Discriminative High Order SVD: Adaptive Tensor Subspace Selection for Image Classification, Clustering, and Retrieval

Dijun Luo, Heng Huang, Chris Ding
The University of Texas at Arlington
701 S Nedderman Drive, Arlington, Texas, USA
dijun.luo@gmail.com, \{heng, chqding\}@uta.edu

Abstract

Tensor based dimensionality reduction has recently attracted attention from computer vision and pattern recognition communities for both feature extraction and data compression. As an unsupervised method, High-Order Singular Value Decomposition (HOSVD) searches for low-rank subspaces such that the low-rank approximation error is minimized. The data projection on subspace can be used as features for different computer vision tasks such as image clustering, classification, and retrieval. However, without the class labels, the discriminative power of selected features is restricted. In this paper, we propose a new unsupervised high-order tensor decomposition approach which employs the strength of discriminative analysis and \textit{K}-means clustering to adaptively select subspaces that improve the clustering, classification, and retrieval capabilities of HOSVD. We provide both theoretical analysis to guarantee that our new method generates more discriminative subspaces and empirical studies on several public computer vision data sets to show the consistent improvement of image clustering, classification, and content-based image retrieval results over existing methods.

1. Introduction

Subspace learning methods such as Principal Component Analysis (PCA), Singular Value Decomposition (SVD), and Linear Discriminant Analysis (LDA) were popular as tools for the analysis of two-dimensional arrays of data in a wide variety of computer vision applications, such as previous research in papers [10, 15, 2, 17]. As subspace analysis approaches, PCA and SVD employ unsupervised ways to learn the low-dimensional structure of high-dimensional data, but LDA is a supervised method and seeks for a set of vectors that maximized Fisher Discriminant Criterion. LDA simultaneously minimizes the within-class scatter while maximizing the between-class scatter in the projective feature vector space.

Although PCA/SVD and LDA work well for data dimension reduction of two-dimensional arrays, they traditionally express the input data as vectors and it is not natural to apply them into higher dimensional data, known as high-order data. For high-order data we can use Tucker decomposition [16]. After adding orthogonality constraints to tensor decomposition, the High-Order Singular Value Decomposition (HOSVD) [11] has been widely used in computer vision applications [18]. Several other tensor based methods have been also proposed, e.g. Two Dimensional PCA (2DPCA) [21], Generalized Low Rank Approximation of Matrices (GLRAM) [23], Two Dimensional Singular Value Decomposition (2DSVD) [6], rank-R tensors [19], dynamic and streaming tensor factorization [14]. Substantial efforts have also been devoted to improve the vector-based classical LDA. Ye \textit{et al.} [24] introduced two dimensional LDA (2DLDA) with an iterative solution algorithm, while paper [12] solved the ambiguity problem in 2DLDA. Both of them consider the projection of the data onto a space which is the tensor product of two vector spaces. Yan \textit{et al.} [20] used tensor representation in discriminative analysis. In previous research, HOSVD usually is used for data compression and classification. In paper [9], the authors pointed out HOSVD simultaneously does subspace selection and tensor clustering. In computer vision applications, more discriminative subspaces usually are desired to be selected to provide better clustering, classification, or retrieval tasks.

In this paper, we introduce discriminative analysis into tensor decomposition to propose a novel Discriminative HOSVD that achieves better clustering/classification/retrieval results. First, an Adaptive 2DLDA method is proposed to find the more discriminative subspaces \(U\) and \(V\). In our Adaptive 2DLDA method, the K-means clustering and 2DLDA algorithms are iteratively used to find class labels and to obtain subspaces \(U\) and \(V\) unsupervisedly. After that, we calculate matrix \(W\) with improved clustering capability based on our theoretical proof. We emphasize that the Discriminative HOSVD is still an...
unsupervised learning method. Therefore, it can relieve human effort of classifying large amount data points. In our extensive experiments on four datasets, our Discriminative HOSVD outperforms the standard unsupervised learning methods (PCA, HOSVD, and GLRAM) in clustering, classification, and retrieval accuracy.

Although many studies on tensor factorization appeared in computer vision areas, to our knowledge, this paper is the first one to propose and prove the feasibility of discriminative analysis, a supervised learning idea, in unsupervised tensor factorization. Our novel Adaptive 2DLDA and Discriminative HOSVD approaches have broad applications in high dimensional image clustering, classification, and image retrieval.

In previous research, there are several papers ever explored the relations between unsupervised dimension reduction and unsupervised learning [3, 25, 4, 5]. But all previous methods only work for vector based input data, not for high dimensional data (e.g. images and videos). The main contributions of this paper are the following:

(a) We show that 2DLDA and tensor clustering are optimizing the same objective function, i.e., they both minimize the within-class scatter and maximize the between-class scatter.

(b) Based on the above theoretical analysis, we show that the objective function for our proposed Adaptive 2DLDA provides a natural generalization which combines the strength of 2DLDA and K-means tensor clustering to select more discriminative subspaces $U$ and $V$ that provide better tensor classification results.

(c) A new Discriminative HOSVD is proposed to improve the clustering capability of standard HOSVD, i.e., a better clustering structure exists in $W$. We also prove Discriminative HOSVD is equivalent to tensor clustering with subspaces selected by Adaptive 2DLDA.

### 2. Related Work

In computer vision areas, HOSVD has been widely used for dimensionality reduction and data compression. In previous work, researchers showed the HOSVD can do simultaneous subspace selection and K-means clustering [9], i.e., the clustering structure exists in matrix $W$ of HOSVD in (2). In this section, we will introduce the related previous research work, including HOSVD and tensor clustering.

#### 2.1. High Order SVD

HOSVD searches low-rank matrices and a core tensor to approximate high order tensors. Given a 3D tensor $X = \{X_1, \cdots, X_n\}$ where each $X_i$ is a 2D matrix of size $n_1 \times n_2$, the standard HOSVD factorizations is to solve:

$$\min_{U, V, W, S} J_1 = \|X - U \otimes_1 V \otimes_2 W \otimes_3 S\|^2$$

subject to $U^T U = I, V^T V = I, W^T W = I$.

where $U, V, W$ are 2D matrices and $S$ is a 3D core tensor: $S \in \mathbb{R}^{k_1 \times k_2 \times k_3}$.

#### 2.2. Tensor Clustering

Given a three dimensional tensor $X$, or, equivalently a set of two dimensional images $X_1, X_2, \cdots, X_n$, the $K$-means tensor clustering minimizes

$$\min_{\{C_k\}} \sum_{i=1}^n \min_{1 \leq k \leq K} \|X_i - C_k\|^2 = \sum_{k=1}^K \sum_{\ell \in C_k} \|X_\ell - C_k\|^2$$

where $C_k$ is the centroid tensor of cluster $C_k$.

After running HOSVD on tensor $X$, we obtain $U$ and $V$. We can get:

$$Y_\ell = U^T X_\ell V.$$

Using the distance relationship, the tensor clustering can be equivalently done in $\{Y_\ell\}$:

$$\min_{\{C_k\}} J_{K\text{means}}(Y) = \sum_{k=1}^K \sum_{\ell \in C_k} \|Y_\ell - C_k^Y\|^2$$

where $C_k^Y = U^T C_k V$.

#### 3. Discriminative High Order SVD

We propose a new tensor factorization method, Discriminative HOSVD, which adaptively selects the subspaces $U$ and $V$ in an unsupervised way so as to improve the clustering capability of $W$, and the clustering and classification capability of embedding space projection by $U, V$. In order to select more discriminative subspaces $U$ and $V$, we first develop the Adaptive 2DLDA (A2DLDA) method in §3.2., which select subspaces by explicitly combining 2DLDA and K-means clustering in a coherent way. After selecting discriminative subspaces $U$ and $V$, we can solve matrix $W$ for Discriminative HOSVD. Compared to the standard HOSVD, our Discriminative HOSVD can find better subspaces for further supervised and unsupervised learning tasks.

#### 3.1. Discriminative Subspace Selection

Suppose performing HOSVD on tensor $X$ and obtaining $U, V$. Our goal is to find $U$ and $V$ such that the clustering structure of data become more apparent in the subspace as (3). Thus, after the subspace projection, the distances of data points within the same class should be minimized and the distances of data points between classes should be maximized. The intuition here is to introduce LDA-like idea into
unsupervised subspace selection. We should notice the data points here have high dimensions and are not vectors used in LDA.

One way to achieve the goal is to maximize the between class scatter: \( \max_{U,V} \) \( \text{Tr} \ S^Y_b \), and minimize the within class scatter: \( \min_{U,V} \) \( \text{Tr} \ S^Y_w \), or another alternative way is to use objective function as:

\[
\max_{U,V} \frac{\text{Tr} \ S^Y_b}{\text{Tr} \ S^Y_w},
\]

where scatter matrices in projection subspaces of (3) as

\[
S^Y_b = \sum_{j=1}^n n_j \left( C^Y_j - C^Y \right) \left( C^Y_j - C^Y \right)^T, \tag{6}
\]

\[
S^Y_w = \sum_{j=1}^n \sum_{i \in \Pi_j} \left( Y_i - C^Y_j \right) \left( Y_i - C^Y_j \right)^T, \tag{7}
\]

where \( Y_i \) is the subspace projection result of the data point \( X_i \) by \( U \) and \( V \) as (3), \( C^Y_j \) is the centroid of cluster \( j \), \( n_j \) is the total number of data points in cluster \( j \), and \( C^Y \) is the centroid of the whole dataset.

We use (5) as the objective function to select the discriminative subspaces \( U \) and \( V \). The advantage of using (5) is to be revealed by Lemma 1.

**Lemma 1** Tensor clustering in (4) is equivalent to minimize the within-class scatter matrix \( S^Y_w \) or maximize the between-class scatter matrix \( S^Y_b \) in projection subspaces.

**Proof:**

\[
\text{Tr} \ S^Y_w = \text{Tr} \left( \sum_{j=1}^n \sum_{i \in \Pi_j} U^T (X_i - C_j) V V^T (X_i - C_j)^T U \right)
\]

\[
= \sum_{j=1}^n \sum_{i \in \Pi_j} \text{Tr} \left( U^T (X_i - C_j) V V^T (X_i - C_j)^T U \right)
\]

\[
= \sum_{j=1}^n \sum_{i \in \Pi_j} ||U^T X_i V - U^T C_j V||^2
\]

\[
= \sum_{j=1}^n \sum_{i \in \Pi_j} ||Y_i - C^Y_j||^2
\]

\[
= J_{\text{Kmeans}}(Y)
\]

Therefore, K-means (tensor) clustering minimizes the within-class scatter matrix \( S^Y_w \) in projection subspaces, or maximizes the between-class scatter matrix \( S^Y_b \) since the total scatter matrix is a constant. \( \square \)

LDA is widely used to select the subspace which has the maximal discriminant power. However, LDA is a supervised learning method which requires we know the class label for each data point as prior knowledge.

Since 2DLDA and K-means (tensor) clustering both minimize \( S^Y_w \) and maximize \( S^Y_b \), we propose the following main algorithmic approach to optimize the objective function (5) and select discriminative subspaces \( U \) and \( V \) in an unsupervised way:

(A1) Tensor clustering in (4) provides class labels for data points;

(A2) With these labels, (5) can be solved as 2DLDA.

Iteratively repeating these two steps is the main approach of our new Adaptive 2DLDA framework as following.

### 3.2. Adaptive 2DLDA

The principle of Adaptive 2DLDA is to do tensor clustering and 2DLDA iteratively. Once we have the K-means clustering results, we obtain the clustering labels for each data point. Thus, we can perform 2DLDA on the original data space to derive more discriminative subspaces according to the clustering labels. With the discriminative subspace, we can find better clustering labels. Till converge, we will get discriminative subspaces \( U \) and \( V \).

We use 3D Tensor decomposition as example to show how our Adaptive 2DLDA framework works. They can be easily generalized to higher dimension tensors. The goal of our Adaptive 2DLDA method is to optimize (5) in an unsupervised way. Because

\[
\text{Tr} \ S^Y_b = \text{Tr} \ U^T S^Y_b U = \text{Tr} \ V^T S^Y_b V
\]

and

\[
\text{Tr} \ S^Y_w = \text{Tr} \ U^T S^Y_w U = \text{Tr} \ V^T S^Y_w V,
\]

(5) can be written as:

\[
\max_{U,V} \frac{\text{Tr} \ U^T S^Y_b U}{\text{Tr} \ U^T S^Y_w U} \quad \text{or} \quad \max_{U,V} \frac{\text{Tr} \ V^T S^Y_w V}{\text{Tr} \ V^T S^Y_b V}, \tag{8}
\]

where

\[
S^Y_b = \sum_{j=1}^n n_j (C_j - C) V V^T (C_j - C)^T
\]

\[
S^Y_w = \sum_{j=1}^n \sum_{i \in \Pi_j} (X_i - C_j) V V^T (X_i - C_j)^T
\]

\[
S^U_b = \sum_{j=1}^n n_j (C_j - C)^T U U^T (C_j - C)
\]

\[
S^U_w = \sum_{j=1}^n \sum_{i \in \Pi_j} (X_i - C_j)^T U U^T (X_i - C_j).
\]

The labels for \( \{X_i\} \) can be found by running K-means clustering on \( \{Y_i\} \). Using these labels, similar to standard LDA
solutions, (8) can be solved as a generalized eigenvalue problem. $U$ can be calculated as the first $k_1$ eigenvectors of $(S_w^V)^{-1} S_w Y$, and $V$ can be calculated as the first $k_2$ eigenvectors of $(S_w^U)^{-1} S_v$.

We summarize the Adaptive 2DLDA algorithm in Alg. 1.

**Algorithm 1 Adaptive 2DLDA Algorithm**

**Input**

$X = \{X_1, \cdots, X_{n_3}\}$, parameters $k_1, k_2$, iteration number $t$

**Initialize:**

Compute $U$ and $V$ as eigenvectors of $F = \sum_{i=1}^N X_i X_i^T$ and $G = \sum_{i=1}^N X_i^T X_i$.

$Y_i = U^T X_i V, i = 1, 2, \ldots, n_3$.

**Do**

\[ a) \text{Run K-means on } \{Y_i\} \text{ to obtain class label for each } X_i \]

\[ b) \text{Seek discriminative subspaces } U \text{ and } V \text{ using class labels of } \{X_i\} \text{ and repeating iterations:} \]

\[ i) \text{Calculate } U \text{ as first } k_1 \text{ eigenvectors of } (S_w^V)^{-1} S_w Y \]

\[ ii) \text{Calculate } V \text{ as first } k_2 \text{ eigenvectors of } (S_w^U)^{-1} S_v \]

**Until convergence**

**Output** Cluster indicator $Q$ as class labels, and discriminative subspaces $U, V$.

### 3.3. Discriminative HOSVD Algorithm

After getting $U$ and $V$ from §3.1, we can obtain $W$ [11] by

$$\max_W \text{Tr}(W^T H W) \quad \text{s.t.} \quad W^T W = I,$$

where

$$H_{kk'} = \sum_{ij} X_{ijk} X_{i'j'k'} (UU^T)_{ii'} (VV^T)_{jj'}.$$ (10)

The solution $W$ of (9) including eigenvectors of matrix $H$. These $U$, $V$, and $W$ are the final results of Discriminative HOSVD.

Our Discriminative HOSVD algorithm is described in Alg. 2. The subspaces $U$ and $V$ are adaptively selected. With the discriminative $U$ and $V$, the clustering capability of matrix $W$ is also improved accordingly. We theoretically prove that in Theorem 1 in next section.

### 3.4. Demonstration of Discriminative HOSVD

Fig. 1 illustrates the other example to demonstrate the discriminative power of DHOSVD by clustering application. The real world Yale image dataset is used (detailed experimental settings can be found in §4). We randomly select 4 subjects with 40 images totally. ($U, V$) are obtained from HOSVD and DHOSVD. In each figure, the data points are plotted using two components (there is no difference for two or three components and we randomly select this number): $u_1^T X_{i1}$ and $u_2^T X_{i2}$, where $u_1, u_2$ are the first two columns of $U$ and $v_1, v_2$ are first two columns of $V$.

![Figure 1. Comparison of Discriminative HOSVD and HOSVD.](image)

Both methods are performed on 40 randomly selected images for four people from Yale dataset. The two coordinates are $(u_1^T X_{i1}, u_2^T X_{i2})$, where $u_1, u_2$ are the first two columns of $U$ and $v_1, v_2$ are first two columns of $V$. Images from the same subject use the same color.

All four subjects in HOSVD result mix to each other and the within-class distances are large. The clustering result is poor. But DHOSVD can unsupervisedly select good discriminative subspaces for these subjects to improve the clustering results. After two iterations, the within-class distances are already small. After five iterations, the between-class distances are increased substantially. Therefore, results shown in Fig. 1 demonstrate the discriminative power of DHOSVD subspaces.

### 3.5. Analysis of Discriminative HOSVD

In Discriminative HOSVD, $U$ and $V$ are calculated by Adaptive 2DLDA algorithm and improve the clustering capability of data points after subspace projection. In following Theorem 1, we will prove Discriminative HOSVD is equivalent to tensor clustering with subspaces $U$ and $V$, and the clustering structure exists in matrix $W$. Because $U$ and $V$ improve the tensor clustering, the clustering capability of matrix $W$ is also improved.

![Algorithm 2 Discriminative HOSVD Algorithm](image)

**Input**

$X = \{X_1, \cdots, X_{n_3}\}$, Parameters $k_1, k_2, k_3$

**Do**

1. Set $(U, V)$ as the results of Adaptive 2DLDA (see §3.1).
2. Solve $W$ using (10) where $H$ is calculated using $U$ and $V$ from the results of Adaptive 2DLDA.
3. $S = U^T \otimes_1 V^T \otimes_2 W^T \otimes_3 X$.

**Output** $U, V, W, S$. 

(a) 40 images are projected onto subspaces

(b) Images used in (a)
Theorem 1 Discriminative HOSVD is equivalent to tensor clustering of (5) with $U$ and $V$ selected by Adaptive 2DLDA and the clustering structure exists in matrix $W$.

**Proof:** In Lemma 1, we have proved

$$
\text{Tr}(S_{ii'}^Y) = \sum_{i=1}^{n} \sum_{i' \in C_j} ||Y_i - C_j^Y||^2 \\
= \sum_{i=1}^{n} \sum_{i' \in C_j} ||Y_i||^2 - \sum_{j=1}^{n} \frac{1}{n_k} \sum_{i', i \in C_j} \text{Tr}(Y_i^T Y_{i'})
$$

where $C_j^Y$ is centroid the $j$th cluster in the $U, V$ subspace. We introduce the clustering indicator matrix

$$Q = (q_k, \cdots, q_K), \quad q_i^T q_j = \delta_{ij},
$$

where

$$q_j = (0, 0, 1, \cdots, 1, 0, \cdots, 0)^T/\sqrt{n_j} (12)$$

It is obvious that $Q^T Q = I$. Since $\sum_{i=1}^{n} ||Y_i||^2$ is constant, minimizing $\text{Tr}(S_{ii'}^Y)$ is equivalent to solve

$$\max_{Q} \sum_{j=1}^{n} \frac{1}{n_j} \sum_{i', i \in C_j^Y} \text{Tr}(Y_i^T Y_{i'}) = \text{Tr}(Q^T X Q) (13)$$

where

$$X_{ii'} = \text{Tr}(Y_i^T Y_{i'}) = \text{Tr}(V^T X_i^T U U^T X_{i'} V),
$$

and

$$(QQ^T)_{ii'} = \begin{cases} 0 & \text{if } Y_i \text{ or } Y_{i'} \notin C_j^Y \\ 1/n_j & \text{if } Y_i \text{ and } Y_{i'} \in C_j^Y \end{cases} (15)$$

Obviously, (14) is the same as (10). As a result, (13) s.t. $Q^T Q = I$ is equivalent to the objective function to solve matrix $W$ in HOSVD decomposition, also in our Discriminative HOSVD:

$$\max_{W} \text{Tr}(W^T H W) \text{ s.t. } W^T W = I. (16)$$

Thus, solution $W$ in (16) is exact the same as the solution $Q$ in (13). $\square$

### 3.6. DHOSVD vs Adaptive 2DLDA

Although in [9], the authors proved the clustering capability of HOSVD. In fact, they only focus on matrix $W'$ at that time. Here we clarify that $U$ and $V$ can be also used for clustering via embedding space projection as (3). Especially after selecting discriminative subspaces $U$ and $V$, the clustering results from $W$ in DHOSVD is better than standard one in HOSVD. We have proved that in Theorem 1 and will also demonstrate that by empirical experiments in §4.

In §4, the tensor clustering results based on $U$ and $V$ will also be further discussed. If we use $U$ and $V$ in clustering, DHOSVD has the same results as Adaptive 2DLDA. When we use DHOSVD, we usually emphasize $W$; when we use Adaptive 2DLDA, we emphasize $U$ and $V$.

### 4. Experimental Results

In order to evaluate the discriminative power of our method, we perform clustering and classification experiments on four commonly used datasets: AT&T [13], Binary Alphadigitis (BinAlpha) [1], UMIST [8], and YaleB [7]. We also validate the quality of the subspaces using a content-based image retrieval problem on the People Playing Musical Instrument (PPMI) data set [22].

In the benchmark AT&T face database, images from 40 distinct persons were taken at different times, different facial expression, and facial details. We use the whole AT&T dataset and resize all images (400 images from 40 persons, each of which has 10 images in similar illumination conditions) from $112 \times 92$ to $28 \times 23$. The second dataset is Binary Alphadigitis corpus. This is a small vocabulary task which is fairly challenging. The corpus has a vocabulary of 36 characters: 26 letters (from ‘A’ to ‘Z’) and 10 digits (from ‘0’ to ‘9’). For each word, we have 39 images (20 $\times$ 16). In our experiments, we use the whole dataset and the original size. UMIST dataset includes multi-view images of human faces and consists of 564 images of 20 individuals (mixed race, gender, and appearance). Each individual is shown in a range of poses from profile to frontal views. The image resolution is $220 \times 220$ and we resize them to $28 \times 23$. YaleB face images database used in our experiment is the combination of extended and original Yale database B [7]. We resize the images from $192 \times 168$ to $24 \times 21$. Because there is a set of images which are corrupted during the image acquisition [7], we have 31 subjects in total without any corrupted images. We randomly select ten illumination conditions for all 31 subjects to create the experimental data set with 310 images. The PPMI dataset contains images of humans interacting with twelve different musical instruments, bassoon, cello, clarinet, erhu, flute, French horn, guitar, harp, recorder, saxophone, trumpet, and violin.

In order to compare the interpretability and discriminability of DHOSVD and HOSVD, we first use all four data sets to evaluate the clustering structures existing in matrix $W$ of DHOSVD and HOSVD.

After that, we use two data sets to show the performance of discriminative subspaces $U$ and $V$ selected by A2DLDA. Both clustering and classification applications are used. As we discussed in §3.6, DHOSVD is the same as A2DLDA if
we only use $U$ and $V$ to do projection and clustering.

### 4.1. Clustering Capability Enhancement in Matrix $W$ of DHOSVD

Our goal of proposing Discriminative HOSVD is to improve the clustering capability of HOSVD. Thus, we compare the clustering performance (clustering accuracy and normalized mutual information) between our DHOSVD and standard HOSVD. Four standard image data sets are employed. Both HOSVD and DHOSVD are performed on data set to get matrix $W$ and K-means clustering method is utilized to find the clustering structure in $W$.

We use two standard metrics to measure the quality of clustering results: clustering accuracy and normalized mutual information. The clustering accuracy is defined by the ratio of the number of images that are clustered into the default subject clusters and the number of total images. Because clustering is an unsupervised method, the subject label is not identical to the cluster label. For each cluster in clustering results, we find the largest group of images from the same subject and use their label for this cluster. If two clusters are marked with the same label, for the cluster whose largest group of images has a smaller number, we change its label to the second largest group of images. We iteratively label the clusters, till convergence.

The Normalized Mutual Information (NMI) is calculated by:

$$ NMI = \frac{MI(C, C')}{\max(H(C), H(C'))}, $$

where $C$ is a set of clusters obtained from the true labels and $C'$ is a set of clusters obtained from the clustering algorithm. $MI(C, C')$ is the mutual information metric, and $H(C)$ and $H(C')$ are the entropies of $C$ and $C'$, respectively. NMI is between 0 and 1. A larger NMI value indicates a better performance.

The clustering results using DHOSVD and HOSVD are compared in Fig. 2. Our new DHOSVD provides better clustering capability than standard HOSVD in all four data sets. Thus, DHOSVD decomposition results are more discriminative.

### 4.2. Clustering Performance Evaluation of Adaptive 2DLDA

In § 3.6, we discussed there are two ways to compare the clustering results of DHOSVD and HOSVD. In Fig. 2, the clustering results using $W$ of both DHOSVD and HOSVD have been shown. Now we emphasize $U$ and $V$ to compare the clustering results. In this way, DHOSVD has the same results as Adaptive 2DLDA.

In Fig. 3, we compare Adaptive 2DLDA to four other unsupervised methods: HOSVD + K-means, GLRAM [23] + K-means, PCA + K-means, and K-means. Our method has better clustering results than others in terms of both clustering accuracy and normalized mutual information.

For Adaptive 2DLDA, HOSVD, and GLRAM, we first project image $X_i$ into their $U$ and $V$ subspaces as (3). We then run K-means on $\{Y_i\}$.

For PCA+K-means, we always set the reduced dimension $K_{pca} = k_1k_2$ and perform K-means on the $K_{pca}$ principle components subspace. For K-means, we do K-means on the original space treating an image as a one line vector. The clustering accuracy is calculated by the way described in §4.1.

Since K-means clustering is used in all 5 methods with different subspaces, we are unable to start K-means clustering in all methods with the same centroid. In order to generate more reliable comparison results, we start K-means using the same random partition of data points for all methods in each round.

For each method, we run 30 rounds. In each round, 30 times K-means are performed and the result with minimum K-means clustering objective function error is selected (totally 900 K-means are computed for each method). We use the average results of all 30 rounds as the final results which are shown in Fig. 3.

### 4.3. Classification Performance Evaluation of A2DLDA

Since Adaptive 2DLDA selects more discriminative subspaces, we are interested whether it will help classification. We compare subspaces selected by Adaptive 2DLDA to subspaces selected by PCA, GLRAM, and HOSVD, and the original space. In all cases, we use nearest neighbor classifier (1-NN) and 5-fold cross validation scheme. For A2DLDA, GLRAM, HOSVD, the distance between 2-D objects can be interpreted as a bilinear metric:

$$ d(X_i, X_j) = \|X_i - X_j\|_{U,V} = \|U^T X_i V - U^T X_j V\| $$

(17)

We can use this metric to perform 1-NN classification. For PCA, the distance is the Euclidean distance of the principal component subspace. For the original space, $d(i, j) = \|X_i - X_j\|$.
The query image is a woman playing violin. One can visually conclude that our method retrieval much more reasonable images.

5. Conclusions

In this paper, we present a new tensor approach using adaptive subspace learning, in which high order objects are discriminatively separatively. We first proved K-means tensor clustering and 2DLDA use the same objective functions to minimize the within-class scatter and maximize the between-class scatter. Based on our theoretical analysis, we proposed an adaptive 2DLDA framework to iteratively do K-means clustering and select discriminative subspaces. Combining the strength of K-means clustering and 2DLDA, we unsupervisedly selected more discriminative subspaces $U$ and $V$. Moreover, we proved Discriminative HOSVD is equivalent to tensor clustering with subspaces $U$ and $V$ selected by Adaptive 2DLDA. Because $U$ and $V$ are more discriminative and improve the tensor clustering, the clustering capability of matrix $W$ in DHOSVD is enhanced. Experimental results on clustering, classification, and image retrieval indicate that the quality of the subspaces selected by our methods is much higher than all the other methods.

Acknowledgment This research is partially supported by NSF-CCF-0830780, NSF-DMS-0915228, NSF-CCF-0917274.

References


Figure 5. One query example for the four distance matrix measurements: Euclidean, Mahalanobis, PCA, and our method. For each method, the first image is the query image and the following 10 images are the top 10 retrieved images using the corresponding method. Our method makes two errors in the top 10 retrieval images: the 4-th and 10-th in which the person is playing guitar. This error is reasonable since the shapes of violin and guitar are similar.


