Fusion Tensor Subspace Transformation Framework

Su-Jing Wang1,2,3*, Chun-Guang Zhou2,3, Xiaolan Fu1

1 State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China, 2 College of Computer Science and Technology, Jilin University, Changchun, Jilin, China, 3 Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun, P.R. China

Abstract

Tensor subspace transformation, a commonly used subspace transformation technique, has gained more and more popularity over the past few years because many objects in the real world can be naturally represented as multidimensional arrays, i.e., tensors. For example, a RGB facial image can be represented as a three-dimensional array (or 3rd-order tensor). The first two dimensionalities (or modes) represent the facial spatial information and the third dimensionality (or mode) represents the color space information. Each mode of the tensor may express a different semantic meaning. Thus different transformation strategies should be applied to different modes of the tensor according to their semantic meanings to obtain the best performance. To the best of our knowledge, there are no existing tensor subspace transformation algorithm which implements different transformation strategies on different modes of a tensor accordingly. In this paper, we propose a fusion tensor subspace transformation framework, a novel idea where different transformation strategies are implemented on separate modes of a tensor. Under the framework, we propose the Fusion Tensor Color Space (FTCS) model for face recognition.

Citation: Wang S-J, Zhou C-G, Fu X (2013) Fusion Tensor Subspace Transformation Framework. PLoS ONE 8(7): e66647. doi:10.1371/journal.pone.0066647

Editor: Randen Lee Patterson, UC Davis School of Medicine, United States of America

Received February 19, 2013; Accepted May 8, 2013; Published July 1, 2013

Copyright: © 2013 Wang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported in part by grants from 973 Program (2011CB302201), the National Natural Science Foundation of China (61075042, 61175023), China Postdoctoral Science Foundation funded project (2012M520428) and the open project program (93K172013K04) of Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: wangsujing@psych.cn.cn

Introduction

Subspace transformation (or subspace analysis [1]), a main type of feature extraction, has gained huge popularity over the past few years. Principal Component Analysis (PCA) [2] seeks the optimal projection directions according to maximal variances. Linear Discriminant Analysis (LDA) [3] uses discriminant information to search for the directions which are most effective for discrimination by maximizing the ratio between the between-class and within-class scatters. Both PCA and LDA aim to preserve global structures of the samples. Locality Preserving Projections (LPP) [4] aims to preserve the local structure of the original space in the projective subspace. Discriminant Locality Preserving Projections (DLPP) [5] encodes discriminant information into LPP to further improve the discriminant performance of LPP for face recognition. These algorithms need to vectorize the objects (samples).

In the real world, however, many objects are naturally represented by multidimensional arrays, i.e., tensors, such as a color facial image used in face recognition (see Fig. 1). If these objects are vectorized, their natural structure information will be lost [6]. As such, a great deal of interests are aroused in the field of tensor analysis [7][8][9][10][11]. Among the subspace transformation techniques, tensor subspace transformation has also become a highly discussed topic. Multilinear Principal Component Analysis (MPCA) [12], a tensor version of PCA, applies PCA transformation on each mode (or dimensionality) of tensors. Similarly, Discriminant Analysis with Tensor Representation (DATER) [13], General Tensor Discriminant Analysis (GTDA) [14], Tensor Subspace Analysis (TSA) [15], and Discriminant Tensor Subspace Analysis (DTSA) [16] apply LDA, Maximum Scatter Difference (MSD) [17], LPP, and DLPP to transform each mode of tensors, respectively. These tensor subspace transformation methods use a certain vector subspace transformation method to transform every modes of tensors.

However, each mode of tensors may express a different semantic meaning. For example, a color facial image can be treated as a 3rd-order tensor, where mode-1 and mode-2 represent the facial spatial information and mode-3 representing the color space information (see Fig. 1). The facial spatial information and color space information are two different types of information, which should be handled by two different transformations to obtain better performance. In other words, for color facial images, we should implement a transformation strategy on the first two modes and another transformation strategy on the third mode. As such, each type of information should implement a transformation strategy best suited for the semantic meaning.

To the best of our knowledge, there are no existing tensor subspace transformation algorithm, which implements different transformation strategies on different modes of tensors according to their semantic meanings. To address this problem, we propose the fusion tensor subspace transformation framework, which shows an novel idea that different transformation strategies can be implemented on different modes of tensors. Under the framework, we propose the Fusion Tensor Color Space (FTCS) model for face recognition.
Materials and Methods
Tensor Fundamentals and Denotations
A tensor is a multidimensional array. It is the higher-order generalization of scalar (zero-order tensor), vector (1st-order tensor), and matrix (2nd-order tensor). In this paper, lowercase italic letters (a, b, ...) denote scalars, bold lowercase letters (\(\mathbf{a}, \mathbf{b}, \ldots\)) denote vectors, bold uppercase letters (\(\mathbf{A}, \mathbf{B}, \ldots\)) denote matrices, and calligraphic uppercase letters (\(\mathcal{A}, \mathcal{B}, \ldots\)) denote tensors. The formal definition is given below:[18]:

\[
\mathbf{A} = \begin{pmatrix}
A_{11} & A_{12} & \cdots & A_{1n} \\
A_{21} & A_{22} & \cdots & A_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
A_{m1} & A_{m2} & \cdots & A_{mn}
\end{pmatrix}
\]

Table 1. Fusion tensor subspace transformation framework.

| INPUT: M N-order tensors \(X_i \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_N}, i = 1, 2, \ldots, M\). |
|-----------------------------------------------|
| OUTPUT: \(U_{\mathcal{A}} \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_N}\) |
| Initialize \(U_{\mathcal{A}}\) with a set of identity matrices; |
| for \(t = 1\) to \(T_{\text{max}}\) do |
| for \(n = 1\) to \(N\) do |
| for \(i = 1\) to \(M\) do |
| \(Y_{\text{def}} = X_i \times_1 U_1^T \times_2 U_2^T \times_3 U_3^T \times \cdots \times_N U_N^T\); |
| \(Y_{\text{def}} \leftarrow \text{the mode-}n\text{-unfolding matrix of } Y_{\text{def}}\); |
| end for |
| end for |
| end for |
| if \(t > 2\) and \(\|U_{\mathcal{A}n} - U_{\mathcal{A}n-1}\|_F < \varepsilon\) then |
| end if |
| end for |

Under the framework, we developed the Fusion Tensor Color Space (FTCS) model for face recognition.


doi:10.1371/journal.pone.0066647.t001

Figure 1. The tensorial represents of a color facial image.
doi:10.1371/journal.pone.0066647.g001

Figure 2. Sample images of one individual from the AR database.
doi:10.1371/journal.pone.0066647.g002

Definition 1. The order of a tensor \(\mathbf{A} \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_N}\) is \(N\). An element of \(\mathbf{A}\) is denoted by \(A_{i_1i_2i_3 \ldots i_N}\) or \(a_{i_1i_2i_3 \ldots i_N}\), where \(1 \leq i_n \leq I_n, n = 1, 2, \ldots, N\).

Definition 2. The mode-\(n\) product of a tensor \(\mathbf{A}\) is the \(I_n\)-dimensional vectors obtained from \(\mathbf{A}\) by fixing every index but index \(i_n\).

Definition 3. The mode-\(n\) unfolding matrix of \(\mathbf{A}\), denoted by \((A_{ij})_{(i_1-1)I_1+1 \times (i_2-1)I_2+1 \times \cdots \times (i_N-1)I_N+1})\), contains the element \(a_{i_1i_2 \ldots i_N}\) at \(i_1\)th row and \(i_2\)th column, where

\[
j = 1 + \sum_{k=1, k \neq n}^{N} (i_k - 1)J_k, \quad \text{with} \quad J_n = \frac{I_n-1}{m-1}, \quad m = 1, 2, \ldots, N. \quad (1)
\]

We can generalize the product of two matrices to the product of a tensor and a matrix.

Definition 4. The mode-\(n\) product of a tensor \(\mathbf{A} \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_N}\) by a matrix \(\mathbf{U} \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_N}\), denoted by \(\mathbf{A} \times_n \mathbf{U}\), is an 

\[
(I_1 \times I_2 \times \cdots \times I_N) \times \mathbf{U} \quad \text{tensor of which the entries are given by:}
\]

\[
(A \times_n U)_{i_1 \times \cdots \times i_N} \equiv \sum_{m=1}^{I_N} a_{i_1 \times \cdots \times i_m \times j_1 \times \cdots \times j_{N-1}} U_{j_1 \times \cdots \times j_{N-1}}. \quad (2)
\]

Definition 5. The scalar product of two tensors \(\mathbf{A} \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_N}\) and \(\mathbf{B} \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_N}\), denoted by \(\langle \mathbf{A}, \mathbf{B} \rangle\), is defined in a straightforward way as

\[
\langle \mathbf{A}, \mathbf{B} \rangle \equiv \sum_{i_1} \sum_{i_2} \cdots \sum_{i_N} a_{i_1 \times \cdots \times i_N} b_{i_1 \times \cdots \times i_N}. \quad (3)
\]

The Frobenius norm of a tensor \(\mathbf{A} \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_N}\) is then defined as

\[
\|\mathbf{A}\|_F = \sqrt{\langle \mathbf{A}, \mathbf{A} \rangle}. \quad (4)
\]

From the definition of the mode-\(n\) unfolding matrix, we have

\[
\|\mathbf{A}\|_F = \|\langle \mathbf{A}, \mathbf{A} \rangle\|_F \quad (5)
\]

By using tensor decomposition, any tensor \(\mathbf{A}\) can be expressed as the product

\[
\mathbf{A} = \mathbf{C} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3 \times \cdots \times_N \mathbf{U}_N \quad (6)
\]

where \(\mathbf{U}_n \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N} \) is an orthonormal matrix and contains the ordered principal components for the \(n\)th mode; \(\mathbf{C}\) is called the core tensor. Unfolding the above equation, we have

\[
\mathbf{A}_{(i_1)} = \mathbf{U}_n \mathbf{C}_{(i_1)} \mathbf{U}_{n+1} \otimes \cdots \otimes \mathbf{U}_{n+1} \otimes \mathbf{U}_{n-1} \otimes \cdots \otimes \mathbf{U}_1^T \quad (7)
\]

where \(\otimes\) is the Kronecker product of the matrices.

The Connection among PCA, 2D-PCA and MPCA
Before introducing the fusion tensor subspace transformation framework, we firstly investigate the connection among PCA, 2D-
PCA [19] and MPCA. From the previous section, we know that a tensor is the higher-order generalization of scalar, vector and matrix. Similarly, MPCA is the higher-order generalization of PCA and 2D-PCA.

Suppose there are $M$ $N$-order tensors $\mathbf{X}_i \in \mathbb{R}^{l_1 \times l_2 \times \ldots \times l_N}$, $i = 1, 2, \ldots, M$. MPCA seeks $N$ projection matrices $\mathbf{U}_1 \in \mathbb{R}^{l_1 \times L_1}$, $\mathbf{U}_2 \in \mathbb{R}^{l_2 \times L_2}$, $\ldots$, $\mathbf{U}_N \in \mathbb{R}^{l_N \times L_N}$ in order to transform $\mathbf{X}_i$ as

$$Y_i = \mathbf{X}_i \times_1 \mathbf{U}_1^T \times_2 \mathbf{U}_2^T \times_3 \ldots \times_N \mathbf{U}_N^T,$$

such that $Y_i, (i = 1, 2, \ldots, M)$ captures most of the variations observed in the original tensor objects $\mathbf{X}_i$.
In Eq. (7),
\[
\Phi^{(o)} = \sum_{m=1}^{M} (X_{m(o)} - \bar{X}_{(o)}) U_{q_m(o)}^T (X_{m(o)} - \bar{X}_{(o)})^T
\]

where \(\bar{X}_{(o)}\) denotes the mode-n unfolding matrix of the mean values of \(X\), \(X_{m(o)}\) denotes the mode-n unfolding matrix of the mean values of \(X_m\)

\[
U_{q_m(o)} = (U_{n+1} \otimes U_{n+2} \otimes \ldots \otimes U_1 \otimes \ldots \otimes U_{n-1}).
\]

In Eq. (7), \(U_{q_m(o)}\) is to use the fixed \(N-1\) projection matrices \(U_1, \ldots, U_{n-1}, U_{n+1}, \ldots, U_N\) to transform the corresponding \(N-1\) modes. When \(X^{(i)}(i=1,2,\ldots,M)\) is a 2nd-order tensor, Eq. (6) is simplified to

\[
y_i = x_i \times 1 U_1^T \times 2 U_2^T, \quad i = 1,2,\ldots,M.
\]

If we only transform mode-2, \(U_1\), is an identity matrix of size \(I_1\). Then, \(U_{q(2)}\) is also an identity matrix and \(X_{m(2)} = X^T\). In this case, Eq. (7) is simplified to

\[
\Phi^{(2)} = \sum_{m=1}^{M} (X_m^T - X^T)(X_m - X)^T
\]

\[
= \sum_{m=1}^{M} (X_m - X)^T (X_m - X)
\]

Eq. (10) is exactly the image covariance (scatter) matrix \(G\) in 2D-PCA [19]. So, 2D-PCA is a special case of MPCA. When objects are represented by matrices and only rows of matrices need to be transformed, MPCA degenerates into 2D-PCA.

When \(X^{(i)}(i=1,2,\ldots,M)\) are 1st-order tensors, Eq. (6) is simplified to

\[
y_i = x_i \times 1 U_1^T, \quad i = 1,2,\ldots,M.
\]

In this case, Eq (7) is simplified to

\[
\Phi^{(1)} = \sum_{m=1}^{M} (X_m - x)(X_m - x)^T
\]

The above equation is exactly the scatter matrix in PCA. So, PCA is a special case of MPCA. When objects are represented by vectors, MPCA degenerates into PCA.

Similarly, through the above analysis, one can notice that DATER also applies LDA transformation on all modes of tensors. Likewise, GTDA, TSA and DTS also apply MSD, LPP and DLPP transformation on all modes of tensors respectively. There are several other tensor subspace transformation methods that also apply a single type of transformation on all modes of tensors, however due to page limit we chose to only mention a portion of these algorithms.

Fusion tensor subspace transformation framework

Tensor subspace transformation method firstly initializes \(N\) projection matrices \(U_1, \ldots, U_{n-1}, U_n, U_{n+1}, \ldots, U_N\) as identity matrices or random matrices, then fixes \(N-1\) projection matrices \(U_1, \ldots, U_{n-1}, U_{n+1}, \ldots, U_N\). Following, the matrices are used to transform \(X^{(i)}\), and the transformed results are unfolded on mode- \(n\). Finally, \(U_n\) is obtained by implementing a certain transformation on the mode-\(n\) unfolding matrices. We can see that the solution of \(U_n\) depends on the other projection matrices. \(N\) projection matrices are solved by constructing an iterative procedure.

Existing tensor subspace transformation methods only implement one transformation strategy on all modes. In the real world, each mode of tensors may represent a different type of information. We should implement different transformation strategies on different modes according to their semantic meaning. So we propose a Fusion Tensor Subspace Transformation (FTSA) framework, which is described in Table 1. In Table 1, the statement denoted [*] is the core statement in the framework. For a certain represented object, different transformations are used on different modes according to their semantic meaning.

In the algorithm, we use the maximal iterative times \(T_{\text{max}}\) to deal with the problem that the algorithms may be not convergent. Actually, the convergence of many tensor subspace analysis algorithms cannot be generally proved, the classification results based on these algorithms show to be stable after rounds of iterations as illustrated in these papers (e.g. DATER, 2D LDA [20]). The convergence of FTSA depends on the specific transformations.
Fusion tensor color space model

Recently, researches showed that color information may help to improve the face recognition accuracy. While, the R, G, and B component images in the RGB color space are correlated. Decorrelation among the components of these images helps reduce redundancy and is an important strategy to improve the accuracy of subsequent recognition method [21]. Liu [22] proposed the Uncorrelated Color Space (UCS), the Independent Color Space (ICS), and the Discriminating Color Space (DCS). Specifically, the UCS applies PCA to decorrelate the R, G, and B component images. The ICS and DCS further enhance the discriminating power for the subsequent recognition method by means of Independent Component Analysis (ICA [23]) and LDA, respectively. The experimental results showed that ICS obtains the best color space because its components are not only uncorrelated but also independent.

Many papers have reported that the discriminant analysis methods on facial images can enhance subsequent recognition method [3][5]. Color Image Discriminant model (CID) [24], borrowing the idea of LDA, aims to seek an optimal color space and an effective recognition method of color images in a unified framework. Tensor Discriminant Color Space (TDCS) [25] model, borrowed the idea of DATER [13], seeks two discriminant projection matrices $U_1$, $U_2$ corresponding to the facial spatial information and one color space transformation matrix $U_3$ corresponding to the color space. Actually, TDCS uses LDA

Figure 6. Illustration of R, G, and B color components and the various components generated by CID, TDCS and FTCS on the LFW face database.
doi:10.1371/journal.pone.0066647.g006

Figure 7. ROC curves of FTCS, TDCS, CID, 2D-LDA and LDA on the LFW face database. Here, $N$ projection matrices $U_n$ ($n = 1, 2, \ldots, N$) need to be determined. Generally, we fix $N-1$ projection matrices $U_1, U_2, \ldots, U_{n-1}, U_{n+1}, \ldots, U_N$ to solve for $U_n$. In MPCA, $U_n$ consists of the $L_n$ eigenvectors corresponding to the largest $L_n$ eigenvalues of the matrix
doi:10.1371/journal.pone.0066647.g007
transformation on both facial spatial information and color space information. They [26] also used elastic net to propose Sparse Tensor Discriminant Color Space (STDCS).

For color space information, however, ICA transformation is better than LDA transformation [22]. Motivated by the insights, we explore a Fusion Tensor Color Space (FTCS) model which applies discriminant analysis on the facial spatial information and applies ICA on the color space information.

A color facial image is naturally represented by a 3rd-order tensor, where mode-1 and mode-2 of a tensor are facial spatial information and mode-3 of the color tensor is the color information. For instance, a RGB image with size $I_1 \times I_2$ is represented as a tensor $A \in \mathbb{R}^{I_1 \times I_2 \times L_3}$, where $I_3 = 3$. The mode-3 of $A$ is the color variable in the RGB color space which has 3 components corresponding to $R$, $G$ and $B$ in RGB space. FTCS uses LDA on the first two modes and ICA on the third mode.

Assuming $C$ is the number of individuals, $A_{i,c}$ is the $c$th color facial image of the $i$th individual, and $M_i$ is the number of color facial images of the $i$th individual, where $M = M_1 + M_2 + \ldots + M_C$. The FTCS algorithm seeks two discriminant projection matrices $U_{i,c} \in \mathbb{R}^{I_1 \times L_1}$, $U_c \in \mathbb{R}^{I_2 \times L_2}$ and a color space transformation matrix $U_c \in \mathbb{R}^{L_3 \times L_2}$ (usually $L_1 < I_1$, $L_2 < I_2$ and $L_3 \leq I_3$) for transformation

$$Y_i = U_i^T \times 2 U_i^T \times 3 U_i^T,$$

where $U_1$ and $U_2$ are obtained by using discriminant analysis and $U_3$ is obtained by using ICA.

The mean image of the $c$-th individual and the mean image of all individuals are defined by:

$$\mathcal{X}_i = \frac{1}{M_i} \sum_{i=1}^{M_i} A_{i,c}^c$$

and

$$\mathcal{X} = \frac{1}{C} \sum_{c=1}^{C} \mathcal{X}_c$$

The between-class scatter and within-class of color images are defined as:

$$\Psi_b(\mathcal{X}) = \sum_{c=1}^{C} \| \mathcal{X} - \mathcal{X}_c \|_F^2$$

and

$$\Psi_w(\mathcal{X}) = \sum_{i=1}^{C} \sum_{c=1}^{M_i} \| \mathcal{X}_i^c - \mathcal{X}_c \|_F^2.$$  

We can define mode-$n$ between-class scatter matrix $S_{b}^{(n)}$ and mode-$n$ within-class scatter matrix $S_{w}^{(n)}$ as:

$$S_{b}^{(n)} = \sum_{c=1}^{C} \sum_{i=1}^{M_i} (X_{i,c}^c - X_{i,c}^w) U_i^T U_i (X_{i,c}^c - X_{i,c}^w)^T$$

and

$$S_{w}^{(n)} = \sum_{c=1}^{C} \sum_{i=1}^{M_i} (X_{i,c}^c - X_{i,c}^w) U_i^T U_i (X_{i,c}^c - X_{i,c}^w)^T,$$

where $U_i = U_{n+1} \otimes U_{n+2} \otimes U_{n+3} \otimes \ldots \otimes U_{n}$, $n = 1,2,3$.

Then, the between-class scatter of the projected tensors $\Psi_b(Y)$ and the within-class scatter of the projected tensors $\Psi_w(Y)$ can be rewritten as follows:

$$\Psi_b(Y) = tr \left( U_b^T S_{b}^{(n)} U_b \right)$$

and

$$\Psi_w(Y) = tr \left( U_w^T S_{w}^{(n)} U_w \right).$$

So, given $U_2$ and $U_3$ (or $U_1, U_2$), $U_1$ (or $U_2$) can be obtained by the following discriminant analysis:

$$\max_{U} \frac{tr \left( U_b^T S_{b}^{(n)} U_b \right)}{tr \left( U_w^T S_{w}^{(n)} U_w \right)} = n, 1,2$$

According to Rayleigh quotient, Eq. (21) is maximized if and only if the matrix $U_b$ consists of $L_n$ generalized eigenvectors, which corresponds to the largest $L_n$ generalized eigenvalues of the matrix pencil $(S_{b}^{(n)}, S_{w}^{(n)})$, which satisfies:

$$S_{b}^{(n)} u = \lambda S_{w}^{(n)} u \quad n, 1,2$$

Since $S_{b}^{(n)}$ and $S_{w}^{(n)}$ are dependent on $U_1, \ldots, U_{n-1}, U_{n+1}, \ldots, U_{n}$, we can see that the optimization of $U_b$ depends on the projections of other modes.

In order to obtain $U_3$, we use ICA (For ICA operations, we used Hyvarinen’s fixed-point algorithm http://www.cis.hut.fi/projects/ica/fastica/) to decorrelate the RGB color space, we use $U_1$ and $U_2$, which are obtained through the above discriminant analysis, to transform:

$$Y_i^{(3)} = X_i^T \times 1 U_1^T \times 2 U_2^T,$$

where $Y_i^{(3)} \in \mathbb{R}^{L_1 \times L_2 \times L_3}$. $M$ 3rd-order tensor $Y_i^{(3)}$ are concatenated to a 4th-order tensor $F \in \mathbb{R}^{L_1 \times L_2 \times L_3 \times M}$. The mode-3 unfolding matrix $F_{(3)}$ is a $3 \times K$ matrix, where $K = L_1 \times L_2 \times M$ and the three rows of $F_{(3)}$ corresponding to the three components in RGB space, respectively.

The color space transformation matrix $U_3$ may be derived using ICA on $F_{(3)}$. The ICA of $F_{(3)}$ factors the covariance matrix $\Sigma_F$ into the following form:

$$\Sigma_F = U_3^{-1} \Sigma U_3^T$$
where $V\in\mathbb{R}^{3 \times 3}$ is diagonal real positive and $U_3$ transforms RGB color space to a new color space whose three components are independent or the most independent three component possible. The $U_3$ in Eq. (24) may be derived using Comon’s ICA algorithm by calculating mutual information and high-order statistics. As a result, an iterative procedure can be constructed to obtain $U_1$, $U_2$ and $U_3$.

**Results**

Experiments and results on the AR database

The AR database contains over 4,000 color facial images of 126 people. Each individual participated in two photo sessions. In both sessions, the pictures were taken under identical requirements and conditions. In our experiments, we selected 100 people, where 14 images of each individual are selected and occluded face images are excluded. These facial images have been cropped [27] and can be downloaded from the AR face database official web (http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html). All images are cropped and resized to $32 \times 32$ pixels. The sample images for one individual of the AR database are shown in Fig. 2, where the images on the top row are from the first session as the training set, and the images on the bottom row are from the second session as the testing set.

In this experiment, we trained FTCS, TDCS (Although STDCS [26] is better than TDCS, we still did not compare FTCS to STDCS. Because the motivation of the paper is to implement different transformations on different modes of a tensor.) and CID. The convergence threshold $\varepsilon$ was set as 0.1 and $x_i$ was initialized as $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}^T$. In this case, we got three color space transformation matrices:

$$U^{CID} = \begin{bmatrix} 0.4126 & -0.2107 & -0.5558 \\ -0.0261 & -0.4683 & 1.0536 \\ 1.0000 & 0.9739 & -0.5524 \end{bmatrix},$$  \quad (25)$$

$$U^{TDCS} = \begin{bmatrix} 0.1267 & -0.2084 & 0.3358 \\ -0.2128 & -0.4168 & -0.7897 \\ 0.9689 & 0.8848 & 0.5134 \end{bmatrix},$$  \quad (26)$$

and

$$U^{FTCS} = \begin{bmatrix} -0.2220 & -0.1853 & 0.6189 \\ -0.6473 & 0.0507 & 0.2239 \\ 0.7292 & -0.9814 & 0.7529 \end{bmatrix}.$$  \quad (27)$$

Using these three matrices, we obtained three color components $\mathbf{D}^1, \mathbf{D}^2, \mathbf{D}^3$ of CID; three color components $\mathbf{T}^1, \mathbf{T}^2, \mathbf{T}^3$ of TDCS and three color components $\mathbf{F}^1, \mathbf{F}^2, \mathbf{F}^3$ of FTCS (see Fig. 3).

Meanwhile, we carried out LDA and 2D-LDA on corresponding gray images. In LDA and CID, only 99 discriminant projection basis vectors were extracted. For 2D-LDA, TDCS and FTCS, the spatial dimensions of the two modes are both reduced to 10. The score matrices were generated by Manhattan distance and Euclidean distance, respectively. The ROC curves of the five methods are shown in Fig. 4. The results indicate that the performance of FTCS with Manhattan distance obtains the best performance. However, the space between two curves of FTCS is narrower than the space between two curves of TDCS. This shows that FTCS is more robust to the type of distance used and results of both Manhattan and Euclidean distance produces closer results than those of TDCS.

In the five algorithms, LDA and CID are used on vectorized face images. Overall, their performances are poorer than the other three algorithms based on tensorized face images. This shows that the facial spatial structure information is important to face recognition. Specially, when the false accept rate is less than 0.03, the performance of 2D-LDA outperforms that of TDCS with the color information. In the case, the color information fails to work for face recognition. This is due to the fact that the color space information is transformed by LDA, which is not an optimal transformation for color space information in comparison to ICA. Whereas, FTCS uses ICA on the color space information. As a results, FTCS obtains the best performance.

Table 2 list the verification rates of the five methods with 0.1 FAR. In both cases of Manhattan distance and Euclidean distance, FTCS gets the best verification rate among the five methods. For Manhattan distance, 2D-LDA without the color information is better than TDCS with the color information. Whereas, FTCS outperforms 2D-LDA. This also shows that ICA is better than LDA to transform the color space information.

Experiments and results on the LFW face database

$$U^{CID} = \begin{bmatrix} 1.0000 & -1.0089 & -0.3600 \\ 0.3166 & 0.2805 & 0.9981 \\ -0.2236 & 0.9078 & -0.6621 \end{bmatrix},$$  \quad (28)$$

$$U^{TDCS} = \begin{bmatrix} -0.9103 & 0.6923 & 0.1885 \\ 0.4085 & -0.7043 & -0.7688 \\ -0.0673 & -0.1571 & 0.6111 \end{bmatrix},$$  \quad (29)$$

and

$$U^{FTCS} = \begin{bmatrix} -0.3602 & 0.8170 & -0.9108 \\ 0.9032 & -0.5766 & 0.1319 \\ 0.2336 & -0.0017 & -0.3912 \end{bmatrix}.$$  \quad (30)$$

These components are illustrated in Fig. 6.

These three matrices are not the same as Eq. (25), Eq. (26) and Eq. (27) due to the different training sets. Using these three matrices, we got three color components $\mathbf{D}^1, \mathbf{D}^2, \mathbf{D}^3$ of CID; three color components $\mathbf{T}^1, \mathbf{T}^2, \mathbf{T}^3$ of TDCS and three color components $\mathbf{F}^1, \mathbf{F}^2, \mathbf{F}^3$ of FTCS.
Discussion

Recently, tensor subspace transformation is a highly mentioned topic, because many real objects can be represented by tensors. For different objects, the semantic meaning of tensorial modes are different. Even when the objects are the same, each mode of tensors may express a different semantic meaning. To the best of our knowledge, there aren’t any existing tensor subspace transformation algorithms which implements different transformation strategies on different mode of tensors according to their semantic meaning. In this paper, we propose the fusion tensor subspace analysis framework, which shows an novel idea that different transformation strategies can applied on different modes of tensor. Under the framework, we propose FTCS for face recognition. The experimental results show the performances of the proposed algorithm is better than existing tensor subspace transformation algorithms. FTCS is only an example of fusion tensor subspace transformation framework. Under the framework, many algorithms can be developed for action recognition, micro-expression recognition, EEG recognition and so on.

Author Contributions
Conceived and designed the experiments: SW CZ XF. Performed the experiments: SW. Analyzed the data: SW. Wrote the paper: SW.

References
1. Wang X, Tang X (2004) A unified framework for subspace face recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 26: 1222–1228.
2. Turk M, Pentland A (1991) Eigenfaces for recognition. Journal of Cognitive Neuroscience 3: 71–86.
3. Belhumeur PN, Hespanha JP, Kriegman DJ (1997) Eigenfaces vs. Fisherefaces: recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence 19: 711–720.
4. He XF, Niyogi P (2004) Locality preserving projections. Advances In Neural Information Processing Systems 16: 153–160.
5. Yu WW, Teng XL, Liu CQ (2006) Face recognition using discriminant locality preserving projections. Image and Vision Computing 24: 239–248.
6. Lu H, Plataniotis K, Venetsanopoulos A (2011) A survey of multilinear subspace learning for tensor data. Pattern Recognition 44: 1540–1551.
7. Tao D, Li X, Wu X, Hu W, Maybank SJ (2007) Supervised tensor learning. Knowledge and Information Systems 13: 1–42.
8. Zhang L, Zhang L, Tao D, Huang X (2011) A multifeature tensor for remote-sensing target recognition. Geoscience and Remote Sensing Letters, IEEE 8: 374–378.
9. Tao D, Li X, Wu X, Maybank S (2008) Tensor rank one discriminant analysis – convergent method for discriminative multilinear subspace selection. Neurocomputing 71: 1866–1882.
10. Tao D, Song M, Li X, Shen J, Sun J, et al. (2008) Bayesian tensor approach for 3-d face modeling. Circuits and Systems for Video Technology, IEEE Transactions on 18: 1397–1410.
11. Sun J, Tao D, Psapartitiou S, Yu PS, Faloutsos C (2008) Incremental tensor analysis: Theory and applications. ACM Transactions on Knowledge Discovery from Data (TKDD) 2: 11.
12. Lu HP, Konstantinos NP, Venetsanopoulos AN (2008) MPCA: Multilinear principal component analysis of tensor objects. IEEE Transactions on Neural Networks 19: 18–39.
13. Yan S, Xu D, Yang Q, Zhang L, Tang X, et al. (2005) discriminant analysis with tensor representation. CVPR 2005: 11.