AMSE JOURNALS-2016-Series: Advances B; Vol. 59; N°1; pp 91-112
Submitted July 2016; Revised Oct. 15, 2016, Accepted Dec. 10, 2016

Tensorial Biometric Signal Recognition Based on Multilinear PCA Plus GTDA

* Fan Zhang, **Ke Sun

* School of Information Engineering, North China University of Water Resources and Electric Power, Zhengzhou, 450045, China, (zfgh8221@163.com)
** Mathematics and Computational Science College, Guilin University of Electronic Technology, Guilin, 541000, China, (1065517483@qq.com)

Abstract

Most of the biometric signals are naturally multi-dimensional objects, which are formally known as tensors. There is an increasing interest in the multilinear subspace analysis and many methods have been proposed to operate directly on these tensorial data during the past several years. One of these popular unsupervised multilinear algorithms is Multilinear Principal Component Analysis (MPCA) while another of the supervised multilinear algorithm is Multilinear Discriminant Analysis (MDA). Then a MPCA+MDA method has been introduced to deal with the tensorial signal and a better recognition accuracy can be obtained. However, due to the no convergence of MDA, it is difficult for MPCA+MDA to obtain a precise result. Hence, to overcome this limitation, a new MPCA plus General Tensor Discriminant Analysis (GTDA) solution with well convergence is presented for tensorial biometric signal feature extraction in this paper. Several experiments are carried out to evaluate the performance of MPCA+GTDA on different databases and the results show that this method has the potential to achieve comparative effect as MPCA+MDA. In addition, some basic issues of GTDA such as initialization conditions, convergence and space dimension determination which have not been described clearly before are also discussed in detail in this paper.

Key words

Feature extraction, tensor object, multilinear principal component analysis, general tensor discriminant analysis.

1. Introduction

Feature extraction or dimensionality reduction lies at the heart of several aspects of subspace learning techniques and has attracted growing interest over the past several years. The goal of these methods is to transform a high-dimensional data set into a low-dimensional equivalent representation, while keeping most of the information consistent with the underlying structure of the actual physical phenomenon [1]. Most biometric signals have multi-dimensional representation. For example, two-dimensional biometric signals include gray-level images of fingerprint, palm print, ear, face, and multichannel electroencephalography signals in neuroscience. Three-dimensional data include color biometric images, Gabor faces, silhouette sequences in gait analysis, and gray video sequences in action recognition. A few multi-dimensional biometric signals can also be formed in more than three orders, such as color video sequence surveillance [2]. Tensor provides a natural and efficient way to describe that multi-dimensional data, such as vectors, are first-order tensors, whereas matrices are second-order tensors. The elements of a tensor are to be addressed by a number of indices that are used to define the order of the tensor object. Notably, each index defines a “mode” [3]

Principal component analysis (PCA) [4] and linear discriminative analysis (LDA) [5] are two of the most popular classical subspace learning methods among several existing dimensionality reduction algorithms. PCA seeks an optimal projection direction of maximal variation while disregarding the label information of samples. By contrast, LDA is a supervised method that searches for the optimal discriminative subspace by maximizing the ratio between the between-class scatter and the within-class scatter. Another widely used dimensionality reduction algorithm [6] is proposed, which is a linear extension of Laplacian eigenmaps [7], namely, locality preserving projections (LPP). LPP aims to preserve the local structure of the original space, whereas PCA and LDA aim to preserve global structures of the samples. However, tensors must be reshaped into vectors first before these linear subspace learning methods can be applied on tensor data. Such reshaping generally has two fundamental limitations: high computational burden and a loss of the potential spatial structure information of the original data. To overcome these shortcomings, some researchers have attempted to treat the second-order data as a matrix instead of a vector, such as 2DPCA [8], 2DLDA [9], and 2DLPP [10]. Although such methods effectively deal with second-order data, such as image classification and face recognition, these methods no longer obtain better results when faced with higher-order data. Therefore, several multilinear algorithms, which can directly operate on the original tensorial data without the limitation of order, have been proposed.

Multilinear principal component analysis (MPCA) [11], a tensor version of PCA, performs dimensionality reduction in all tensor modes to capture most variations presented in the original tensors. Several researchers have attempted to extend other conventional transformation approaches for tensor use. For example, by using high order singular value decomposition (HOSVD), the following multilinear algorithms, multilinear discriminant analysis (MDA) [12], general tensor discriminant analysis (GTDA) [13], and tensor subspace analysis (TSA) [14] applied LDA, maximum scatter difference (MSD) [15], and LPP, respectively, to transform each mode of the tensors. Moreover, to obtain better accuracy, a new method based on MPCA+MDA for face recognition has been proposed in reference [16]-[18]. In that paper, MPCA was implemented on the tensorial data for dimensionality reduction, and a new data set with a new dimension was generated. This new data set will be the inputs for the MDA algorithm to learn the most discriminative subspaces of the input samples. Given the use of MDA after applying the MPCA, this method is performed in a much lower-dimension feature space than the traditional methods, such as LDA and PCA, and it can overcome the small sample size problem [19]. However, like 2DLDA, the MDA algorithm does not converge [11], because the optimization algorithm applied to MDA fails to converge. Considering the instability, achieving a stable and precise accuracy is difficult for MDA.

Given the converged alternating projection, GTDA can provide stable recognition accuracy, whereas MDA cannot. By maximizing the between-class variance and minimizing the within-classes variance, GTDA decomposes tensors into core tensors and a series of discriminative matrices over every modality [20], similar to MDA. Moreover, empirical analysis on tensorial samples [21] indicates that GTDA can achieve results comparable to those of MDA. Consequently, GTDA can be used in this study to transform multilinear discriminative subspace from high-dimensional and high-order biometric signals. On the basis of the works briefly reviewed above, we introduce a new MPCA+GTDA algorithm to deal with tensorial data instead of MPCA+MDA, and expect this novel method to be a better choice. Considering that MPCA and GTDA are multilinear algorithms, lower dimensionality dilemma and better correct recognition rate can be captured. To the best of our knowledge, this study is the first to employ MPCA+GTDA on feature extraction and dimensionality reduction for tensor objects. Several issues caused by the iterative nature of the GTDA algorithm are also addressed in this paper.

The rest of the paper is organized as follows: Section 2 provides a brief introduction of multilinear algebra for dimensionality reduction. The algorithm of MPCA and GTDA are initially summarized and discussed in detail, and then a few aspects of GTDA such as initialization conditions, and termination criteria are discussed in section 3. We analyze experiment results on several databases to verify the properties of the proposed method and compare its performance with those of other algorithms in Section 4. The major findings and conclusions are presented in Section 5.

2. Tensor Fundamentals

Before discussing the MPCA and GTDA algorithms, introducing a few basic multilinear operations is necessary. Then, multilinear tensor-to-tensor projection (TTP) is described for dimensionality reduction and feature extraction from tensor objects. Finally, a typical TTP-based multilinear subspace learning algorithm is described in Section 2.3.

2.1 Notations and Basic Multilinear Algebra

The notation in this chapter follows the conventions in the multilinear algebra, such as the notation in [22]. Vectors are denoted by lowercase boldface letters, e.g., **x**; matrices by uppercase boldface, e.g., **U**; and tensors by calligraphic letters, e.g., $\mathcal{A}$. An *N*th-order tensor is denoted as $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$. Their elements are addressed by N indices i_n, $n = 1, \cdots, N$, and each i_n addresses the *n*-mode of $\mathcal{A}$.

The *n*-mode unfolding of $\mathcal{A}$ is defined as the I_n dimensional vectors are denoted as

$$\mathcal{A}_{(n)} \in \mathbb{R}^{I_n \times (I_1 \times \cdots \times I_{n-1} \times I_{n+1} \times \cdots \times I_N)} \quad (1)$$

Where the column vectors of $\mathcal{A}_{(n)}$ are gained from $\mathcal{A}$ by varying its index i_n while keeping all the other indices fixed. Fig.1 (a)-(e) provides a visual illustration of a 3-order tensor $\mathcal{A} \in \mathbb{R}^{3\times4\times3}$ where 1-mode corresponds to the column vector, see Fig.1(b); 2-mode corresponds to row vector, see Fig.1(c); 3-mode corresponds to the depth vector, see Fig.1(d); and 1-mode matricization of $\mathcal{A}$, see Fig.1(e).

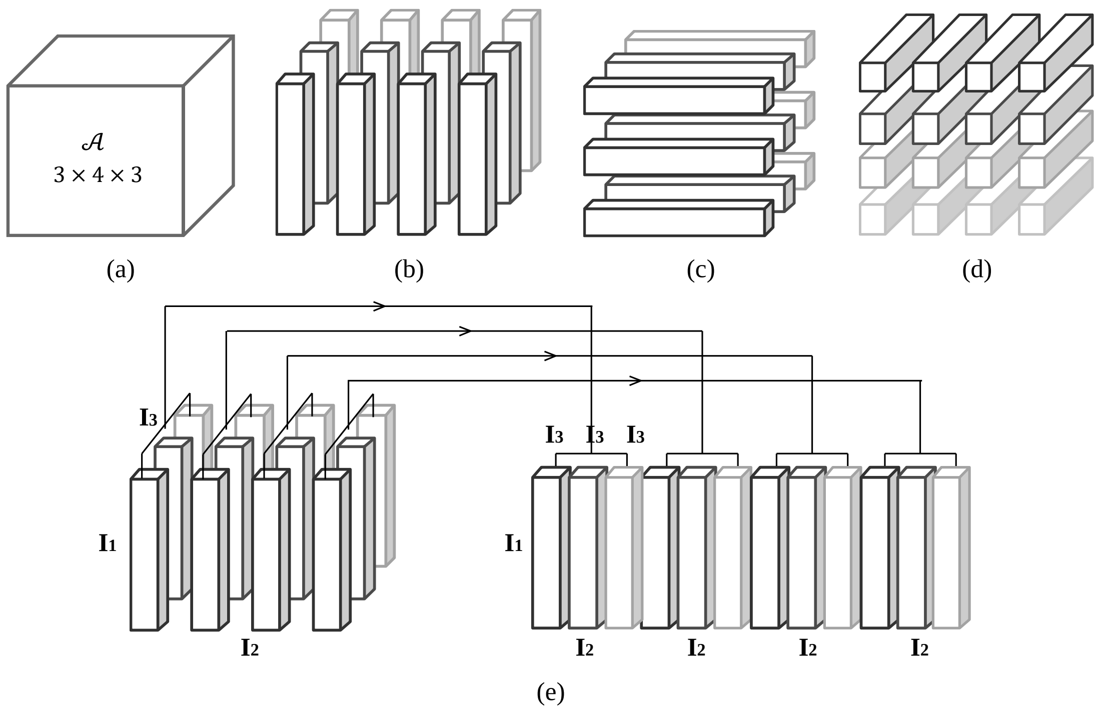


Fig.1. Illustration of the *n*-mode vectors: (a) a tensor $\mathcal{A} \in \mathbb{R}^{3\times4\times3}$, (b) the 1-mode vectors, (c) the 2-mode vectors, (d) the 3-mode vectors, and (e) 1-mode unfolding of a third-order tensor.

The *n*-mode product of a tensor $\mathcal{A}$ by a matrix $\mathbf{U} \in \mathbb{R}^{J_n \times I_n}$, denoted by $\mathcal{A} \times_n \mathbf{U}$, is a tensor defined as

$$(\mathcal{A} \times_n \mathrm{U})(i_1, \cdots, i_{n-1}, j_n, i_{n+1}, \cdots, i_N) = \sum_{i_n} \mathcal{A}(i_1, i_2, \cdots, i_N) \cdot \mathrm{U}(j_n, i_n) \tag{2}$$

In Fig.2, a third-order tensor $\mathcal{A} \in \mathbb{R}^{9\times7\times4}$ is projected in the 1-mode vector space by a projection matrix $\mathbf{B} \in \mathbb{R}^{4\times9}$, resulting in the projected tensor $\mathcal{A} \times_1 \mathbf{B} \in \mathbb{R}^{4\times7\times4}$.

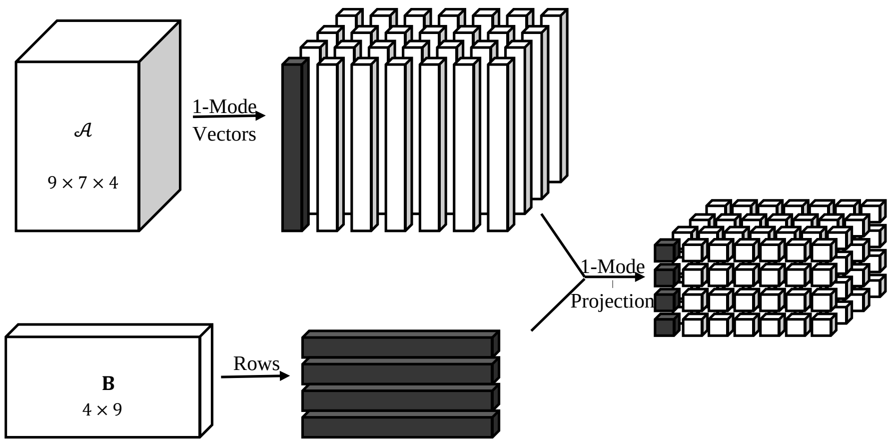


Fig.2. Example of the n-mode (1-mode) multiplication.

One of the most commonly used tensor decompositions is Tucker, which can be regarded as higher-order generalization of the matrix singular value decomposition (SVD). Let $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ denotes an Nth-order tensor, then the Tucker decomposition is defined as follows

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \cdots \times_N \mathbf{U}^{(N)} \tag{3}$$

Where $\mathcal{S} \in \mathbb{R}^{P_1 \times P_2 \times \cdots \times P_N}$ with $P_n < I_n$, denotes the core tensor and $\mathbf{U}^{(n)} = [\mathbf{u}_1^{(n)} \mathbf{u}_2^{(n)} \cdots \mathbf{u}_{P_n}^{(n)}]$ is an $I_n \times P_n$ matrix. The Tucker decomposition is illustrated in Fig.3. When all $\{\mathbf{U}^{(n)}\}_{n=1}^{N}$ are orthonormal and the core tensor is all orthogonal, this model is called HOSVD [23].

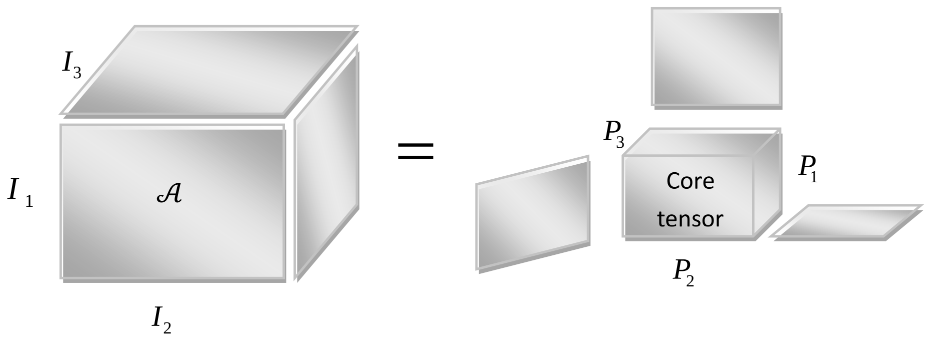


Fig.3. Visualisation of the decomposition of tensor into 3 subspace matrices and a core tensor.

The scalar product of two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is defined as

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i_1} \cdots \sum_{i_N} \mathcal{A}(i_1, i_2, \cdots, i_N) \cdot \mathcal{B}(i_1, i_2, \cdots, i_N) \tag{4}$$

The Frobenius norm of $\mathcal{A}$ is defined as

$$\|\mathcal{A}\| = \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle} = \|\mathcal{A}_{(n)}\|_F = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} \mathrm{a}_{i_1 i_2 \cdots i_N}^2} \tag{5}$$

2.2 Tensor-to-Tensor Projection

The dimension reduction of biometric signals in this paper is applied during multilinear projection from a tensor space to another tensor space, namely TTP. TTP is formulated based on the Tucker decomposition, the projection framework was first introduced in [24] [25]. To project an Nth-order tensor $\mathcal{X}$ in a tensor space $\mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \cdots \otimes \mathbb{R}^{I_N}$, to another tensor $\mathcal{Y}$ in a lower-dimensional tensor space $\mathbb{R}^{P_1} \otimes \mathbb{R}^{P_2} \cdots \otimes \mathbb{R}^{P_N}$, where $P_n \leq I_n$ for all n, N projection matrices are utilized

$$\mathcal{Y} = \mathcal{X} \times_1 \mathbf{U}^{(1)^T} \times_2 \mathbf{U}^{(2)^T} \cdots \times_N \mathbf{U}^{(N)^T} \tag{6}$$

These N projection matrices used for TTP can be concisely written as $\{\mathbf{U}^{(n)}\}$. Fig.4 shows the TTP of a tensor sample $\mathcal{A}$ to a smaller tensor of size $P_1 \times P_2 \times P_3$.

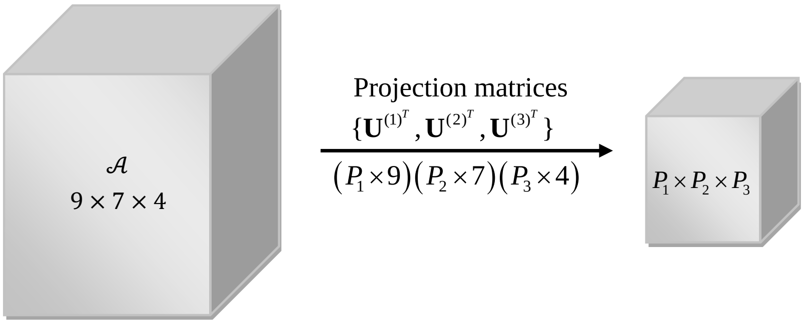


Fig.4. Illustration of tensor-to tensor projection.

2.3 Multilinear Subspace Learning Algorithms Based on TTP

TTP method has been solved to develop several multilinear algorithms, including MPCA, MDA, and GTDA. N sets of parameters for N projection matrices to be solved (one in each mode) during a TTP in multilinear subspace learning (MSL). An iterative procedure called alternating partial projections (APP), which originated from the alternating least squares algorithm [26], is commonly organized to solve for tensor-based projections that alternate with partial projections. Given the iterative nature of the solution, issues such as initialization, termination, and convergence need to be determined in MSL. According to the projection matrices in all the other modes ($j \neq n$), the mode-n projection matrix is solved one by one. The mode-n partial multilinear projection of a tensor $\mathcal{X}$ in TTP using $\{\mathbf{U}^{(j)}, j \neq n\}$ is written as

$$\tilde{\mathcal{Y}}^{(n)} = \mathcal{X} \times_1 \mathbf{U}^{(1)^T} \times_2 \mathbf{U}^{(2)^T} \cdots \times_{(n-1)} \mathbf{U}^{(n-1)^T} \times_{(n+1)} \mathbf{U}^{(n+1)^T} \cdots \times_N \mathbf{U}^{(N)^T} \tag{7}$$

In addition, for TTP-based MSL, testing all the possible combinations of the N values, P_1, P_2, $\cdots$, P_N, for a desired amount of dimensionality reduction is often costly. Thus, the desired subspace dimensions $\{P_1, P_2, \cdots, P_N\}$ for this approach need to be predetermined [27].

Typical algorithmic procedures for TTP-based MSL algorithms are shown in Algorithm 1.

Algorithm 1 A typical TTP-based multilinear subspace learning algorithm

INPUT: a series of *N*th order tensors $\{\mathcal{X}_m \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}, m = 1, \cdots, M\}$

OUTPUT: *N* matrices $\{\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times P_n}, (P_n < I_n, n = 1,2, \cdots, N)\}$

Algorithm:

1: Initialize $\{\mathbf{U}^{(n)}\}$ with a set of identity matrices

2: Local optimization

for $k = 1$ to K **do**

for $n = 1$ to N **do**

for *m*=1 to *M* **do**

Calculate Equation (7) to get the mode-*n* partial multilinear projection of input tensors.

Solve for the mode-n unfolding matrix of $\mathcal{Y}_m^{(n)}$.

end for

Solve for the mode-n $\mathbf{U}^{(n)}$ as a linear problem obtained through the mode-*n* unfolding matrix of $\mathcal{Y}_m^{(n)}$.

end for

If the algorithm converges or a maximum number of iterations *K* attains, break and output the current $\{\mathbf{U}^{(n)}\}$.

end for

3. Multilinear Principal Component Analysis and General Tensor Discriminant Analysis

As a high-order extension of PCA, MPCA is an unsupervised MSL algorithm for general tensors that target variation maximization by solving a TTP. As a multilinear extension of MSD, GTDA is a supervised MSL algorithm for performing discriminant analysis on general tensor inputs by also solving a TTP. Motivated by Fisherface [5] and the MPCA+MDA algorithm, we use MPCA algorithm for tensor objects feature extraction and dimension reduction, and then apply GTDA on the low dimensionality features. Finally, nearest neighbor (NN) classifier is implemented to classify the computed GTDA features.

3.1 Multilinear Principal Component Analysis

In this section, the MPCA algorithm is described in detail based on the analysis introduced in [11]. A set of *M* tensor object samples $\{\mathcal{X}_1, \mathcal{X}_2, \cdots, \mathcal{X}_M\}$ are available for training and each tensor object $\mathcal{X}_m \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$. The MPCA objective is the determination of the *N* projection matrices $\left\{\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times P_n}, n = 1, \cdots, N\right\}$ that maximize the total tensor scatter $\Psi_{\mathcal{Y}}$

$$\left\{\tilde{\mathbf{U}}^{(n)}\right\} = \underset{\{\mathbf{U}^{(n)}\}}{\arg\max}\, \Psi_{\mathcal{Y}} = \underset{\{\mathbf{U}^{(n)}\}}{\arg\max}\, \sum_{m=1}^{M} \|\mathcal{Y}_m - \bar{\mathcal{Y}}\|_F^2 \quad (8)$$

where $\mathcal{Y}_m = \mathcal{X}_m \times \left\{\mathrm{U}^{(n)^T}\right\}_{n=1}^{N}$ according to Equation (6), $\bar{\mathcal{Y}} = \frac{1}{M}\sum_{m=1}^{M} \mathcal{Y}_m$. The dimensionality P_n for each mode $\mathrm{U}^{(n)}$ should be predetermined.

As discussed in Section 2.3, the *N* optimization subproblems are solved through the APP method by finding the mode-*n* projection matrix $\mathbf{U}^{(n)}$ that maximizes the mode-*n* total scatter conditioned on the projection matrices in all the other modes. Let $\left\{\mathbf{U}^{(n)}, n = 1, \cdots, N\right\}$ be the answer to Equation (8), and $\left\{\mathbf{U}^{(1)}, \cdots, \mathbf{U}^{(n-1)}, \mathbf{U}^{(n+1)}, \cdots, \mathbf{U}^{(N)}\right\}$ be all the other known projection matrices, the P_n eigenvectors reside in the matrix $\mathbf{U}^{(n)}$ and correspond to the largest P_n eigenvalues of the matrix $\mathbf{\Phi}^{(n)}$

$$\mathbf{\Phi}^{(n)} = \sum_{m=1}^{M} \left(\mathbf{X}_{m(n)} - \bar{\mathbf{X}}_{(n)}\right) \mathbf{U}_{\Phi^{(n)}} \mathbf{U}_{\Phi^{(n)}}^T \left(\mathbf{X}_{m(n)} - \bar{\mathbf{X}}_{(n)}\right)^T \quad (9)$$

where $\mathbf{X}_{m(n)}$ is the mode-*n* unfolding of $\mathcal{X}_m$ and

$$\mathbf{U}_{\Phi^{(n)}} = \left(\mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n+2)} \otimes \cdots \otimes \mathbf{U}^{(N)} \otimes \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \otimes \cdots \mathbf{U}^{(n-1)}\right) \quad (10)$$

The projection matrices $\left\{\mathbf{U}^{(n)}\right\}$ are initialized through the full projection truncation (FPT) described in [11] and updated one by one with all the others fixed.

3.2 General Tensor Discriminant Analysis

GTDA aims to maximize a multilinear extension of the scatter-difference based discriminant criterion in [14]. GTDA intends to solve for a TTP $\{\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times P_n}, P_n \leq I_n, n = 1, \cdots, N\}$ that projects a tensor $\mathcal{X}_m \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ to a low-dimensional tensor $\mathcal{Y}_m \in \mathbb{R}^{P_1 \times \cdots \times P_N}$.

A set of M labeled tensor data samples $\{\mathcal{X}_1, \mathcal{X}_2, \cdots, \mathcal{X}_M\}$ in $\mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \cdots \otimes \mathbb{R}^{I_N}$ are available for training with class label $\mathbf{c} \in \mathbb{R}^M$. The class label for the mth sample $\mathcal{X}_m$ is $c_m = \mathbf{c}(m)$ and there are C classes in total. The between-class scatter matrix of these tensors is defined as

$$\Psi_{B_\mathcal{Y}} = \sum_{c=1}^{C} M_c \|\bar{\mathcal{Y}}_c - \bar{\mathcal{Y}}\| \tag{11}$$

and the within-class scatter of these tensors is defined as

$$\Psi_{W_\mathcal{Y}} = \sum_{m=1}^{M} \|\mathcal{Y}_m - \bar{\mathcal{Y}}_{c_m}\| \tag{12}$$

Where $\mathcal{Y}_m = \mathcal{X}_m \times \{\mathbf{U}^{(n)^T}\}_{n=1}^{N}$, M_c is the number of samples for class c, c_m is the class label for the mth sample $\mathcal{X}_m$, the overall mean tensor $\bar{\mathcal{Y}} = \frac{1}{M}\sum_{m=1}^{M} \mathcal{Y}_m$ and the class mean tensor $\bar{\mathcal{Y}}_c = \frac{1}{M_c}\sum_{m,c_m=c} \mathcal{Y}_m$. The objective function of GTDA can be written as

$$\{\tilde{\mathbf{U}}^{(n)}\} = \arg\max_{\{\mathbf{U}^{(n)}\}} \Psi_{\mathcal{Y}_{dif}} = \arg\max_{\{\mathbf{U}^{(n)}\}} \Psi_{B_\mathcal{Y}} - \zeta \cdot \Psi_{W_\mathcal{Y}} \tag{13}$$

where ζ is a tuning parameter and is automatically selected during the training procedure according to [14]. The APP method is also employed to utilize for the N projection matrices iteratively. First the input tensors (that are the outputs of MPCA) should be solved with the mode-n projection matrix conditioned on the projection matrices in all the other modes which defined in Equation (7), and then all the new tensors are unfolded into a matrix along the nth-mode.

The mode-n between-class and within-class scatter matrices can be obtained from $\hat{\mathbf{Y}}_{m(n)}$ which is the mode-n unfolding of $\hat{\mathcal{Y}}_m^{(n)}$

$$\mathbf{S}_{B_\mathcal{Y}}^{(n)} = \sum_{c=1}^{C} \mathrm{M}_c \left(\bar{\hat{\mathbf{Y}}}_{c(n)} - \bar{\hat{\mathbf{Y}}}_{(n)}\right)\left(\bar{\hat{\mathbf{Y}}}_{c(n)} - \bar{\hat{\mathbf{Y}}}_{(n)}\right)^T \tag{14}$$

$$\mathbf{S}_{W_\mathcal{Y}}^{(n)} = \sum_{m=1}^{M} \left(\hat{\mathbf{Y}}_{m(n)} - \bar{\hat{\mathbf{Y}}}_{c_m(n)}\right)\left(\hat{\mathbf{Y}}_{m(n)} - \bar{\hat{\mathbf{Y}}}_{c_m(n)}\right)^T \tag{15}$$

respectively, where $\bar{\bar{\mathbf{Y}}}_{(n)} = \frac{1}{M}\sum_{m=1}^{M} \hat{\mathbf{Y}}_{m(n)}$ and $\bar{\bar{\mathbf{Y}}}_{c(n)} = \frac{1}{M_c}\sum_{m=1,c_m=c}^{M} \hat{\mathbf{Y}}_{m(n)}$.

Following Equation (13), the mode-*n* projection matrix $\mathbf{U}^{(n)}$ in this conditional optimization problem are solved as

$$\{\tilde{\mathbf{U}}^{(n)}\} = \underset{\{\mathbf{U}^{(n)}\}}{\arg\max}\ \mathrm{tr}\left(\mathbf{U}^{(n)T}\mathbf{S}_{B_{\mathcal{Y}}}^{(n)}\mathbf{U}^{(n)}\right) - \zeta \cdot \mathrm{tr}\left(\mathbf{U}^{(n)T}\mathbf{S}_{W_{\mathcal{Y}}}^{(n)}\mathbf{U}^{(n)}\right)$$

$$= \underset{\{\mathbf{U}^{(n)}\}}{\arg\max}\ \mathrm{tr}\left(\mathbf{U}^{(n)T}\left(\mathbf{S}_{B_{\mathcal{Y}}}^{(n)} - \zeta \cdot \mathbf{S}_{W_{\mathcal{Y}}}^{(n)}\right)\mathbf{U}^{(n)}\right) \tag{16}$$

This can be treated as an eigenvalue problem and the objective function above is maximized only if $\tilde{\mathbf{U}}^{(n)}$ consists of the P_n eigenvectors of the total tensor scatter-difference $\left(\mathbf{S}_{B_{\mathcal{Y}}}^{(n)} - \zeta \cdot \mathbf{S}_{W_{\mathcal{Y}}}^{(n)}\right)$ associated with the P_n largest eigenvalues.

As described above, similar to MDA, GTDA aims to maximize between-class variation while minimizing within-class variation to achieve the best class separability. The only difference between the two is that MDA maximizes the ratio of the between-class scatter over the within-class scatter, whereas GTDA maximizes a tensor-based scatter difference criterion and the projection matrices in GTDA have orthonormal columns.

The advantages of GTDA are as follows: 1) GTDA operates on each modality of training tensors independently to reduce the small sample size problem (SSS) [20]. SSS means when the number of samples is smaller than the dimensionality of the input samples, the scatter matrix may become singular. The inverse of $\mathbf{S}_{W_{\mathcal{Y}}}^{(n)}$ can't be computed in MDA under SSS, whereas we don't need to solve the inverse of $\mathbf{S}_{W_{\mathcal{Y}}}^{(n)}$ in GTDA (see Equation 16), then the SSS problem can be avoided. 2) GTDA retains the discriminative information in the training tensors by considering the class label information. As a supervised learning algorithm, GTDA extracts the features that make class separation as large as possible by maximizing the difference of between-class variation and within-class variation of input tensors. 3) The optimization algorithm of GTDA converges, and the convergence will be proven in the experiment section.

Several issues that refer to the development and implementation of GTDA algorithm are discussed in the following sections. The initialization method, construction of termination criteria, convergence, and dimension issues, which were not described clearly in [14], are proposed in the following section.

3.3 Initialization, Convergence and Termination of GTDA

Considering the iterative nature of MSL solutions, initial estimations for the projection matrices are necessary at the beginning of the iterative process. Three popular initialization methods are commonly used for TTP, namely, pseudo-identity matrices, random matrices, and FPT [27]. As the description of GTDA in [14], the projection matrices were initialized simply by using pseudo-identity matrices of all ones. In Section 4.1, different initializations have been utilized to select the right one to obtain the best result. The empirical studies show that the result of GTDA is insignificantly influenced by the different initialization methods. However, different choices of initial method can influence the convergence speed for the iterative solution, and the FPT method seems better.

FPT is utilized to initialize the GTDA iterative with the mode-*n* projection matrix, which is achieved by truncating the mode-*n* full projection matrix. The mode-*n* full projection matrix refers to the multilinear projection for GTDA with $P_n = I_n$ fo $n = 1, \cdots, N$. Let $\{\mathbf{U}_n\}$ keeping the first P_n eigenvectors of

$$\Psi_{\mathcal{X}dif}^{(n)^*} = \sum_{c=1}^{C} M_c \left\| \bar{\mathcal{X}}_{c(n)} - \bar{\mathcal{X}}_{(n)} \right\| - \zeta \cdot \sum_{m=1}^{M} \left\| \mathcal{X}_{m(n)} - \bar{\mathcal{X}}_{c(n)} \right\| \tag{17}$$

Where $\mathcal{X}_{m(n)}$is the *n*-mode unfolding of input sample $\mathcal{X}_m$, M_c is the number for class *c*, $\bar{\mathcal{X}}_{(n)}$ and $\bar{\mathcal{X}}_{c(n)}$ equal to the overall mean and the class mean of $\mathcal{X}_{m(n)}$. It is obviously that $\Psi_{\mathcal{X}dif}^{(n)^*}$ is obtained by the input samples only and then $\mathbf{U}_n$ is determined as the eigenvectors of it directly without iteration. Consequently, FPT, which is faster than the other two initial methods, is expected to be a good choice to start the iterations in GTDA.

Unlike 2DLDA and MDA, the alternating projection optimization procedure for GTDA converges. This result is due to the fact that the alternating projection optimization procedure for GTDA is a monotonic increasing procedure. Therefore, the total scatter-difference $\Psi_{\mathcal{Y}_{dif}}$ is a nondecreasing function, and the function value is the lower and upper bound by two limiting values [14]. As demonstrated in Section 4.1, the proposed GTDA algorithm converges very quickly (within four iterations) for face images.

The total scatter difference function $\Psi_{\mathcal{Y}_{dif}}$ is used to determine the termination criterion. Moreover, given that $\Psi_{\mathcal{Y}_{dif(k)}}$and $\Psi_{\mathcal{Y}_{dif(k-1)}}$are the results from the *k*th and (*k* − 1)th iterations, if

$\left(\Psi_{\mathcal{Y}_{dif(k)}} - \Psi_{\mathcal{Y}_{dif(k-1)}}\right) < \varepsilon$, the iterative procedure will be terminated. According to the experiment results on the convergence of GTDA in Section 4.1, the maximum number of iterations *K* can be set with the number under four without affecting the final results.

3.4 Tensor Subspace Dimension Determination

The target dimensionality $\{P_n, n = 1, \cdots, N\}$ has to be determined before solving the GTDA projection. *Q*-based method is a simplified dimensionality determination procedure that needs no iteration and was first introduced in [11]. The ratio is defined

$$Q^{(n)} = \frac{\sum_{i_n=1}^{P_n} \breve{\lambda}_{i_n}^{(n)}}{\sum_{i_n=1}^{I_n} \breve{\lambda}_{i_n}^{(n)}} \tag{18}$$

as the remaining portion of the mode-*n* total scatter difference after truncation of the mode-*n* eigenvectors beyond the P_nth, where $\breve{\lambda}_{i_n}^{(n)}$ is the ith full-projection mode-*n* eigenvalues. In the *Q*-based method, the first P_n eigenvectors are kept in mode *n* so that $Q^{(1)} = Q^{(2)} = \cdots = Q^{(N)} = Q$.

The empirical study in [11] indicated that the *Q*-based method obtains results similar to those obtained by sequential mode truncation. Hence, applying the *Q*-based method on GTDA to reduce costly computations is preferred.

3.5 Feature Extraction and Classification Using MPCA Plus GTDA

In the problem of tensor sample recognition, the input training sample $\{\mathcal{X}_m\}$ in $\mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \cdots \otimes \mathbb{R}^{I_N}$ is projected through MPCA and GTDA which both use the FPT for initialization and use the *Q*-based method for dimension determination. With the learned $\{\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times P_n}\}$, a series of tensor $\mathcal{Y}_m$ in a low-dimensional tensor space $\mathbb{R}^{P_1} \otimes \mathbb{R}^{P_2} \cdots \otimes \mathbb{R}^{P_N}$ can be obtained with $\mathcal{Y}_m = \mathcal{X}_m \times_1 \mathbf{U}^{(1)^T} \times_2 \mathbf{U}^{(2)^T} \cdots \times_N \mathbf{U}^{(N)^T}$.

Due to the distance of two samples is identical whether with the form tensor or vector, the NN network can be used for the final classification. The distance between two arbitrary feature tensors is defined as

$$dist(\mathcal{Y}_i, \mathcal{Y}_j) = \|\mathcal{Y}_i - \mathcal{Y}_j\|_F = \|vec(\mathcal{Y}_i) - vec(\mathcal{Y}_j)\|_2 \tag{19}$$

where the Frobenius norm equals the Euclidean distance. Consider that a set of *M* tensor feature samples $\{\mathcal{Y}_1, \mathcal{Y}_2, \cdots, \mathcal{Y}_M\}$ are available for training, and *C* classes exist in total. The class label for the *m*th samples $\mathcal{Y}_m$ is $c_m = \mathbf{c}(m)$. If a testing feature sample $\mathcal{Y}$ has

$$dist(\mathcal{Y}, \mathcal{Y}_m) = \min_{i=1,\dots,M}\{dist(\mathcal{Y}, \mathcal{Y}_i)\}, \mathcal{Y}_m \in c_m \quad (20)$$

then we can judge $\mathcal{Y} \in c_m$. Although the classification accuracy is expected to improve if a more complicated classifier, such as support vector machine (SVM) [28] or artificial neural network (ANN) [29], is applied, NN classifier is preferred in this paper because of the focus is on the performance mainly contributed by MPCA + GTDA based feature extraction.

4. Experiments and Discussion

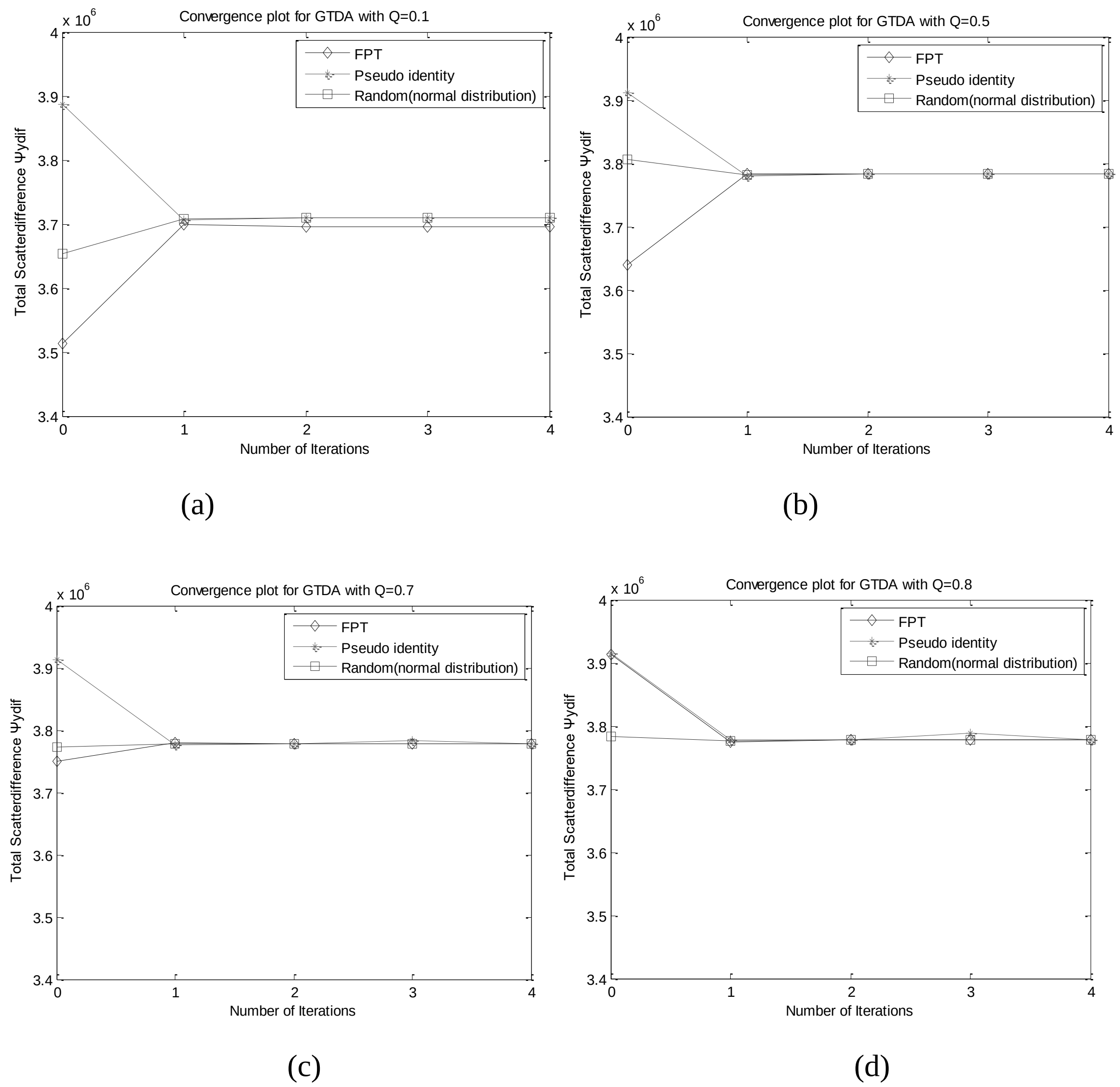
In this section, several experiments are designed to evaluate the performance of the proposed algorithm. The first one uses ORL face database which obtains 400 face images to discuss the effects of the initial conditions of GTDA. The following experiments which are carried out on gait, and color face databases illustrate the efficacy of the proposed MPCA+GTDA in tensor object recognition and compare its performance against state-of-the-art algorithms.

4.1 Initialization of GTDA on ORL Database

To study the GTDA properties on biometric data of different characteristics, the ORL database with a total of 400 different face images is applied. The gray-level face images from the ORL database have a resolution of 112 × 92, and 40 subjects with 10 images each included in the database. Images of individuals have been selected based on different characteristics, such as with or without glasses, different facial expressions, and facial details.

First, the effects of the initial conditions of GTDA are examined by using the face images from ORL. The distribution of random matrices initialization is standard normal distribution. Next, all three normal initialization methods (random matrices, pseudo-identity matrices, and full projection truncation) have been considered. Based on the simulation studies in Fig.5, the algorithm converges to the same point (obtaining the same total scatter difference $\Psi_{\mathcal{Y}\mathrm{dif}}$) within four iterations when $Q \geq 0.5$, despite the different initializations. Furthermore, with a small value of *Q* (= 0.1) for GTDA, the algorithm that uses FPT as initialization converges to a point slightly lower than the point that is converged by using the other two initialization methods [Fig.5 (a)].

This result suggested that the initialization method could affect the final results of GTDA on biometric data when a small number is used by Q. However, in most cases, most variations or energy in original data is needed for application in pattern recognition field, and $Q \geq 0.5$ can be easily satisfied. Considering that GTDA seems insensitive to the choice of initial method with $Q \geq 0.5$, the FPT, which needs no iteration and has a fast calculation speed, is employed for initialization in the following work.



(a) (b) (c) (d)

Fig.5. Convergence plot for GTDA with different initializations: (a) Convergence plot with Q=0.1; (b) Convergence plot with Q=0.5; (c) Convergence plot with Q=0.7; (d) Convergence plot with Q=0.8.

4.2 MPCA Plus GTDA Based Gait Recognition

Gait video is naturally a third-order tensor with the column, row, and time mode. In this subsection, the experiment is conducted on the gait base and the correct classification recognition (CCR) is employed to evaluate the performance of algorithms. The gait database we selected is one of the USF HumanID "gait challenge" data sets [30] which was used in [11]. The selected gait database contains a total of 71 subjects and 731 human gait samples with a size of $32 \times 22 \times 10$ for preliminary evaluation. An average of roughly 10 samples is selected for each subject. The first four samples from each subject (284 in total) are used for training, and the remaining 447 samples are used for testing. In this test, we use the first *P* numbers of the extracted features from each mode of the training samples to calculate the performance of each method. The scale of *P* is from 2 to 10, and the best recognition results are shown in bold. The CCR of MPCA+GTDA, MPCA, and GTDA on gait recognition increases as the number of extracted features increases (Fig.6). This monotonicity is due to the convergence of MPCA and GTDA. By contrast, the curves of MDA and MPCA+MDA demonstrate the instability of MDA. Moreover, applying GTDA after MPCA can obtain higher CCR than using MPCA only for feature extraction.

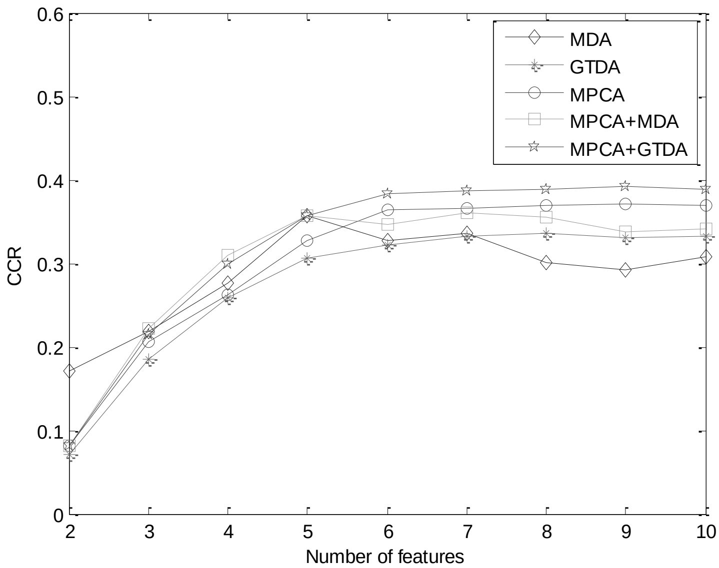


Fig.6. The CCR over dimensions of feature on the gait database.

4.3 MPCA Plus GTDA Based Color Face Recognition

Color facial images in the RGB color space can be viewed as third-order objects with column, row, and color modes [31]. Let an RGB image of size $I_1 \times I_2$ be represented as a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, where the mode-3 of $\mathcal{X}$ is the color variables that correspond to **R**, **G**, and **B** with

$I_3 = 3$ [32] (see Fig.7). Several algorithms have been proposed to deal with the color images as tensor objects. For example, tensor discriminant color space (TDCS) [32] seeks two discriminant transformation matrices $\mathbf{U}_1$, $\mathbf{U}_2$ that correspond to facial spatial information and one color space projection matrix $\mathbf{U}_3$ corresponding to the color space. In fact, TDCS applies LDA transformation on each mode of the third-order color image similar to MDA. In [33], fusion tensor subspace transformation (FTSA) which applies LDA on the facial spatial information and applies independent color analysis (ICA) on the color space information, is proposed. Considering the small dimension of mode-3 ($I_3 = 3$), applying dimensionality reduction on the mode-3 of input tensor $\mathcal{X}$ seems unnecessary. To study the performance of the proposed MPCA+GTDA method in the color face database, we initially apply MPCA on the mode-1 and mode-2 of color images to reduce the dimensionality of the facial spatial information. Then, GTDA is used for all modes of the processed third-order tensor. For a color face data $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, the proposed method aims to seek two linear projection matrices $\mathbf{U}_1 \in \mathbb{R}^{I_1 \times P_1}$, $\mathbf{U}_2 \in \mathbb{R}^{I_2 \times P_2}$ and a color transformation matrix $\mathbf{U}_3 \in \mathbb{R}^{I_3 \times P_3}$ ($P_1 < I_1,\ P_2 < I_2 and\ P_3 \le I_3$) for calculation

$$\mathcal{Y} = \mathcal{X} \times_1 \mathbf{U}_1^T \times_2 \mathbf{U}_2^T \times_3 \mathbf{U}_3^T \tag{21}$$

where $\mathbf{U}_1$ and $\mathbf{U}_2$ are gained by applying MPCA plus GTDA, $\mathbf{U}_3$ is achieved by using GTDA.

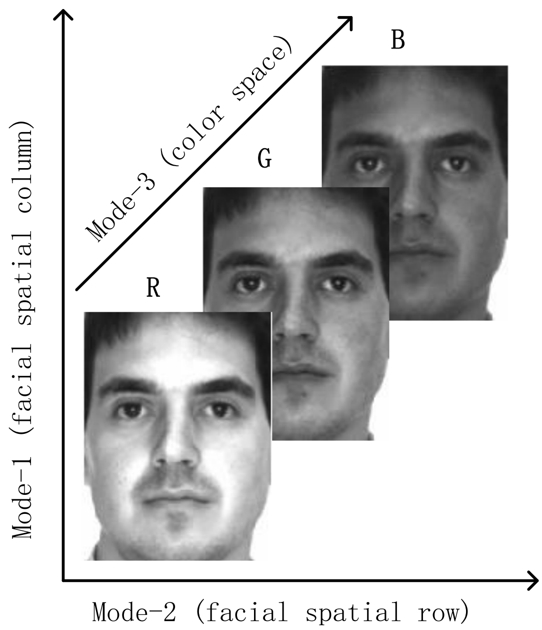


Fig.7. Illustration of color component image.

We conducted the experiment on the widely used color face database AR [34]. The AR color face database consists of over 4000 color facial images of 126 people. In this test, 1400 face images from 100 people (50 women and 50 men) were selected. These images are selected with different facial expressions (open/closed eyes, smiling/not smiling, open/closed mouth) and facial details (glasses/no glasses). There are no restrictions on wear, hair style, occlusion, etc. All color face images are manually cropped to $54 \times 40 \times 3$ pixels. Figure 8 depicts sample face images from one individual in this AR database. Like FERET database, 7 images were randomly selected from each subject to form the training set and the remaining images were used for testing. We calculated the average recognition rate by over ten such random repetitions.

Fig.8. Example images of one individual in AR database

In this experiment, we trained MPCA+GTDA, MPCA+MDA, GTDA, MDA, and MPCA. The transformation matrices $\mathbf{U}_1$ and $\mathbf{U}_2$ were initialized by FPT and $\mathbf{U}_3$ was initialized by a matrix $\mathbf{U} \in \mathbb{R}^{3\times3}$ of all $\frac{1}{3}$s. The maximum number of iterations K was set to be three. Figure 9 shows the recognition rates versus the variations of the dimensions. Again, MPCA+GTDA and MPCA+MDA perform better than the other methods on the color face database and the recognition rates of MPCA+GTDA and MPCA+MDA are quite similar. Therefore, the proposed MPCA+GTDA algorithm can improve the performance on tensorial signal recognition.

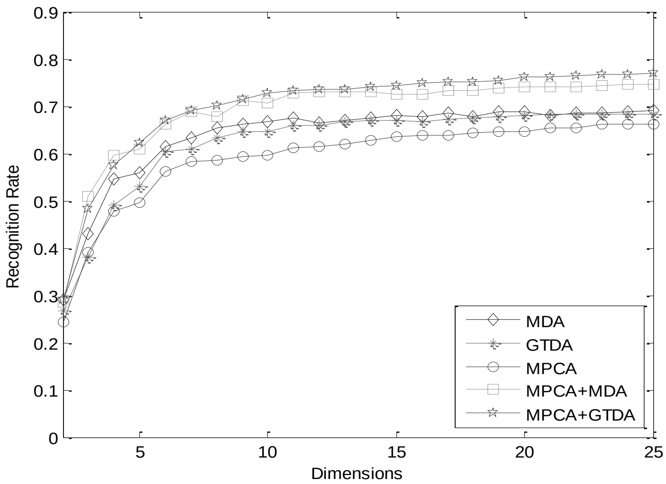


Fig.9. The recognition rates versus the dimension on the AR database.

5. Conclusion

This paper proposed a new MPCA+GTDA framework to supervise dimensionality reduction in tensorial biometric signal. This algorithm overcomes the limitation of MPCA+MDA and captures better convergence and recognition accuracy. Moreover, a few fundamental issues of GTDA, including initialization, termination, and dimension determination, are discussed in detail. The experimental results of the comparison with the state-of-the-art algorithms indicated that the MPCA+GTDA method is a promising tool for biometric signal in research and applications.

Acknowledgements

The research of this paper is sponsored jointly by the program for innovative research team i n Science and Technology in Universities of Henan Province under Grant No. 13IRTSTHN023, National Major Science and Technology Project under Grant No. 2014ZX03005001, and Zhengzhou Science and Technology Research Plan under Grant no. 153PKJGG113.

References

1. M.H.C. Law, A.K. Jain, Incremental nonlinear dimensionality reduction by manifold learning, 2006, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 3, pp. 377-391.
2. C. Liu, W. Xu, Q. Wu, TKPCA: Tensorial kernel principal component analysis for action recognition, 2013, Mathematical Problems in Engineering, article ID. 816836.

3. L. De Lathauwer, B. De Moor, J. Vandewalle, On the best rank-1 and rank-($R_1, R_2, \cdots, R_N$) approximation of higher-order tensors, 2000, SIAM Journal on Matrix Analysis and Applications, vol. 21, no. 4, pp. 1324-1342.
4. I. Jolliffe, Principal component analysis, Wiley Online Library, 2005.
5. P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. fisherfaces: Recognition using class specific linear projection, 1997, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720.
6. X. He, S. Yan, Y. Hu, P. Niyogi, H.J. Zhang, Face recognition using Laplacianfaces, 2005, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 3, pp. 328-340.
7. M. Belkin, P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, 2003, Neural Computation, vol. 15, no. 6, pp. 1373–1396.
8. J. Yang, D. Zhang, A.F. Frangi, J.Y. Yang, Two-dimensional PCA: a new approach to appearance-based face representation and recognition, 2004, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 1, pp. 131-137.
9. M. Li, B.Z. Yuan, 2D-LDA: a statistical linear discriminant analysis for image matrix, 2005, Pattern Recognition Letters, vol. 26, no. 5, pp. 527-532.
10. S. Chen, H. Zhao, M. Kong, B. Luo, 2D-LPP: a two-dimensional extension of locality preserving projections, 2007, Neurocomputing, vol. 70, no. 4, pp. 912-921.
11. H. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, MPCA: Multilinear principal component analysis of tensor objects, 2008, IEEE Transactions on Neural Networks, vol. 19, no. 1, pp. 18–39.
12. S. Yan, D. Xu, Q. Yang, L. Zhang, X. Tang, H.J. Zhang, Multilinear discriminant analysis for face recognition, 2007, IEEE Transactions on Image Processing, vol. 16, no. 1, pp. 212–220.
13. D. Tao, X. Li, X. Wu, S. Maybank, General tensor discriminant analysis and gabor features for gait recognition, 2007, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 10, pp. 1700–1715.
14. X. He, D. Cai, P. Niyogi, Tensor subspace analysis, 2005, Advances in Neural Information Processing Systems, pp. 499-506.
15. F. Song, D. Zhang, D. Mei, Z. Guo, A multiple maximum scatter difference discriminant criterion for facial feature extraction, 2007, IEEE Transactions on Systems, Man, and Cybernetics, Part B:Cybernetics, vol. 37, no. 6, pp. 1599-1606.
16. S. Kong, D. Wang, A report on multilinear PCA plus multilinear LDA to deal with tensorial data: visual classification as an example, arXiv preprint, arXiv. 1203.0744v1, 2012.

17. A.A.S. Baboli, G.Rezai-rad, A.S. Baboli, MPCA+DATER: a novel approach for face recognition based on tensor objects, 2010, 18th Telecommunications forum.
18. A.S. Baboli, S.M.H. Nia, A.A.S. Baboli, A new method based on MDA to enhance the face recognition performance, 2011, International Journal of Image Processing, vol. 5, no. 1, pp. 69-77.
19. S.M. Hosseyninia, F. Roosta, A.A.S. Baboli, Gholamali, Improving the performance of MPCA+MDA for face recognition, 2011, 19th Iranian Conference on Electrical Engineering.
20. J. Li, L. Zhang, D. Tao, H. Sun, Q. Zhao, A prior neurophysiologic knowledge free tensor-based scheme for single trial EEG classification, 2009, IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 17, no. 2, pp. 107-115.
21. H. Lu, K.N. Plataniotis, A. N. Venetsanopoulos, A taxonomy of emerging multilinear discriminant analysis solutions for biometric signal recognition, 2009, Biometrics: Theory, Methods, and Application, Wiley/IEEE, pp. 21-45.
22. T.G. Kolda, B.W. Bader, Tensor decompositions and applications, 2009, SIAM Review, vol. 51, no. 3, pp. 455-500.
23. Q. Zhao, C.F. Caiafa, D.P. Mandic, L.Zhang, T.Ball, A.SchulzeBonhage, A.S. Cichocki, Multilinear subspace regression: an orthogonal tensor decomposition approach, 2011, Advances in Neural Information Processing Systems, pp.1269-1277.
24. H. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, A survey of multilinear subspace learning for tensor data, 2011, Pattern Recognition, vol. 44, no. 7, pp. 1540-1551.
25. L. de Lathauwer, B. de Moor, J. Vandewalle, A multilinear singular value decomposition, 2000, SIAM Journal on Matrix Analysis and Applications, vol. 21, no. 4, pp. 1253-1278.
26. P. Kroonenberg, J. Leeuw, Principal component analysis of three-mode data by means of alternating least squares algorithms, 1980, Psychometrika , vol. 45, no. 1, pp. 69-97.
27. H. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, Multilinear Subspace Learning: Dimensionality Reduction of Multidimensional Data, 2013, CRC press, pp. 93-96.
28. J. Qin, Z.S. He, A SVM face recognition method based on Gabor-featured key points, Processing of 4th IEEE Conference on Machine Learning and Cybernetics, 2005, pp. 5144-5149.
29. M. J. Er et al, Face recognition with radial basis function (RBF) neural networks, 2002, IEEE Transactions on Neural Networks, vol. 13, no. 3, pp. 697-710.

30. S. Sarkar, P.J. Phillips, Z. Liu et al, The human ID gait challenge problem: Data sets, performance, and analysis, 2005, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 2, pp. 162-177.
31. J. Yang, C. Liu, Color image discriminant models and algorithms for face recognition, 2008, IEEE Transactions on Neural Networks, vol. 19, no. 12, pp. 2088-2098.
32. S.J. Wang, J. Yang, Tensor discriminant color space for face recognition, 2011, IEEE Transactions on Image Processing, vol. 20, no. 9, pp. 2490-2501.
33. S.J. Wang, C.G Zhou, X. Fu, Fusion tensor subspace transformation framework, 2013, Plos One, vol. 8, no. 7, Article ID e66647.
34. A.M. Martinez, R. Benavente, The AR face database, 1998, CVC Technical Report, vol. 24.