

Approximation of N -Way Principal Component Analysis for Organ Data

Hayato Itoh^{1(✉)}, Atsushi Imiya², and Tomoya Sakai³

¹ School of Advanced Integration Science, Chiba University, Chiba, Japan
`hayato-ito@graduate.chiba-u.jp`

² Institute of Management and Information Technologies,
Chiba University, Chiba, Japan

³ Graduate School of Engineering, Nagasaki University, Nagasaki, Japan

Abstract. We apply multilinear principal component analysis to dimension reduction and classification of human volumetric organ data, which are expressed as multiway array data. For the decomposition of multiway array data, tensor-based principal component analysis extracts multilinear structure of the data. We numerically clarify that low-pass filtering after the multidimensional discrete cosine transform efficiently approximates data dimension reduction procedure based on the tensor principal component analysis.

1 Introduction

For computer assisted diagnosis, inspection and biopsy in precision medicine, abnormality detection based on pattern recognition is a fundamental technique. From cell to human body, medical data used for diagnosis are multiway array data. This paper aims to apply the multiway principal component analysis (mPCA) to volumetric data analysis and processing. The method allows dimension reduction and retrieval of volumetric data preserving both geometric and statistic properties as objects and textures, respectively, of volumetric objects. Furthermore, these PCA-based data dimension reduction for multiway data array allows to extract outline volumetric shapes. Moreover, we develop multiclass classifier based on the multilinearity of volumetric data.

Organs, cells in the organ and microstructure in cells, which are dealt with in biomedical image analysis, are spatial texture. Furthermore, these biological objects possess spatial geometric and topological properties of volumetric structure as three-dimensional objects. Although, local structures of them as volumetric data are computed from geometric and topological properties, texture allows to estimate both local and global statistical properties of these objects. For the data analysis of these volumetric data, the methods simultaneously process geometrical and topological structures and texture properties.

Human bodies and organs in the body are three-dimensional volumetric objects. Since these medical objects are large-size volumetric data and sequence of volumetric data, this paper introduces a pattern recognition method for

dynamics volumetric data using tensor PCA. Toward as abnormality detection of beating-heart sequences, we develop an algorithm to identify individual using a volumetric image from a collection of sequences of a dynamic hearts. Since it is possible to model the beating of a heart in sequence of images as geometrically periodic small deformations to the geometric average-shape of a beating sequence, a recognition method is stable and robust against geometric perturbation of the shape. Mutual subspace method processes stable and robust recognition properties against small geometric perturbations [1]. Therefore, we develop the mutual subspace method for three-way volumetric data.

For numerical computation in pattern recognition, we deal with sampled patterns. In traditional pattern recognition, these sampled patterns are embedded in an appropriate-dimensional Euclidean space as vectors. The other way is to deal with sampled patterns as multiway array data [2–9]. These multiway array data are expressed as tensors to preserve multilinearity of in the original pattern space, since tensors allows to express multiway array data in multilinear forms. Furthermore, for applications of modern pattern recognition techniques such as deep learning [10] and machine learning [11] for big data, we are mathematically and numerically required to evaluate the performance of tensor-based pattern recognition of multilinear data. Importantly, for fast image pattern recognition, a compact representation of these image data is desirable. Therefore, we need dimension-reduction procedure that reduce size of tensor without significant loss of the recognition rate. Tensor expressions fulfill these requirements in applications of pattern recognition of multidimensional array data. By adopting a tensor representation, we can use the spatio-temporal and temporal structure of volumetric data for dimension reduction and classification of the data. Furthermore, the tensor representation of volumetric data does not require the procedure of transforming the data to high-dimensional vectors required for vector-based dimension-reduction and recognition methods.

We develop an approximate closed form for the Tucker-3 tensor PCA, though the Tucker-3 tensor decomposition is achieved by solving variational optimisation iteratively. Our method solves a system of variational optimisation problems derived from the original expression of the Tucker-3 decomposition with the orthogonal constraints for solutions. We also numerically clarify that data reduction by discrete cosine transform (DCT) [12] efficiently approximates dimension reduction based on the tensor PCA, since DCT approximates Karhunen-Loéva (KL) transform [13]. For the validation of these approximation in the recognition of volumetric data, we develop mutual subspace method. Using multilinearity, this method can classify volumetric data without vectorisation of data.

2 Tensor Decomposition

2.1 Tensor Representation for N -way Arrays

We briefly summarise the multilinear projection for N -dimensional arrays from ref. [14]. A N th-order tensor \mathcal{X} defined in $\mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ is expressed as

$$\mathcal{X} = (x_{i_1, i_2, \dots, i_N}) \quad (1)$$

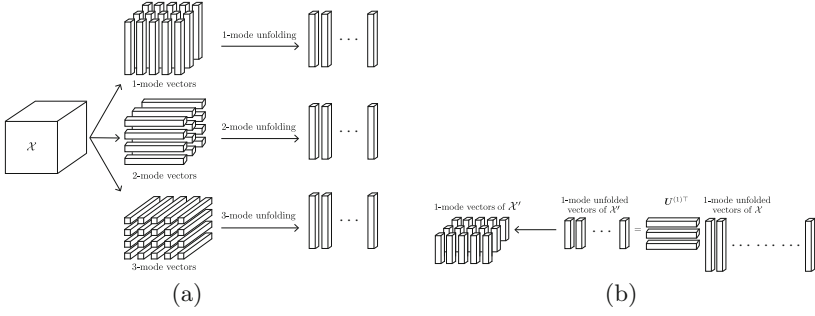


Fig. 1. Unfolding and n -mode projection of a tensor \mathcal{X} . (a) Unfoldings for a third-order tensor \mathcal{X} . For a tensor $\mathcal{X} \in \mathbb{R}^{4 \times 5 \times 3}$, unfoldings for 1-, 2- and 3-modes give 15 1-mode vectors, 12 2-mode vectors and 20 3-mode vectors, respectively. (b) 1-mode product for \mathcal{X} represented by a linear projection $\mathcal{X}'_{(1)} = \mathbf{U}^{(1)\top} \mathcal{X}_{(1)}$. Lower subscription denote 1-mode unfolding of a third-order tensor.

for $x_{i_1, i_2, \dots, i_N} \in \mathbb{R}$, using N indices i_n . Each subscript n denotes the n -mode of \mathcal{X} . For \mathcal{X} , the n -mode vectors, $n = 1, 2, \dots, N$, are defined as the I_n -dimensional vectors obtained from \mathcal{X} by varying this index i_n while fixing all the other indices. The unfolding of \mathcal{X} along the n -mode vectors of \mathcal{X} is defined as

$$\mathcal{X}_{(n)} \in \mathbb{R}^{I_n \times (I_1 \times I_2 \times \dots \times I_{n-1} \times I_{n+1} \times \dots \times I_N)}, \quad (2)$$

where the column vectors of $\mathcal{X}_{(n)}$ are the n -mode vectors of \mathcal{X} . Figure 1(a) illustrates unfoldings for a third-order tensor as an example of unfolding of N th-order tensor. The n -mode product $\mathcal{X} \times_n \mathbf{U}$ of a matrix $\mathbf{U} \in \mathbb{R}^{J_n \times I_n}$ and a tensor \mathcal{X} is a tensor $\mathcal{G} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_{n-1} \times J_n \times I_{n+1} \times \dots \times I_N}$, with elements

$$g_{i_1, i_2, \dots, i_{n-1}, j_n, i_{n+1}, \dots, i_N} = \sum_{i_n=1}^{I_n} x_{i_1, i_2, \dots, i_N} u_{j_n, i_n}, \quad (3)$$

by the manner in ref. [15]. A linear projection form of n -mode product in Eq. (3) is given by

$$\mathcal{G}_{(n)} = \mathbf{U} \mathcal{X}_{(n)}. \quad (4)$$

Figure 1(b) shows a linear projection form of 1-mode projection for a third-order tensor. For the m - and n -mode product by matrices \mathbf{U} and \mathbf{V} , we have

$$\mathcal{X} \times_m \mathbf{U} \times_n \mathbf{V} = \mathcal{X} \times_n \mathbf{V} \times_m \mathbf{U} \quad (5)$$

since n -mode projections are commutative [15]. We define the inner product of two tensors $\mathcal{X} = (x_{i_1, i_2, \dots, i_N})$, $\mathcal{Y} = (y_{i_1, i_2, \dots, i_N}) \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ by

$$\langle \mathcal{X}, \mathcal{Y} \rangle = \sum_{i_1} \sum_{i_2} \dots \sum_{i_N} x_{i_1, i_2, \dots, i_N} y_{i_1, i_2, \dots, i_N}. \quad (6)$$

Using this inner product, we have the Frobenius norm of a tensor \mathcal{X} by $\|\mathcal{X}\|_F = \sqrt{\langle \mathcal{X}, \mathcal{X} \rangle}$. For the Frobenius norm of a tensor, we have $\|\mathcal{X}\|_F = \|\text{vec } \mathcal{X}\|_2$, where vec and $\|\cdot\|_2$ are the vectorisation operator for a tensor and Euclidean norm for a vector, respectively. For the two tensors \mathcal{X}_1 and \mathcal{X}_2 , we define the distance between them by

$$d(\mathcal{X}_1, \mathcal{X}_2) = \|\mathcal{X}_1 - \mathcal{X}_2\|_F. \quad (7)$$

Although this definition is a tensor-based measure, this distance is equivalent to the Euclidean distance between the vectorised tensors \mathcal{X}_1 and \mathcal{X}_2 .

For a tensor, a multilinear projection maps the input tensor data from one space to another space. We have three basic multilinear projections, that is, the vector-to-vector projection (VVP), tensor-to-vector projection (TVP) and tensor-to-tensor projection (TTP). The VVP is a linear projection from a vector to another vector. To use the VVP for tensors, we need to reshape tensors into vectors before the projection. The TVP, which is also referred to as the rank-one projection [16–18], consists of elementary multilinear projections (EMPs). An EMP projects a tensor to a scalar. Using d EMPs, the TVP obtains a d -dimensional vector projected from a tensor. The TTP projects a tensor to another tensor of the same order. In this paper, we focus on methods of finding the optimal projection for the TTP.

As the tensor \mathcal{X} is in the tensor space $\mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \otimes \cdots \otimes \mathbb{R}^{I_N}$, the tensor space can be interpreted as the Kronecker product of N vector spaces $\mathbb{R}^{I_1}, \mathbb{R}^{I_2}, \dots, \mathbb{R}^{I_N}$. To project $\mathcal{X} \in \mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \otimes \cdots \otimes \mathbb{R}^{I_N}$ to another tensor \mathcal{Y} in a lower-dimensional tensor space $\mathbb{R}^{P_1} \otimes \mathbb{R}^{P_2} \otimes \cdots \otimes \mathbb{R}^{P_N}$, where $P_n \leq I_n$ for $n = 1, 2, \dots, N$, we need N projection matrices $\{\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times P_n}\}_{n=1}^N$. Using the N projection matrices, the TTP is given by

$$\mathcal{Y} = \mathcal{X} \times_1 \mathbf{U}^{(1)\top} \times_2 \mathbf{U}^{(2)\top} \cdots \times_N \mathbf{U}^{(N)\top}. \quad (8)$$

This projection is established in N steps, where at the n th step, each n -mode vector is projected to a P_n -dimensional space by $\mathbf{U}^{(n)}$. We call this operation the orthogonal projection of \mathcal{X} to \mathcal{Y} .

2.2 Principal Component Analysis for N -way Arrays

A N th-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, which is the array $\mathbf{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, is denoted as a set of indices (i_1, i_2, \dots, i_N) . Here we summarise the higher-order singular value decomposition (HOSVD) for N th-order tensors since N -way principal component is numerically computed by HOSVD. The HOSVD is the Tucker-3 decomposition with orthogonal constraints. For a collection of tensors $\{\mathcal{X}_i\}_{i=1}^M \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ satisfying the zero expectation condition $E(\mathcal{X}_i) = 0$, we compute the

$$\hat{\mathcal{X}}_i = \mathcal{X}_i \times_1 \mathbf{U}^{(1)\top} \times_2 \mathbf{U}^{(2)\top} \cdots \times_N \mathbf{U}^{(N)\top}, \quad (9)$$

where $\mathbf{U}^{(j)} = [\mathbf{u}_1^{(j)}, \dots, \mathbf{u}_{I_j}^{(j)}]$, that minimises the criterion

$$J_- = \mathbb{E} \left(\|\mathcal{X}_i - \hat{\mathcal{X}}_i \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \cdots \times_N \mathbf{U}^{(N)}\|_{\text{F}}^2 \right) \quad (10)$$

and maximises the criterion

$$J_+ = \mathbb{E} \left(\|\hat{\mathcal{X}}_i\|_{\text{F}}^2 \right), \quad (11)$$

with respect to the conditions

$$\mathbf{U}^{(j)\top} \mathbf{U}^{(j)} = \mathbf{I}_j, \quad (12)$$

where \mathbf{I}_j , $j = 1, 2, \dots, N$ are the identity matrices in $\mathbb{R}^{I_j \times I_j}$. By fixing $\{\mathbf{U}^{(j)}\}_{j=1}^N$ except $\mathbf{U}^{(j')}$, $j' \in \{1, 2, \dots, N\}$, we have

$$J_j = \mathbb{E} \left(\|\mathbf{U}^{(j)\top} \mathcal{X}_{i,(j)} \mathcal{X}_{i,(j)}^\top \mathbf{U}^{(j)}\|_{\text{F}}^2 \right), \quad (13)$$

where $\mathcal{X}_{i,(j)}$, $j = 1, 2, \dots, N$, are the j -mode unfolded tensors of \mathcal{X}_i .

Eigendecomposition problems are derived by computing the extremals of

$$E_j = J_j + \text{tr}((\mathbf{I}_j - \mathbf{U}^{(j)\top} \mathbf{U}^{(j)}) \boldsymbol{\Sigma}^{(j)}), \quad j = 1, 2, \dots, N. \quad (14)$$

For matrices $\mathbf{M}^{(j)} = \frac{1}{N} \sum_{i=1}^N \mathcal{X}_{i,(j)} \mathcal{X}_{i,(j)}^\top$, $j = 1, 2, \dots, N$, the optimisation of J_- and J_+ derives the eigenvalue decomposition

$$\mathbf{M}^{(j)} \mathbf{U}^{(j)} = \mathbf{U}^{(j)} \boldsymbol{\Sigma}^{(j)}, \quad (15)$$

where $\boldsymbol{\Sigma}^{(j)} \in \mathbb{R}^{I_j \times I_j}$, $j = 1, 2, \dots, N$, are diagonal matrices satisfying the relationships $\sigma_k^{(j)} = \sigma_k^{(j')}$, $k \in \{1, 2, \dots, K\}$ for

$$\boldsymbol{\Sigma}^{(j)} = \text{diag}(\lambda_1^{(j)}, \lambda_2^{(j)} \dots, \lambda_K^{(j)}, 0 \dots, 0). \quad (16)$$

The optimisation of each J_j derives the eigendecomposition problems in Eq. (15). However, for the optimisation of $\{J_j\}_{j=1}^N$, there is no closed-form solution to this maximisation problem [19, 20]. Algorithm 1 is the iterative procedure of the N -way PCA [2]. This algorithm is one of alternating-least-square (ALS) algorithms for tensors. In Algorithm 1 of K iteration, time complexity is $\mathcal{O}(KI_k^2)$, where $I_j \leq I_k$, $j \neq k$, due to the eigendecomposition problem. For the N -way PCA for volumetric data represented in $\mathbb{R}^{n \times n \times n}$, we then have time complexity $\mathcal{O}(Kn^2)$.

For Algorithm 1, we have the following property.

Property 1. *The N -way PCA without iteration in Algorithm 1 is equivalent to the HOSVD for N th-order tensors if dimensions of a projected tensor are coincident to ones of each mode of an original tensor.*

For N th-order tensors, there are $N!$ combinations in selecting the order of modes in a tensor-to-tensor projection for Algorithm 1. For the selection of combinations for Algorithm 1, we have the following property [2].

Algorithm 1. Iterative method for N -way PCA (ALS algorithm)**Input:** A set of tensors $\{\mathcal{X}_i\}_{i=1}^M$. Dimension of projected tensors $\{k_j\}_{j=1}^N$ A maximum number of iteration K . A sufficiently small number η **Output:** A set of projection matrices $\{\mathbf{U}^{(j)}\}_{j=1}^N$

1: Compute the eigendecomposition of a covariant matrix

 $\mathbf{M}^{(j)} = \frac{1}{M} \sum_{i=1}^M \mathcal{X}_{i,(j)} \mathcal{X}_{i,(j)}^\top$, where $\mathcal{X}_{i,(j)}$ is an j -mode unfolded \mathcal{X}_i , for $j = 1, 2, \dots, N$

2: Construct projection matrices by selecting eigenvectors

corresponding to the k_j largest eigenvalues for $j = 1, 2, \dots, N$ 3: Compute $\Psi_0 = \sum_{i=1}^M \|\mathcal{X}_i \times_1 \mathbf{U}^{(1)\top} \times_2 \mathbf{U}^{(2)\top} \dots \times_N \mathbf{U}^{(N)\top}\|_F$

4: Iteratively Compute the following procedure

for $k = 1, 2, \dots, K$ for $j = 1, 2, \dots, N$ Update $\mathbf{U}^{(j)}$ by decomposing matrix $\sum_{i=1}^M \mathbf{W}_{i,(j)}^{(-j)} \mathbf{W}_{i,(j)}^{(-j)\top}$, where $\mathbf{W}_{i,(j)}^{(-j)}$ is an unfolding of $\mathcal{X}_i \times_1 \mathbf{U}^{(1)} \dots \times_{j-1} \mathbf{U}^{(j-1)} \times_{j+1} \mathbf{U}^{(j+1)} \dots \times_N \mathbf{U}^{(N)}$ for a mode j

end

Compute $\Psi_k = \sum_{i=1}^M \|\mathcal{X}_i \times_1 \mathbf{U}^{(1)\top} \times_2 \mathbf{U}^{(2)\top} \dots \times_N \mathbf{U}^{(N)\top}\|_F$ if $|\Psi_k - \Psi_{k-1}| < \eta$

break

end

Property 2. For N th-order tensors, the selection of order of modes does not effect to the results of a tensor-to-tensor projection since n -mode projection is cumulative.

From these two properties, we adopt Algorithm 1 [2] to solve the optimisation of $\{J_j\}_{j=1}^N$. For a set of orthonormal vectors $\{\mathbf{e}_k\}_{k=1}^K$, $\mathbf{e}_i^\top \mathbf{e}_j = \delta_{ij}$, we set orthogonal projection matrices $\mathbf{P}^{(j)} = \sum_{k=1}^{k_j} \mathbf{e}_k \mathbf{e}_k^\top$ for $j = 1, 2, 3$. Using these $\{\mathbf{P}^{(j)}\}_{j=1}^N$, the low-rank tensor approximation [20] is achieved by

$$\mathcal{Y} = \mathcal{X} \times_1 (\mathbf{P}^{(1)} \mathbf{U}^{(1)}) \times_2 (\mathbf{P}^{(2)} \mathbf{U}^{(2)}) \dots \times_N (\mathbf{P}^{(N)} \mathbf{U}^{(N)}), \quad (17)$$

where $\mathbf{P}^{(j)}$ selects k_j bases of orthogonal matrices $\mathbf{U}^{(j)}$. The low-rank approximation using Eq. (17) is used for compression in the N -way PCA.

For the HOSVD for N th-order tensors, we have the following theorem.

Theorem 1. The HOSVD method is equivalent to the vector PCA method in the compression of N th-order tensors.

(Proof). The equation

$$\mathcal{X} \times_1 (\mathbf{P}^{(1)} \mathbf{U}^{(1)})^\top \times_2 (\mathbf{P}^{(2)} \mathbf{U}^{(2)})^\top \dots \times_N (\mathbf{P}^{(N)} \mathbf{U}^{(N)})^\top = \mathcal{Y} \quad (18)$$

is equivalent to

$$(\mathbf{P}^{(N)} \mathbf{U}^{(N)} \otimes \dots \mathbf{P}^{(2)} \mathbf{U}^{(2)} \otimes \mathbf{P}^{(1)} \mathbf{U}^{(1)})^\top \text{vec} \mathcal{X} = \text{vec} \mathcal{Y}. \quad (19)$$

(Q.E.D.)

This theorem implies that N -dimensional discrete cosine transform (NDDCT) is an acceptable approximation of the HOSVD for N th-order tensors since this is the analogy of the approximation of the PCA of two-dimensional images by the 2DDCT [13, 21].

Furthermore, we have the following theorem.

Theorem 2. *The compression of N th-order tensors computed by the HOSVD is equivalent to the compression computed by the TPCA.*

(Proof). The projection that selects $K = k_1 k_2 \dots k_N$ bases of the tensor space spanned by $u_{i_1}^{(1)} \circ u_{i_2}^{(2)} \circ u_{i_3}^{(3)}, i_j = 1, 2, \dots, k_j$ for $j = 1, 2, \dots, N$, is

$$\begin{aligned} & (\mathbf{P}^{(N)} \mathbf{U}^{(N)} \otimes \dots \mathbf{P}^{(2)} \mathbf{U}^{(2)} \otimes \mathbf{P}^{(1)} \mathbf{U}^{(1)}) \\ &= (\mathbf{P}^{(N)} \otimes \dots \mathbf{P}^{(2)} \otimes \mathbf{P}^{(1)}) (\mathbf{U}^{(N)} \otimes \dots \mathbf{U}^{(2)} \otimes \mathbf{U}^{(1)}) = \mathbf{P} \mathbf{W}, \end{aligned} \quad (20)$$

where \mathbf{W} and \mathbf{P} are an orthogonal matrix and the orthogonal projection matrix, respectively. Therefore, HOSVD is equivalent to TPCA for third-order tensors.

2.3 N -Dimensional Discrete Cosine Transform

For a N th-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$, we have DCT matrices

$$\mathbf{D}^{(k)} = ((d_{ij}^{(k)})), \quad d_{ij}^{(k)} = \cos \left(\frac{\pi}{I_k} \left((j-1) + \frac{1}{2} \right) (i-1) \right). \quad (21)$$

for mode k , $i = 1, 2, \dots, I_k$, $j = 1, 2, \dots, N^{(-k)}$, where $N^{(-k)} = I_1 \times I_2 \times \dots \times I_{k-1} \times I_{k+1} \times \dots \times I_N$. For each mode k , we set projection matrices $\mathbf{P}^{(k)} \in \mathbb{R}^{l_k \times I_k}$. These projection matrices select l_k bases of DCT for mode k . Using DCT matrices and projection matrices, we define compression by NDDCT for N th-order tensors as

$$\mathcal{Y} = \mathcal{X} \times_1 (\mathbf{P}^{(1)} \mathbf{D}^{(1)})^\top \times_2 (\mathbf{P}^{(2)} \mathbf{D}^{(2)})^\top \dots \times_N (\mathbf{P}^{(N)} \mathbf{D}^{(N)})^\top. \quad (22)$$

This NDDCT is an acceptable approximation for the NDTPCA.

In our application for volumetric data, an $n \times n \times n$ digital array is directly compressed by the NDDCT-II with order $\mathcal{O}(n^3)$. If we apply the fast Fourier transform to the computation of the NDDCT-II, the computational complexity is $\mathcal{O}(n \log n)$.

3 Classification of Tensor Data

Tensor Subspace of Categories. Setting $\{\mathbf{U}_k^{(j)}\}_{j=1}^N$ to be orthogonal matrices of a tensor projection for N th-order tensors, we have a tensor subspace spanned by $\{\mathbf{U}_k^{(j)}\}_{j=1}^N$ for k th category. Therefore, we can define a tensor subspace of a category by

$$\mathcal{C}_k = \{\mathcal{X} \mid \mathcal{X} \times_1 \mathbf{U}_k^{(1)\top} \times_2 \mathbf{U}_k^{(2)\top} \dots \times_n \mathbf{U}_k^{(N)} = \mathcal{X}\}. \quad (23)$$

Since a pattern represented by tensors contains perturbation, we define k th category by

$$\mathcal{C}_k(\delta) = \{\mathcal{X} \mid \|\mathcal{X} \times_1 \mathbf{U}^{(1)\top} \times_2 \mathbf{U}^{(2)\top} \cdots \times_N \mathbf{U}^{(N)} - \mathcal{X}\|_F \ll \delta\}, \quad (24)$$

where a positive constant δ is the bound for a small perturbation to a pattern. Therefore, by defining similarity and dissimilarity between a tensor subspace and query, we can construct tensor-subspace-based classifiers that are robust and stable against small perturbations to patterns.

Tensor Subspace Method. As an extension of the subspace method [22, 23] for N -way data, we introduce a new linear tensor subspace method for N th-order tensors. This method is a N -dimensional version of the 2DTSM [24].

For a N th-order tensor \mathcal{X} , we set $\mathbf{U}^{(j)}$, $j = 1, 2, \dots, N$, to be projection matrices of the tensor-to-tensor projection of \mathcal{X} to \mathcal{Y} . For a collection of normalised tensors $\{\mathcal{X}_i\}_{i=1}^M$, such that $\mathcal{X}_i \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, $\|\mathcal{X}_i\|_F = 1$ and $E(\mathcal{X}_i) = 0$, the solutions of

$$\{\mathbf{U}^{(j)}\}_{j=1}^N = \arg \max E \left(\|\mathcal{X} \times_1 \mathbf{U}^{(1)\top} \times_2 \mathbf{U}^{(2)\top} \cdots \times_N \mathbf{U}^{(N)\top}\|_F / \|\mathcal{X}_i\|_F \right) \quad (25)$$

with respect to $\mathbf{U}^{(j)\top} \mathbf{U}^{(j)} = \mathbf{I}$ for $j = 1, 2, \dots, N$ define a multilinear subspace that approximates $\{\mathcal{X}_i\}_{i=1}^M$. Therefore, using projection matrices $\{\mathbf{U}_k^{(j)}\}_{j=1}^N$ obtained as the solutions of Eq. (25) for k th category, if a query tensor \mathcal{G} satisfies the condition

$$\arg \left(\max_l \|\mathcal{G} \times_1 \mathbf{U}_l^{(1)\top} \times_2 \mathbf{U}_l^{(2)\top} \cdots \times_N \mathbf{U}_l^{(N)\top}\|_F / \|\mathcal{G}\|_F \right) = \{\mathbf{U}_k^{(j)}\}_{j=1}^N, \quad (26)$$

we conclude that $\mathcal{G} \in \mathcal{C}_k$, $k, l = 1, 2, \dots, N_C$, where \mathcal{C}_k and N_C are the tensor subspace of k th category and the number of categories, respectively.

Tensor Subspace of Queries. We have a collection of query tensors $\{\mathcal{G}_{i'}\}_{i'=1}^{M'}$ normalised by $\mathcal{G}_{i'} / \|\mathcal{G}_{i'}\|_F$. We assume that these queries belong to the same category. Then, using multiway PCA for a collection of queries $\{\mathcal{G}_{i'}\}_{i'=1}^{M'}$, we obtain orthogonal matrices $\{\mathbf{V}^{(j)}\}_{j=1}^N$, which are given by Eq. (25) for each mode. The column vectors of orthogonal matrices $\{\mathbf{V}^{(j)}\}_{j=1}^N$ are eigenvectors of a set of unfolded tensors for each mode. We note that even if we have only one query \mathcal{G} we can obtain eigenvectors of j -mode by the multiway PCA, since j -mode unfolding gives a set of column vectors of mode- j as shown in Fig. 1. That is, for a query \mathcal{G} , we have orthogonal matrices by the decomposition [19]

$$\mathcal{G} = \mathcal{A} \times_1 \mathbf{V}^{(1)} \times_2 \mathbf{V}^{(2)} \cdots \times_N \mathbf{V}^{(N)}. \quad (27)$$

Mutual Tensor Subspace Method. As the extension of mutual subspace method for vector data [1], we define a classifier for two tensor subspaces. For

each of N_C categories of tensor data, we set a collection of normalised N th-order tensors $\{\mathcal{X}_i\}_{i=1}^M$, such that $\mathcal{X}_i \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$, $\|\mathcal{X}_i\|_F = 1$ and $E(\mathcal{X}_i) = 0$. For the k th category in N_C categories, we have an orthogonal matrices $\{\mathbf{U}_k^{(j)}\}_{j=1}^N$, which satisfy Eq. (25). The orthogonal matrices for k th category span a tensor subspaces \mathcal{C}_k of the category.

We have a collection of query tensors $\{\mathcal{G}_{i'}\}_{i'=1}^{M'}$ normalised by $\mathcal{G}_{i'}/\|\mathcal{G}_{i'}\|_F$. We assume that these queries belong to the same category. Then, using multiway PCA for a collection of queries $\{\mathcal{G}_{i'}\}_{i'=1}^{M'}$, we obtain orthogonal matrices $\{\mathbf{V}^{(j)}\}_{j=1}^N$. The obtained eigenvectors of all modes span a tensor subspace \mathcal{C}_q for queries. For the classification of queries, we measure the dissimilarity between a category subspace \mathcal{C}_k and a query subspace \mathcal{C}_q . Since \mathcal{C}_k and \mathcal{C}_q represent patterns with perturbations, we can robustly recognise queries against the pattern perturbations by measuring the dissimilarity between two tensor subspace.

Using orthogonal matrices $\{\mathbf{U}_k^{(j)}\}_{j=1}^N$ for k th category, we have a projected tensor in a category subspace \mathcal{C}_k by

$$\mathcal{A}_{i'} = \mathcal{G}_{i'} \times_1 \mathbf{U}_k^{(1)\top} \times_2 \mathbf{U}_k^{(2)\top} \dots \times_N \mathbf{U}_k^{(N)\top}. \quad (28)$$

Furthermore, using orthogonal matrices $\{\mathbf{V}^{(j)}\}_{j=1}^N$, we have a projected tensor in a query subspace \mathcal{C}_q by

$$\mathcal{B}_{i'} = \mathcal{G}_{i'} \times_1 \mathbf{V}^{(1)\top} \times_2 \mathbf{V}^{(2)\top} \dots \times_N \mathbf{V}^{(N)\top}. \quad (29)$$

For a tensor subspaces \mathcal{C}_k and \mathcal{C}_q , we define the dissimilarity of subspaces $d(\mathcal{C}_k, \mathcal{C}_q)$ by

$$E\left(\|\mathcal{A}_{i'} \times_1 \mathbf{P}\mathbf{U}_k^{(1)} \dots \times_N \mathbf{P}\mathbf{U}_k^{(N)} - \mathcal{B}_{i'} \times_1 \mathbf{P}\mathbf{V}^{(1)} \dots \times_N \mathbf{P}\mathbf{V}^{(N)}\|_F^2\right), \quad (30)$$

where a projection matrix \mathbf{P} selects bases for each mode of tensors. Therefore, using the dissimilarity given by Eq. (30), if queries $\{\mathcal{G}_{i'}\}_{i'=1}^{M'}$ satisfy the condition

$$\arg\left(\min_l d(\mathcal{C}_l, \mathcal{C}_q)\right) = \mathcal{C}_k, \quad (31)$$

we conclude that $\{\mathcal{G}_{i'}\}_{i'=1}^{M'} \in \mathcal{C}_k(\delta)$ for $k, l = 1, 2, \dots, N_C$.

4 Numerical Examples

We comparatively evaluate performance of the TPCA and the 3D-DCT in dimension reduction and classification of volumetric data. For this evaluation, we adopt the following steps.

1. Extract the volumetric data of the left ventricles from cardiac MRI dataset.
2. Reduce the dimension of the extracted data by using four methods.
3. Divide the dimension-reduced data to taring and test data.

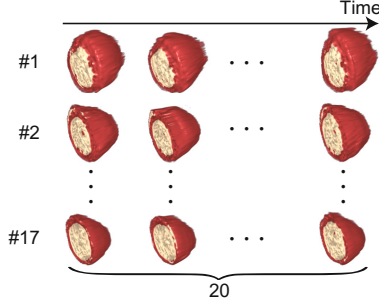


Fig. 2. Illustration of extracted cardiac MRI dataset. These sequences of volumetric data are extracted from cardiac MRI dataset with landmarks of endocardium of left ventricles [25]. As shown in Table 1, we have 17 sequences of volumetric data of left ventricle for 17 patients. Each sequence of volumetric data represents one cardiac beat by 20 frames. Every sequence starts with maximally expanded state. Red and white parts of volume rendering of the data represent muscle and inner space of left ventricles. We set the center of the first sagittal slice of each volume data to the center of the slice. (Color figure online)

4. Construct tensor subspace of each categories by applying the TPCA to the training data.
5. Classify a query (or query subspace) of a category in test data by using tensor-based classifiers.

We extract sequences of volumetric data of left ventricle from cardiac MRI dataset [25], since we need to validate classification of normal organs before abnormality detection. For the extraction, we use the landmarks of endocardium of left ventricle. These landmarks are manually given and provided as the part of the dataset. Table 1 summarises the number and the size of the extracted volumetric data at all phases. Figure 2 illustrates the extracted sequences of volumetric data for 17 patients. Since a beating heart deforms its volumetric shape, we obtain third-order tensors representing shape of heart with geometrical perturbation from a sequence.

For the dimension reduction of volumetric data, we use the TPCA and 3D-DCT. For the practical computation of the TPCA, we use the HOSVD and MPCA. In the MPCA, if we set the number of bases to the size of the original

Table 1. Sizes and number of volumetric data of left ventricles. #category represents the number of individuals. #data/category represents the number of frames in one sequence of left ventricles. The data size is the original size of the volumetric data. The reduced data size is the size of the volumetric data after reduction. We set $d \in \{8, 16, 32\}$.

	#category	#data/category	Data size [voxel]	Reduced data size [voxel]
Volumetric data	17	20	$81 \times 81 \times 63$	$d \times d \times d$

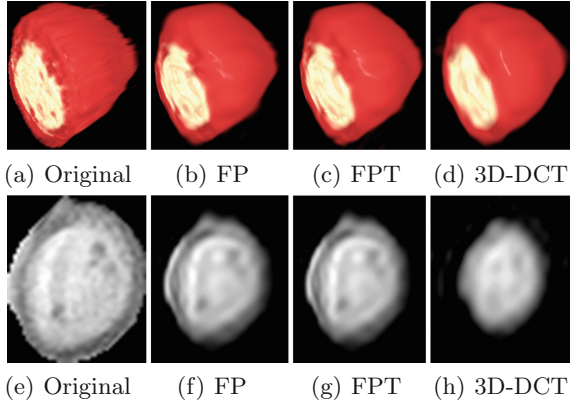


Fig. 3. Shape and inner texture of reconstructed volume data of left ventricle from compressed data. Upper and lower rows show volume rendering and sagittal slice of the volumetric data, respectively. In (a)–(d), red and white parts depict the muscle of heart and inner of heart, respectively, for original and approximation by the FP, the FPT and the 3D-DCT. In these approximation, the data are reduced to the size $16 \times 16 \times 16$. (Color figure online)

tensors in Algorithm 1, we call the method full projection (FP). If we set the number of bases to fewer than the size of the original tensors in Algorithm 1, we call the method the full-projection truncation (FPT). For the dimension reduction by the HOSVD, FP and FPT, we apply these methods to all the extracted volumetric data in all categories. For the evaluation the robustness and stabilizes of methods with respect to the sizes of the data, we set the sizes of the dimension-reduced data to $8 \times 8 \times 8$, $16 \times 16 \times 16$ and $32 \times 32 \times 32$.

Figure 3 illustrates the comparison between original and dimension-reduced data by the three-methods. In Fig. 3(a)–(d), volume rendering of the original and reconstructed volume data are presented. For the data reduced by the FP and FPT, the shapes of volumetric data reconstructed from the compressed data are almost the same in their appearances. The reconstructed data from the data reduced by the 3DDCT is the closest shape to the shape of original volumetric data. In Figs. 3(e)–(h), the differences of appearances between the sagittal slices of reconstructed data and original shape are compared. Compared to the original data shown in Fig. 3(a), the 3D-DCT gives blurred inner texture as shown in Fig. 3(h). As shown in Figs. 3(f) and (g), the dimension reduction by the FP and FPT extract outline shapes of ventricle without inner texture. Figure 4 illustrates reconstructed data from principal components of dimension-reduced volume data. This result show that the principal components of the dimension-reduced volume data are almost the same as shown in Table 2.

In the dimension-reduced data, each sequences consist from 20 frames. We use odd and even frames in dimension reduced data as training and test data, respectively. Applying the FP to training data of each category, we construct

Table 2. Reconstruction error of volumetric data. The reconstruction error is given by distance between tensors of the original and reconstructed volumetric data.

	FP	FPT	3D-DCT
Reconstruction error	11.4×10^3	11.4×10^3	8.23×10^3

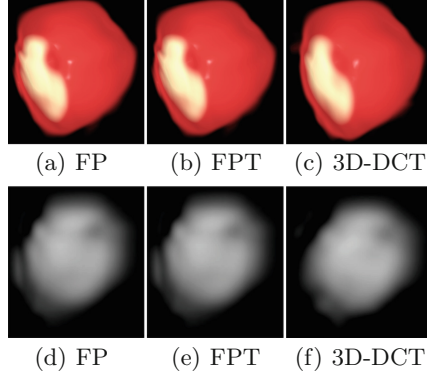


Fig. 4. Extracted principal components of dimension-reduced volume data. For the data dimension reduced by the FP, FPT and 3D-DCT, we apply the FP. Using the extracted principal component, we reconstruct volumetric data. For the extraction, we select the 20 principal eigenvectors of ones of three modes.

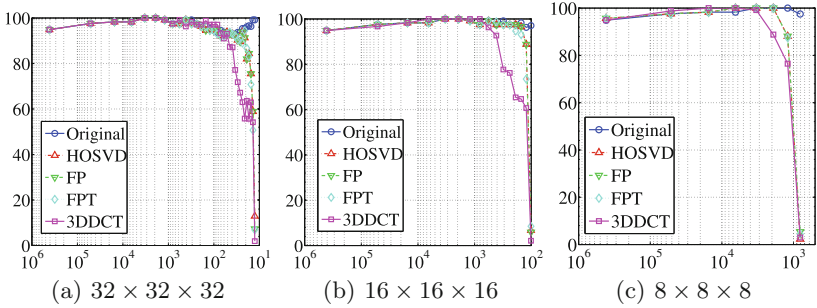


Fig. 5. Recognition rates of the left ventricles for original and compressed tensors. We use tensor subspace method as classifier. The data are reduced to $32 \times 32 \times 32$, $16 \times 16 \times 16$ and $8 \times 8 \times 8$. The HOSVD, FP, FPT and 3D-DCT are used for the reduction. Vertical and horizontal axes represent recognition rate and compression ratio, respectively. For the original size $D = 81 \times 81 \times 63$ and reduced size $K = k \times k' \times k'$, the compression ratio is given by D/K .

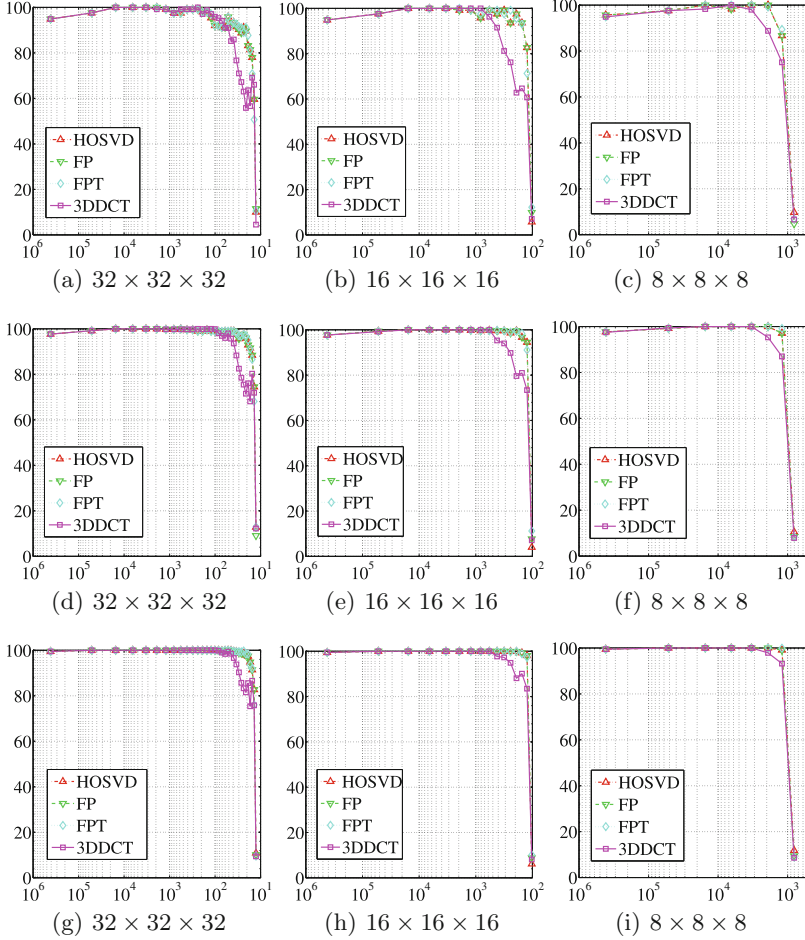


Fig. 6. Recognition rates of left ventricles for compressed tensors. We adopt the reduces sizes of $32 \times 32 \times 32$, $16 \times 16 \times 16$ and $8 \times 8 \times 8$. For compression, we use the HOSVD, FP, FPT and 3D-DCT. In the mutual tensor subspace method, input is a query subspace. The query subspace is spanned by a few queries. To construct a query subspace, we use one, two and three queries. Top, middle and bottom row show recognition rates for the case of one, two and three queries, respectively. Vertical and horizontal axes represent recognition rate and compression ratio, respectively. For the original size $D = 81 \times 81 \times 63$ and reduces size $K = k \times k' \times k''$, the compression ratio is given by D/K .

tensor subspace of 17 categories for the TSM and MTSM. The TSM and MTSM are robust classification methods against geometrical perturbations. Therefore, we use only odd frame for construction of tensor subspaces of categories to evaluate these robustness. If the TSM and MTSM classify a category of even

frame, we conclude that these classifiers are robust to the small geometrical change between frames.

The recognition rate is defined as the successful classification ration of individuals in 1000 classifications. We use the TSM and MTSM as classifiers. In the selection of a query for the TSM, we randomly select one of 17 individuals and randomly select one of test data of the individual. From a collection of input data of the heart sequence of a patient, we constructed the subspace of queries for the MTSM. After randomly select an individual, from one to three queries are randomly selected from test data of the selected individual. Applying the FP to the selected queries, we construct a query tensor subspace for the MTSM. Figures 5 and 6 show recognition rate for left ventricles by the TSM and the MTSM, respectively.

In Fig. 5, the profiles of recognition curves for the HOSVD, FP, FPT and 3D-DCT are almost the same in the higher compression ratio than 10^3 . Furthermore, for the higher compression ratio than 10^3 , the dimension reduced data by four methods derive almost same recognition rates. These recognition rates are the same recognition rate of the tensors of original size. Moreover, the TSM with major five eigenvectors in each modes processes accurate recognition. Figure 6(a)–(c) show that the MTSM for the query subspace spanned by one query. The results show that the recognition properties are almost the same for data with $8 \times 8 \times 8$, $16 \times 16 \times 16$ and $32 \times 32 \times 32$. The results in Figs. 6(d)–(i) show that the MSM achieve more robust recognition against small geometrical perturbations by using a query subspace than the TSM, since a query subspace spanned by a few queries with geometrical perturbations.

These numerical examples show that the 3DDCT accurately approximates performance of the TPCA. Furthermore, recognition of three-way data by the TSM and MTSM is accurate and robust for volumetric data contain geometrical perturbations as temporal deformation.

5 Conclusions

We developed an approximate closed form for the Tucker decomposition for N th-order tensors. Our method solves a system of variational optimisation problems derived from the original expression of the Tucker decomposition with the orthogonal constraints for solutions. Furthermore, we developed tensor-based multilinear classifiers, the TSM and MTSM, for order- N tensors. In the numerical examples, we evaluated the performance of the dimension reduction by the HOSVD, FP, FPT and 3D-DCT for the recognition of the individual of left ventricles. In the reduction, the FP and FPT extract the outline shapes of left ventricles. The results of the evaluations showed that the dimension reduction by the HOSVD gives the same performance by FP and FPT without iterative computation procedures. The results also numerically clarified that data compression by the DCT efficiently approximates data compression procedure based on the tensor PCA. In the evaluations, the TSM and MTSM accurately recognised individual of left ventricles though cardiac MRI include geometrical

perturbations. For the MTSM, using a query subspace spanned by more than one frame, the MTSM achieved more robust and stable recognition to geometrical perturbations than the TSM.

This research was supported by “Multidisciplinary Computational Anatomy and Its Application to Highly Intelligent Diagnosis and Therapy” project funded by a Grant-in-Aid for Scientific Research on Innovative Areas from MEXT, Japan, and by Grants-in-Aid for Scientific Research funded by the Japan Society for the Promotion of Science.

References

1. Maeda, K.: From the subspace methods to the mutual subspace method. In: Cipolla, R., Battiato, S., Farinella, G.M. (eds.) *Computer Vision*, vol. 285, pp. 135–156. Springer, Heidelberg (2010). doi:[10.1007/978-3-642-12848-6_5](https://doi.org/10.1007/978-3-642-12848-6_5)
2. Lu, H., Plataniotis, K., Venetsanopoulos, A.: MPCA: multilinear principal component analysis of tensor objects. *IEEE Trans. Neural Netw.* **19**, 18–39 (2008)
3. Jiang, B., Ma, S., Zhang, S.: Tensor principal component analysis via convex optimization. *Math. Program.* **150**, 423–457 (2014)
4. Lu, H., Plataniotis, K., Venetsanopoulos, A.: Uncorrelated multilinear principal component analysis for unsupervised multilinear subspace learning. *IEEE Trans. Neural Netw.* **20**, 1820–1836 (2009)
5. Shen, H., Huang, J.Z.: Sparse principal component analysis via regularized low rank matrix approximation. *J. Multivar. Anal.* **99**, 1015–1034 (2008)
6. Lai, Z., Xu, Y., Chen, Q., Yang, J., Zhang, D.: Multilinear sparse principal component analysis. *IEEE Trans. Neural Netw. Learn. Syst.* **25**, 1942–1950 (2014)
7. Panagakis, Y., Kotropoulos, C., Arce, G.R.: Non-negative multilinear principal component analysis of auditory temporal modulations for music genre classification. *IEEE Trans. Audio Speech Lang. Process.* **18**, 576–588 (2010)
8. Vasilescu, M.A.O., Terzopoulos, D.: Multilinear (Tensor) ICA and dimensionality reduction. In: Davies, M.E., James, C.J., Abdallah, S.A., Plumbley, M.D. (eds.) *ICA 2007*. LNCS, vol. 4666, pp. 818–826. Springer, Heidelberg (2007). doi:[10.1007/978-3-540-74494-8_102](https://doi.org/10.1007/978-3-540-74494-8_102)
9. Bro, R.: PARAFAC. Tutorial and applications. *Chemometr. Intell. Lab. Syst.* **38**, 149–171 (1997)
10. Dean, J., Corrado, G., Monga, R., Chen, K., Devin, M., Mao, M., Ranzato, M., Senior, A., Tucker, P., Yang, K., Le, Q.V., Ng, A.Y.: Large scale distributed deep networks. In: *Proceedings of the Conference on Neural Information Processing Systems*, pp. 1232–1240 (2012)
11. Cohen, N., Shashua, A.: Simnets: a generalization of convolutional networks. In: *Proceedings NIPS Workshop on Deep Learning* (2014)
12. Hamidi, M., Pearl, J.: Comparison of the cosine and fourier transforms of Markov-1 signals. *IEEE Trans Acoust. Speech Sig. Process.* **24**, 428–429 (1976)
13. Oja, E.: *Subspace Methods of Pattern Recognition*. Research Studies Press, Brighton (1983)
14. Lu, H., Plataniotis, K., Venetsanopoulos, A.: A survey of multilinear subspace learning for tensor data. *Pattern Recogn.* **44**, 1540–1551 (2011)
15. Cichoki, A., Zdunek, R., Phan, A.H., Amari, S.: *Nonnegative Matrix and Tensor Factorizations*. Wiley, Hoboken (2009)

16. Wang, Y., Gong, S.: Tensor discriminant analysis for view-based object recognition. *Proc. Int. Conf. Pattern Recogn.* **3**, 33–36 (2006)
17. Tao, D., Li, X., Wu, X., Maybank, S.: Elapsed time in human gait recognition: a new approach. *Proc. Int. Conf. Acoust. Speech Sig. Process.* **2**, II (2006). <http://ieeexplore.ieee.org/document/1660308/>
18. Hua, G., Viola, P., Drucker, S.: Face recognition using discriminatively trained orthogonal rank one tensor projections. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2007)
19. Lathauwer, L., Moor, B., Vandewalle, J.: A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.* **21**, 1253–1278 (2000)
20. Lathauwer, L.D., Moor, B.D., Vandewalle, J.: On the best rank-1 and rank- (r_1, r_2, r_n) approximation of higher-order tensors. *SIAM J. Matrix Anal. Appl.* **21**, 1324–1342 (2000)
21. Itoh, H., Imiya, A., Sakai, T.: Low-dimensional tensor principle component analysis. *Proc. Int. Conf. Comput. Anal. Images Patterns Part I* **9256**, 223–235 (2015)
22. Iijima, T.: Theory of pattern recognition. *Electron. Commun. Jpn.* **1**, 123–134 (1963)
23. Watanabe, S., Pakvasa, N.: Subspace method of pattern recognition. In: *Proceedings of the 1st International Joint Conference of Pattern Recognition* (1973)
24. Itoh, H., Sakai, T., Kawamoto, K., Imiya, A.: Topology-preserving dimension-reduction methods for image pattern recognition. In: Kämäräinen, J.-K., Koskela, M. (eds.) *SCIA 2013. LNCS*, vol. 7944, pp. 195–204. Springer, Heidelberg (2013). doi:[10.1007/978-3-642-38886-6_19](https://doi.org/10.1007/978-3-642-38886-6_19)
25. Andreopoulos, A., Tsotsos, J.K.: Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI. *Med. Image Anal.* **12**, 335–357 (2008)

Computer Vision – ACCV 2016 Workshops

ACCV 2016 International Workshops, Taipei, Taiwan,

November 20-24, 2016, Revised Selected Papers, Part

III

Chen, C.-S.; Lu, J.; Ma, K.-K. (Eds.)

2017, XV, 660 p. 319 illus., Softcover

ISBN: 978-3-319-54525-7