Neurocomputing 173 (2016) 245–255

Contents lists available at ScienceDirect

Neurocomputing

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

Immersive visualization of visual data using nonnegative matrix factorization

Mohammadreza Babaee [a,*], Stefanos Tsoukalas [a], Gerhard Rigoll [a], Mihai Datcu [b]

[a] Institute for Human–Machine Communication, Technische Universität München, Germany
[b] Institute for Remote Sensing Technology, German Aerospace Center, Wessling, Germany

ARTICLE INFO

Article history:
Received 30 June 2014
Received in revised form
23 February 2015
Accepted 19 March 2015
Available online 5 September 2015

Keywords:
Nonnegative matrix factorization
Dimensionality reduction
Immersive visualization

ABSTRACT

Over the last two decades, dimension reduction for visualization has gained a high amount of attention in visual data mining where the data is represented by high-dimensional features. Basically, this approach leads to an unbalanced and occluded distribution of visual data in display space, giving rise to difficulties in browsing the data. In this paper we propose an approach for the visualization of image collections in such a way as (1) images are not occluded by each other, and the provided space is used as much as possible; (2) the similar images are positioned close together; (3) an overview of data is feasible. To fulfill these requirements, we propose to use regularized Nonnegative Matrix Factorization (NMF) controlled by parameters to reduce the dimensionality of data. Experiments performed on optical and radar images confirm the flexibility of proposed method in visualizing large-scale visual data. Finally, an immersive 3D virtual environment is suggested, to visualize the images, to allow the user to navigate and explore the data.

1. Introduction

The world is dealing with a massive amount of collected data, where much of it is visual data (e.g., images and videos). Facebook reports six billion photo uploads per month. The amount of Earth Observation (EO) images is in the order of several Terra bytes per day. Therefore, browsing and visualizing visual data could help to design new Visual Data Mining (VDM) systems. In such systems, usually, the content of each image (e.g., color, texture, shape) is represented by high-dimensional feature vectors [1,2], where the similarity relationship between images is measured based on the distance between feature points. In VDM a query image might be fetched into the system and the resulting similar images are visualized as thumbnails in a 2D or 3D display space. In interactive VDM [3–5], the interface between the human and the machine plays a key role in enhancing the performance of the system. The interface should provide the user the ability to gain a deep understanding of the data by its visualization.

Visualization of the high-dimensional data (e.g., images represented by high dimensional features) has been always a challenging problem in the area of information mining and visualization. Perhaps the most common way to tackle this problem is to utilize Dimensionality Reduction (DR) techniques, to map high dimensional data to 2D or 3D for visualization. During the last 20 years numerous methods, such as linear or nonlinear, have been proposed to reduce the dimensionality [6–11]. The most common linear methods are Principal Component Analysis (PCA) [12] and Multidimensional Scaling (MDS) [13]. Nonlinear methods assume that the data points are coming from a manifold embedded in the high-dimensional space. Depending on whether to preserve the local or global structure of the manifold, they can be categorized, typically, in local and global methods. Local methods like Locally Linear Embedding (LLE) [14] and Laplacian Eigenmap (LE) [15] emphasize to preserve the locality of data points in contrast to global methods like Stochastic Neighbor Embedding (SNE) [16] and Isomap [17], which emphasize on preserving the global structure of data points.

The NMF was introduced in [18] as a method for dimensionality reduction, where the low-dimensional structure is presented by two nonnegative matrices. Due to the non-negativity of the matrices, the vectors in the low-dimensional representation are the additive combinations of the basis vectors, which leads to a part-based representation. This representation has been shown to correspond to the way images are represented in the human brain [19]. As an extension to the NMF, Graph regularized NMF (GNMF) was proposed that tries to preserve the similarity of the feature vectors in the low-dimensional space [20]. This technique applies the manifold learning technique LLE [14] and adds an additional term to the main objective function.

* Corresponding author.
E-mail addresses: reza.babaee@tum.de (M. Babaee), s.tsoukalas@googlemail.com (S. Tsoukalas), rigoll@tum.de (G. Rigoll), mihai.datcu@dlr.de (M. Datcu).

http://dx.doi.org/10.1016/j.neucom.2015.03.121
0925-2312/

Dimensionality reduction is widely employed to determine the position of images [15,16] in 2D or 3D display space. However, the images are mostly occluded and much of the display space is not used, giving rise to difficulties for the user to gain understanding from the data. To address this issue, some works have been proposed to arrange the images in display space based on optimizing a predefined cost function [21–23]. In [22,23] the authors estimate the two-dimensional locations of the images by minimizing the overlap between images. They defined a cost function as a compromise between similarity preserving and overlap minimization and used gradient descent method to optimized this cost function. Additionally, the authors in [21] proposed an algorithm that spreads images equally in a given area. This is achieved by minimizing a cost function, which consists of a structure-preserving term, an entropy term and a term that penalizes locations of images outside the predefined layout. All the aforementioned methods first reduce the dimensionality of the data and then change the position of the data points to fulfill the other requirements. In summary, a good visualization of images should fulfill three main requirements listed as follows [23]:

(i) *Structure preservation*: the relations between images, mainly similarity and dissimilarity, should be preserved.
(ii) *Visibility*: All displayed images should be visible by the user (i.e. less overlap between images).
(iii) *Overview*: the user should be able to gain an overview of the distribution of images.

In this paper we propose a technique for arranging image collections in 2D/3D display space for the task of image retrieval. The main contribution of our work is (1) developing a novel regularized NMF to position image collections by taking into account the three aforementioned requirements; (2) developing an immersive 3D virtual environment to visualize the images [24,25] to allow the user to navigate inside the data and explore it. Basically, there is no harm in non-negativity constraint of NMF, since each image is represented by a Bag-of-Words (BoW) model of local features [26], which has nonnegative values. In BoW, an image is treated as a document and its local features as words. The extracted feature from all images are pooled and clustered. Then, a histogram of extracted local features from each image is constructed based on cluster centers to represent the image. In our work, we propose a regularization term for each aforementioned requirements controlled by some parameters. Precisely, we add one regularization for structure preservation requirement, one for overview requirement and one for visibility requirement. For structure preservation we consider the sum of locality (similarity) preserving [20] and farness preserving [27]. The Renyi entropy is used to define the visibility regularizer. Finally, the result of clustering in the original space is selected to define the overview regularization. These regularization terms, controlled by some parameters, are added to the main NMF function to define the main formulation of image positioning. To visualize the images in a 3D virtual environment, we utilize Virtual Reality (VR) technology to build an immersive 3D virtual environment, namely Cave Automated Virtual Environment (CAVE). Evidently, a good visualization of images does not depend on only the position of images, but also a proper display space can deliver much more information about the data to the user.

The rest of paper is organized as follows: Section 2 provides a review of NMF method and its updating rules. In Section 3 we explain our proposed regularized NMF for the visualization of image collections. The details of the immersive visualization system are provided in Sections 4. Experimental validations are represented in Section 5. Finally, in Section 6 we draw our conclusions.

2. A review of NMF

We consider a data matrix $X=[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N]\in\mathbb{R}^{M\times N}$, where x_i is a feature vector, N is the number of samples and M is the dimension of the feature vectors. Given a new reduced dimension K, the NMF algorithm approximates the matrix X by a product of two non-negative matrices $U=[u_{ik}]\in\mathbb{R}^{M\times K}$ and $V=[v_{jk}]\in\mathbb{R}^{N\times K}$

$$X\approx UV^{\mathrm{T}}. \tag{1}$$

Thereby, in the new representation, U can be considered as a set of basis vectors and $V=[\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n]^{\mathrm{T}}$ as the coordinates of each sample with respect to these basis vectors. Perhaps, the two most popular cost functions, that quantify the quality of the approximation, are the square of the Frobenius norm of the matrix differences and the divergence between the two matrices [28]. During the rest of the paper we will focus on the Frobenius cost function, for which the NMF can be stated as the following minimization problem:

$$\min O_F=\|X-UV^{\mathrm{T}}\|^2 \tag{2}$$
$$\text{s.t. } U=[u_{ik}]\geq 0$$
$$V=[v_{jk}]\geq 0.$$

While the minimization problem is convex with respect to U and with respect to V, it is not convex in both variables together. Therefore there exist many local minima. In [28] Lee and Seung presented an update rule that finds a local minimum as follows:

$$u_{ik}\leftarrow u_{ik}\frac{(XV)_{ik}}{(UV^{\mathrm{T}}V)_{ik}},\quad v_{jk}\leftarrow v_{jk}\frac{(X^{\mathrm{T}}U)_{jk}}{(VU^{\mathrm{T}}U)_{jk}} \tag{3}$$

It is proved [28] that the update rules find a local minimum for the minimization objective (2).

3. Regularized NMF for visualization

In order to achieve a good visualization, we require the following constraints for the low-dimensional representation:

3.1. Structure preservation

For the preservation of structure, we require that similar images are placed close to each other and dissimilar images far away from each other. The constraint for similarity, which was introduced with the GNMF-algorithm in [20], forces samples which are close to each other in the high dimensional space, to be also close to each other in the low-dimensional representation. This constraint is achieved with the help of a weight matrix W, which represents the internal manifold structure of the high-dimensional data. This matrix is based on the construction of a nearest neighbor graph, where for each point $\boldsymbol{x}_j$ we find its k nearest neighbors and put an edge between $\boldsymbol{x}_j$ and each neighbor. Based on this graph, there are many possibilities to construct the matrix W. In this paper we adopt the heat kernel weighting, where

$$W_{jl}=e^{-((\|\boldsymbol{x}_j-\boldsymbol{x}_l\|^2)/\sigma)}\quad \text{s. t. } \sigma>0 \tag{4}$$

if nodes j and l are connected and 0 otherwise. Based on W the authors of [20] introduced the following term for similarity preservation in the NMF-objective:

$$O_s=\frac{1}{2}\sum_{j,l}\|\boldsymbol{v}_j-\boldsymbol{v}_l\|^2W_{jl}$$
$$=\mathrm{Tr}(V^{\mathrm{T}}LV) \tag{5}$$

where $L=D-W$ and D is a diagonal matrix, whose entries are column sums of W, $D_{jj}=\sum_l W_{jl}$.