

HYPERSPECTRAL BIOLOGICAL IMAGES COMPRESSION BASED ON MULTIWAY TENSOR PROJECTION

Bo Du¹, Mengfei Zhang¹, Lefei Zhang^{1*}, Xuelong Li²

¹ Computer School, Wuhan University

² Center for OPTIMAL, State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences

Email: {remoteking,mzhang,zhanglefei}@whu.edu.cn, xuelong_li@opt.ac.cn

ABSTRACT

Since the hyperspectral images (HSI) could provide much more useful discriminative information that cannot be obtained by the conventional imaging techniques, the hyperspectral imaging technology was widely used in remote sensing area and recently used in many other aspects, such as the biological images recognition. However, most of the time, the size of hyperspectral data is so large that to process these data is both time-consuming and space-consuming. In this paper, a multiway tensor projection (MTP) algorithm is proposed as an extension to the conventional PCA for hyperspectral data compression and reconstruction. Technologically speaking, MTP carries out a tensor data compression in all the modes simultaneously to seek a projection matrix along each order to make sure that the projected core tensor can preserve most of the information present in the original tensor. Since the MTP algorithm uses the arbitrary order tensor as the input, it can preserve the structure information not only among the rows and columns but also among the spectral channels as much as possible and without vectorization. Numerous experiments on hyperspectral biological databases show that the MTP algorithm has better compression performance than PCA in many aspects.

Index Terms— Hyperspectral images, tensor projection, principle component analysis, image compression

1. INTRODUCTION

Hyperspectral imaging was previously used in remote sensing area and recently introduced in many other aspects [1]. Compared to the conventional gray level images which are commonly presented as 2D matrices, the hyperspectral images (HSI) appear as 3D data cubes which have two spatial dimensions (width and height) and a spectral dimension [2, 3]. Therefore, for the pattern recognition tasks such as

remote sensing image classification and human face recognition, the HSI could provide much more useful discriminative information that cannot be obtained by the conventional imaging methods [4]. To this end, some hyperspectral biological image databases such as hyperspectral face database and hyperspectral palm database are recently released in the literature [5].

Unfortunately, most of the time, the size of hyperspectral data is so large that to process these data becomes a time-consuming and space-consuming task. In order to address this problem, similar to the framework of grey level image processing, the image compression is therefore considered as an important preprocessing step before the subsequent recognition. The most frequency employed approach for the grey level image compression is JPEG coding and wavelet transform coding (WTC) [6]. These methods are reported to work well in most of the time. However, since they are originally designed for gray level images, when it comes to the HSI, they have to process each channel of the data cube independently [7]. By this way, the correlation among the spectral channels of HSI would be lost during the compression and cause the potential inaccurate in the subsequent recognition.

Apart from the aforementioned image coding perspective, the data dimensionality reduction (DR) techniques could also be adopted for grey level image as well as HSI compression, the most standard algorithm of which is principle component analysis (PCA) [8]. The basic engine of PCA algorithm is the Karhunen-Loeve transform which finds the principal components of the input data in accordance with the maximum variance. Note that the input data of PCA must be a set of 1D-vectors, which indicates that each image (2D grey level image or 3D HSI) have to be vectorized in order to compute the covariance matrix. It is worth noting that such vectorization obviously breaks the data structure of the data matrix or data cube, which is useful information for data reconstruction. On the other hand, the size of the covariance matrix using the vectorized data is always very large (more than 10^4), thus it is time-consuming and space-consuming to perform the eigen-decomposition.

*Corresponding author: L. Zhang. This work is supported by the National Basic Research Program of China (973 program) (Grant Nos: 2012CB719905 and 2011CB707105), and the National Natural Science Foundation of China (Grant Nos: 61125106, 61102128).

Some improvements of the PCA algorithm are as following [9]. Zhang *et al.* put forward a matrix K-L transform to reduce the dimension of the covariance matrix and the method can improve the speed compared to standard PCA [10]. Yang *et al.* proposed a 2DPCA which could process the 2D matrices directly without vectorization, so the data structure is partly preserved in the 2D representation and the dimension of the covariance matrix is much smaller than PCA as well [11]. Since the 2DPCA only considers the correlations only among the row vector, Zhang *et al.* further suggested the (2D)²PCA which considers the coefficient both among the row vectors and column vectors [12]. Although the abovementioned extensions of PCA can be directly used for gray level images, as these methods do not consider the relation among the adjacent spectral channels so far, they still cant deal with the HSI directly and efficiently.

In this paper, we treat the hyperspectral biological images as 3-order-tensors and thus does not destroy the intrinsic spatial-spectral relationship of the data structure information within the data cubes. Based on the tensor representation, a multiway tensor projection (MTP) algorithm is proposed as an extension to the conventional PCA for hyperspectral data cubes compression and reconstruction. In MTP, the tensor object is compressed in all the modes simultaneously by using a projection matrix in each order and thus the projected core tensor could preserve most of the information present in the original tensor. Since the MTP algorithm could accept the arbitrary order tensor as inputs, it can preserve the structure information that not only between rows and columns but also between bands as much as possible and without vectorization.

The remainder of this paper is organized as follows: in Section 2, we briefly provide the background of related tensor algebra, and then present the proposed MTP algorithm in detail. The experiment results on the two public hyperspectral biological image databases are discussed in Section 3, followed by the conclusions in Section 4.

2. MULTIWAY TENSOR PROJECTION

In this paper, the vectors are denoted by lowercase bold italics, e.g., \mathbf{a} , the matrices are denoted by uppercase bold italics, e.g., \mathbf{A} , and the tensors are denoted by calligraphic letter, e.g., \mathcal{A} . The elements of an object (including the vector, matrix, and tensor) are denoted with indices in blankets where the indices are denoted by lowercase letters, e.g., $\mathcal{A}(i, j)$. Before introducing the MTP algorithm, we give some related tensor algebra as follows, further details are available in papers [13–16].

Suppose we have an N -order tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$, which is addressed by N indices $i_n, n = 1, \dots, N$, while i_n addresses the size of n th-order of tensor \mathcal{A} . An element of tensor \mathcal{A} is expressed as is $\mathcal{A}(i_1, i_2, \dots, i_N)$ where $1 \leq i_n \leq I_n$. Particularly, a scalar is a 0-order tensor, a vector is a 1-order tensor, and a matrix is a 2-order-tensor. In this paper, the

hyperspectral data cube is therefore represented as a 3-order-tensor where all the spatial-spectral constraints are well preserved. The n th-order tensor-matrix product of a tensor \mathcal{A} by matrix $\mathbf{U} \in \mathbb{R}^{J_n \times I_n}$ is denoted by $\mathcal{A} \times_n \mathbf{U}$, in which the entries $\mathcal{A} \times_n \mathbf{U}(i_1, i_2, \dots, i_{n-1}, j_n, i_{n+1}, \dots, i_N) = \sum_{i_n} \mathcal{A}(i_1, i_2, \dots, i_n) \mathbf{U}(j_n, i_n)$. It can be inferred that there is a precondition that the width of the matrix \mathbf{U} must be the same as the n th-order of the original tensor, and such the multiplication result is a tensor of the same as the original tensor except for the n th-order which is changed to the height of the matrix \mathbf{U} . The scalar product of two tensor $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ is defined as $\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i_1} \sum_{i_2} \dots \sum_{i_N} \mathcal{A}(i_1, i_2, \dots, i_N) \mathcal{B}(i_1, i_2, \dots, i_N)$, and the Frobenius norm of a tensor \mathcal{A} is defined as $\|\mathcal{A}\| = \sqrt{\|\mathcal{A}\|_F}$. The matrix unfolding $\mathcal{A}_{(n)} \in \mathbb{R}^{I_n \times (I_1 I_2 \dots I_{n-1} I_{n+1} \dots I_N)}$ of \mathcal{A} along the n th order contains the element $\mathbf{a}_{i_1 i_2 \dots i_N}$ at the position with row number i_n and column number equal to $1 + \sum_{k=1, k \neq n}^N (i_k - 1) \prod_{j=k+1, j \neq n}^N I_j$. The column vectors of $\mathcal{A}_{(n)}$ are the n th-order vectors of \mathcal{A} , while an n th-order vectors is defined as an I_n -dimensional vector obtained from \mathcal{A} by varying the index i_n while keeping all the other indices fixed. The rank-1 tensor is a tensor that consists of the outer product of a number of vectors. According to definition of tensor-matrix product, it can be inferred that any tensor \mathcal{A} can be approximately expressed as the core tensor \mathcal{C} multiplied by a series of projection matrices, i.e., $\mathcal{A} = \mathcal{C} \times_1 \mathbf{U}^{(1)T} \times_2 \mathbf{U}_n^{(2)T} \times \dots \times_N \mathbf{U}^{(N)T}$, where $\mathcal{C} = \mathcal{A} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times \dots \times_N \mathbf{U}^{(N)}$ and $\mathbf{U}^{(n)} = (\mathbf{u}_1^n, \mathbf{u}_2^n, \dots, \mathbf{u}_{I_n}^n)$ is an orthogonal $I_n \times I_n$ matrix. The matrix representation of the aforementioned tensor-matrix product can be obtained by unfolding \mathcal{A} and \mathcal{C} on the n th order, i.e., $\mathcal{C}_{(n)} = \mathbf{U}^{(n)} \mathcal{A}_{(n)} (\mathbf{U}^{(n+1)})^T \otimes \mathbf{U}^{(n+2)} \otimes \dots \otimes \mathbf{U}^{(N)} \otimes \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \otimes \dots \otimes \mathbf{U}^{(N-1)T}$, in which \otimes denotes the Kronecker product [17].

Different from the previous works which dealt with 2-D matrices input data (gray-level images), the proposed MTP algorithm aims to process the hyperspectral images which are in form of 3-order-tensors. Compared to PCA, the MTP algorithm is another way for data dimension reduction which could find the most representative multiway subspace for high order data representation. The compressed core tensor has much smaller size in each order.

Here we firstly give a brief flowchart of PCA algorithm for images compression and reconstruction, which is highly related to our proposed method. In PCA, the input images are vectorized into 1-D vectors. Let $\{\mathbf{A}_m \in \mathbb{R}^{L_1 \times L_2}\}$ to be the input image set and $\{\mathbf{a}_m \in \mathbb{R}^L, m = 1, \dots, M\}$ to be the vectorized inputs $L = L_1 \times L_2$. PCA finds a projection matrix $\mathbf{W} \in \mathbb{R}^{L \times L}$ to linear transform the data from \mathbf{a}_m to \mathbf{b}_m in which $\{\mathbf{b}_m \in \mathbb{R}^{L'}, m = 1, \dots, M\}$ is the compressed data vectors in the low dimensionality feature space, i.e., $L' < L$, and thus the images are compressed. This linear projection could be written as:

$$\mathbf{b}_m = \mathbf{a}_m \mathbf{W}^T, m = 1, \dots, M \quad (1)$$

After that, the reconstructed data vectors $\{\mathbf{r}_m \in \mathbb{R}^L, m = 1, \dots, M\}$ are obtained by the following backward projection:

$$\mathbf{r}_m = \mathbf{b}_m \mathbf{W}^T, m = 1, \dots, M \quad (2)$$

It is known that the mapping matrix \mathbf{W} is composed by the leading eigenvectors corresponding to the L' largest eigenvalues of the covariance matrix $\mathbf{C} = \sum_{m=1}^M (\mathbf{a}_m - \bar{\mathbf{a}})(\mathbf{a}_m - \bar{\mathbf{a}})^T$, where $\bar{\mathbf{a}}$ is the mean vector of the image set. PCA algorithm follows the assumption that although the observed images lie in a high dimensionality space \mathbb{R}^L , they are substantially embedded in a much lower dimensionality feature space $\mathbb{R}^{L'}$, thus the original images can be represented by the smaller variables while keeping the most of important information. In fact, since the HSI could be viewed as 3-order-tensors without vectorization, the original tensor data should be embedded in a lower dimensional tensor subspace which captures the most of the variation in the original tensor space. According to this fact, we have the following MTP algorithm which directly processes the high-order-tensors and thus the original tensor data is compressed on all orders simultaneously.

Suppose $\mathcal{A}_m (m = 1, \dots, M)$ is the tensor data set resides in the tensor space $\mathbb{R}^{L_1} \otimes \mathbb{R}^{L_2} \otimes \dots \otimes \mathbb{R}^{L_N}$, MTP aims to compress it to the reduced tensor data set $\mathcal{B}_m (m = 1, \dots, M)$, which resides in the tensor space $\mathbb{R}^{L'_1} \otimes \mathbb{R}^{L'_2} \otimes \dots \otimes \mathbb{R}^{L'_N}$, $L'_n < L_n, n = 1, 2, \dots, N$. L'_n for each mode is determined in practice. The proposed multiway tensor compression can be expressed as the following multi-linear projection:

$$\mathcal{B}_m = \mathcal{A}_m \times_1 \mathbf{W}_1^T \times_2 \mathbf{W}_2^T \times \dots \times_N \mathbf{W}_N^T, m = 1, \dots, M \quad (3)$$

where $\mathbf{W}_n^T (n = 1, \dots, N)$ is a $L_n \times L'_n$ matrix containing L'_n orthonormal columns. The reconstructed tensors $\mathcal{R}_m \in \mathbb{R}^{L_1 \times L_2 \times \dots \times L_N}$ can be computed by:

$$\mathcal{R}_m = \mathcal{B}_m \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times \dots \times_N \mathbf{W}_N, m = 1, \dots, M \quad (4)$$

In order to obtain the optimal MTP matrices $\mathbf{W}_N, n = 1, \dots, N$, we need to further define the scatter of a tensor data set. Here we take the reduced tensor data set $\mathcal{B}_m \in \mathbb{R}^{L'_1 \times L'_2 \times \dots \times L'_N}, m = 1, \dots, M$ as an example. Similar to the definitions in the vector space, the mean tensor of $\mathcal{B}_m, m = 1, \dots, M$ is $\bar{\mathcal{B}} = (1/M) \sum_{m=1}^M \mathcal{B}_m$, and the total scatter of the data set is defined as $\omega_{\mathcal{B}} = \sum_{m=1}^M \|\mathcal{B}_m - \bar{\mathcal{B}}\|^2$. In the n th mode, n th-order total scatter matrix is defined as $S_{\mathcal{B}}^{(n)} = \sum_{m=1}^M (\mathcal{B}_{m(n)} - \bar{\mathcal{B}}_{m(n)})(\mathcal{B}_{m(n)} - \bar{\mathcal{B}}_{m(n)})^T$ in which $\mathcal{B}_{m(n)}$ is the n th-mode unfolded matrix of \mathcal{B}_m .

Now we provide the objective function of the proposed MTP algorithm. The basic idea of MTP algorithm is that the larger of total scatter in the compressed subspace, the more discriminative information is preserved in the reduced core tensor subsets. Followed by this idea, the multilinear core transformation matrices optimized by MTP should try to maximize the total tensor scatter in the reduced tensor space:

$$\{\mathbf{W}_n, n = 1, \dots, N\} = \arg \max_{\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_N} \omega_{\mathcal{B}} \quad (5)$$

By putting (3) into (5), we have:

$$\{\mathbf{W}_n, n = 1, \dots, N\} = \arg \max_{\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_N} \sum_{m=1}^M \|\mathcal{B}_m - \bar{\mathcal{B}}\|^2 \quad (6)$$

To the best of our knowledge, there exists no closed form solution for the maximization (6). In this paper, we propose to optimize (6) with an iterative approach according to the higher order tensor decomposition. It is known that if $(\mathbf{W}_1, \dots, \mathbf{W}_{n-1}, \mathbf{W}_{n+1}, \dots, \mathbf{W}_N)$ are given, then the only left projection matrix \mathbf{W}_n could be globally optimized by combine the top L'_n eigenvectors which corresponding to the largest L'_n eigenvalues of the covariance matrix $\mathbf{C}^n = \sum_{m=1}^M (\mathbf{A}_{m(n)} - \bar{\mathbf{A}}_{(n)})(\mathbf{A}_{m(n)} - \bar{\mathbf{A}}_{(n)})^T$, where $\mathbf{A}_{m(n)}$ is the n -order unfolded matrix of $\mathcal{A}_m^{(n)}$, and $\mathcal{A}_m^{(n)}$ is defined as $\mathcal{A}_m^{(n)} = \mathcal{A}_m \times_1 \mathbf{W}_1^T \times \dots \times_{n-1} \mathbf{W}_{n-1}^T \times_{n+1} \mathbf{W}_{n+1}^T \times \dots \times_N \mathbf{W}_N^T$, i.e., $\mathcal{A}_m^{(n)}$ is a tensor that all the orders of the original tensor are compressed except the n -mode. $\bar{\mathbf{A}}_{(n)}$ is the unfolded matrix of $\bar{\mathcal{A}}$, while $\bar{\mathcal{A}}$ is the mean tensor of the input M tensors. This fact indicates that if we set a suitable initial value of the projection matrices, the objective function (6) could be locally optimized. Furthermore, the convergence of such iterative approach is guaranteed as reported in some tensor papers [18, 19]. The procedure of MTP algorithm is summarized as follows:

According to the above procedure, MTP is the tensor extension of PCA which performs PCA on each order simultaneously. If the input data is a set of 1-order tensors, i.e., a set of vectors, then MTP and PCA are equivalent. Compared to conventional PCA, the proposed MTP has the following advantages: (1) MTP uses much less parameters than PCA to construct the projection model. Suppose a tensor $\mathcal{A} \in \mathbb{R}^{L_1 \times L_2 \times \dots \times L_N}$, the compressed small size tensor is $\mathcal{B} \in \mathbb{R}^{L'_1 \times L'_2 \times \dots \times L'_N}$, the projection matrices are $\{\mathbf{W}_n \in \mathbb{R}^{L_n \times L'_n}, n = 1, \dots, N\}$. MTP algorithm is directly process on \mathcal{A} , while PCA firstly reshape the tensor data into a vector, i.e., $\mathbf{a} \in \mathbb{R}^{L_1 L_2 \dots L_N}$, and then compress the data to the reduced feature vector $\mathbf{b} \in \mathbb{R}^{L'_1 L'_2 \dots L'_N}$. The projection matrix of PCA is $\mathbf{W} \in \mathbb{R}^{(L_1 L_2 \dots L_N) \times (L'_1 L'_2 \dots L'_N)}$. Obviously, the size of \mathbf{W} is much larger than the sum of size $\{\mathbf{W}_n, n = 1, \dots, N\}$. In other words, the MTP algorithm needs much less parameters than PCA to build the projection model, thus MTP can avoid the small sample size problem in practice. (2) MTP algorithm can preserve much more spatial-spectral constraints of the HSI. It is mentioned above that PCA algorithm have to firstly transform the HSI into an 1D vector, which destroys the structure information of HSI, since the adjacent rows, columns and spectral channels are related and the potential information behind these relationship are useful for optimize the projection

Algorithm 1 procedure of MTP algorithm

Input: Given a tensor data set $\{\mathcal{A}_m \in \mathbb{R}^{L_1 \times L_2 \times \dots \times L_N}, m = 1, \dots, M$, the compressed lower dimension $L'_n, n = 1, \dots, N$, and the maximal iteration number T_m ;

Initialize For $n = 1, \dots, N$, calculate the eigen-decomposition of the matrix $\mathbf{C}^0 = \sum_{m=1}^M (\mathcal{A}_{m(n)} - \bar{\mathcal{A}}_{(n)}) (\mathcal{A}_{m(n)} - \bar{\mathcal{A}}_{(n)})^T$, where $\mathcal{A}_{m(n)}$ is the n -order unfolded matrix of tensor \mathcal{A}_m . Then sort the eigenvalues of this matrix and choose L_n eigenvectors corresponding to the leading eigenvalues as \mathbf{W}_n^0 .

repeat

- For $n = 1$ to N

 Compute $\mathcal{A}_m^{(n)} = \mathcal{A}_m \times_1 \mathbf{W}_1^T \times \dots \times_{n-1} \mathbf{W}_{n-1}^T \times_{n+1} \mathbf{W}_{n+1}^T \times \dots \times_N \mathbf{W}_N^T$;

 Unfold $\mathcal{A}_m^{(n)}$ to $\mathcal{A}_{m(n)}$;

 Compute $\mathbf{C}^n = \sum_{m=1}^M \mathcal{A}_{m(n)} \mathcal{A}_{m(n)}^T$;

 Select the top L'_n eigenvectors to compose the column vectors of \mathbf{W}_n^t .

- End

until Convergence or the iteration number t reaches T_m

Output: The multiway tensor projection matrices $\{\mathbf{W}_n, n = 1, \dots, N\}$, and the compressed tensor data set $\{\mathcal{B}_m \in \mathbb{R}^{L'_1 \times L'_2 \times \dots \times L'_N}, m = 1, \dots, M\}$

model. (3) MTP is less space and time consuming. For PCA, the size of each sample after vectorization is $\prod_{n=1}^N L_n$, and thus the size of covariance matrix is $\prod_{n=1}^N L_n \times \prod_{n=1}^N L_n$; while for MTP, we need to compute N covariance matrices but the size of each is only $L_n \times L_n$, which is much smaller than the covariance matrix of PCA. On the other hand, it is known that the time complexity of eigen-decomposition on a $L \times L$ matrix is $\mathcal{O}(L^3)$, so PCA will cost much more time than MTP especially when the size of original image is very large.

3. EXPERIMENTS AND ANALYSIS

In this section, two public available hyperspectral biological databases provided by Hong Kong Polytechnic University (HKPU) are used to demonstrate the effectiveness of the proposed MTP algorithm for HSI compression. The first database is the hyperspectral face database, which includes 400 hyperspectral face images from 25 volunteers whose age range from 21 to 33 (8 female and 17 male). There are three kinds of human face images: the frontal view face images (151 images), the left hand view face images (124 images) and the right hand view images (125 images). For each hyperspectral face image, there are 33 spectral channels per HSI and the spectral range is from 400nm to 720nm with a step wavelength of 10nm. The original spatial size of the image is 220

$\times 180$ pixels, in the experiment, in order to reduce the calculation cost, we down sampled each image to 44×36 .

The second database is the hyperspectral palm print database, which also provided by HKPU. The palm print images in this database were captured from 250 volunteers, including 195 males and 55 females. Since each volunteer has 2 palms captured, thus there are totally 500 HSI in this database. The original spatial size of each palm print image is 128×128 pixels, similar to the hyperspectral face image, we down sampled each image to 32×32 in the experiments.

Fig. 1 shows the first hyperspectral face image in all 33 spectral channels. It can be seen from Fig. 1 that some bands of images are with high visual quality while some bands appeared very low signal-noise-ratio. Therefore, in the experiment we choose the HSI from bands 11 to 25 for compression and thus the size of each HSI is $44 \times 36 \times 15$. The inputs of MTP algorithm are 400 3-order-tensors, i.e., $\{\mathcal{X}_m \in \mathbb{R}^{44 \times 36 \times 15}, m = 1, \dots, 400\}$. While in PCA, because the feature dimensionality is too large for our computer memory to process (the feature size is $44 \times 36 \times 15 = 23760$ after vectorization), as an alternative, each band of 44×36 grey level image is reshaped to a vector and then considered as an independent sample, thus we have $15 \times 400 = 6000$ vectors as algorithm inputs, the feature size of each sample is reduced to $44 \times 36 = 1584$, i.e., $\{\mathbf{x}_m \in \mathbb{R}^{1584}, m = 1, \dots, 6000\}$.

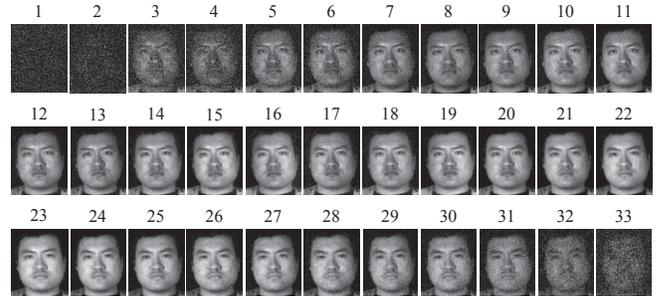


Fig. 1. The first hyperspectral face image with 33 spectral bands in the database 1.

In this paper, we compare the HSI compression performance of MTP and PCA by the measure the quality of the reconstructed HSI under a fix compression rate. The quality of the reconstructed HSI is evaluated by the Peak Signal Noise Ratio (PSNR) value. The compression rate is defined as $r = (L'_1 L'_2 L'_3 / L_1 L_2 L_3)$ for MTP and $r = L' / L$ for PCA. Here we take the result of frontal view face image of the 1st volunteer as a case study. Table 1 shows some selected spectral bands of the reconstructed images of MTP and PCA compared to the original face image, and the corresponding PSNR values are listed below each image. The related parameters are $[L'_1, L'_2, L'_3] = [15, 11, 2]$ in MTP and $L = 22$ in PCA, and thus for both of the two algorithms, compression rate is equal to 1.39 %.

From Table 1 it can be seen that images after compres-

Table 1. The HSI reconstruction performance of MTP and PCA algorithms in the database 1.

	band 1	band 4	band 7	band 10	band 13
Original Face Image					
MTP					
PSNR/dB	20.8893	23.5676	20.4299	24.7129	21.7087
PCA					
PSNR/dB	15.0516	13.1610	17.1202	9.6673	13.2289

sion and reconstruction has more or less information losses, but compared to PCA, the proposed MTP algorithm preserves more spatial information of the original images. For all the selected images, the reconstructed faces are similar to the original faces and the reconstructed PSNR of MTP algorithm is significantly higher than PCA as well, this phenomenon suggests that the compression and reconstruction performance of MTP is better than PCA.

Since Table 1 just provide the results of a few selected bands of images from the single person, Fig. 2 presents more comprehensive experimental results by compare the PSNR values of the whole 15 bands of HSI of 4 different volunteers, i.e., the 11th, 14th, 16th, and 20th volunteers, respectively, in various of view conditions. The compression parameters set-up is equal to the above experiment, i.e., at the compression rate of 1.39%. It can be seen from Fig. 2 that for all of the 15 bands, the PSNR value of MTP algorithm is better than PCA.

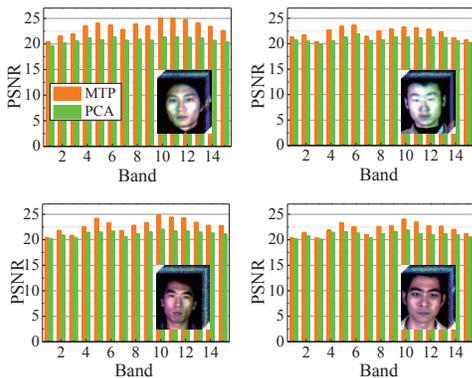


Fig. 2. The PSNR values of the whole 15 bands of HSI of 4 different volunteers in the database 1.

In order to show that the proposed MTP algorithm could preserve the spectral signature of the HSI in the reconstruction data as well, Fig. 3 provide some observations on the

spectral curve of the certain pixel in the 1st volunteers reconstruction image at the compression ratio of 1.39%. The focused pixel is the center pixel of the volunteers eye (location [18, 11]), as highlighted in Fig. 3(a). The black line in Figs. 3(b) and 3(c) give the original 15 bands spectral curve of the HSI, while the reds lines show the curves that extracted from the reconstructed HSI by MTP and PCA, respectively. From Figs. 3 (b) and (c) it is clear that the error of MTP is much smaller than PCA. Since the smaller error in Fig. 3 suggests the better compression performance on the spectral domain, thus all the above experiments demonstrate that the proposed MTP algorithm performs better than PCA both in the spatial and the spectral view.

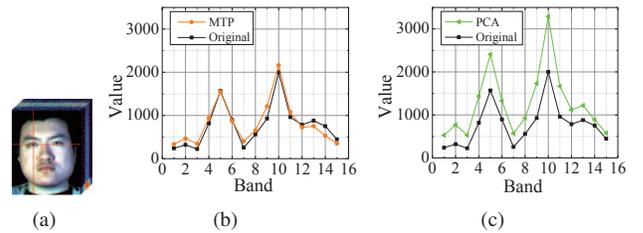


Fig. 3. (a) The location of analyzed pixel; (b) and (c) the spectral curves in the reconstructed HSI by MTP and PCA compared to original HSI in the database 1.

Each hyperspectral palm print has 12 spectral channels. Since all the images look fine and have high signal-noise-ratio, we use all the bands of images in the experiment. Thus the size of each HSI is $32 \times 32 \times 12$. Similar to the first experiment, the inputs of MTP algorithm are $\{\mathcal{X}_m \in \mathbb{R}^{32 \times 32 \times 12}, m = 1, \dots, 500\}$ while the inputs of PCA are $\{\mathbf{x}_m \in \mathbb{R}^{1024}, m = 1, \dots, 6000\}$.

Table 2 show the visual result of selected bands of images and the corresponding PSNR values on the database 2 before compression and after reconstruction of MTP and PCA, respectively. The related parameters are $[L'_1, L'_2, L'_3] = [12, 20, 2]$ in MTP and $L=22$ in PCA, thus the compression rate is 3.91%. It is obvious that the proposed MTP preserves more original information in the compressed data and thus the reconstructed HSI is much more similar to the original HSI.

In order to show the HSI compression performance with different compression rates, Fig. 4 investigates the compression rate respect to reconstruction error of the MTP and PCA algorithm on the hyperspectral palm print database. In the Fig. 4, the compression rates of the two algorithms varies from 0.01 to 0.86. As we know, for the same compression rate, the algorithm with lower reconstruction error indicates the better performance. From Fig. 4, we learn that for most of the time, the reconstruction error of MTP is less than PCA, this observation suggests that the MTP performs better than PCA on various of compression rates. We also notice that there are 3 points where the PCA outperforms MTP, this phe-

Table 2. The HSI reconstruction performance of MTP and PCA algorithms in the database 2.

	band 1	band 2	band 4	band 10
Original Palm Print Image				
MTP				
PSNR/dB	27.6692	27.8539	27.6230	25.2977
PCA				
PSNR/dB	18.2148	21.2948	18.8195	18.4171

nomenon may due to that MTP has to set three independent parameters, which would decrease the robust performance of the MTP algorithm.

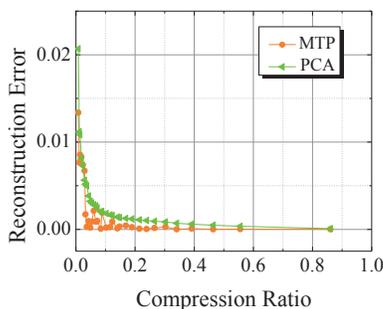


Fig. 4. Compression rate respect to reconstruction error of the MTP and PCA algorithms in database 2.

4. CONCLUSION

In this paper, we proposed a multiway tensor projection (MTP) algorithm for biological HSI database compression and reconstruction. The MTP algorithm employ the arbitrary order tensor as the inputs directly without the vectorization, and using an iterative procedure to find the optimal projection matrix in each order to make sure that the projected core tensor could preserve most of the information present in the original tensor. Nevertheless, there may still be room for improvement of the proposed approach in practice, e.g., to provide some criteria for estimating the parameters in MTP. This will be explored in our future work.

References

[1] A. Goetz, G. Vane, J. Solomon, and B. Rock, "Imaging spectrometry for earth remote sensing," *Science*, vol. 228, no. 4704, pp. 1147–1153,

1985.

- [2] B. Du and L. Zhang, "Random-selection-based anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 5, pp. 1578–1589, 2011.
- [3] L. Zhang, L. Zhang, D. Tao, and X. Huang, "Tensor discriminative locality alignment for hyperspectral image spectral-spatial feature extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 242–256, 2013.
- [4] Z. Pan, G. Healey, M. Prasad, and B. Tromberg, "Face recognition in hyperspectral images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1552–1560, 2003.
- [5] M. Uzair, A. Mahmood, and A. Mian, "Hyperspectral face recognition using 3D-DCT and partial least squares," in *British Machine Vision Conference*, 2013.
- [6] G. Wallace, "The JPEG still picture compression standard," *Commun. ACM*, vol. 34, no. 4, pp. 30–44, 1991.
- [7] Q. Du and J. Fowler, "Hyperspectral image compression using JPEG2000 and principal component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 2, pp. 201–205, 2007.
- [8] I. Jolliffe, *Principle Component Analysis, second edition*. Springer Series in Statistics, 2002.
- [9] Y. Pang, D. Tao, Y. Yuan, and X. Li, "Binary two-dimensional PCA," *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, vol. 38, no. 4, pp. 1176–1180, 2008.
- [10] D. Zhang and S. Chen, "Fast image compression using matrix K-L transform," *Neurocomputing*, vol. 68, pp. 258–266, 2005.
- [11] J. Yang, D. Zhang, A. Frangi, and J. Yang, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 131–137, 2004.
- [12] D. Zhang and Z. Zhou, "(2D)²PCA: Two-directional two-dimensional PCA for efficient face representation and recognition," *Neurocomputing*, vol. 69, pp. 224–231, 2005.
- [13] H. Lu, K. Plataniotis, and A. Venetsanopoulos, "MPCA: Multilinear principal component analysis of tensor objects," *IEEE Trans. Neural Networks*, vol. 19, no. 1, pp. 18–39, 2008.
- [14] D. Xu, S. Yan, L. Zhang, S. Lin, H. Zhang, and T. Huang, "Reconstruction and recognition of tensor-based objects with concurrent subspaces analysis," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 1, pp. 36–47, 2008.
- [15] D. Tao, X. Li, X. Wu, W. Hu, and S. Maybank, "Supervised tensor learning," *Knowl. Inf. Syst.*, vol. 13, no. 1, pp. 1–42, 2007.
- [16] D. Tao, X. Li, X. Wu, and S. Maybank, "General tensor discriminant analysis and gabor features for gait recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 10, pp. 1700–1715, 2007.
- [17] T. Kolda and B. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.
- [18] X. Li, S. Lin, S. Yan, and D. Xu, "Discriminant locally linear embedding with high-order tensor data," *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, vol. 38, no. 2, pp. 342–352, 2008.
- [19] D. Tao, X. Li, X. Wu, and S. Maybank, "Tensor rank one discriminant analysis-A convergent method for discriminative multilinear subspace selection," *Neurocomputing*, vol. 71, no. 10–12, pp. 1866–1882, 2008.