Orthogonal Tensor Sparse Neighborhood Preserving Embedding for Two-dimensional Image

MINGMING QI
Tongji University
Department of Computer science and technology
Shanghai, 201804
Also: Yuanpei College, Shaoxing University
Shaoxing, Zhejiang 312000
P. R. China
Corresponding author: Webqmm1974@sohu.com

YANG XIANG
Tongji University
Department of Computer science and technology
Shanghai, 201804
P. R. China
xiangyang@tongji.edu.cn

Abstract: Orthogonal Tensor Neighborhood Preserving Embedding (OTNPE) is an efficient dimensionality reduction algorithm for two-dimensional images. However, insufficiencies of the robustness performance and deficiencies of supervised discriminant information are remained. Motivated by the sparse learning, an algorithm called Orthogonal Tensor Sparse Neighborhood Embedding (OTSNPE) for two-dimensional images is proposed in the paper. The algorithm firstly regards two-dimensional images as points in the second-order tensor space, then, the neighborhood reconstruction of samples within the same class is achieved with sparse reconstruction. Finally, projections are gotten to preserve local sparse reconstruction relation and neighborhood relation within the same class and spatial relation of pixels in an image. Experiments operated on Yale, YaleB and AR databases show, in contrast to the existing typical tensor dimensionality reduction algorithms, the algorithm can improve the accuracy rate of classification algorithms based on the shortest Euclidean distance.

Key–Words: Dimensionality Reduction, Tensor Image, Neighborhood Preserving Embedding, Sparse Reconstruction, Supervised Discriminant Information, Face Recognition.

1 Introduction

With rapid development on the information technology, there are plenty of high-dimensional data in practical applications of machine learning and pattern recognition, such as face images, gene representation data, and text data etc. It is expensive cost to do with these high-dimensional data directly and is easy to produce classification information redundancies, namely curse of dimensionality [1]. According to certain rule, dimensionality reduction transforms high-dimensional data into meaningful low-dimensional data [2]. Therefore, dimensionality reduction can eliminate the problem and is the main research topic of data mining.

Principal Component Analysis (PCA) [3] and Linear Discriminant Analysis (LDA) [4] are initial linear dimensionality reduction methods. PCA is a common dimensionality reduction method based on variable covariance matrix, which aims to find essential linear dimensionality by estimating statistics properties of data. As a supervised dimensionality reduction method, the purpose of LDA is projecting high-dimensional data into the most distinguishable vector space to exact classification information and reduce feature space dimensionality, which guarantees more between-class distance and less within-class distance of input samples in the new projected subspace. PCA and LDA are not successfully applied in nonlinear data because they are linear dimensionality reduction methods. Therefore, Kernel Principal Component Analysis(KPCA)[5] and Kernel Fisher Discriminant Analysis (KFDA) [6] are proposed and successfully applied in actual nonlinear data mining practice. However kernel functions and parameters are difficult to be selected.

To solve the problem of kernel versions, manifold learning-based dimensionality reduction algorithms are specially proposed in nonlinear data, exposing inherent structure hidden datasets and preserving local geometric characteristics that are embedded in a high-dimensional space. Representative algorithms include Locally Linear Embedding (LLE) [7], Isometric Feature Mapping (ISOMAP) [8], Laplacian Eigenmaps (LE) [9] , Hessian-based Locally Linear Embedding (HLLE)[10], Maximum Variance Unfolding (MVU) [11], Local Tangent Space Alignment (LTSA) [12,13], Riemannian Manifold Learning(RML) [14,15], and Local Spline Embedding (LSE)[16],etc. Each manifold learning algorithm attempts to preserve a different geometrical property of the underlying manifold. However, implicit mappings are defined training samples and fail to directly reflect new data points, which are created by these algorithms between high-dimensional space and low-dimensional space. In order to solve this problem, Locality Preserving Projections (LPP) [17,18] and Neighborhood Preserving Embedding (NPE) [19] have been proposed. As unsupervised dimensionality reduction algorithms, NPE aims at finding a low-dimensional embedding that optimally preserves the local neighborhood structure on the original data manifold. Though experiments on real face database have shown the effectiveness of the NPE algorithm, there are some disadvantages should be paid close attention to and extend versions of NPE have been proposed. For overcoming the out-of-sample problem, Orthogonal Neighborhood Preserving Projections (ONPP) [20] and Discriminative Orthogonal Neighborhood Preserving Projections (DONPP) [21] are proposed. Complete Neighborhood Preserving Embedding (CNPE) [22] transforms the singular generalized eigen-system computation into two eigenvalue decomposition problems.

Above algorithms adopt the matrix-to-vector way to gain features descriptions. The way transforms $n1 \times n2$ matrix into $n1 \times n2$ dimensional vector, which not only costs more computational time but also has great passive impact on the evaluation of the covariance matrix. In view of these, Yang et al. [23] proposed a two-dimensional PCA(2DPCA) algorithm, which enables us to directly use a feature input of 2D image matrices rather than 1D vectors. 2DPCA achieves promising results in contrast to the traditional PCA in terms of face recognition rate and training time. Motivated by 2DPCA, Li and Yuan [24] designed 2DLDA and Chen et al. [25] proposed 2DLPP algorithms. Two-dimensional Neighborhood Preserving Projection (2DNPP) [26] is proposed for appearance-based face representation and recognition.

However, dimensionality reduction algorithms based on two-dimensional are only limited on dimensionality reduction on the row or the rank, which fail to take into account the spatial relation of image pixels. Tensor-based dimensionality reduction algorithms represent sample data with second-order tensor and preserve local information in image pixels spaces, which can describe structure information in image. [27] not only extend traditional PCA and LDA into second-order tensor space but also make deep analysis of the reason for more performances. [28] extend traditional LPP into tensor LPP. [29] proposes an Uncorrelated Multi-linear Discriminant Analysis (UMLDA) framework for the recognition of multidimensional tensor objects. Dai and Yeung [30] proposed tensor NPE (TNPE). [31] proposed Orthogonal Tensor Neighborhood Preserving Embedding (OTNPE) extend the TNPE to an orthogonal version by our orthogonal tensor subspace model and apply it to the prototypic facial expressions recognition. [32] proposed to represent an image as a local descriptor tensor and used a Multi-linear Supervised Neighborhood Embedding (MSNE) for discriminant feature extraction for subject or scene recognition. However, these is an important problem remained in algorithms on NPE, namely, local approximation based on Euler distance is far from preserving local geometric structure when serious external disturbs happen on images.

Recently sparse learning is applied successfully in machine learning and pattern recognition, including object detection [33,35], classification [34,36] and so on. Sparse representation-based classification algorithms reconstruct one of training samples with rest of training samples. Recent researches [35,36] demonstrate that sparse representation-based classification algorithms are available of great performance of robustness on face images with external disturbance, including deformity, dressing and shelter.[37] is representative sparse learning algorithm, which proposed graphic l_1 construction method based on sparse represent to achieve the reconstruction of every training sample with training samples, ensuring relations among training samples. The theory of sparse learning has generalized into the field of dimensionality reduction. These algorithms include (VS-SSVM) [38], Sparse Principal Component Analysis (SPCA)[39], Sparsity Preserving Projections (SPP) [40].

Motivated by sparse learning, we propose a Orthogonal Tensor Sparsity Neighborhood Preserving Embedding (OTSNPE) algorithm for dimensionality reduction on two-dimensional images. The algorithm preserves local sparse reconstruction of samples within the same class on images in the process of dimensionality reduction. Data in the new projected data spaces not only preserve neighborhood geometric structure information of input data and spatial relation of pixels in images but also preserves supervised discriminant information based on class label. Experiments on AR, Yale and YaleB face databases demonstrate the effectiveness of the algorithm.

OTSNPE is available of excellencies as follows:

(1) To traditional NPE, the neighborhood parameter k or ε is required to set for selecting neighbourhood samples while the problem is free in OTSNPE.

(2) Sparse neighborhood reconstructions within the same class introduce supervised discriminant information based on class label.

(3) It is expensive to achieve sparse reconstruction of all training samples when the number of samples is larger, which do not happen on OTSNPE be-

cause of the number of the same class is limited.

(4) Powerful robustness in sparse reconstruction guarantees OTSNPE to adapt to image with various lighting conditions ,deformities, dressings and shelters.

The rest of this paper is organized as follows: In Section 2 we will introduce basic knowledge and related works. The objective function and steps of OTSNPE are given in Section 3. The experimental results for applying our method to face recognition will be presented in Section 4, followed by the conclusions in Section 5.

2 Basic knowledge and related works

In this section, we first introduce basic knowledge, including basic theory of tensor and tensor-based image representation in 2.1; and then in 2.2 we review related works on sparse representation and sparse reconstruction, NPE as well OTNPE.

2.1 Basic knowledge

2.1.1 Basic theory of tensor

Tensor is multiple linear maps of a serial vector spaces. Generally speaking, the tensor is regarded as the extend of the matrix. A vector is a first order tensor and a matrix is a second order tensor. If several matrices that have same dimensions are stacked into a array, the array becomes a third order tensor.

Let $\chi \in R^{N_1 \times N_2 \times \cdots \times N_M}$ be a tensor. The order of χ is M. One point of χ is defined as $\chi_{n_1, n_2, \cdots, n_M} (1 \leq n_i \leq N, 1 \leq i \leq M)$ and the size of the i-th dimension of is N_i. Fig.1 shows a third order tensor.

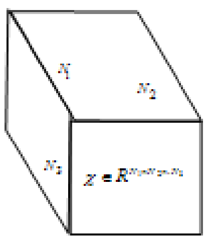


Figure 1: A third order tensor $\chi \in R^{N_1 \times N_2 \times N_3}$

Definition 1 *d-mode matrix unfolding of M order tensor $\chi \in R^{N_1 \times N_2 \times \cdots \times N_M}$, namely $X_d = R^{(\Pi_{i \neq d} N_i) \times N_d}$.*

Definition 2 *The scalar product of two tensor $\chi, \gamma \in R^{N_1 \times N_2 \times \cdots \times N_M}$ is defined as $< \chi, \gamma >= \Sigma_{i_1} \cdots \Sigma_{i_k} \chi_{i_1 \cdots i_k} \gamma^*_{i_1 \cdots i_k}$, where $*$ denotes complex conjugation.*

Definition 3 *The i-mode product of a tensor $\chi \in R^{N_1 \times N_2 \times \cdots \times N_M}$ and a matrix $U \in R^{N_i \times N'}$ is an $N_1 \times N_2 \times \cdots \times N_{i-1} \times N' \times N_{i+1} \times \cdots \times N_M$ tensor denoted as $\chi \times_d U$, namely,*

$$\chi \times_d U(N_1 \times N_2 \times \cdots \times N_{i-1} \times N' \times N_{i+1} \times \cdots \times N_M) = \Sigma_{i_d=1}^{N_i} \chi(i_1, \cdots, i_{d-1}, j, i_{d+1}, \cdots, i_M) U(i_d, j) \tag{1}$$

2.1.2 Image representation based on second-order tensor

A matrix image is intrinsically a second-order tensor. The relationship between the rows vectors of the matrix and the column vectors of the matrix might be important for finding projections, especially when the number of training samples is small.

Based on the algebra of higher order tensor, we introduce coordinate transformation in the tensor space. Let $\{u_{ki}\}_{k=1}^{n_1}$ be the standard orthogonal base of the tensor space R^{n_1} and $\{v_{li}\}_{l=1}^{n_2}$ be the standard orthogonal base of the tensor space R^{n_2}. Then

$$\begin{aligned} u_k = (u_{k1}, \cdots, u_{kn_1})^T = \sum_i u_{ki} \varepsilon_i \\ v_k = (v_{l1}, \cdots, v_{ln_2})^T = \sum_j v_{lj} \varepsilon_j \end{aligned} \tag{2}$$

According to Eq.(1) and Eq.(2), we can get as follows:

$$\varepsilon_i = \sum_{k=1}^{n_1} u_{ki} U_k, \quad \varepsilon_j = \sum_{l=1}^{n_2} u_{li} U_l \tag{3}$$

Therefore, an image is represented by new basis, namely,

$$\begin{aligned} T &= \sum_{ij} T_{ij} \varepsilon_i \bigotimes \varepsilon_j \\ &= \sum_{kl} (\sum_{ij} T_{ij} u_{ki} v_{tj}) U_K \bigotimes V_t \\ &= \sum_{kl} (U_K^T T V_T) U_K \bigotimes V_t \end{aligned} \tag{4}$$

An image of size $n_1 \times n_2$ pixels is naturally regarded as one point in the tensor space that is the inner product space of two vectors. The base of $u_i \times v_j$ is regard as tensor base of the image.

2.2 Related works

2.2.1 Sparse representation and sparse reconstruction

Given a set of training samples $X = \{x_1, x_2, x_3, \cdots, x_n\} \in R^{d \times n}$, sparse representation aims to reconstruct each sample x_i with else sample, using as few samples as possible, namely, seek a sparse reconstructive weight vector S_i for each x_i through the following minimization problem:

$$\begin{aligned} &\min_{S_i} \|S_i\|_0 \\ &s.t. \quad x_i = XS_i \end{aligned} \tag{5}$$

where S_{ij} denotes the contribution of each x_j to reconstructing x_i. $\|S_i\|_0$ is the pseudo- l_0 norm which is equal to the number of non-zero components in S. However, Eq.(5) is NP-hard. The solution of l_0 minimization problem is equal to the solution of l_1 minimization problem [33-35]. Therefore, this difficulty can be bypassed by transforming the problem and solving as follows:

$$\begin{aligned} &\min_{S_i} \|S_i\|_1 \\ &s.t. \quad x_i = XS_i \end{aligned} \tag{6}$$

However, in many practical applications, the signal X is generally noisy, Eq.(6) does not always hold. Two robust extensions can be used to handle this problem [39]: 1)$\|x_i - XS_i\| \prec \varepsilon$, where ε denotes an error tolerance. 2) replace X with $[X, I]$, where I denotes a n-order identity matrix.

Sparse reconstruction seeks a sparse reconstructive weight vector S_i for each x_i through the following modified l_1 minimization problem:

$$\begin{aligned} &\min_{S_i} \|S_i\|_1 \\ &s.t. \quad x_i = XS_i \\ &\quad 1 = 1^T s_i \end{aligned} \tag{7}$$

where $\|S_i\|_1$ denotes the l_1 normal of S_i, $S_i = [S_{i1}, \cdots, S_{ii-1}, 0, S_{ii+1}, \cdots, S_{in}]^T \in R^n$ is a vector in which S_{ij} denotes the contribution of each x_j to reconstructing x_i and $1 \in R^n$ is a vector of all ones.

$$\begin{aligned} x_i = S_{i1}x_1 + \cdots + S_{ii-1}x_{i-1} + S_{ii+1}x_{i+1} \\ + \cdots + S_{in}x_n \end{aligned} \tag{8}$$

The sparse reconstruction matrix $S = [S_1, S_2, \cdots, S_n]^T$ is attained through calculating S_i.

2.2.2 Neighborhood preserving embedding (NPE)

Given samples $X = \{x_1, x_2, x_3, \cdots, x_n\} \in R^{d\times n}$, NPE attempts to seek an optimal transformation matrix T to map high-dimensional data X into low-dimensional data $Y = T^T X$, in which the local neighborhood structure of X can be preserved. There are some basic steps of NPE as follows:

(1) Construct neighborhood adjacent graphic G. The adjacent graphic G is composed of N nodes. Node i corresponds to sample x_i . If sample x_j is the neighborhood of sample x_i, there is a line between x_j and x_i. Common methods for construct neighborhood adjacent graphic G are k-neighborhood and ε-neighborhood.

(2) Calculate the sparse reconstructive weights. According to G, each sample in training samples is reconstructed through the linear combination of neighborhood nodes of the sample as follows

$$\begin{aligned} &\min_T \sum_i \| x_i - \sum_{j=1}^{k} w_{ij}x_j \|^2 \\ &s.t. \sum_{j=1}^{k} w_{ij} = 1 \end{aligned} \tag{9}$$

(3) projected low-dimensional data Y satisfy:

$$\begin{aligned} &\min_T \sum_i \| y_i - \sum_{j=1}^{k} w_{ij}y_j \|^2 \\ &= \min_T \| Y(I-W) \|^2 \\ &= \min_T (Y(I-W)(I-W)^T Y^T) \\ &= \min_T (TX(I-W)(I-W)^T XT^T) \\ &= \min_T (TXMXT^T) \end{aligned} \tag{10}$$

where $M = (I-W)(I-W)^T$, I represents an identity matrix. In order to make the optimization problem well-posed, constrain conditions are introduced as follows:

$$\begin{aligned} &\sum_{i=1}^{N} y_i = 0 \\ &\tfrac{1}{N-1} \sum_{i=1}^{N} y_i^T y_i = I \end{aligned} \tag{11}$$

According to $Y = T^T X$, we replace Y with $T^T X$ in Eq. (10) and Eq. (11) and get the objective function:

$$\begin{aligned} &\min_T T^T XMX^T T \\ &T^T XX^T T = I \end{aligned} \tag{12}$$

2.2.3 Orthogonal Tensor neighborhood preserving embedding (OTNPE)

OTNPE can process directly data based on high-order tensors without unfolding them to vectors. Given n original data points $x_1, \cdots, x_n$ in tensor space $R^{I_1\times I_2\times\cdots\times I_k}$. In order to preserve the local structure explicitly, OTNPE is to find k transformation matrices $U^i \in R^{m_i \times m_i'}(m_i > m_i', i = 1, \cdots, k)$ to achieve projections on the basic of TNPE, which is defined by the following objective function based on the Frobenius norm of a tensor:

$$\begin{aligned} &\min \sum_i \| x_i \times_1 U^1 \cdots \times_k U^k \\ &- \sum_j M_{i,j} x_j \times_1 U^1 \cdots \times_k U^k \|_F^2 \\ &s.t. \ \sum_i \| x_i \times_1 U^1 \cdots \times_k U^k \|_F^2 = 1 \\ &\quad (U^m)^T (U^m) = I_m (m = 1 \cdots k) \end{aligned} \tag{13}$$

By employing the Alternative Least Squares, $U^1, U^2, \cdots, U^k$ converge to local optimal solutions technique, and then the low dimensional representations are obtained by $Y = X \times_1 U^1 \cdots \times_k U^k$ for X.

3 Orthogonal Tensor Sparse Neighborhood Preserving Embedding (OTSNPE)

In this section, we first introduce the basic idea of our algorithm and then the objective function is obtained; finally we give steps of our algorithm.

3.1 Basic idea

Although TNPE and OTNPE have been successfully applied in dimensionality reduces on two-dimensional images, there are some following shortcomings:

(1) The setting of the parameter of k or ε is difficult to deal with because of the lack of mature mathematical models.

(2) Deficiencies of supervised discriminant information require plenty training samples, which is unfit for practical application.

(3) Approximate nonlinear reconstruction not only fails to reflect fully local intrinsic geometric structures on images with nonlinear structure but also has poor robustness performance on disturbed images.

According to above analysis, there are some methods for these shortcomings as follows:

(1) Sparse reconstruction relations reflect intrinsic geometric properties of the data and contain natural discriminant information. Besides, sparse reconstruction has been proved to have strong robustness performance on disturbed images.

(2) The reconstruction of neighborhood within the same class provides supervised discriminant information based on class label.

Therefore, it is viable to fuse these methods into OTNPE, which share advantages of them.

3.2 Objective function

Given samples $X = \{x_i | x_i \in R^{n_1 \times n_2}, 1 \le i \le n\}$ and classes $C = \{1, 2, 3, \cdots, m\}$. There are h samples in each class. Firstly, the sparse construction of neighbourhoods within the same class is defined as follows:

$$\begin{aligned} &\min_{s_i^l} \| S_i^l \|_1 \\ &s.t. \ x_i^l = \sum_{j=1}^k S_{ij}^l x_j^l \\ &1 = 1^T S_i^l \end{aligned} \tag{14}$$

where x_i^l denotes a sample x_i of the class $l \in C(1 \le l \le m)$, S_{ij}^l denotes the contribution of each x_j to reconstructing x_i in the class l and S_i^l denotes the sparse reconstruction vector of x_i in the class l. In contrast to Eq.(7), Eq.(14) gives us description of sparse reconstructive weights of x_i in neighborhood samples within the same class, which make the improved reconstruction way not only fuses supervised discriminant information from class label but also fuses sparse reconstruction to preserve local intrinsic geometric property. The weight matrix W^S of training samples can be gotten by the way of neighborhood sparse reconstructions.

Secondly, a two-dimensional image is regarded as a second-order tensor image. Given a group data $X = (x_1, x_2, x_3, \cdots, x_n)$ of the tensor space $R^{n_1} \otimes R^{n_2}$, the destination of dimensionality reduction based on second-order tensor representation is to search two transformational matrices $U(n_1 \times l_1)$ and $V(n_2 \times l_2)$ to get a new group of data $Y = (y_1, y_2, y_3, \cdots, y_n)$ in the new tensor space $R^{l_1} \otimes R^{l_2}, l_1 < n_1, l_2 < n_2$, namely $Y = U^T X V$.

According to Eq.(13) and Eq.(14), we can draw the following objective function:

$$\begin{aligned} &\min_{U,V} \sum_{i=1}^n \| y_i - \sum_{j=1}^h W_{i,j}^S y_j \|_F^2 \\ &= \min_{U,V} \sum_{i=1}^n \| U^T x_i \times V - \sum_{j=1}^h W_{i,j}^S U^T x_j \times V \|_F^2 \\ &s.t. \ \sum_i \| X_i \times_1 U \times_2 V \|_F^2 = 1 \\ &\quad U^T U = I, V^T V = I \end{aligned} \tag{15}$$

where $M^S = (I - W^S)(I - W^S)^T$.

The process of computing transformation matrices U and V is as follow:

(1) According to $\|A\|_F^2 = tr(AA^T)$, we can get:

$$\begin{aligned}
&\sum_{i=1}^{n} \|y_i - \sum_{j=1}^{h} W_{i,j}^S y_j\|_F^2 \\
&= \sum_{i=1}^{n} \|y_i(I - W^S)\|_F^2 \\
&= \sum_{i=1}^{n} tr[(y_i(I - W^S))(y_i(I - W^S))^T] \\
&= \sum_{i=1}^{n} tr[(U^T x_i V(I - W^S))(U^T x_i V(I - W^S))^T] \\
&= \sum_{i=1}^{n} tr[U^T(x_i V)(I - W^S)(I - W^S)^T(x_i V)^T U] \\
&= \sum_{i=1}^{n} tr[U^T(x_i V)M^S(x_i V)^T U] \\
&= tr[U^T \sum_{i=1}^{n}(x_i V)M^S(x_i V)^T U]
\end{aligned} \tag{16}$$

Let $x_i^\nu = x_i V (1 \le i \le n)$. In order to get U, we define the objective function as follows:

$$\begin{aligned}
&\min_U tr[U^T \sum_{i=1}^{n}(x_i^\nu)M^S(x_i^\nu)^T U] \\
&s.t. \quad \sum_i \|x_i^\nu \times U\|_F^2 = 1 \\
&\qquad U^T U = I
\end{aligned} \tag{17}$$

Eq.(17) is converted to a generalized eigenvalue problem:

$$X^\nu M^S (X^\nu)^T u_h = \lambda_h X^\nu (X^\nu)^T u_h \tag{18}$$

where $u_h (1 \le h \le l_1)$ is generalized eigenvector, λ_h is generalized eigenvalue and $\lambda_1 > \lambda_2 > \lambda_3 > \cdots > \lambda_{l_1}$ is required to be satisfied. We can attain corresponding generalized eigenvector $U = \{u_1, u_2, \cdots, u_{l_1}\}$.

(2) Similarly, according to $\|A\|_F^2 = tr(A^T A)$, we can draw:

$$\begin{aligned}
&\sum_{i=1}^{n} \|y_i - \sum_{j=1}^{h} W_{i,j}^S y_j\|_F^2 \\
&= tr[V^T \sum_{i=1}^{n}(x_i^T U)M^S(x_i^T U)^T V]
\end{aligned} \tag{19}$$

Let $x_i^u = x_i^T U (1 \le i \le n)$, in order to get V, we define the objective function as follows:

$$\begin{aligned}
&\min_V \sum_{i=1}^{n} tr[V^T x_i^u M^S (x_i^u)^T V] \\
&s.t. \quad \sum_i \|X_i^u \times V\|_F^2 = 1 \\
&\qquad V^T V = I
\end{aligned} \tag{20}$$

Eq.(20) can be further to convert into a generalized eigenvalue problem:

$$X^u M^S (X^u)^T v_h = \lambda_h X^u (X^u)^T v_h \tag{21}$$

Similarly, we can attain corresponding generalized eigenvector $V = \{v_1, v_2, \cdots, v_{l_1}\}$.

To ensure U and V converged, we iterate the procedure for several times until error conditions are satisfied.

3.3 Algorithm steps

Input: Training samples $X = \{x_1, x_2, x_3, \cdots, x_n\}$ of the tensor space $R^{n_1} \otimes R^{n_2}$ and $x_i \in R^{n_1} \otimes R^{n_2}$, class $C = \{c_1, c_2, c_3, \cdots, c_m\}$, k samples in each class, error ε.

Output: two project matrix $U(n_1 \times l_1)$ and $V(n_2 \times l_2)$.

Steps:

(1) Initial setting: $U_{(1)} = I(n_1, l_1)$,$V_{(1)} = I(n_2, l_2)$, $t = 1$, $U_{(1)}$ and $V_{(1)}$ denote respectively unit matrix, $X^\nu = \{x_i^\nu | x_i^\nu \in R^{n_1 \times l_2}, 1 \le i \le n\}$, $X^u = \{x_i^u | x_i^u \in R^{n_2 \times l_1}, 1 \le i \le n\}$, elements of Y and Z are zero.

(2) According to Eq.(14), calculate the sparse reconstructive weight S^l of samples in the each class. The weight matrix W^S is attained with S^l.

(3) Set iteration circle variation $t = 1$.

(4) Calculate $x_i^\nu = x_i V (1 \le i \le n)$.

(5) According to Eq. (18) calculate $U_{(t)}$ and $U_{(t)}$ is normalized.

(6) Calculate $x_i^u = x_i^T U (1 \le i \le n)$.

(7) According to Eq. (21) calculate $V_{(t)}$ and $V_{(t)}$ is normalized.

(8) If $\| U_{(t)} - U_{(t-1)} \| \prec \varepsilon$ and $\| V_{(t)} - V_{(t-1)} \| \prec \varepsilon$, then jump into step (9), else $t = t + 1$ and jump into step (4).

(9) Obtain projections matrix $U = U_{(t)}$ and $V = V_{(t)}$.

4 Experiments and analyses

In this section, in order to evaluate the performance of our proposed algorithm, we apply our OTSNPE

Figure 2: A group of face images on AR

into face recognition and compare it with the TPCA, TLDA, TLPP,TNPE and OTNPE on AR, Yale and YaleB two-dimensional face databases which are disturbed by external environment.

4.1 Experimental settings

TPCA, TLDA, TLPP,TNPE and OTNPE are compared with the proposed OTSNPE for analyses on performances of OTSNPE. Apart from error ε set to 0.01, the detail settings of other parameters on algorithms are shown in Table 1.

Table 1: The detail settings of parameters on algorithms

Algorithms name	Parameters settings
TPCA	no
TLDA	no
TLPP	$\kappa = 7$
TNPE	$\kappa = 7$
OTNPE	$\kappa = 7$
OTSNPE	no

where the parameter k denotes the neighborhood size.

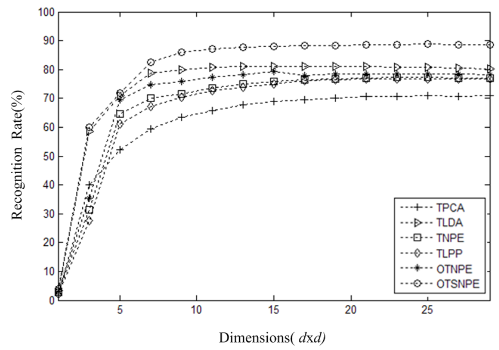
The simplest Nearest Neighbor Classifier(NNC) is applied in the experiment. For computational convenience, we resize face images to different pixels according to different face databases. We select randomly the certain number of images from each group of faces for training samples and remains for testing on face databases.

4.2 Experimental results

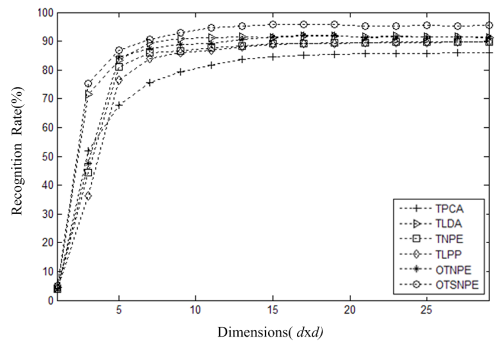
4.2.1 Experimental results on AR

AR face database consists of over 4000 face images of 126 individuals. For each individual, 26 pictures were taken in two sessions that separated by two weeks and each section contains 13 images, which include front view of faces with different expressions, illuminations and occlusions. Fig.2 shows a group of face images on AR.

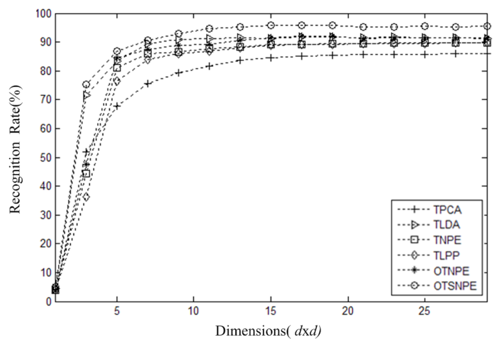
In our experiment, we firstly resize theses face images to 30×30 pixels. Furthermore ,we set $L(5, 10, 15)$ images of each group faces for training samples and remains for testing.



(a) $L = 5$



(b) $L = 10$



(c) $L = 15$

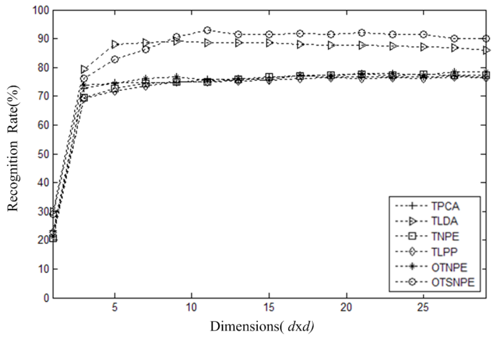
Figure 3: The recognition rate VS. the number of dimensions under L training samples on AR

In order to evaluate the performance of our algorithm, we select respectively different number of retained feature dimension with increment of twenty and calculate corresponding recognition accuracy rate. Concrete experimental results on AR are shown in Fig.3.

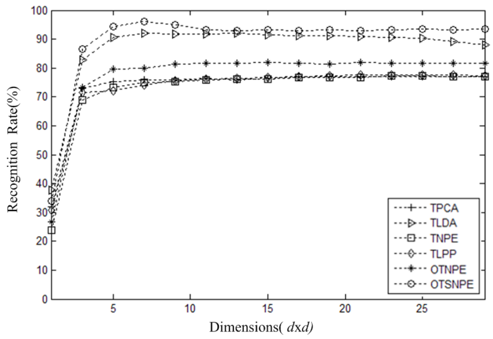
Figure 4: A group of face images on Yale

4.2.2 Experimental results on Yale

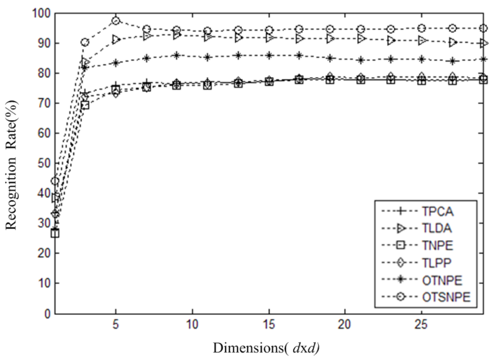
Yale face database contains 165 face images of 15 individuals. There are 11 images per subject, and these 11 images are respectively, under the following different facial expression or configuration: center-light, wearing glasses, happy, left-light, wearing no glasses, normal, right-light, sad, sleepy, surprised and wink. A group of face images on Yale are shown in Fig.4.



(a) $L = 5$



(b) $L = 6$



(c) $L = 7$

Figure 5: The recognition rate VS. the number of dimensions under L training samples on Yale

In the experiment image are resized to 32×32 pixels. we get $L(5, 6, 7)$ images of each group faces for training samples and remains for testing. Similarly, we select respectively different number of feature dimension and calculate corresponding recognition accuracy rate. Fig.5 gives us concrete experimental results on Yale.

4.2.3 Experimental results on YaleB

YaleB face database contains 2414 front-view face images of 38 individuals. For each individual, about 64 pictures were taken under various laboratory-controlled lighting conditions. A group face images of YaleB are shown in Fig.6.

Firstly images are resized to 32×32 pixels. we set $L(5, 10, 15, 20)$ images of each group faces for training samples and remains for testing. Similarly, we respectively select different number of feature dimension and calculate corresponding recognition accuracy rate. Fig.7 shows concrete experimental results.

4.3 Experimental analyses

From Fig.3, Fig.5 and Fig.7, we can draw conclusions as follows:

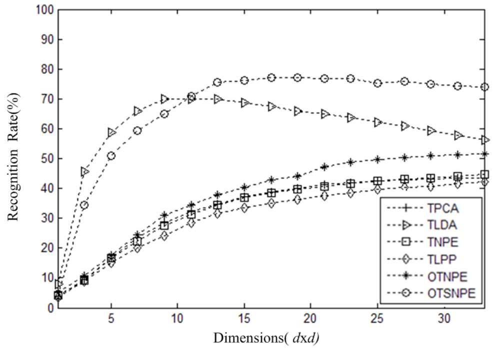
(1) In contrast to TPCA, Our OTSNPE algorithm outperforms TPCA obviously. This is mainly because TPCA only sees the global Euclidean structure of data, which is obviously not fit for AR, Yale and YaleB with nonlinear properties.

(2) TLDA also is a linear dimensionality reduction, our OTSNPE algorithm is not always superior to TLDA in different retained subspace dimension. TLDA outperform OTSNPE when the number of retained subspace dimension is small and OTSNPE outperform TLDA when the number of retained subspace dimension is more than the certain number. This is illuminated by that the global supervised linear discriminant information is still limited by linear structure and is more easy to result in over-fitting problem.

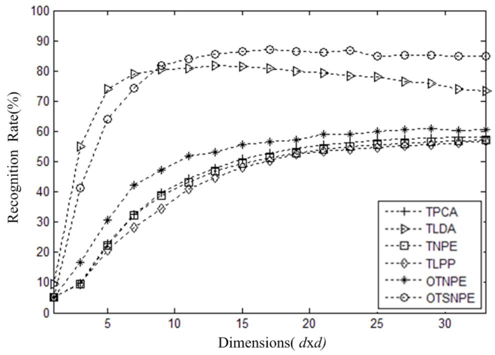
(3) Our OTSNPE algorithm is obviously superior to TLPP which preserve the local nonlinear structure of data. However, TLPP is difficult to attain accurate local geometric structure discriminant information on AR, Yale and YaleB with various lighting conditions, deformity, dressing and shelter, which demonstrates the power robustness performance of OTSNPE.

(4) OTSNPE algorithm is superior to TNPE and OTNPE, which is illuminated by that TNPE and OTNPE aim to preserve the local manifold structure of

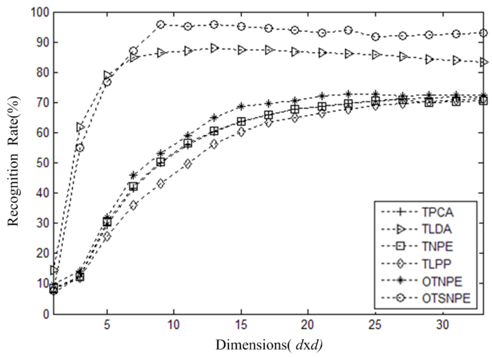
Figure 6: A group of face images on YaleB



(a) $L = 5$



(b) $L = 10$



(c) $L = 20$

Figure 7: The recognition rate VS. the number of dimensions under L training samples on YaleB

data with the approximate nonlinear way that is poorer in robustness on disturbed images than OTSNPE.

(5) With increment on the number of retained subspace dimension, the recognition accuracy of TPCA, TLPP, TNPE and OTNPE promote continually. However, when the number of retained feature dimension attains the threshold, there is a downtrend in the recognition accuracy on TLDA and smoothness on OTSNPE. This illustrates that supervised discriminant information based on class labels play an important role in OTSNPE.

To sum up, our proposed OTSNPE has the best performance among all of the six dimensionality reduction approaches. This is probably due to the facts that OTSNPE considers explicitly the within-class sparse reconstruction by the objective function (15), which captures local nonlinear structure properties with neighborhood reconstruction based on sparse representations that have strong robustness on images. Moreover, the supervised discriminant information detected by the objective function (15) is helpful to make OTSNPE get most accuracy rate in the low-dimensional subspace.

5 Conclusion

In the paper, aiming to solving the problem of existing tensor neighborhood preserving embedding on two-dimensional images, we propose Orthogonal Tensor Sparse Neighborhood Preserving Embedding(OTSNPE) for dimensionality reduction on the basic of sparse reconstruction and tensor representation. The algorithm not only preserves local within-class sparsity reconstruction but also preserves spatial relations of pixels in images in the projected subspace. Experimental results on AR, Yale and YaleB demonstrate that our algorithm outperform TPCA,TLDA,TLPP,TNPE and OTNPE.

However, each approach has its own advantages and disadvantages. In order to capture local geometric characteristics through sparse reconstructions within the same class, the certain number of training samples in each class is required in our proposed algorithm. The algorithm also ignores global geometric characteristics of training data. How to absorb global nonlinear discriminant information into the algorithm to further improve its performance is an important work. Moreover, the proposed algorithm is for two-dimensional images. How to extend it to more order tensor to deal with data with more than three dimension is also the future work.

Acknowledgements: The research is supported by NSF of China (grant No. 71171148) and NSF of Zhejiang province(grant No.LQ12F02007, Y201122544).

References:

[1] R. O. Duda , P. E. Hart and D. G. Stork, *Pattern Classification. 2nd ed.*, New York: John Wiley & Sons, 2001.

[2] L. Maaten, E. Postma and J. Herik, Dimension reduction: A comparative review, Technical Report, *TiCC-TR 2009-005, Tilburg University*, 2009.

[3] H. Hotelling, Analysis of a complex of statistical variables into principal components, *Journal of Educational Psychology*, 24(6), 1933, pp.417-441.

[4] R. A. Fisher, The use of multiple measurements in taxonomic problems, *Annals of Eugenics*, 7(2), 1936, pp.179-188.

[5] B. Scholkopf, A. J. Smola and K. R. Muller, Nonlinear component analysis as a kernel eigenvalue problem, *Neural Computation*, 10(5), 1998, pp.1299-1319.

[6] G. Baudat and F. E. Anouar, Generalized discriminant analysis using a kernel approach, *Neural Computation*, 12(10), 2000, pp.2385-2404.

[7] S. T. Roweis and L. K. Saul, Nonlnear dimensionality reduction by locally linear embedding, *Science*, 290 (5500), 2000, pp.2323-2326.

[8] J. B. Tenenbaum, V. Silva and L. JC, A global geometric framework for nonlinear dimensionality reduction, *Science*, 290(5500), 2000, pp.2319-2323.

[9] M. Belkin and P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, *Neural Computation*,15(6), 2003, pp.1373-1396.

[10] D. L. Donoho and C. Grimes, Hessian eigenmaps: locally linear embedding techniques for high-dimensional data, *Proceedings of the National Academy of Sciences of the United States of America*, 100(10), 2003, pp.5591-5596.

[11] D. K. Saxena and K. Deb, Non-linear dimensionality reduction procedures for certain large-dimensional multi-objective optimization problems: employing correntropy and a novel maximum variance unfolding, *Evolutionary Multi-Criterion Optimization*, 4403, 2007, pp. 772-787.

[12] Z. Y. Zhang and H. Y. Zha, Nonlinear dimension reduction via local tangent space alignment, *Intelligent Data Engineering and Automated Learning*, 2690, 2003, pp.477-481.

[13] Y. B. Zhan and J .P. Yin, Robust local tangent space alignment, *Neural Information Processing*, 5863, 2009, pp.293-301.

[14] T. Lin and H. B. Zha, Riemannian manifold learning, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(5), 2008, pp.796-809.

[15] T. Lin, H. B. Zha and S. U. Lee, Riemannian manifold learning for nonlinear dimensionality reduction, *Proceeding ECCV'06 Proceedings of the 9th European conference on Computer Vision*, 1, 2006, pp.44-55.

[16] S. M. Xiang, F. P. Nie, C. S. Zhang and C. X. Zhang, Nonlinear dimensionality reduction with local spline embedding, *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 2009, pp.1285-1298.

[17] X. F. He, S. C. Yan, Y. X. Hu, P. Niyogi and H. J. Zhang, Face recognition using Laplacian faces, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3), 2005, pp.328-340.

[18] X. F. He and P. Niyogi, Locality Preserving Projections, *Proceedings of Conference on Advances in Neural Information Processing Systems*, 2003, pp.153-160.

[19] X. F. He, D. Cai and S. C. Yan, Neighborhood Preserving Embedding, *Proceedings of the 10th IEEE International Conference on Computer Vision. Washington: IEEE Computer Society Press*, 2005, pp.1208-1213.

[20] E. Kokiopoulou and Y. Saad, Orthogonal neighborhood preserving projections: A projection-based dimensionality reduction technique, *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(12), 2007, pp. 2143-2156.

[21] T. Zhang, K. Huang, X. Li, J. Yang and D. Tao, Discriminative orthogonal neighborhood-preserving projections for classification, *IEEE Trans Syst Man Cybern B Cybern*, 40(1), 2010, pp.253-263.

[22] Y. Wang and Y. Wu, Complete neighborhood preserving embedding for face recognition, *Pattern Recognition*,43(1), 2010, pp.1008-1015.

[23] J. Yang, D. Zhang, A. F. Fran 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*, 26(1), 2004, pp.131-137.

[24] M. Li and B. Yuan, 2D-LDA: a statistical linear discriminant analysis for image matrix, *Pattern Recognition Letters*, 26(5), 2005, pp.527-532.

[25] S. Chen, H. F. Zhao, M. Kong and B.Luo, 2D-LPP: a two-dimensional extension of locality preserving projections, *Neuro computing*, 70(4), 2007, pp.912-921.

[26] Y. Wang and Y. Wu, A two-dimensional Neighborhood Preserving Projection for appearance-based face recognition, *Pattern Recognition*, 45(5), 2012, pp.1866-1876.

[27] X. He, D. Cai and P. Niyogi, Tensor subspace analysis, *Advances in Neural Information Processing Systems*, 2006.

[28] Y. T. Wei, H. Li and L. Q. Li, Tensor locality sensitive discriminant analysis and its complexity, *International Journal of Wavelets Multiresolution and Information Processing*, 7 (11), 2009, pp.865-880.

[29] H. P. Lu, K. N. Plataniotis and A. N. Venetsanopoulos, Uncorrelated multilinear discriminant analysis with regularization and aggregation for tensor object recognition, *IEEE Transactions on Neural Networks*, 20(1), 2009, pp.103-123.

[30] G. Dai and D. Yeung, Tensor embedding methods, *Proceedings of the National Conference on Artificial Intelligence*, 1, 2006, pp.330-335.

[31] Y. Wang and Y. Wu, Orthogonal Tensor Neighborhood Preserving Embedding for facial expression recognition, *Pattern Recognition*, 44(7), 2011, pp.1497-1513.

[32] X. H. Han, Y. W. Chen, X. Ruan, Multilinear Supervised Neighborhood Embedding of a Local Descriptor Tensor for Scene/Object Recognition, *IEEE Transaction on Image Processing*, 21(3), 2012, pp.1314-1326.

[33] M. Davenport, M. Duarte, M. Wakin, D. Takhar, K. Kelly and R Baraniuk, The smashed filter for compressive classification and target recognition, *Proc. IS&T/SPIE Symposiumon Electronic Imaging: Computational Imaging*, 2007.

[34] K. Huang and S. Aviyente, Sparse representation for signal classification, *Advances in Neural Information Processing Systems*, 2006, pp.609-616.

[35] A. Yang, J. Wright, Y. Ma and S. Sastry, Feature selection in face recognition: A sparse representation perspective, *UC Berkeley Tech Report UCB/EECS-2007-99*, 2007.

[36] J. Wright, A. Yang, S. Sastry and Y. Ma, Robust face recognition via sparse representation, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 31(2), 2009, pp.210-227.

[37] J. Rodrguez, L. Kuncheva and C. Alonso, Rotation forest: A new classifier ensemble method, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 28 (10), 2006, pp. 1619-1630.

[38] J. Bi, K. Bennett, M. Embrechts, C. Breneman and M. Song, Dimensionality reduction via sparse support vector machines, *Journal of Machine Learning Research*, 3(1), 2003, pp.1229-1243.

[39] H. Zou, T. Hastie and R. Tibshirani, *Sparse principal component analysis*, Technical report, 2004.

[40] L. Qiao, S. Chen, and X. Tan, Sparsity preserving projections with applications to face recognition, *Pattern Recognition*, 43(1), 2010, pp.331-341.