Fast Approximate Karhunen-Loève Transform for Three-Way Array Data

Hayato Itoh
Nagoya University
Furo-cho, Chikusa-ku,
Nagoya 464-8601, Japan

Atsushi Imiya
IMIT, Chiba University
Yayoi-cho 1-33, Inage-ku,
Chiba 263-8522, Japan
imiya@faculty.chiba-u.jp

Tomoya Sakai
Nagasaki University
Bunkyo 1-14, 852-8521
Nagasaki, Japan

Abstract

Organs, cells and microstructures in cells dealt with in biomedical image analysis are volumetric data. We are required to process and analyse these data as volumetric data without embedding into higher-dimensional vector space from the viewpoints of object oriented data analysis. Sampled values of volumetric data are expressed as three-way array data. Therefore, principal component analysis of multi-way data is an essential technique for subspace-based pattern recognition, data retrievals and data compression of volumetric data. For one-way array (the vector form) problem the discrete cosine transform matrix is a good relaxed solution of the eigenmatrix for principal component analysis. This algebraic property of principal component analysis, derives an approximate fast algorithm for PCA of three-way data arrays.

1. Introduction

Principal component analysis (PCA) is a fundamental methodology [1, 2, 3, 4, 5, 6] in pattern recognition, computer vision, physiology and many natural and social sciences used for data processing. The aim of the paper is derivation of a first approximation method for the computation of tensor PCA. There is no closed form for the computation of PCA for three-way array data [7, 8, 9, 10], although there is closed form for 1- and 2-way problems. For 1-way array (vector form) problem the discrete cosine transform (DCT) is an approximate solution to Karhunen-Loève transform [11, 5]. Employing this algebraic property of Karhunen-Loève transform, we construct a fast approximate algorithm for PCA of three-way array data.

Using a metric for the collection of trees, the mean and principal components are computed. Space of binary trees with tree edit distance as the metric is a Riemannian space with negative curvature. The mean is computed as Fratch mean using the metric [14]. Furthermore, the first measure principal component is the geodesic which pass through the mean. This data processing is called principal geodesic analysis (PGA) [15]. GCPCA for phylogenetic trees computes the mean of the trees in the data space. In geodesic PCA (GPCA), the curvature of spaces is extended from zero to non-zero. GPCA in a shape space is used for longitudinal analysis (follow-up analysis) of cancers in organs.

The subspace method based on Karhunen-Loève transform is a fundamental technique in pattern recognition. Modern pattern recognition techniques for sampled value of patterns are described using linear algebra for sampled value embedded in vector space. Organs, cells and microstructures in cells dealt with in biomedical image analysis are volumetric data. We are required to process and analyse these data as volumetric data without embedding sampled values in vector space from the viewpoints of object oriented data analysis [13].

We express sampled values of volumetric data as three-way array data. These three-way array data are processed as the third order tensor. This expression of data requires to develop subspace method for tensor data. For dealing with sampled organs as three-way data from viewpoint of OODA, we introduce PCA for three-way data arrays employing the Tucker3 tensor decomposition. In biopsychometry, the correlation of three-way arrays are studied. For instance, correlations among gender, age and internet usage time can be investigated by three-way PCA.

2. Mathematical Preliminaries

2.1. Subspace Method in Vector Space [1, 2, 5]

A volumetric pattern is assumed to be a square integrable function in a linear space and to be defined on a finite support in three-dimensional Euclidean space [1, 16, 17] such that $\int_\Omega |f|^2 d\boldsymbol{x} < \infty$ for $\Omega \subset \mathbb{R}^3$. Furthermore, we assume $\int_\Omega |\nabla f|^2 d\boldsymbol{x} < \infty$ and $\int_\Omega tr\{(\nabla\nabla^\top f)^\top(\nabla\nabla^\top f)\}d\boldsymbol{x} < \infty$, where $\nabla\nabla^\top f$ is the Hessian matrix of f. For an orthogonal projection $\boldsymbol{P}_\perp = \boldsymbol{I} - \boldsymbol{P}$, $f^\| = \boldsymbol{P}f$ and $f^\perp = \boldsymbol{P}_\perp f$ are the canonical element and canonical form of f with respect to $\boldsymbol{P}$ and $\boldsymbol{P}_\perp$, respectively. If $\boldsymbol{P}$ is the projection to the space spanned by the constant element, the operation $\boldsymbol{P}_{\perp} f$ is called the constant canonicalisation. Let $\boldsymbol{P}_i$ be the orthogonal projection to the linear subspace corresponding to the category C_i. For a pattern f, if $|\boldsymbol{P}_{i^*}(f/|f|)| \leq \delta$ for an appropriately small positive number δ, we conclude that $f \in C_{i^*}$.

Setting (f, g) to be the inner product in Hilbert space (H), the relation $|f|^2 = (f, f)$ is satisfied. Let θ be the canonical angle between a pair of linear subspaces L_1 and L_2. Setting $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ to be the orthogonal projections to L_1 and L_2, respectively, $\cos^2\theta$ is the maximiser of $(\boldsymbol{P}_1 f, \boldsymbol{P}_2 g)^2$ with respect to the conditions $|f| = 1$, $|g| = 1$ $\boldsymbol{P}_1 f = f$ and $\boldsymbol{P}_2 g = g$. The relation $\cos^2\theta = \lambda^2_{\max}$ is satisfied, where $\lambda_{\max}$ is the maximal singular value of $\boldsymbol{P}_2\boldsymbol{P}_1$.

Since, in mutual subspace method, a query f is expressed by using a set of local bases, we set that $\boldsymbol{Q}_f$ is the orthogonal projection to linear subspace expressing the query f. Then, if the canonical angle between $\boldsymbol{Q}_f$ and $\boldsymbol{P}_i$ satisfies the relation $\angle(\boldsymbol{Q}_f, \boldsymbol{P}_i) < \angle(\boldsymbol{Q}_f, \boldsymbol{P}_i^*)$ for all C_i, we conclude that $f \in C_{i^*}$.

Setting $\boldsymbol{P}_i$ to be the orthogonal projection to linear subspace $\mathfrak{L}_i$ corresponding to the category C_i, the orthogonal projection which maximises the criterion $J = \sum_{i=1}^{n} |\boldsymbol{Q}\boldsymbol{P}_i|_2^2$with respect to the condition $\boldsymbol{Q}^*\boldsymbol{Q} = \boldsymbol{I}$ where $\boldsymbol{Q}^*$ is the conjugate of $\boldsymbol{Q}$ and $|\boldsymbol{A}|$ is the trace norm of the operator $\boldsymbol{A}$ in Hilbert space $\mathfrak{H}$. Though operation $\boldsymbol{Q}f$ removes common part for all categories from f, $(\boldsymbol{I} - \boldsymbol{Q})f$ preserves essentially significant parts for pattern recognition of f.

In traditional pattern recognition, these sampled patterns are embedded in an appropriate-dimensional Euclidean space as vectors. For $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{X} \in \mathbb{R}^{n \times n}$, $|\boldsymbol{x}|_2$ and $|\boldsymbol{X}|_F$ are the vector norm and Frobenius norm of $\boldsymbol{x}$ and $\boldsymbol{X}$, respectively.

Setting the data matrix $\boldsymbol{X}$ to be $\boldsymbol{X} = (\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_m)$ for data vectors $\{\boldsymbol{f}_i\}_{i=1}^m$ in $\mathbb{R}^N$, whose mean is zero, the Karhunen-Loève transform is established by computing $\hat{\boldsymbol{f}}_i = \boldsymbol{U}\boldsymbol{f}_i$ for $\boldsymbol{U}$ which minimises $J_1 = |\boldsymbol{U}\boldsymbol{X}|_F^2$with the condition $\boldsymbol{U}^\top\boldsymbol{U} = \boldsymbol{I}_N$. The orthogonal matrix $\boldsymbol{U}$ is the minimiser of

$$J_{11} = |\boldsymbol{U}\boldsymbol{X}|_F^2 + \langle(\boldsymbol{U}^\top\boldsymbol{U} - \boldsymbol{I})\boldsymbol{\Lambda}\rangle \tag{1}$$

where

$$\boldsymbol{\Lambda} = Diag(\lambda_1, \lambda_2, \cdots, \lambda_N) \tag{2}$$

for

$$\lambda_1 \geq \lambda_2 \geq \lambda_2 \geq \cdots \geq \lambda_N \geq 0. \tag{3}$$

The minimiser of eq. (1) is the solution of the eigenmatrix problem

$$\boldsymbol{M}\boldsymbol{U} = \boldsymbol{U}\boldsymbol{\Lambda},\ \boldsymbol{M} = \boldsymbol{X}\boldsymbol{X}^\top \tag{4}$$

The row vectors of $\boldsymbol{U}$ are the principal components.

2.2. PCA of Two-Way Array Data[16]

For a collection of matrices $\{\boldsymbol{F}_i\}_{i=1}^N \in \mathbb{R}^{m \times n}$ satisfying $\mathrm{E}_i(\boldsymbol{F}_i) = 0$, the orthogonal-projection-based data reduction

$$\hat{F}_i = \boldsymbol{U}^\top\boldsymbol{F}_i\boldsymbol{V} \tag{5}$$

is performed by maximising

$$J_2(\boldsymbol{U}, \boldsymbol{V}) = \mathrm{E}_i\left(|\boldsymbol{U}\hat{\boldsymbol{F}}_i\boldsymbol{V}^\top|_{\mathrm{F}}^2\right)$$

with respect to the conditions $\boldsymbol{U}^\top\boldsymbol{U} = \boldsymbol{I}_m$ and $\boldsymbol{V}^\top\boldsymbol{V} = \boldsymbol{I}_n$. The solutions are the minimiser of the Euler-Lagrange equation

$$\begin{aligned} J_{22}(\boldsymbol{U}, \boldsymbol{V}) &= \mathrm{E}\left(|\boldsymbol{U}\hat{\boldsymbol{F}}_i\boldsymbol{V}^\top|_{\mathrm{F}}^2\right) \\ &\quad + \langle(\boldsymbol{I}_m - \boldsymbol{U}^\top\boldsymbol{U}), \boldsymbol{\Sigma}\rangle + \langle(\boldsymbol{I}_n - \boldsymbol{V}^\top\boldsymbol{V}), \boldsymbol{\Lambda}\rangle \end{aligned} \tag{6}$$

for diagonal matrices $\boldsymbol{\Lambda}$ and $\boldsymbol{\Sigma}$.

Setting

$$\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{F}_i^\top\boldsymbol{F}_i = \boldsymbol{M},\ \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{F}_i\boldsymbol{F}_i^\top = \boldsymbol{N}, \tag{7}$$

$\boldsymbol{U}$ and $\boldsymbol{V}$ are the solutions of the eigendecomposition problems

$$\boldsymbol{M}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{\Lambda},\ \boldsymbol{N}\boldsymbol{U} = \boldsymbol{U}\boldsymbol{\Sigma}, \tag{8}$$

where $\boldsymbol{\Sigma} \in \mathbb{R}^{m \times m}$ and $\boldsymbol{\Lambda} \in \mathbb{R}^{n \times n}$ are diagonal matrices satisfying the relationships $\lambda_i = \sigma_i$ for

$$\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_K, 0, \ldots, 0), \tag{9}$$

$$\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_K, 0, \ldots, 0). \tag{10}$$

2.3. Data Analysis of Three-Way Array Data[18, 12]

For the triplet of positive integers I_1, I_2 and I_3, the third-order tensor $\mathbb{R}^{I_1 \times I_2 \times I_3}$ is expressed as $\mathcal{X} = ((x_{ijk}))$ Indices i, j and k are called the 1-mode, 2-mode and 3-mode of $\mathcal{X}$, respectively. The tensor space $\mathbb{R}^{I_1 \times I_2 \times I_3}$ is interpreted as the Kronecker product of three vector spaces $\mathbb{R}^{I_1}$, $\mathbb{R}^{I_2}$ and $\mathbb{R}^{I_3}$ such that $\mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \otimes \mathbb{R}^{I_3}$. We set $I = \max(I_1, I_2, I_3)$.

Samples $Sf(\Delta\mathbf{z})$ of $f(\boldsymbol{x})$ for $\mathbf{z} \in \mathbf{Z}^3$ and $|\mathbf{z}|_\infty \leq I$ defines an $I \times I \times I$ three-way array $\mathbf{F}$. To preserve the multi-linearity of the function $f(\boldsymbol{x})$, we deal with the array $\mathbf{F}$ as a third-order tensor $\mathcal{F}$. The operation $vec\mathcal{F}$ derives a vector $\boldsymbol{f} \in \mathbb{R}^{I_{123}}$ for $I_{123} = I_2 \cdot I_2 \cdot I_3$. We can reconstruct f from $\mathcal{F}$ using an interpolation procedure.

For $\mathcal{X}$, the n-mode vectors, $n = 1, 2, 3$, 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 matrices such that

$$\mathcal{X}_{(1)} \in \mathbb{R}^{I_1 \times I_{23}},\ \mathcal{X}_{(2)} \in \mathbb{R}^{I_2 \times I_{13}},\ \mathcal{X}_{(3)} \in \mathbb{R}^{I_3 \times I_{12}} \tag{11}$$

for $I_{12} = I_1 \cdot I_2$, $I_{23} = I_2 \cdot I_3$ and $I_{13} = I_1 \cdot I_3$, where the column vectors of $\mathcal{X}_{(j)}$ are the j-mode vectors of $\mathcal{X}$ for $i = 1, 2, 3$. We express the j-mode unfolding of $\mathcal{X}_i$ as $\mathcal{X}_{i,(j)}$.

For matrices

$$\begin{aligned} \boldsymbol{U} &= ((u_{ii'})) \in \mathbb{R}^{I_1 \times I_1}, \\ \boldsymbol{V} &= ((v_{jj'})) \in \mathbb{R}^{I_2 \times I_2}, \\ \boldsymbol{W} &= ((w_{kk'})) \in \mathbb{R}^{I_3 \times I_3}, \end{aligned} \tag{12}$$

the n-mode products for $n = 1, 2, 3$ of a tensor $\mathcal{X}$ are the tensors with entries

$$\begin{aligned} (\mathcal{X} \times_1 \boldsymbol{U})_{ijk} &= \sum_{i'=1}^{I_1} x_{i'jk} u_{i'i}, \\ (\mathcal{X} \times_2 \boldsymbol{V})_{ijk} &= \sum_{j'=1}^{I_2} x_{ij'k} v_{j'j}, \\ (\mathcal{X} \times_3 \boldsymbol{W})_{ijk} &= \sum_{k'=1}^{I_3} x_{ijk'} w_{k'k}, \end{aligned} \tag{13}$$

where $(\mathcal{X})_{ijk} = x_{ijk}$ is the ijk-th element of the tensor $\mathcal{X}$. The inner product of two tensors $\mathcal{X}$ and $\mathcal{Y}$ in $\mathbb{R}^{I_1 \times I_2 \times I_3}$ is

$$\langle \mathcal{X}, \mathcal{Y} \rangle = \sum_{i=1}^{I_1} \sum_{j=1}^{I_2} \sum_{k=1}^{I_3} x_{ijk} y_{ijk}. \tag{14}$$

Using this inner product, we have the Frobenius norm of a tensor $\mathcal{X}$ as $|\mathcal{X}|_F = \sqrt{\langle \mathcal{X}, \mathcal{X} \rangle}$. The Frobenius norm $|\mathcal{X}|_F$ of the tensor $\mathcal{X}$ satisfies the relation $|\mathcal{X}|_F = |\boldsymbol{f}|_2$, where $|\boldsymbol{f}|_2$ is the Euclidean norm of the vector $\boldsymbol{f}$.

3. PCA of Three-Way Array Data

To project a tensor $\mathcal{X}$ in $\mathbb{R}^{I_1 \times I_2 \times I_3}$ to the tensor $\mathcal{Y}$ in a lower-dimensional tensor space $\mathbb{R}^{P_1 \times P_2 \times P_3}$, where $P_n \leq I_n$, three projection matrices $\{\boldsymbol{U}^{(n)}\}_{n=1}^3$, for $\boldsymbol{U}^{(n)} \in \mathbb{R}^{I_n \times P_n}$ are required for $n = 1, 2, 3$. Using these three projection matrices, we have the tensor orthogonal projection

$$\mathcal{Y} = \mathcal{X} \times_1 \boldsymbol{U}^{(1)\top} \times_2 \boldsymbol{U}^{(2)\top} \times_3 \boldsymbol{U}^{(3)\top}. \tag{15}$$

This projection is established in three steps, where in each step, each n-mode vector is projected to a P_n-dimensional space by $\boldsymbol{U}^{(n)}$ for $n = 1, 2, 3$.

Using the three projection matrices $\boldsymbol{U}^{(i)}$ for $i = 1, 2, 3$, we have the tensor orthogonal projection for a third-order tensor as

$$\mathcal{Y} = \mathcal{X} \times_1 \boldsymbol{U}^{(1)\top} \times_2 \boldsymbol{U}^{(2)\top} \times_3 \boldsymbol{U}^{(3)\top}. \tag{16}$$

For a collection $\{\mathcal{X}_k\}_{k=1}^m$ of third-order tensors, the orthogonal-projection-based dimension reduction procedure is achieved by maximising the criterion

$$J_3 = E_k(|\mathcal{X}_k \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)}|_F^2) \tag{17}$$

with respect to the conditions $\boldsymbol{U}^{(i)\top} \boldsymbol{U}^{(i)} = \boldsymbol{I}$ for $i = 1, 2, 3$. The Euler-Lagrange equation of this conditional optimisation problem is

$$\begin{aligned} & J_{33}(\boldsymbol{U}^{(1)}, \boldsymbol{U}^{(2)}, \boldsymbol{U}^{(3)}) \\ = \; & E_k(|\mathcal{X}_k \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)}|_F^2) \\ & + \sum_{i=1}^{3} \langle (\boldsymbol{I} - \boldsymbol{U}^{(i)\top} \boldsymbol{U}^{(i)}), \boldsymbol{\Lambda}^{(i)} \rangle. \end{aligned} \tag{18}$$

This minimisation problem is solved by an iteration procedure [12].

Setting $\boldsymbol{P}^{(j)}$ to be an orthogonal projection in the linear space $\mathcal{L}(\{\boldsymbol{u}_i^{(j)}\}_{i=1}^{I_j})$ spanned by the column vectors of $\boldsymbol{U}^{(j)}$, the data reduction is computed by

$$\mathcal{Y} = \mathcal{X} \times_1 \boldsymbol{P}^{(1)} \boldsymbol{U}^{(1)} \times_2 \boldsymbol{P}^{(2)} \boldsymbol{U}^{(2)} \times_3 \boldsymbol{P}^{(3)} \boldsymbol{U}^{(3)}. \tag{19}$$

This expression is equivalent to the vector form

$$vec\mathcal{Y} = (\boldsymbol{P}^{(3)} \otimes \boldsymbol{P}^{(2)} \otimes \boldsymbol{P}^{(1)})(\boldsymbol{U}^{(3)} \otimes \boldsymbol{U}^{(2)} \otimes \boldsymbol{U}^{(1)}) vec\mathcal{X}, \tag{20}$$

The two-dimensional analogue of the problem has a closed form for the solution as a system of eigenvalue problems. The optimisation criterion is, however, trilinear, The three-dimensional problem has no closed form for its solution, although it is a convex optimisation problem. An efficient method to solve the three-dimensional problem is an alternative optimisation of the triplet of the matrices. Since, in each step of the iteration, singular value decomposition is achieved on each mode of a tensor, the total time complexity of the alternative optimisation is $\mathcal{O}(K \times 3 \times m^3) = \mathcal{O}(m^3)$, where K is the total number of iteration for convergence of the solutions and m is the dimennsion of the array.

4. Approximate Karhunen-Loève Transform

4.1. Non-Iteration Relaxation [19]

The non-iteration relaxation algorithm proposed in ref. [19] is achieved by solving the system of optimisation problems

$$J_j = E(|\boldsymbol{U}^{(j)\top} \mathcal{X}_{i,(j)} \boldsymbol{U}^{(j)}|_F^2) + \langle (\boldsymbol{U}^{(j)\top} \boldsymbol{U}^{(j)} - \boldsymbol{I}_j), \boldsymbol{\Lambda}^{(j)} \rangle \tag{21}$$

for $i = 1, 2, 3$, where $\mathcal{X}_{i,(j)}$ is the ith column vector of the unfolding matrix $\mathcal{X}_{(j)}$. These optimisation problems derive a system of eigenmatrix problems

$$\boldsymbol{M}^{(j)} \boldsymbol{U}^{(j)} = \boldsymbol{U}^{(j)} \boldsymbol{\Lambda}^{(j)}, \;\; \boldsymbol{M}^{(j)} = \frac{1}{N} \sum_{i=1}^{N} \mathcal{X}_{i,(j)} \mathcal{X}_{i,(j)}^\top \tag{22}$$

for $j = 1, 2, 3$. This system of equations decomposes the minimisation problem for PCA of three-way array data to the triplet of singular value decomposition for two-way array data. Therefore, we have an approximate closed form approximation for PCA of three-way array data. Moreover, the total time complexity of this relaxed method is $\mathcal{O}(3 \times m^3) = \mathcal{O}(m^3)$, where m is the dimennsion of the array.

4.2. DCT Relaxation

For data observed from a first-order Markov model in a vector space, DCT-II matrix efficiently approximates Karhunen-Loéve transform (KLT) [11, 5]. This property implies that among the DCT-matrix, KLT matrix $\boldsymbol{U}$, which is derived by PCA, and an orthogonal projection $\boldsymbol{P}_N$ to a linear subspace spanned by an appropriate number of column vectors of $\boldsymbol{U}$, the relation

$$E_i|\boldsymbol{P}_N\boldsymbol{\Phi} f_i - \boldsymbol{P}_N\boldsymbol{U} f_i|^2 \ll \varepsilon \tag{23}$$

is satisfied for a sufficiently small positive number ε, where $E_i(\phi(x_i))$ is the value of $\phi(x_i)$ for $\{x_i\}_{i=1}^n$ in a data space.

Equation (23) suggests that the triplet of the DCT-II matrices [20] are acceptable relaxed solutions for tensor PCA to derive the KLT. Therefore, the dimension reduction by PCA is relaxed to the process using the DCT as

$$f^n_{ijk} = \sum_{i'j'k'=0}^{N-1} a_{i'j'k'}\varphi_{i'i}\varphi_{j'j}\varphi_{k'k}, \tag{24}$$

$$a_{ijk} = \sum_{i'j'k'=0}^{n-1} f_{i'j'k'}\varphi_{ii'}\varphi_{jj'}\varphi_{kk'} \tag{25}$$

for $n \leq N$, where

$$\boldsymbol{\Phi}_{(n)} = \left(\left(\epsilon \cos\frac{(2j+1)i}{2\pi n}\right)\right) = ((\varphi_{ij})), \tag{26}$$

$$\epsilon = \begin{cases} 1 & \text{if } j = 0 \\ \frac{1}{\sqrt{2}} & \text{otherwise} \end{cases}$$

is the DCT-II matrix of order N. If we apply the fast cosine transform to the computation of the 3D-DCT-II matrix, the computational complexity is $\mathcal{O}(3m \log m)$, where m is the dimennsion of the array.

Since

$$vec(\boldsymbol{u} \circ \boldsymbol{v} \circ \boldsymbol{w}) = \boldsymbol{u} \otimes \boldsymbol{v} \otimes \boldsymbol{w} \tag{27}$$

the outer products of vectors redescribes the DCT-based transform as

$$\mathcal{F} = \sum_{i,j,k=0}^{n-1} a_{ijk}\boldsymbol{\varphi}_i \circ \boldsymbol{\varphi}_j \circ \boldsymbol{\varphi}_k,\ a_{ijk} = \langle \mathcal{F}, (\boldsymbol{\varphi}_i \circ \boldsymbol{\varphi}_j \circ \boldsymbol{\varphi}_k)\rangle, \tag{28}$$

where

$$\boldsymbol{\Phi}_{(n)} = \left(\boldsymbol{\varphi}_0, \boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_{N-1}\right). \tag{29}$$

Property 1 *DCT-based Karhunen-Loève transform is the orthogonal projection from* $\mathfrak{L}(\{\boldsymbol{\varphi}_i \circ \boldsymbol{\varphi}_j \circ \boldsymbol{\varphi}_j^\top\}_{i,j,k=0}^{N-1})$ *to* $\mathfrak{L}(\{\boldsymbol{\varphi}_i \circ \boldsymbol{\varphi}_j \circ \boldsymbol{\varphi}_k\}_{i,j,k=0}^{n-1})$.

The DCT matrix $\boldsymbol{\Phi}_{(n)}$ is the eigenmatrix of the discrete Laplacian with the Neumann boundary condition. We can define the order of the column vectors of DCT matrix using the order of the eigenvalue $\{\lambda_i\}_{i=0}^{n-1}$ of the discrete Laplacian. Since $\lambda_i\lambda_j\lambda_k$ derives the semi-order

$$\begin{aligned} \lambda_i\lambda_j\lambda_k &\geq \lambda_{i+1}\lambda_j\lambda_k, \\ \lambda_i\lambda_j\lambda_k &\geq \lambda_i\lambda_{j+1}\lambda_k, \\ \lambda_i\lambda_j\lambda_k &\geq \lambda_i\lambda_j\lambda_{k+1}, \end{aligned} \tag{30}$$

we define the order for the outer product of the column vectors $\{\boldsymbol{\varphi}_{i=0}^{n-1}\}$

$$\begin{aligned} \boldsymbol{\varphi}_i \circ \boldsymbol{\varphi}_j \circ \boldsymbol{\varphi}_k &\succ \boldsymbol{\varphi}_{i+1} \circ \boldsymbol{\varphi}_j \circ \boldsymbol{\varphi}_k, \\ \boldsymbol{\varphi}_i \circ \boldsymbol{\varphi}_j \circ \boldsymbol{\varphi}_k &\succ \boldsymbol{\varphi}_i \circ \boldsymbol{\varphi}_{j+1} \circ \boldsymbol{\varphi}_k, \\ \boldsymbol{\varphi}_i \circ \boldsymbol{\varphi}_j \circ \boldsymbol{\varphi}_k &\succ \boldsymbol{\varphi}_i \circ \boldsymbol{\varphi}_j \circ \boldsymbol{\varphi}_{k+1}. \end{aligned} \tag{31}$$

This order is used for the definition of the dimension of subspace for the relaxed PCA with DCT. On this order, the kth elements lies on the surface of the oct-sphere of the radius $k-1$ with the l_1-distance. Therefore, this order defines the low-pass filter of which path window is the oct-diamond in discrete space.

Regarding the selection of the dimension of the tensor subspace, the semi-order implies the following theorem.

Theorem 1 *The dimension of the subspace of the tensor space for data compression is* $\frac{1}{6}N(N+1)(N+2)$ *if we select* N *principal components in each mode of three-way array data.*

(*Proof*) For a positive integer N, the number s_N of eigenvalues $\lambda_i\lambda_j\lambda_k$ is

$$\begin{aligned} s_N &= \sum_{i+j+k=0, i,j,k\geq 0}^{N-1} (i+j+k) \\ &= \frac{1}{6}N(N+1)(N+2). \end{aligned} \tag{32}$$

If $N = 1, 2, 3, 4$, we have $N = 1, 4, 10, 20$, respectively, for tensors $\mathcal{X} = ((x_{ijk}))$ in $\mathbb{R}^{I\times I\times I}$.

5. PCA vs. Pyramid Transform

Setting $w_{\pm1} = \frac{1}{4}$ and $w_0 = \frac{1}{2}$, for the sampled function $f_{ij} = f(i,j,k)$, the pyramid transform for the three-way array is

$$Rf_{kmn} = \sum_{p,q,r=-1}^{1} w_p w_q w_r f_{2k-p\ 2m-q\ 2n-p}, \tag{33}$$

where the summation is achieved for $\frac{(k-p)}{2}$, $\frac{(m-q)}{2}$ and $\frac{(n-r)}{2}$ being integers. This operation involve the reduction image sizes. This process in each step is achieved by a weighted average of the image values in a finite small region, which is called the window for the operation. Therefore, image features are extracted in the higher-layer images of the pyramid transform.

Using the second-order differential matrix $\boldsymbol{D}$ with the Neumann boundary condition, we define

$$\begin{aligned}\boldsymbol{W} &= \frac{1}{4}(\boldsymbol{D}+4\boldsymbol{I}) \\ &= \begin{pmatrix} 3 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 2 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 2 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 3 \end{pmatrix}. \end{aligned} \quad (34)$$

The eigenvalues of $\boldsymbol{W}$ is $\theta_i = 1 + \frac{1}{4}\lambda_i$ where $\lambda_i = -4\sin^2(\pi i/n)$ is an eigenvalue of $\boldsymbol{D}$. Furthermore, the eigenmatrix of $\boldsymbol{W}$ is that of $\boldsymbol{D}$.

The downsampling operation [21] is expressed as $\boldsymbol{S} = (\boldsymbol{I} \otimes \boldsymbol{e}_2^\top)$, where $\boldsymbol{e}_2 = (0,1)^\top$. Since the Gaussian pyramid transform is achieved by applying downsampling to the result of the Gaussian convolution for the image array, we have the next lemma.

Lemma 1 *The matrix forms of the pyramid transform is* $\boldsymbol{R} = \boldsymbol{S}\boldsymbol{W}$ *for discrete signal.*

Setting the discrete three-way array to be

$$\boldsymbol{F} = ((f_{ijk})),\ 0 \leq i,j,k \leq 2^n - 1, \quad (35)$$

the pyramid transform is expressed as

$$\hat{\boldsymbol{F}} = \boldsymbol{F} \times_1 \boldsymbol{R} \times_2 \boldsymbol{R} \times_3 \boldsymbol{R}, \quad (36)$$

where $\hat{\boldsymbol{F}} = ((\hat{f}_{ijk}))$ for $0 \leq i,j,k \leq 2^{n/2} - 1$, that is,

$$vec\hat{\boldsymbol{F}} = (\boldsymbol{R} \otimes \boldsymbol{R} \otimes \boldsymbol{R})vec\boldsymbol{F}, \quad (37)$$

Moreover, since the pyramid transform is expressed as

$$\begin{aligned} Rf_{ijk} &= \sum_{i'j'k'=0}^{\frac{n}{2}-1} \sigma_{ijk} a_{2i'2j'2k'} \varphi_{2i'2i} \varphi_{2j'2j} \varphi_{2k'2k}, \\ a_{ijk} &= \sum_{i'j'k'=0}^{n-1} f_{i'j'k'} \varphi_{ii'} \varphi_{jj'} \varphi_{kk'}, \end{aligned} \quad (38)$$

for $\sigma_{ijk} = \theta_{2i'}\theta_{2j'}\theta_{2k'}$ using the column vectors of $\boldsymbol{\Phi}_{(n)}$ and $\boldsymbol{\Phi}_{(\frac{n}{2})}$, we have the following property.

Property 2 *For three-way array, the Gaussian pyramid transform is a linear transform from* $\mathfrak{L}(\{\boldsymbol{\varphi}_i \circ \boldsymbol{\varphi}_j \circ \boldsymbol{\varphi}_j^\top\}_{i,j,k=0}^{2^n-1})$ *to* $\mathfrak{L}(\{\boldsymbol{\varphi}_{2i} \circ \boldsymbol{\varphi}_{2j} \circ \boldsymbol{\varphi}_{2k}\}_{i,j,k=0}^{2^{n-1}-1})$.

This property allows us to compute the pyramid transform using DCT-II. Furthermore, this property shows that the pyramid transform is a lowpass filtering operation.

Although, the PCA-based data compression controls dimension of the reduced data using minimisation criterions of eq. (23) and lemma 1, the pyramid transform reduces the data to the pre-defined dimension based on the transformation. Since the dimension is pre-defined in the pyramid transform, the transform involves filtering operation to enhance leading major components and to suppress minor components.

6. Numerical Examples and Evaluation

The perforamces of tensor principal component analysis (TPCA), the three-dimensional discrete cosine transform (3D-DCT) and the pyramid transform (PT) are compared for the approximation of volumetric data. The volumetric data of human livers in computational anatomy dataset and of human left ventricles in cardiac MRI dataset [22] are used for comparisions. Table 1 summarises the numbers and sizes of the volumetric data. Figures 1 and 2 illustrate the original and approximated volumetric data. Figure 3 summarises the reconstruction errors of the three methods in terms of the compression ratio.

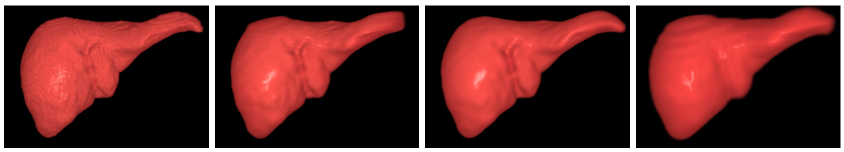
Figures 1 and 2 illustrate that, in terms of appearance, the DCT efficiently approximates the KLT derived by as relaxed tensor PCA for three-way array data. The pyramid transform for volumetric grey-valued images is an acceptable approximation of the KLT in low-compression late. Since the pyramid transform is a convolution operation, the time complexity of the transform is $\mathcal{O}(n\log_2 n)$. However, for high-compression ratio, the pyramid transform looses details of interior texture, although the transform preserves the appearance of outline shapes of the three-way array data.

Figure 4 shows dependencies of reconstruction error and cumulative contribution ratio to the dimensions of the space for reconstruction by using 3D-DCT (3DDCT) as a relaxed tenor PCA. In Fig. 4(a), we have 85, 93 and 71 nonzero eigenvalues for mode 1, 2 and 3, respectively. In Fig. 4 (b), we have 77, 70 and 63 nonzero eigenvalues for mode 1, 2 and 3, respectively. The reconstruction error and the cumulative contribution ratio, respectively, decreases and increases as the number of dimensions of the linear subspace increases.

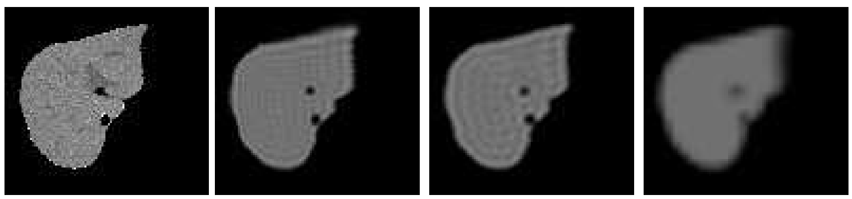
Figures 4(a) and 4(b) show that reconstruction errors and cumulative contribution ratios by using the 3D-DCT (3DDCT), the pyramid transform (PT) as a relaxed tenor PCA (TPCA). are similar properties for the compression ratios are $1/4$ and $1/2$. This property means the 3D-DCT processes the same compression properties with Karhunen-Loéve transform. Therefore, DCT is an acceptable relaxed form to KLT.

Table 1. Size and number of volumetric data. ♯data represents the number of volumetric data. The data size is the original size of the volumetric data. The reduced data size is the size of the volumetric data after tensor-representation-based dimension reduction.

	♯data	data size [voxel]	reduced data size [voxel]
CA dataset	32	$89 \times 97 \times 76$	$32 \times 32 \times 32$
Cardiac MRI dataset	340	$81 \times 81 \times 63$	$16 \times 16 \times 16$



(a) Original (b) TPCA (c) DCT (d) PT



(e) Original (f) TPCA (g) DCT (h) PT

Figure 1. Original and reconstructed volumetric data of a human liver. The left column illustrates the original volumetric data. The other three columns illustrate reconstructed volumetric data from the data compressed by the tensor PCA(TPCA), the three-dimensional DCT (DCT) and the pyramid transform (PT). (a)-(d) show the volume rendering of the volumetric images. (e)-(h) show the 30th axial slice of the volumetric images. The size of the reduced volumetric data is $32 \times 32 \times 32$. The compression ratio is 0.05, that is, the size is 5.0 % of the original size of $89 \times 97 \times 76$.

7. Conclusions

In this paper, we first showed that the PCA-based average computation for three-way data leads to tensor PCA. Secondly, we numerically showed that the triplet of DCT-II matrices are acceptable relaxed solutions for tensor PCA for the third order tensor. Moreover, we expressed the pyramid transform of volumetric data as the mode decomposition of tensor. This mathematical properties of the pyramid transform allow us to geometrically compare data reduction by the pyramid transform with tensor PCA for data compression and reduction using functional analysis.

In traditional method in medical image analysis, outline shapes of objects such as organs and statistical properties of interior textures are independently extracted using separate methods. However, tensor PCA for volumetric data allows us to simultaneously extract both the outline shapes of volumetric objects and the statistical properties of interior textures of volumetric data from data projected onto a low-dimensional linear subspace spanned by tensors.

This research was supported by the "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] T. Iijima: *Pattern Recognition,* Corona-sha, 1974.

[2] S. Watanabe *Knowing and Guessing: Quantitative Study of Inference and Information*, John Wiley and Sons, 1969.

[3] S. Watanabe *Pattern Recognition: Human and Mechanical* John Wiley and Sons, 1985.

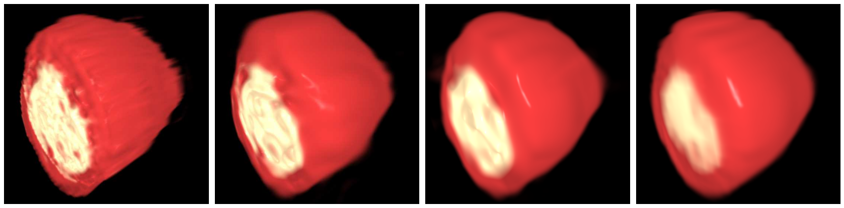
[4] S. Watanabe, Pattern recognition as conceptual morphogenesis, IEEE PAMI, **2**, 161-165 1980.

[5] E. Oja, *Subspace Methods of Pattern Recognition*, Research Studies Press, 1983.

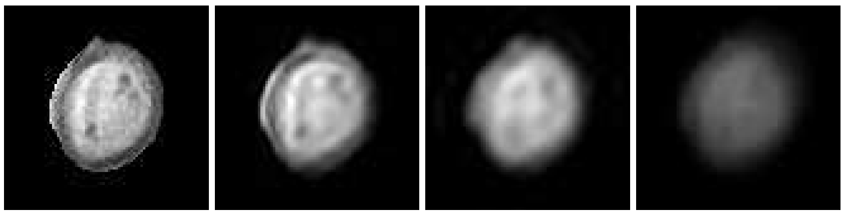
[6] A. Hyvärinen, J. Karhunen, E.Oja, *Independent Component Analysis*, Wiley-Interscience, 2001.

[7] A. Cichocki, R. Zdunek, A.-H. Phan, S. Amari: *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, Wiley, 2009.

[8] M. Itskov: *Tensor Algebra and Tensor Analysis for Engineers*, Springer, 2013.

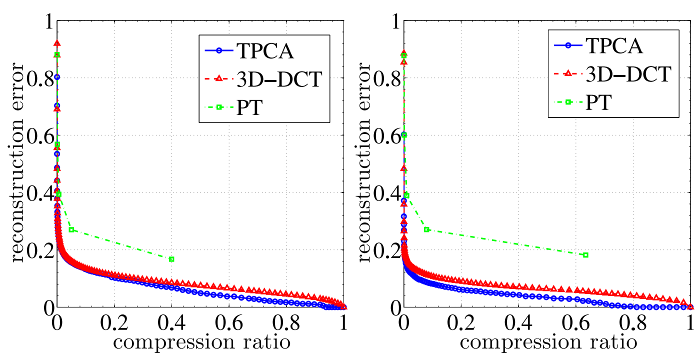


(a) Original (b) TPCA (c) DCT (d) PT



(e) Original (f) TPCA (g) DCT (h) PT

Figure 2. Original and reconstructed volumetric data of a human left ventricle. The left column illustrates the original volumetric data. The other three columns illustrate the reconstructed volumetric data from the data compressed by the tensor PCA (TPCA), the three-dimensional DCT (DCT) and the pyramid transform (PT). (a)-(d) show volume the rendering of the volumetric images. (e)-(h) show the first axial slice of the volumetric images. The size of reduced volumetric data is $16 \times 16 \times 16$. The compression ratio is 0.01, that is, the size is 1.0 % of the original size $81 \times 81 \times 63$.



(a) Liver (b) Heart

Figure 3. Reconstruction error vs. compression ratio. The reconstruction error is given as a relative error $\|\boldsymbol{X} - \hat{\boldsymbol{X}}\|_{\rm F}/\|\boldsymbol{X}\|_{\rm F}$, where $\boldsymbol{X}$ and $\hat{\boldsymbol{X}}$ are original and dimension-reduced volume data, respectively. The compression ratio is given as d/n, where d and n are the reduced size and the original size, respectively. In (a) and (b) the original sizes are $86 \times 97 \times 76$ and $81 \times 81 \times 63$, respectively. The reduced size are given by $i \times j \times k$, for the original size $I \times J \times K$, and $i = 1, 2, \dots I$, $j = 1, 2, \dots J$ and $k = 1, 2, \dots K$ for the tensor PCA (TPCA) and three-dimensional DCT (3D-DCT). For the pyramid transform (PT), reduced sizes are $l \times l \times l$ for $l = 4, 8, 16, 32, 64$.

[9] M. Mørup: Applications of tensor (multiway array) factorizations and decompositions in data mining, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, **1**, 24-40, 2011.

[10] T. G. Kolda, B. W. Bader: Tensor decompositions and applications, SIAM Review, **51**, 455-500. 2009.

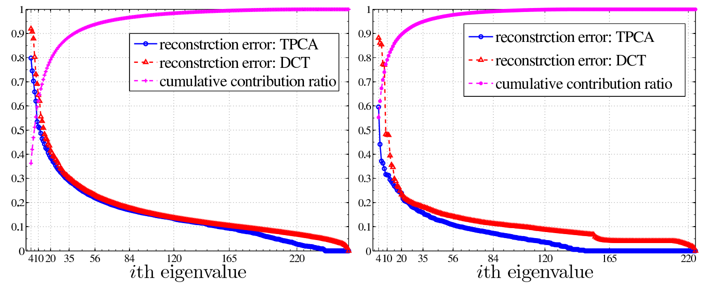
[11] M. Hamidi, J. Pearl: Comparison of the cosine Fourier transform of Markov-1 signals, IEEE ASSP, **24**, 428-429, 1976.

[12] P. M. Kroonenberg: *Applied Multiway Data Analysis*, Wiley, 2008.

[13] J. M. Marron, A. M. Alonso: Overview of object oriented data analysis, Biometrical Journal, **56**, 732-753, 2014.

[14] M. Ferrer, E. Valveny, F. Serratosa, K. Riesen, H. Bunke: Generalized median graph computation by means of graph

[tbh]



(a) Liver (b) Heart

Figure 4. Reconstruction errors and cumulative contribution ratios by using the 3D-DCT (3DDCT), the pyramid transform (PT) as a relaxed tenor PCA (TPCA). The reconstruction error is given as a relative error $\|\boldsymbol{X} - \hat{\boldsymbol{X}}\|_{\mathrm{F}}/\|\boldsymbol{X}\|_{\mathrm{F}}$, where $\boldsymbol{X}$ and $\hat{\boldsymbol{X}}$ are the original and the dimension-reduced volume data. From the 1st to nth frequency, cumulative contribution ratio is computed by $\sum_{l=1}^{n} \lambda_l / \sum_{l=1}^{N} \lambda_l$, where λ_l is the eigenvalues of three-modes in the descending order. In (a) and (b) the original sizes are $76 \times 86 \times 97$ and $81 \times 81 \times 63$, respectively. The reduced size are given by $i \times j \times k$, for the original size $I \times J \times K$, and $i = 1, 2, \ldots I$, $j = 1, 2, \ldots J$ and $k = 1, 2, \ldots K$.

embedding in vector spaces, Pattern Recognition, **43**, 1642-1655, 2010.

[15] T. M. W. Nye: Principal component analysis in the space of phylogenetic trees, The Annals of Statistics, **39**, 2716-2739, 2011.

[16] N. Otsu: *Mathematical Studies on Feature Extraction in Pattern Recognition,* Researches of The Electrotechnical Laboratory, **818**, 1981 (in Japanese).

[17] U. Grenander, M. Miller: *Pattern Theory: From Representation to Inference*, OUP, 2007.

[18] A. Malcev:*Foundations of Linear Algebra*, (in Russian edition) 1948, (English translation W.H. Freeman and Company, 1963).

[19] H. Itoh, A. Imiya, T. Sakai, Pattern recognition in multilinear space and its applications: mathematics, computational algorithms and numerical validations, Mach. Vis. Appl. **27**, 1259-1273, 2016.

[20] G. Strang, T. Nguyen: *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1996.

[21] G. Strang, *Computational Science and Engineering*, Wellesley-Cambridge Press, 2007.

[22] A. Andreopoulos, J. K. Tsotsos: Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI, Medical Image Analysis, **12**, 335-357, 2008.

Appendix:Alternative Projection Method

The Euler-Lagrange equation

$$
\begin{aligned}
&J_{33}(\boldsymbol{U}^{(1)}, \boldsymbol{U}^{(2)}, \boldsymbol{U}^{(3)}) \\
&= E_k(|\mathcal{X}_k \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \times_3 \boldsymbol{U}^{(3)}|_F^2) \\
&+ \sum_{i=1}^{3} |(\boldsymbol{I} - \boldsymbol{U}^{(i)^\top} \boldsymbol{U}^{(i)}) \boldsymbol{\Lambda}^{(i)}|_F^2.
\end{aligned}
$$

is solved by the following iteration procedure.

1: $\boldsymbol{U}_0^{(i)} := \boldsymbol{Q}^{(i)}$ such that $\boldsymbol{Q}^{(i)\top} \boldsymbol{Q}^{(i)} = \boldsymbol{I}$ and $\alpha = 0$.

2: $\boldsymbol{U}_{(\alpha+1)}^{(1)} = arg\min J_{33}(\boldsymbol{U}^{(1)}, \boldsymbol{U}_{(\alpha)}^{(2)}, \boldsymbol{U}_{(\alpha)}^{(3)})$.

3: $\boldsymbol{U}_{(\alpha+1)}^{(2)} = arg\min J_{33}(\boldsymbol{U}_{(\alpha+1)}^{(1)}, \boldsymbol{U}^{(2)}, \boldsymbol{U}_{(\alpha)}^{(3)})$.

4: $\boldsymbol{U}_{(\alpha+1)}^{(3)} = arg\min J_{33}(\boldsymbol{U}_{(\alpha+1)}^{(1)}, \boldsymbol{U}_{(\alpha+1)}^{(2)}, \boldsymbol{U}^{(3)})$.

5: if $|\boldsymbol{U}_{(\alpha+1)}^{(i)} - \boldsymbol{U}_{(\alpha)}^{(i)}|_F \leq \varepsilon$, then stop, else $\alpha := \alpha + 1$ and go to step 2.