NEIGHBORHOOD PRESERVING NON-NEGATIVE TENSOR FACTORIZATION FOR IMAGE REPRESENTATION

Yu-Xiong Wang, Liang-Yan Gui, Yu-Jin Zhang

Department of Electronic Engineering, Tsinghua University
Tsinghua National Laboratory for Information Science and Technology, Tsinghua University
Beijing 100084, China
albertwyx@gmail.com, guily05@gmail.com, zhang-yj@tsinghua.edu.cn

ABSTRACT

Non-negative Matrix Factorization (NMF) has become a powerful tool for image representation due to its enhanced semantic interpretability under non-negativity. Unfortunately, two types of neighborhood information essential to representation are lost in NMF. For individual image, the local structure information is missing in the vectorization, which can then be avoided by Non-negative Tensor Factorization (NTF). For image data points, they often reside on a low dimensional submanifold embedded in a high dimensional ambient space. NMF and NTF are incapable of encoding the local geometrical information, which can nevertheless be resuscitated by manifold learning. To simultaneously model both of the neighborhood relationship within and among image data, this paper proposes a novel algorithm called Neighborhood Preserving Non-negative Tensor Factorization (NPNTF) by incorporating locally linear embedding regularization into tensor factorization. Experimental results on image clustering show the superior performance of NPNTF with more natural and discriminating representation ability.

***Index Terms*—** Non-negative matrix and tensor factorization, locally linear embedding, manifold regularization, image representation, image clustering

1. INTRODUCTION

Low-rank approximation (LRA), which tries to find a parsimonious representation, has gained substantial attention in image processing and computer vision areas recent years. Many tasks, such as clustering and classification, can be effectively tackled in the reduced low dimensional subspace. The formulation of LRA can be regarded as decomposing the original data matrix into two or three low-rank factor matrices. By imposing the non-negativity constraint, a new LRA paradigm called Non-negative Matrix Factorization (NMF) is initiated with physiological and psychological evidence [1]. Due to the purely additive combination, NMF obtains the parts-based representation and thus enhances the interpretability of the issue.

However, two types of neighborhood information essential to image representation are lost in the basic NMF model. The data to be processed in NMF are treated as vectors in essence, whereas an image is intrinsically a 2-D matrix. Thus the vectorization of image data will unavoidably lose the spatial coherency and local structure information which might be crucial to following processing. The neighborhood information within individual image can be preserved naturally by regarding the image data set as a 3-D cube and extending matrix into tensor factorization, i.e. Non-negative Tensor Factorization (NTF). NTF is also capable of learning parts-based representation with superior decomposition performance to what NMF can provide with respect to degree of sparsity and lack of ghost residue [2, 3]. Besides, some additional constraints can be incorporated into NTF similarly, such as sparse NTF [4] and discriminant NTF [5].

On the other hand, some recent research work suggests that image space is actually a smooth low dimensional submanifold embedded in a high dimensional ambient space [6]. Several manifold learning methods, such as ISOMAP [7], LLE [8], and Laplacian Eigenmaps [9], have been proposed to explicitly explore the intrinsic topological structure. Unfortunately, NMF and NTF still fit the data in Euclidean space. By combining NMF with appropriate manifold learning approaches, the neighborhood relationship among image data can be unearthed. This is the major motivation in Graph Regularized NMF [10], Neighborhood Preserving NMF [11], and NMF on Multiple Manifolds [12].

In order to simultaneously address the above two neighborhood information loss: neighborhood structure within individual image and neighborhood geometry among image data, we propose a novel algorithm called Neighborhood Preserving Non-negative Tensor Factorization (NPNTF) in this paper by incorporating tensor factorization and manifold learning. Laplacian Regularized Non-negative Tensor Factorization (LRNTF) [13] encodes the latter neighborhood information pairwise, i.e. the two points in the mapped low dimensional space should be close enough to each other if they are neighbors in the original image space. Here, we turn to model the underlying geometrical structure patchwise and adopt the locally linear embedding assumption [8]. Since each data point and its neighboring points lie on or close to a locally flat patch of the manifold, the local geometry of these patches are characterized by linear coefficients that reconstruct each data point from its neighbors. Hence the coefficients are the local invariants in the factorization mapping. This is achieved by adding corresponding manifold regularization term in the NTF objective function, upon which an iterative multiplicative updating algorithm with guaranteed convergence is developed to solve the NPNTF optimization problem. A new parts-based image representation is thus learned which identifies the neighborhood information and leads to a more semantic feature space.

The rest of this paper is organized as follows. In Section 2 NMF and NTF are reviewed briefly. The proposed NPNTF algorithm is elaborated in Section 3. Section 4 presents the experimental results on image clustering. Discussions and conclusions are drawn in Section 5.

This work was partially supported by National Nature Science Foundation (NNSF: 61171118).

2. BRIEF REVIEW OF NON-NEGATIVE MATRIX AND TENSOR FACTORIZAITON

Given an M-D random vector $\boldsymbol{x}$ with non-negative elements, whose N observations are denoted as $\boldsymbol{x}_{i,i=1,2,\ldots,N}$, let data matrix be $\boldsymbol{X}= [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N] \in \mathbb{R}_{\geq 0}^{M\times N}$, NMF seeks non-negative basis matrix $\boldsymbol{U}\in \mathbb{R}_{\geq 0}^{M\times L}$ and coefficient matrix $\boldsymbol{V}\in \mathbb{R}_{\geq 0}^{L\times N}$, such that $\boldsymbol{X} \approx \boldsymbol{UV}$. Using the bilinear model, NMF can be rewritten as linear combination of rank-one non-negative matrices expressed by

$$\boldsymbol{X} \approx \sum_{m=1}^{L} \boldsymbol{U}_{\bullet m}\boldsymbol{V}_{m\bullet} = \sum_{m=1}^{L} \boldsymbol{U}_{\bullet m} \otimes (\boldsymbol{V}_{m\bullet})^T \tag{1}$$

where $\boldsymbol{U}_{\bullet m}$ is the m-th column vector of $\boldsymbol{U}$ while $\boldsymbol{V}_{m\bullet}$ is the m-th row vector of $\boldsymbol{V}$, and $\otimes$ denotes the outer product of two vectors. Regarding the general initial condition $L << \min(M, N)$, NMF tries to represent the high dimensional stochastic pattern compactly.

When it comes to multiway data, a natural extension is tensor factorization. There are generally two types of NTF model—Non-negative Tucker Decomposition (NTD) [14] and more restricted Non-negative Parallel Factor Analysis (NPARAFAC) [2], whose main difference lies in the decomposed core factor tensor. Here we mainly focus on NPARAFAC model, which decomposes an N-order tensor in a sum of rank-one non-negative tensors.

For image space, let 3-order tensor $\boldsymbol{X} \in \mathbb{R}_{\geq 0}^{R\times S\times N}$ with entries $\boldsymbol{X}_{r,s,n,r=1,\ldots,R,s=1,\ldots,S,n=1,\ldots,N}$ denote a set of N images M_i of dimension $R \times S$, NTF seeks three factor matrices $\boldsymbol{U} = [\boldsymbol{u}^1, \boldsymbol{u}^2, \ldots, \boldsymbol{u}^L] \in \mathbb{R}_{\geq 0}^{R\times L}$, $\boldsymbol{V} = [\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^L] \in \mathbb{R}_{\geq 0}^{S\times L}$, and $\boldsymbol{H} = [\boldsymbol{h}^1, \boldsymbol{h}^2, \ldots, \boldsymbol{h}^L] \in \mathbb{R}_{\geq 0}^{N\times L}$, such that $\boldsymbol{X}$ can be approximated as a sum of L rank-one non-negative tensors $\boldsymbol{X} \approx \sum_{m=1}^{L} \boldsymbol{u}^m \otimes \boldsymbol{v}^m \otimes \boldsymbol{h}^m$, i.e. $\boldsymbol{X}_{r,s,n} \approx \sum_{m=1}^{L} \boldsymbol{u}_r^m \boldsymbol{v}_s^m \boldsymbol{h}_n^m$. Here, the rank-one matrix $\boldsymbol{u}^m \otimes \boldsymbol{v}^m$, $m = 1, ..., L$ spans the basis image space, and each image M_i, i.e. the frontal slice $\boldsymbol{X}_i \triangleq \boldsymbol{X}_{\bullet\bullet i}$, is represented as a superposition of these basis images with weight coefficient $\boldsymbol{H}_{i\bullet} = [\boldsymbol{h}_i^1, \ldots, \boldsymbol{h}_i^L]$. Thus the original 2-D image matrices are mapped into new feature space as 1-D coefficient vectors.

In order to quantify the difference between the original data and the approximation, the square Frobenius norm (i.e. Square of Euclidian Distance)$\|\bullet\|_F^2$ is employed as the objective function:

$$O_{NTF} = \frac{1}{2}\left\|\boldsymbol{X} - \sum_{m=1}^{L} \boldsymbol{u}^m \otimes \boldsymbol{v}^m \otimes \boldsymbol{h}^m\right\|_F^2 \tag{2}$$

3. NEIGHBORHOOD PRESERVING NON-NEGATIVE TENSOR FACTORIZATION

By using tensor factorization, the neighborhood information within individual image is preserved, and the parts-based representation is learned by NTF in Euclidean space. However, the image data are often sampled from a nonlinear low dimensional submanifold embedded in a high dimensional ambient space. Thus Neighborhood Preserving Non-negative Tensor Factorization (NPNTF) is introduced to identify and maintain the intrinsic geometrical structure or neighborhood relationship among image data by explicitly incorporating additional manifold regularization term.

3.1. Locally linear embedding assumption

A natural and intuitive local topology in modeling the manifold structure adopted by locally linear embedding (LLE) assumes that a data point generated as a linear combination of several nearby points on a specific manifold in the original space should be reconstructed from its neighbors in a similar way or by the same combination coefficients in the reduced low dimensional subspace. LLE is capable of recovering global nonlinear structure from locally linear fits [8].

For each image data point $\boldsymbol{X}_i$, let $\mathcal{N}_k(\boldsymbol{X}_i)$ be its k-nearest neighborhood. The local topology around $\boldsymbol{X}_i$ is characterized by linear coefficients $\boldsymbol{S}_{i\bullet}$ that best reconstruct $\boldsymbol{X}_i$ from its neighbors. The reconstruction coefficients can be obtained by solving the following constrained least-squares problem:

$$\begin{aligned} &\min \quad \left\|\boldsymbol{X}_i - \sum\nolimits_j \boldsymbol{S}_{ij}\boldsymbol{X}_j\right\|_F^2, \quad i = 1, ..., N \\ &s.t. \ \sum\nolimits_j \boldsymbol{S}_{ij} = 1 \ \& \ \boldsymbol{S}_{ij} = 0, \ for \ \boldsymbol{X}_j \notin \mathcal{N}_k(\boldsymbol{X}_i) \end{aligned} \tag{3}$$

where $\boldsymbol{S} \in \mathbb{R}^{N\times N}$ is the weight coefficient matrix.

In the mapped low dimensional subspace, the new feature vector $\boldsymbol{H}_{i\bullet}$ is supposed to satisfy the similar reconstruction mode with fixed $\boldsymbol{S}$, which can be achieved by minimizing

$$\begin{aligned} R \ &= \ \sum\nolimits_i \|\boldsymbol{H}_{i\bullet} - \boldsymbol{S}_{i\bullet}\boldsymbol{H}\|^2 = \|\boldsymbol{H} - \boldsymbol{SH}\|_F^2 = \|(\boldsymbol{I} - \boldsymbol{S})\,\boldsymbol{H}\|_F^2 \\ &= \ Tr\left\{\boldsymbol{H}^T(\boldsymbol{I} - \boldsymbol{S})^T(\boldsymbol{I} - \boldsymbol{S})\,\boldsymbol{H}\right\} = Tr\left(\boldsymbol{H}^T\boldsymbol{GH}\right) \end{aligned} \tag{4}$$

where $Tr(\bullet)$ denotes the trace of a matrix, $\boldsymbol{I} \in \mathbb{R}^{N\times N}$ is the identity matrix, and $\boldsymbol{G} = (\boldsymbol{I} - \boldsymbol{S})^T(\boldsymbol{I} - \boldsymbol{S})$ is symmetric. Hence the neighborhood relationship among data points is encoded in R.

3.2. NTF with manifold regularization

When incorporating R as the manifold regularization term into the original NTF objective function in (2), Neighborhood Preserving Non-negative Tensor Factorization (NPNTF) is obtained which minimizes the following objective function

$$O_{NPNTF} = \frac{1}{2}\left\|\boldsymbol{X} - \sum_{m=1}^{L} \boldsymbol{u}^m \otimes \boldsymbol{v}^m \otimes \boldsymbol{h}^m\right\|_F^2 + \frac{1}{2}\eta Tr\left(\boldsymbol{H}^T\boldsymbol{GH}\right) \tag{5}$$

where η is the regularization parameter balancing the trade-off between the fitting goodness and the manifold constraint.

Both of the aforementioned neighborhood information is thus involved in (5). To avoid the scaling ambiguity, additional L_2 normalization on columns of $\boldsymbol{U}$ and $\boldsymbol{V}$ (i.e. $\|\boldsymbol{u}^m\|_2 = \|\boldsymbol{v}^m\|_2 = 1, \ m = 1, \ldots, L$) is introduced apart from the non-negativity constraints.

3.3. Update rules

While O_{NPNTF} is not jointly convex in $\boldsymbol{U}$, $\boldsymbol{V}$, and $\boldsymbol{H}$, it is separately convex in $\boldsymbol{U}$, $\boldsymbol{V}$, or $\boldsymbol{H}$. Alternating multiplicative update rules similar to NMF and NTF are derived here for NPNTF, which can be viewed as adaptive rescaled gradient descent approach with non-negativity preservation. Here, we mainly focus on the update rules for the entries in $\boldsymbol{H}$.

Let $\langle \boldsymbol{A}, \boldsymbol{B}\rangle$ denote the inner product of two tensors of the same order. Since the differential with respect to inner product satisfies $d\langle \boldsymbol{A}, \boldsymbol{A}\rangle = 2\langle \boldsymbol{A}, d\boldsymbol{A}\rangle$, it follows that

$$\begin{aligned} d(O_{NPNTF}) = \frac{1}{2}d\Bigg\langle &\boldsymbol{X} - \sum_{m=1}^{L} \boldsymbol{u}^m \otimes \boldsymbol{v}^m \otimes \boldsymbol{h}^m, \\ &\boldsymbol{X} - \sum_{m=1}^{L} \boldsymbol{u}^m \otimes \boldsymbol{v}^m \otimes \boldsymbol{h}^m\Bigg\rangle + \frac{1}{2}\eta d\left\{Tr\left(\boldsymbol{H}^T\boldsymbol{GH}\right)\right\} \end{aligned}$$

$$= \left\langle \boldsymbol{X} - \sum_{m=1}^{L} \boldsymbol{u}^m \otimes \boldsymbol{v}^m \otimes \boldsymbol{h}^m, d\left(\boldsymbol{X} - \sum_{m=1}^{L} \boldsymbol{u}^m \otimes \boldsymbol{v}^m \otimes \boldsymbol{h}^m\right)\right\rangle + \frac{1}{2}\eta d\left\{\sum_{m=1}^{L} (\boldsymbol{h}^m)^T \boldsymbol{G}\boldsymbol{h}^m\right\} \tag{6}$$

Noting that $\boldsymbol{G}$ is symmetric, the differential with respect to $\boldsymbol{h}^j$ is

$$d(O_{NPNTF}(\boldsymbol{h}^j)) = \left\langle \sum_{m=1}^{L} \boldsymbol{u}^m \otimes \boldsymbol{v}^m \otimes \boldsymbol{h}^m, \boldsymbol{u}^j \otimes \boldsymbol{v}^j \otimes d(\boldsymbol{h}^j)\right\rangle - \left\langle \boldsymbol{X}, \boldsymbol{u}^j \otimes \boldsymbol{v}^j \otimes d(\boldsymbol{h}^j)\right\rangle + \eta \boldsymbol{G}\boldsymbol{h}^j d(\boldsymbol{h}^j) \tag{7}$$

Taking the differential with respect to the i-th element of $\boldsymbol{h}^j$, it becomes

$$d(O_{NPNTF}(\boldsymbol{h}_i^j)) = \left\langle \sum_{m=1}^{L} \boldsymbol{u}^m \otimes \boldsymbol{v}^m \otimes \boldsymbol{h}^m, \boldsymbol{u}^j \otimes \boldsymbol{v}^j \otimes \boldsymbol{e}_i\right\rangle - \left\langle \boldsymbol{X}, \boldsymbol{u}^j \otimes \boldsymbol{v}^j \otimes \boldsymbol{e}_i\right\rangle + \eta \boldsymbol{G}\boldsymbol{h}^j \boldsymbol{e}_i \tag{8}$$

where $\boldsymbol{e}_i$ is the i-th column of the $N \times N$ identity matrix. Let $\boldsymbol{G} = \boldsymbol{G}^+ - \boldsymbol{G}^-$, $\boldsymbol{G}_{in}^+ = (|\boldsymbol{G}_{in}| + \boldsymbol{G}_{in})/2$, $\boldsymbol{G}_{in}^- = (|\boldsymbol{G}_{in}| - \boldsymbol{G}_{in})/2$. Using the identity $\langle \boldsymbol{x}_1 \otimes \boldsymbol{y}_1, \boldsymbol{x}_2 \otimes \boldsymbol{y}_2\rangle = \langle \boldsymbol{x}_1, \boldsymbol{x}_2\rangle \langle \boldsymbol{y}_1, \boldsymbol{y}_2\rangle$, the partial derivative with respect to $\boldsymbol{h}_i^j$ is as

$$\begin{aligned}\frac{\partial O_{NPNTF}}{\partial \boldsymbol{h}_i^j} &= \sum_{m=1}^{L} \boldsymbol{h}_i^m \left\langle \boldsymbol{u}^m, \boldsymbol{u}^j\right\rangle \left\langle \boldsymbol{v}^m, \boldsymbol{v}^j\right\rangle - \sum_{r,s} \boldsymbol{X}_{r,s,i} \boldsymbol{u}_r^j \boldsymbol{v}_s^j \\ &\quad + \eta \sum_{n=1}^{N} \boldsymbol{G}_{in}^+ \boldsymbol{h}_n^j - \eta \sum_{n=1}^{N} \boldsymbol{G}_{in}^- \boldsymbol{h}_n^j \\ &= \left\{\sum_{m=1}^{L} \boldsymbol{h}_i^m \left\langle \boldsymbol{u}^m, \boldsymbol{u}^j\right\rangle \left\langle \boldsymbol{v}^m, \boldsymbol{v}^j\right\rangle + \eta \sum_{n=1}^{N} \boldsymbol{G}_{in}^+ \boldsymbol{h}_n^j\right\} \\ &\quad - \left\{\sum_{r,s} \boldsymbol{X}_{r,s,i} \boldsymbol{u}_r^j \boldsymbol{v}_s^j + \eta \sum_{n=1}^{N} \boldsymbol{G}_{in}^- \boldsymbol{h}_n^j\right\} \\ &\triangleq \nabla_+ - \nabla_-\end{aligned} \tag{9}$$

where $\nabla_+ = \sum_{m=1}^{L} \boldsymbol{h}_i^m \langle \boldsymbol{u}^m, \boldsymbol{u}^j\rangle \langle \boldsymbol{v}^m, \boldsymbol{v}^j\rangle + \eta \sum_{n=1}^{N} \boldsymbol{G}_{in}^+ \boldsymbol{h}_n^j$, $\nabla_- = \sum_{r,s} \boldsymbol{X}_{r,s,i} \boldsymbol{u}_r^j \boldsymbol{v}_s^j + \eta \sum_{n=1}^{N} \boldsymbol{G}_{in}^- \boldsymbol{h}_n^j$.

Hence the multiplicative update rule for $\boldsymbol{h}_i^j$ with non-negativity preservation can be obtained as follows:

$$\boldsymbol{h}_i^j \leftarrow \boldsymbol{h}_i^j \frac{\nabla_-}{\nabla_+} = \frac{\boldsymbol{h}_i^j \left\{\sum_{r,s} \boldsymbol{X}_{r,s,i} \boldsymbol{u}_r^j \boldsymbol{v}_s^j + \eta \sum_{n=1}^{N} \boldsymbol{G}_{in}^- \boldsymbol{h}_n^j\right\}}{\sum_{m=1}^{L} \boldsymbol{h}_i^m \langle \boldsymbol{u}^m, \boldsymbol{u}^j\rangle \langle \boldsymbol{v}^m, \boldsymbol{v}^j\rangle + \eta \sum_{n=1}^{N} \boldsymbol{G}_{in}^+ \boldsymbol{h}_n^j} \tag{10}$$

This is equivalent to setting the learning rate $\mu(\boldsymbol{h}_i^j)$ in the gradient descent formula $\boldsymbol{h}_i^j \leftarrow \boldsymbol{h}_i^j - \mu(\boldsymbol{h}_i^j) \frac{\partial O_{NPNTF}}{\partial \boldsymbol{h}_i^j}$ as

$$\mu(\boldsymbol{h}_i^j) = \frac{\boldsymbol{h}_i^j}{\nabla_+} = \frac{\boldsymbol{h}_i^j}{\sum_{m=1}^{L} \boldsymbol{h}_i^m \langle \boldsymbol{u}^m, \boldsymbol{u}^j\rangle \langle \boldsymbol{v}^m, \boldsymbol{v}^j\rangle + \eta \sum_{n=1}^{N} \boldsymbol{G}_{in}^+ \boldsymbol{h}_n^j} \tag{11}$$

Similarly, the update rules for $\boldsymbol{u}_i^j$ and $\boldsymbol{v}_i^j$ are as follows:

$$\boldsymbol{u}_i^j \leftarrow \frac{\boldsymbol{u}_i^j \sum_{s,n} \boldsymbol{X}_{i,s,n} \boldsymbol{v}_s^j \boldsymbol{h}_n^j}{\sum_{m=1}^{L} \boldsymbol{u}_i^m \langle \boldsymbol{v}^m, \boldsymbol{v}^j\rangle \langle \boldsymbol{h}^m, \boldsymbol{h}^j\rangle} \tag{12}$$

$$\boldsymbol{v}_i^j \leftarrow \frac{\boldsymbol{v}_i^j \sum_{r,n} \boldsymbol{X}_{r,i,n} \boldsymbol{u}_r^j \boldsymbol{h}_n^j}{\sum_{m=1}^{L} \boldsymbol{v}_i^m \langle \boldsymbol{u}^m, \boldsymbol{u}^j\rangle \langle \boldsymbol{h}^m, \boldsymbol{h}^j\rangle} \tag{13}$$

which are easier cases without manifold regularizer. Besides, the convergence of the proposed algorithm can be proved by using the Gauss-Seidel fashion similar to [2, 13].

4. EXPERIMENTAL RESULTS

The parts-based representation along with the identified neighborhood information learned by NPNTF leads to a more semantic feature space. Image clustering can then be performed in this subspace.

4.1. Data sets and evaluation metrics

Two image data sets are used to evaluate the clustering performance of the proposed NPNTF algorithm: the COIL20 image library and the CMU PIE face database.

The COIL20 image library contains 72 gray scale images of size 32×32 for each of 20 objects taken from varying angles. The CMU PIE face database contains 41,368 facial images of 68 persons under different poses, illumination conditions, and expressions. Here we select one near frontal pose ($C27$) subset, which contains 42 gray scale images resized to 32×32 for each person under different light and illumination conditions.

To demonstrate the effectiveness of the proposed NPNTF algorithm, we also make comparisons among the following popular clustering algorithms:

(1) Canonical K-means clustering (K-means in short).

(2) Non-negative Matrix Factorization based clustering (NMF in short). K-means clustering is implemented after NMF.

(3) Non-negative Tensor Factorization based clustering (NTF in short). K-means clustering is implemented after NTF.

(4) Neighborhood Preserving Non-negative Tensor Factorization based clustering (NPNTF in short). K-means clustering is implemented after NPNTF.

Two metrics are used to measure the clustering performance: accuracy and normalized mutual information (NMI), which are based on comparing the obtained cluster label and the ground truth. For the definition of these two metrics one may refer to [12].

4.2. Clustering results

For each data set, the evaluations are conducted with different cluster numbers. For each given cluster number L, 10 test runs are conducted on different randomly selected clusters. The regularization parameter η is set by searching the grid {0.01, 0.1, 1, 10, 100, 1000}. The neighborhood size k is set to 5. Table 1 and 2 show the clustering results in term of accuracy and NMI on the COIL20 and PIE data sets, respectively. The mean of the performance are reported in the tables.

Table 1. Clustering Accuracy (%) and NMI(%) on COIL20

L	Accuracy(%)			
	K-means	NMF	NTF	NPNTF
2	90.6	91.2	87.1	**91.4**
3	85.5	83.5	86.8	**94.5**
4	73.6	74.4	78.4	**90.7**
5	66.9	68.8	69.5	**87.3**
6	69.3	69.8	71.1	**87.5**
7	63.2	65.7	65.4	**83.5**
8	59.9	63.3	63.2	**80.6**
9	62.6	64.1	64.5	**77.6**
10	60.4	63	65.5	**78.1**
Avg.	70.2	71.5	72.4	**85.7**
L	NMI(%)			
	K-means	NMF	NTF	NPNTF
2	**70.3**	**70.3**	59.5	68.7
3	77.4	73.6	77.4	**88.2**
4	71.1	69.8	74.2	**85.9**
5	65.8	66.6	67.7	**81.8**
6	70.1	70.0	70.8	**85.6**
7	67.1	67.9	66.8	**80.1**
8	66.0	66.8	66.6	**78.7**
9	67.7	67.7	68.1	**76.6**
10	68.6	68.9	70.9	**79.3**
Avg.	69.3	69.1	69.1	**80.5**

Table 2. Clustering Accuracy(%) and NMI(%) on PIE

L	Accuracy(%)			
	K-means	NMF	NTF	NPNTF
10	30.9	56.5	53.6	**65.9**
20	26.9	56.8	54.0	**63.6**
30	26.3	57.1	57.5	**66.0**
40	25.6	58.6	59.6	**68.2**
50	25.2	58.6	61.4	**69.4**
60	24.6	59.2	60.3	**69.0**
Avg.	26.6	57.8	57.7	**67.0**
L	NMI(%)			
	K-means	NMF	NTF	NPNTF
10	37.1	**66.7**	55.5	64.3
20	43.5	74.4	68.3	**77.1**
30	48.5	78.0	77.1	**85.2**
40	51.1	80.1	81.6	**89.1**
50	52.4	81.3	83.0	**89.8**
60	53.4	82.6	83.1	**89.5**
Avg.	47.7	77.2	74.7	**82.5**

Table 1 and 2 demonstrate that our proposed NPNTF algorithm outperforms all other three methods on both the data sets, which implies an enhanced semantic and discriminating representation capability by encoding the neighborhood information.

5. CONCLUSIONS

Based on the observation of two types of neighborhood information loss, this paper has proposed a novel algorithm called Neighborhood Preserving Non-negative Tensor Factorization (NPNTF) for more effective image representation. By using tensor factorization formula, NPNTF avoids local structure information loss within individual image due to vectorization. Meanwhile, it encodes the local geometrical information among image data by considering the linear reconstruction coefficients with respect to each data point and its neighbors patchwise. Although we handle them in different ways, both of the neighborhood information can be in essence traced back to the strong correlations between neighboring pixels in images. Experimental results on standard image data sets show that NPNTF leads to more powerful image representation and achieves superior clustering performance. Moreover, this general framework is also applicable to higher order tensor apart from image space discussed here. In the future, non-negative tensor factorization on multiple manifolds will be considered.

6. REFERENCES

[1] D.D. Lee and H.S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[2] T. Hazan, S. Polak, and A. Shashua, "Sparse image coding using a 3d non-negative tensor factorization," in *ICCV*, 2005, vol. 1, pp. 50–57.

[3] A. Shashua and T. Hazan, "Non-negative tensor factorization with applications to statistics and computer vision," in *ICML*, 2005, pp. 792–799.

[4] M. Heiler and C. Schnörr, "Controlling sparseness in non-negative tensor factorization," in *ECCV*, 2006, pp. 56–67.

[5] S. Zafeiriou, "Discriminant nonnegative tensor factorization algorithms," *IEEE Trans. Neural Networks,*, vol. 20, no. 2, pp. 217–235, 2009.

[6] H.S. Seung and D.D. Lee, "The manifold ways of perception," *Science*, vol. 290, no. 5500, pp. 2268–2269, 2000.

[7] J.B. Tenenbaum, V. Silva, and J.C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319C–2323, 2000.

[8] S.T. Roweis and L.K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[9] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural computation*, vol. 15, no. 6, pp. 1373–1396, 2003.

[10] D. Cai, X. He, J. Han, and T.S. Huang, "Graph regularized nonnegative matrix factorization for data representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1548–1560, 2011.

[11] Q. Gu and J. Zhou, "Neighborhood preserving nonnegative matrix factorization," in *20th British Machine Vision Conference*, 2009.

[12] B. Shen and L. Si, "Nonnegative matrix factorization clustering on multiple manifolds," in *24th AAAI Conf. Artificial Intelligence*, 2010, pp. 575–580.

[13] C. Wang, X. He, J. Bu, Z. Chen, C. Chen, and Z. Guan, "Image representation using laplacian regularized nonnegative tensor factorization," *Pattern Recognition*, vol. 44, no. 10-11, pp. 2516–2526, 2011.

[14] M. Mørup, L.K. Hansen, and S.M. Arnfred, "Algorithms for sparse nonnegative tucker decompositions," *Neural computation*, vol. 20, no. 8, pp. 2112–2131, 2008.