}_u\mathbf{V}\}}{\mathrm{tr}\{\mathbf{V}^T\mathbf{D}_u\mathbf{V}\}}. \tag{17}$$

While (16) can be computed by solving the generalized eigenvalue problem $(\mathbf{D}_v - \mathbf{S}_v)\mathbf{u} = \lambda \mathbf{D}_v \mathbf{u}$, (17) can be computed by solving the generalized eigenvalue problem $(\mathbf{D}_u - \mathbf{S}_u)\mathbf{v} = \lambda \mathbf{D}_u \mathbf{v}$. For classification, NNC can be applied in the tensor subspace.

Although tensor LPP as reviewed above exploits the 2D nature of face images, one-shot similarity values S_{ij} calculated on poor training images may adversely affect the representation of local data relationships. In addition, simple NNC does not effectively utilize the data representation in the tensor subspace. In the following section, we propose a new two-stage tensor LPP scheme followed by a new stronger classifier based on non-negative sparse representation.

III. The Proposed Procedure

A. A New Two-stage Tensor LPP Scheme

We propose to train two sets of projection matrices, $\mathbf{U}_1 \in \mathbb{R}^{m\times d_1}$, $\mathbf{V}_1 \in \mathbb{R}^{n\times d_1}$ and $\mathbf{U}_2 \in \mathbb{R}^{d_1\times d}$, $\mathbf{V}_2 \in \mathbb{R}^{d_1\times d}$ for dimensionality reduction, $d \le d_1 \le \min\{m,n\}$, among which $\mathbf{U}_1$ and $\mathbf{V}_1$ are the first-stage projection matrices that reduce the face image dimension from $m\times n$ to $d_1 \times d_1$ and $\mathbf{U}_2$ and $\mathbf{V}_2$ are the second-stage projection matrices that further reduce the signal dimension from $d_1\times d_1$ to $d\times d$. With a set of training images $\mathbf{X}_i$, $i = 1,...,N$, we first obtain the pairwise image similarity scores S_{ij} by (1) and compute the first-stage tensor LPP projection matrices $\mathbf{U}_1$ and $\mathbf{V}_1$ by (16) and (17). The resulting first-stage tensor subspace representation of training samples $\mathbf{X}_i$ is then computed as $\mathbf{Y}_i^{(1)} = \mathbf{U}_1^T \mathbf{X}_i \mathbf{V}_1 \in \mathbb{R}^{d_1\times d_1}$, $i = 1,...,N$. We then perform a second-stage tensor LPP in the computed $\mathbb{R}^{d_1} \otimes \mathbb{R}^{d_1}$ tensor subspace, in which $\mathbf{Y}_i^{(1)}$, $i = 1,...,N$ serve as the data samples. We compute the second-stage similarity scores $S_{ij}^{(2)}$ using the pairwise distance of $\mathbf{Y}_i^{(1)}$ and $\mathbf{Y}_j^{(1)}$, i.e.

$$S_{ij}^{(2)} = \exp(-\|\mathbf{Y}_i^{(1)} - \mathbf{Y}_j^{(1)}\|_2^2/t). \tag{18}$$

Accordingly, the second-stage diagonal degree matrix $\mathbf{D}^{(2)}$ is obtained by $D_{ii}^{(2)} = \sum_j S_{ij}^{(2)}$ and the second-stage Laplacian matrix is obtained by $\mathbf{L}^{(2)} = \mathbf{D}^{(2)} - \mathbf{S}^{(2)}$. Afterwards, the second-stage tensor LPP projection matrices $\mathbf{U}_2$ and $\mathbf{V}_2$ can be computed by iteratively solving for

$$\max_{\mathbf{U}_2\in\mathbb{R}^{d_1\times d}} \frac{\mathrm{tr}\{\mathbf{U}_2^T \mathbf{S}_v^{(2)} \mathbf{U}_2\}}{\mathrm{tr}\{\mathbf{U}_2^T \mathbf{D}_v^{(2)} \mathbf{U}_2\}} \tag{19}$$

when $\mathbf{V}_2$ is fixed, and

$$\max_{\mathbf{V}_2\in\mathbb{R}^{d_1\times d}} \frac{\mathrm{tr}\{\mathbf{V}_2^T \mathbf{S}_u^{(2)} \mathbf{V}_2\}}{\mathrm{tr}\{\mathbf{V}_2^T \mathbf{D}_u^{(2)} \mathbf{V}_2\}} \tag{20}$$

when $\mathbf{U}_2$ is fixed where

$$\mathbf{D}_u^{(2)} \triangleq \sum_{i=1}^N D_{ii}^{(2)} \mathbf{Y}_i^{(1)\,T} \mathbf{U}_2 \mathbf{U}_2^T \mathbf{Y}_i^{(1)}, \tag{21}$$

$$\mathbf{S}_u^{(2)} \triangleq \sum_{i,j=1}^N S_{ij}^{(2)} \mathbf{Y}_i^{(1)\,T} \mathbf{U}_2 \mathbf{U}_2^T \mathbf{Y}_j^{(1)}, \tag{22}$$

$$\mathbf{D}_v^{(2)} \triangleq \sum_{i=1}^N D_{ii}^{(2)} \mathbf{Y}_i^{(1)} \mathbf{V}_2 \mathbf{V}_2^T \mathbf{Y}_i^{(1)\,T}, \tag{23}$$

and

$$\mathbf{S}_v^{(2)} \triangleq \sum_{i,j=1}^N S_{ij}^{(2)} \mathbf{Y}_i^{(1)} \mathbf{V}_2 \mathbf{V}_2^T \mathbf{Y}_j^{(1)\,T}. \tag{24}$$

Then, the second-stage tensor subspace representation for the training images $\mathbf{X}_i$ is $\mathbf{Y}_i^{(2)} = \mathbf{U}_2^T(\mathbf{U}_1^T \mathbf{X}_i \mathbf{V}_1)\mathbf{V}_2 \in \mathbb{R}^{d\times d}$, $i = 1,...,N$.

B. Non-negative Sparse Representation Classifier in Tensor Subspace

In this section, we develop a novel non-negative sparse representation classifier (NNSRC) that operates in the computed tensor subspace. Given a test image $\mathbf{X}$, we compute its second-stage tensor subspace representation

$$\mathbf{Y}^{(2)} = \mathbf{U}_2^T \times \mathbf{U}_1^T \times \mathbf{X} \times \mathbf{V}_1 \times \mathbf{V}_2 \tag{25}$$

and vectorize $\mathbf{Y}^{(2)}$ by column concatenation

$$\mathbf{y}^{(2)} = \mathrm{vec}(\mathbf{Y}^{(2)}). \tag{26}$$

For each training sample $\mathbf{X}_i$ that has a known class label ℓ_i, $i = 1,...,N$, we perform the same second-stage tensor subspace projection and obtain the resulting subspace representation $\mathbf{y}_i^{(2)} = \mathrm{vec}(\mathbf{Y}_i^{(2)})$. We define the data representation matrix $\mathbf{A}^{(2)} \triangleq [\mathbf{y}_1^{(2)} \cdots \mathbf{y}_N^{(2)}] \in \mathbb{R}^{d^2\times N}$ and calculate for the test vector $\mathbf{y}^{(2)}$

$$\alpha^{(2)} = \arg\min_{\alpha\in\mathbb{R}^N} \frac{1}{2}\|\mathbf{A}^{(2)}\alpha - \mathbf{y}^{(2)}\|^2 + \lambda\|\alpha\|_1, \quad \text{s.t.} \quad \alpha \succeq \mathbf{0}. \tag{27}$$

Let $\alpha_c^{(2)} \in \mathbb{R}^N$ be such that $\alpha_c^{(2)}[i] = \alpha^{(2)}[i]$ if $\ell_i = c$ and $\alpha_c^{(2)}[i] = 0$ otherwise. Then, the proposed classification rule is

$$\hat{\ell} = \arg\min_c \|\mathbf{A}^{(2)}\alpha_c^{(2)} - \mathbf{y}^{(2)}\|^2. \tag{28}$$

IV. Experimental Results and Performance Analysis

In this section, we carry out experimental studies on two face data sets to demonstrate the efficacy of our proposed algorithm. We compare our algorithm with the familiar state-of-the-art LDA, LPP, and (single-stage) tensor LPP procedures, all with nearest-neighbor classification.

For the first experiment we utilize the Cropped Yale data set [13], in which the face images have different illuminance conditions and are manually aligned, cropped, and re-sized to 168×192 images. We select a subset that has 8 subjects (classes) and each subject has 64 face images. We down-sample the images to dimension 48×42 and randomly select 6 images per subject as the training data set. The remaining 58 images per subject are used for testing. For both the training and testing data sets, 25% randomly selected images per subject are corrupted by outliers. The outlier pattern is a salt-and-pepper square noise patch with size varying from 15×15 to 30×30. Examples of the training images are shown in Fig. 1. To examine our proposed method, we train a two-stage tensor LPP, in which the first stage projection reduces the image dimension to $d_1 \times d_1 = 30 \times 30$ and the second stage projection further reduces the image dimension from $d_1 \times d_1$ to $d \times d, 1 \leq d \leq 15$.

We compare the face recognition error rates of the following methods: (i) LDA; (ii) LPP; (iii) (single-stage) tensor LPP; (iv) our proposed two-stage tensor LPP scheme with NNSRC (schemes (i), (ii), and (iii) execute conventional nearest-neighbor classification). It can be observed in Fig. 2 that tensor-space dimensionality reduction is more effective than vector-space dimensionality reduction. The proposed two-stage tensor LPP with NNSRC offers significant error rate improvement reaching the lowest error rate of 12.7% when the second stage tensor representation is of size $d \times d = 12 \times 12$.

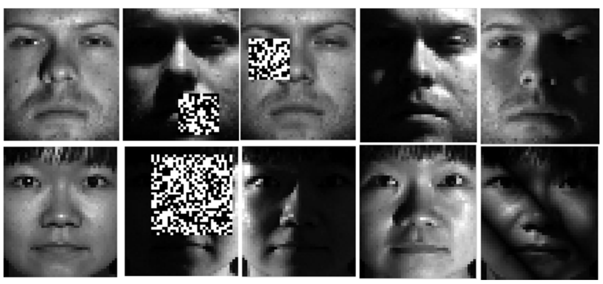

Fig. 1. Cropped Yale data set sample images.

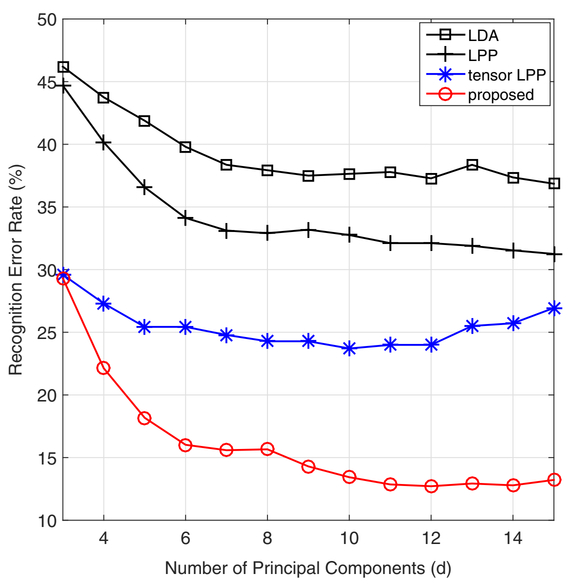

Fig. 2. Recognition error rate of the Cropped Yale data set.

For added demonstration credibility, we experiment also on the AR data set [14], in which the face images have varying facial expressions, illumination conditions, and occlusions. We select a subset that contains 30 subjects (classes). Each subject has 26 face images. We down-sample the images to dimension

Fig. 3. AR data set sample images.

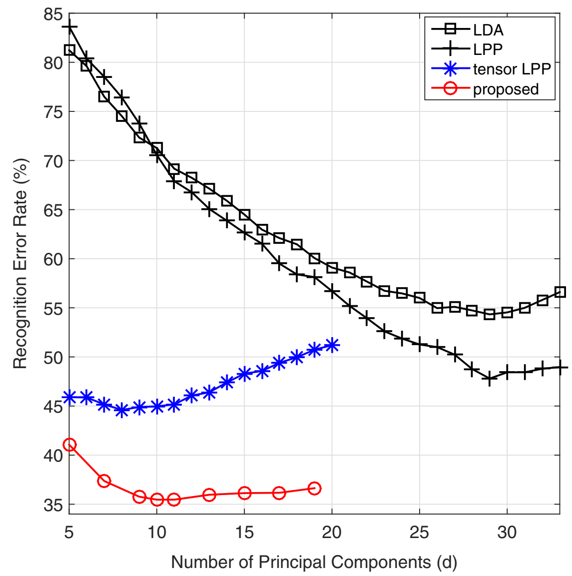

Fig. 4. Recognition error rate of the AR data set.

55×40, randomly select 2 images per subject to form the training data set, and use the remaining 24 images per subject for testing. 25% randomly chosen images per subject in the testing data set are corrupted by outliers. The outlier pattern is again salt-and-pepper square noise of size varying from 15×15 to 30×30. Fig. 3 shows image examples. We train a two-stage tensor LPP scheme in which the first stage projection reduces the data dimension to $d_1 \times d_1 = 30 \times 30$. Then, the second stage projection further reduces the image dimension to $d \times d, 5 \leq d \leq 19$.

In Fig. 4, as in Fig. 2, we compare the performance of the four face recognition methods LDA, LPP, tensor LPP, and proposed. Similar conclusions to Fig. 2 can be drawn that show the significant performance improvement offered by the proposed method.

V. Conclusions

We proposed a two-stage tensor locality-preserving projection method and a subsequent tensor-subspace non-negative sparse representation classifier for face recognition.

To effectively exploit the local structure of the face images, the developed two-stage tensor subspace embedding algorithm computes two sets of row-wise and column-wise projection matrices with an updated pairwise sample similarity measure. The overall row (column) projection matrices are then given by the product of the row (column) projection matrices at the first and second stages.

Then, we developed a novel classifier that seeks a non-negative sparse representation of the unknown sample in the reduced-dimension tensor subspace by using the tensor-subspace training samples as the dictionary. Classification is performed according to the minimum sparse representation error. Experimental studies demonstrated that the proposed two-stage tensor LPP scheme with non-negative sparse representation classification outperforms existing state-of-the-art face recognition methods.

References

[1] M. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Maui, HI, June 1991, pp. 586-591.

[2] H. Murase and S. K. Nayar, "Visual learning and recognition of 3-D objects from appearance," *Int. J. Computer Vision*, vol. 14, pp. 5-24, 1995.

[3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, pp. 711-720, July 1997.

[4] X. He and P. Niyogi, "Locality Preserving Projections," in *Proc. Conf. Advances in Neural Information Processing Systems (NIPS)*, Vancouver, British Columbia, Canada, Dec. 2003, pp. 153-160.

[5] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using Laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, pp. 328-340, Mar. 2005.

[6] 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.*, vol. 26, pp. 131-137, Jan. 2004.

[7] S. Chen, H. Zhao, M. Kong, and B. Luo, "2D-LPP: a two-dimensional extension of locality preserving projections," *Neurocomputing*, vol. 70, pp. 912-921, Jan. 2007.

[8] X. He, D. Cai, and P. Niyogi, "Tensor subspace analysis," in *Proc. Conf. Advances in Neural Information Processing Systems (NIPS)*, Vancouver, British Columbia, Canada, Dec. 2005, pp. 499-506.

[9] H. Kong, L. Wang, E. K. Teoh, X. Li, J-G Wang, and R. Venkateswarlu, "Generalized 2D principal component analysis for face image representation and recognition," *Neural Networks*, vol. 18, pp. 585-594, July-Aug. 2005.

[10] D. Cai, X. He, and J. Han, "Subspace learning based on tensor analysis," *Technical report*, Computer Science Department, UIUC, UIUCDCS-R-2005-2572, May 2005.

[11] X. He, D. Cai, H. Liu, and J. Han, "Image clustering with tensor representation," in *Proc. ACM Conf. on Multimedia*, Singapore, Nov. 2005, pp. 132-140.

[12] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, pp. 210-227, Feb. 2009.

[13] K. C. Lee, J. Ho, and D. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, pp. 684-698, May 2005.

[14] A. M. Martinez, and R. Benavente, "The AR face database," *Computer Vision Center Technical Report*, no. 24, June 1998, Purdue University, West Lafayette, IN.