Pattern Recognition 44 (2011) 2516–2526

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Image representation using Laplacian regularized nonnegative tensor factorization

Can Wang *, Xiaofei He, Jiajun Bu, Zhengguang Chen, Chun Chen, Ziyu Guan

College of Computer Science, Zhejiang University, Hangzhou 310027, China

ARTICLE INFO

Article history:
Received 23 August 2010
Received in revised form
3 February 2011
Accepted 16 March 2011
Available online 25 March 2011

Keywords:
Image representation
Image clustering
Tensor
Graph Laplacian
Manifold

ABSTRACT

Tensor provides a better representation for image space by avoiding information loss in vectorization. Nonnegative tensor factorization (NTF), whose objective is to express an n-way tensor as a sum of k rank-1 tensors under nonnegative constraints, has recently attracted a lot of attentions for its efficient and meaningful representation. However, NTF only sees Euclidean structures in data space and is not optimized for image representation as image space is believed to be a sub-manifold embedded in high-dimensional ambient space. To avoid the limitation of NTF, we propose a novel Laplacian regularized nonnegative tensor factorization (LRNTF) method for image representation and clustering in this paper. In LRNTF, the image space is represented as a 3-way tensor and we explicitly consider the manifold structure of the image space in factorization. That is, two data points that are close to each other in the intrinsic geometry of image space shall also be close to each other under the factorized basis. To evaluate the performance of LRNTF in image representation and clustering, we compare our algorithm with NMF, NTF, NCut and GNMF methods on three standard image databases. Experimental results demonstrate that LRNTF achieves better image clustering performance, while being more insensitive to noise.

1. Introduction

Previous work has shown that the image variations of many objects under variable lighting conditions can be effectively modeled by low dimensional linear spaces [1]. To learn the image subspace, matrix factorization techniques are actively explored in recent years to find two or more lower rank matrices whose product approximates the original matrix. Standard matrix factorization techniques include principle component analysis (PCA), singular value decomposition (SVD), nonnegative matrix factorization (NMF) [2,3] etc. In each of these methods, an $n_1 \times n_2$ image is unfolded into a high dimensional vector in $\mathbb{R}^{n_1 \times n_2}$ and the image space is represented as a high-dimensional matrix.

However, an image is intrinsically a two-dimensional matrix. The vectorized representation fails to take into consideration the spatial locality of pixels in an image and thus will suffer from information loss [4] and usually lead to the *curse of dimensionality* problem [5]. To better discover the inherent structures in image space, it is important to retain the multidimensional structure of the image data. Therefore, tensors or multidimensional arrays become a natural choice for representing image space and tensor decomposition techniques are exploited to gain more insight into image data. Recently, nonnegative tensor factorization [6,7], which is an extension to NMF, has gained much research interest due to its efficient and meaningful representation. NTF algorithms are applied in various domains to discover the latent structures in data sets, including: computer vision [7], EEG data analysis [8,9], discussion tracking in email [10], image representation [11,12,6] etc. Friedlander and Hatzdemonstrate [12] proposed a NTF algorithm with sparseness constraints and demonstrated the effectiveness of their approach on image data. Hazan et al. [11] used NTF to decompose a set of images represented as a 3-way tensor. Their experiments demonstrated NTF could generate more meaningful decomposition and more efficient compression for images than NMF. Comprehensive surveys of NTF algorithms and applications can be found in [13,14].

By imposing nonnegative constraints in factorization, NTF generates a *parts-based* representation for a data set that allows only additive combination of the parts to form the whole. This corresponds to our intuition about physical data sets and hence NTF is considered as an efficient technique for learning the parts of objects such as a set of images. The major limitation of NTF is that it only sees Euclidean structure in data space. However, many researchers have recently shown that the image space is

* Corresponding author.
E-mail addresses: wcan@zju.edu.cn (C. Wang), xiaofeihe@cad.zju.edu.cn (X. He), bjj@cs.zju.edu.cn (J. Bu), cerror@zju.edu.cn (Z. Chen), chenc@cs.zju.edu.cn (C. Chen), guanzh@zju.edu.cn (Z. Guan).

0031-3203/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2011.03.021

generally a nonlinear manifold embedded in the high dimensional ambient space [15–19]. Uncovering the manifold structure of the image space is essential for effective image representation and clustering tasks.

In this paper, we propose a novel Laplacian regularized nonnegative tensor factorization (LRNTF) method for image representation that explicitly takes into account the underlying manifold structure of the image space. Given some images sampled from the image manifold, we can build an adjacency graph to model the local geometrical structure of the manifold. LRNTF finds a factorization that respects this graph structure. That is, data points that are close in the intrinsic geometry of the image space shall thus be close to each other under the factorized tensor basis. We then perform image clustering in this reduced tensor subspace. In this work, we apply k-means in the tensor subspace for clustering. k-means has been frequently used for image clustering for its simplicity. However, its performance will drop significantly when the dimensionality of the image space is high [20]. Furthermore, manifold structure of the image space is not considered in k-means and consequently its performance in clustering data points sampled from nonlinear manifolds is limited. By combining LRNTF and k-means, both limitations of K-means can be overcome.

It is worthwhile to highlight several aspects of the LRNTF proposed in this paper:

(1) LRNTF represents the image space as a 3-way tensor and thus avoid information loss in vectorizing two dimensional images data into one dimensional vector representation. Therefore, LRNTF is expected to achieve better performance in image representation and clustering tasks.

(2) The computation of LRNTF is efficient. Compared with the original NTF algorithm, the only extra cost of LRNFT is computing the Laplacian Regularization term, which is almost negligible. Also, the tensor representation requires few parameters to be independently estimated in clustering, so performance in small data sets is good.

(3) LRNTF explicitly takes into account the underlying manifold structure of the image space, which is modeled by an adjacency graph. Although there exist some previous work that takes into account the underlying manifold structure in nonnegative matrix factorization [21–23], no such work has been done in NTF. Meanwhile, by preserving the neighborhood information in the image manifold, LRNTF is less sensitive to noise and outliers. This result is confirmed in our experiment.

(4) The work in this paper primarily focuses on the 3-way tensor representation of image space. However, the algorithm and analysis presented here can be naturally extended to higher order tensors.

The rest of the paper is organized as follows. In Section 2, we give a brief review of NMF and NTF. We describe our LRNTF algorithm in Section 3. The experimental results are presented in Section 4. Finally, we give concluding discussions and suggestion for future work in Section 5.

2. A brief review of NMF and NTF

NTF is usually considered as a generalization of NMF. To better explain NTF, we give a brief description of both NMF and NTF in this section.

2.1. NMF

The nonnegative matrix factorization (NMF) introduced by Lee and Seung in [2] can be stated as follows: Given a data Matrix $V \in \mathbb{R}^{n\times m}$ with nonnegative elements, find two nonnegative matrices $W \in \mathbb{R}^{n\times r}$ and $H \in \mathbb{R}^{r\times m}$ whose product best approximates W:

$$V \approx WH \tag{1}$$

The rank r of the factorization is generally chosen so that $(n+m)r < nm$, leading to a compressed representation of data in V. In this factorization, each data point $\mathbf{v_j} \in V$ can be viewed as linear combination of the columns of W as follows:

$$\mathbf{v_i} \approx \sum_{j=1}^{r} \mathbf{w_j} h_{ij} \tag{2}$$

Thus the r columns of W can be taken as a basis optimized for this linear approximation and each column of H becomes the new encoding of each data point in this new basis W. Another view of the factorization is to represent NMF in (1) with the following bilinear model:

$$V \approx WH = WB^T = \sum_{i=1}^{r} \mathbf{w_i} \otimes \mathbf{b_i} \tag{3}$$

where B is the transpose of H and $\otimes$ denotes outer products. That is, the data matrix $V \in \mathbb{R}^{n\times m}$ is approximated by a sum of linear combination of rank-one nonnegative matrices $\mathbf{w_i} \otimes \mathbf{b_i}$. As we are going to see, this scheme can be easily extended to the tensor case.

The nonnegative constraints enable a parts-based representation because they allow only additive, not subtractive combinations. NMF corresponds to the parts-based representation in human cognition, as suggested by psychological and physiological evidence in [24,25]. For this reason, NMF is widely applied in many practical problems such as face analysis [26], document clustering [27] and DNA gene expression analysis [28].

Lee and Seung proposed a solution to NMF by minimizing the least-squares error in the approximation shown as follows [3]:

$$\min \|V - WH\|_F^2 \tag{4}$$

where $\|\cdot\|_F^2$ denotes matrix *Frobenius norm*.

Although the above objective function is convex in W only or H only, it is not convex in both variables together. Therefore, it is unrealistic to expect an algorithm to find global minima in Eq. (4). Instead, Lee and Seung presented a "*multiplicative update*" algorithm to find a local minimum for the objective function in Eq. (4) as follows:

$$w_{ij}^{t+1} \leftarrow w_{ij}^{t+1} \frac{(VH)_{ij}}{(VH^T H)_{ij}} \tag{5}$$

$$h_{ij}^{t+1} \leftarrow h_{ij}^{t+1} \frac{(VW)_{ij}}{(VW^T W)_{ij}} \tag{6}$$

Proofs of convergence for this multiplicative update algorithm is presented in [3].

2.2. NTF

Similar to NMF, NTF factorizes data into a lower dimensional space by introducing a more compact basis. If set up appropriately, this new basis can describe the original data in a concise manner, introduce some immunity to noise and facilitate generalization [7]. NTF sees an advantage over NMF in various applications including image representation and clustering, in which the image space is modeled as a 3-way tensor and each image is represented as a 2-way tensor (matrix). This avoids information loss resulted by vectorization as used in NMF [29].

The canonical tensor factorization techniques include the Tucker model [30] that is a higher-order form of principal component analysis, and the CANDECOMP/PARAFAC (CP) model [31,32] that decomposes a tensor as a sum of rank-one tensors. The NTF discussed in this paper is an extension to the CP model.