IMMERSIVE VISUALIZATION OF SAR IMAGES USING NONNEGATIVE MATRIX FACTORIZATION

Mohammadreza Babaee[2,3], *Stefanos Tsoukalas*[2], *Simon Schenk*[2], *Gerhard Rigoll*[2], *Mihai Datcu*[1,3]

[1]Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen
[2]Institute for Human-Machine Communication, TU München, Germany
[3]Munich Aerospace Faculty, Germany

ABSTRACT

Interactive visual data mining, where the user plays a key role in learning process, has gained high attention in data mining and human-machine communication. However, this approach needs Dimensionality Reduction (DR) techniques to visualize image collections. Although the main focus of DR techniques lays on preserving the structure of the data, the occlusion of images and inefficient usage of display space are their two main drawbacks. In this work, we propose to use Non-negative Matrix Factorization (NMF) to reduce the dimensionality of images for immersive visualization. The proposed method aims to preserve the structure of data and at the same time reduce the occlusion between images by defining regularization terms for NMF. Experimental validations performed on two sets of image collections show the efficiency of the proposed method in respect to controlling the trade-off between structure preserving and less occluded visualization.

Index Terms— Immersive visualization, Non-negative matrix factorization, entropy

1. INTRODUCTION

Interactive data mining has gained high attention in Human-Machine Communication (HMC) for pattern recognition problems. One of the main challenges here is to develop an interface between the human and machine to interact efficiently. Immersive visualization of visual data might be a good option to enable the user to explore a large set of images and navigate inside the data [1, 2]. However, the main problem is how to arrange the images in this environment such that:

- Similar images are close together (i.e., data structure is preserved);
- Images are less occluded by each other;
- The display space is used as much as possible;

To address this problem, we propose to formulate a NMF algorithm that takes into account the aforementioned requirements. There is no harm in non-negativity constraint of NMF, since basically each image is represented by a Bag-of-Words (BoW) model of local features. Therefore, the representative feature vectors are non-negative values. In our approach, we compute the graph of similarity in high dimensional space and use it as the structure of the data to be preserved. It has been shown that the Laplacian of the graph as a regularization term can significantly preserve the similarity in a NMF framework [3]. To decrease the overlap among the images, we propose to use entropy as the second regularization term of NMF. A good visualization is a trade-of between similarity and occlusion. This trade-of is controlled by a parameter λ. By increasing the entropy, we have less occlusion but the structure of data vanishes.

The rest of paper is organized as follow: Section 2 reviews related works in the area of visualization of image collection. Sections 3.1 briefly explains the concept of immersive visualization. In Section 4 we explain the proposed NMF for dimensionality reduction. Experimental validations are represented in Section 5. Finally, in Section 6 we draw our conclusions.

2. RELATED WORK

The NMF was introduced in [4] as method for dimensionality-reduction, where the low-dimensional structure is presented by two non-negative 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 parts-based representation. This representation has been shown to correspond to the way images are represented in the human brain. 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 [3]. This technique applies the manifold learning technique LLE [5] and adds an additional term to the main objective function.

In the area of image visualization, previous work has been done in [6], where the authors choose the two-dimensional locations of the images by minimizing a cost function consisting of a structure-preserving and an overlap term. For the overlap all images are represented as circles with the same radius and the overlapping area of the circles is computed. Additionally, in [7] an algorithm has been proposed, 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. In summary, 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.

3. APPROACH

3.1. Immersive Visualization

For the visualization of the image collections we utilize a Cave Automated Virtual Environment (CAVE). The cave consists of four room-sized walls, which play the role of four display screens. They are aligned to form a cube-shape space. This configuration allows the user to have a 180 degree horizontal view. The computer generated scene is projected onto the walls, using two projectors per wall, in order to have stereoscopic scenarios. Furthermore, a real-time tracking system comprising six infrared cameras, mounted on top of the walls, computes the position and orientation of the user inside the cube. This system provides the user with the ability of exploring and interactivity with the data. The user is allowed to navigate inside the data and gain valuable information about the structure of data. A snapshot of immersive visualization of image collection in the CAVE is depicted in Figure. 1.

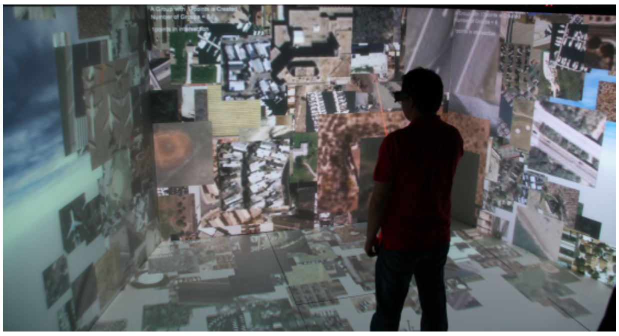

Fig. 1. Immersive visualization of image collections in a Cave Automated Virtual Environment

4. REGULARIZED NMF

We consider a data matrix $X = [\boldsymbol{x}_1, ..., \boldsymbol{x}_N] \in \mathbb{R}^{M\times N}$, 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 as the coordinates of each sample with respect to these basis vectors. There are two cost functions, that quantify the quality of the approximation, the square of the Frobenius norm of the matrix differences and and the divergence between the two matrices [8]. During the rest of the paper we will focus on the divergence cost function:

$$O_D = \sum_{i,j}\left(x_{ij}\log\frac{x_{ij}}{y_{ij}} - x_{ij} + y_{ij}\right) \tag{2}$$

where

$$Y = [y_{ij}] = UV^{\mathrm{T}} \tag{3}$$

Many extensions have been proposed to the NMF algorithm, which introduce different regularizers into the objective function in order to enforce additional favorable properties for the new representation. For example, the GNMF adds a similarity term to the objective in order to preserve the local neighbor structure in the new representation:

$$G = \sum_{i=1}^{M}\sum_{j=1}^{N}\left(x_{ij}\log\frac{x_{ij}}{\sum_{k=1}^{K}u_{ik}v_{jk}} - x_{ij} + \sum_{k=1}^{K}u_{ik}v_{jk}\right) + \frac{\lambda_1}{2}\sum_{j=1}^{N}\sum_{l=1}^{N}\sum_{k=1}^{K}\left(v_{jk}\log\frac{v_{jk}}{v_{lk}} + v_{lk}\log\frac{v_{lk}}{v_{jk}}\right)W_{jl} \tag{4}$$

with the regularization parameter λ_1 and the weight matrix W. While the GNMF algorithm makes it possible to reduce the dimensionality of the data for visualization and keep similar objects close to each other, it has the disadvantage that the resulting distribution does not use the entire available space and therefore the occlusion between the images is still high. To remedy this, we propose to extend the GNMF with an entropy term, which enforces the points in the new representation to be spread across the available space. The entropy term we use, is the generalized entropy term.

$$-H = \sum_{j=1}^{N}\sum_{k=1}^{K}(v_{jk}\log v_{jk} - v_{jk}) \tag{5}$$

which reduces to the Shannon entropy, when the condition $\sum_{jk} v_{jk} = 1$ holds true. In order to maximize the entropy of the system, we introduce the negative of the entropy term into the minimization objective (4), which leads to the new objective:

$$O = \sum_{i=1}^{M} \sum_{j=1}^{N} \left(x_{ij} \log \frac{x_{ij}}{\sum_{k=1}^{K} u_{ik} v_{jk}} - x_{ij} + \sum_{k=1}^{K} u_{ik} v_{jk} \right)$$
$$+ \frac{\lambda_1}{2} \sum_{j=1}^{N} \sum_{l=1}^{N} \sum_{k=1}^{K} \left(v_{jk} \log \frac{v_{jk}}{v_{lk}} + v_{lk} \log \frac{v_{lk}}{v_{jk}} \right) W_{jl}$$
$$+ \lambda_2 \sum_{j=1}^{N} \sum_{k=1}^{K} \left(v_{jk} \log v_{jk} - v_{jk} \right), \tag{6}$$

with the regularization parameter λ_2. In order to find a local minimum of this objective function we follow the principles proposed in [8] to derive update rules for U and V. Introducing the Lagrange multipliers ψ_{ik} and ϕ_{ik} for the constraints $u_{ik} \geq 0$ and $v_{ik} \geq 0$, respectively, and setting $\Psi = [\psi_{ik}]$ and $\Phi = [\phi_{ik}]$ leads to the Lagrange $\mathcal{L}$:

$$\mathcal{L} = O + \text{Tr}(\Psi U^{\text{T}}) + \text{Tr}(\Phi V^{\text{T}}) \tag{7}$$

with the KKT conditions

$$\frac{\partial \mathcal{L}}{\partial U} = 0, \; \frac{\partial \mathcal{L}}{\partial V} = 0 \tag{8}$$

and

$$\psi_{ik} u_{ik} = 0, \; \phi_{ik} u_{ik} = 0 \tag{9}$$

we come up with the following update rules for U and V:

$$u_{ik} \leftarrow u_{ik} \frac{\sum_j \left(x_{ij} v_{jk} / \sum_k u_{ik} v_{jk} \right)}{\sum_j v_{jk}} \tag{10}$$

$$\boldsymbol{v}_k \leftarrow \left(\left(\sum_i u_{ik} + \lambda_2 \right) I + \lambda_1 L \right)^{-1} R \tag{11}$$

with

$$R = \begin{bmatrix} v_{1k} \sum_i \left(x_{i1} u_{ik} / \sum_k u_{ik} v_{1k} \right) + \lambda_2 \\ \vdots \\ v_{Nk} \sum_i \left(x_{iN} u_{ik} / \sum_k u_{ik} v_{Nk} \right) + \lambda_2 \end{bmatrix} \tag{12}$$

where $L = D - W$ and D is a diagonal matrix, whose entries are column sums of W, $D_{jj} = \sum_l W_{jl}$. For the derivation of the update rule for V (11) we used an approximation for the resulting logarithmic functions, which is based on the 1st order Taylor expansion:

$$\log x \approx 1 - \frac{1}{x}. \tag{13}$$

The update rule for U remains the same as in the original algorithm [8], since the newly introduced terms in the objective (6) depend only on the variable V.

5. EXPERIMENTS

5.1. Datasets

The two real datasets used in our experiments are: 1) SAR dataset; 2) Merced dataset.
SAR dataset contains 3434 SAR (Synthetic Aperture Radar) images of the size 160x160 in 15 classes, such as presences of forests, water, roads and urban area density. From this dataset we extracted 64-dimensional SIFT [9]feature vectors.
Merced dataset contains 2100 images in 21 different groups. From these images we also extracted SIFT feature vectors, leading to 64-dimensional feature vectors.

5.2. Setup

For the experiments we selected a random subset of 1000 samples from each dataset. We run the optimization algorithm 10 times with different random starting points for each experiment and selected the best result. In order to analyze the trade-off between similarity and entropy maximization we introduced the parameters α and λ and set $\lambda_1 = \alpha\lambda$, $\lambda_2 = \alpha(1 - \lambda)$. In this way, α represents the scaling of the similarity and entropy terms relative to the NMF-term and λ represents the trade-off between similarity and entropy, where $\lambda = 1$ corresponds to max. similarity and $\lambda = 0$ to max. entropy. For the weighting matrix W we used a heat kernel matrix with the number of neighbors $k = 7$. We compute the similarity preservation of the resulting representation by using the inverse of the similarity term in the cost function (6). For the occlusion we represent all images as cubes with edge size $s = \max(b, h)$, where b and h are the width and height of the images, respectively, to measure the overlapping volume of all cubes.

5.3. Results

Figures 2a and 3a depict the 3D-visualization of the images for $\alpha = 100$ and $\lambda = 1$, figures 2b and 3b the 3D-visualization for $\alpha = 100$ and $\lambda = 0.2$ for the different datasets. As can be seen, with a decreasing value of λ the images are better distributed in the available space, while similar images are still placed close to each other. The behavior of the algorithm is analyzed in figures 2c and 3c. The plots show, that in general for a value of $\alpha \geq 100$ good results can be achieved in respect of similarity and occlusion. Figures 2d and 3d show, that the algorithm converges fast and finds the local minimum for the two datasets in about 40 iterations.

6. CONCLUSIONS AND FUTURE WORK

We presented a novel method to find a low-dimensional representation of an image dataset for 3D-visualization in an immersive virtual environment. By extending the GNMF algorithm with an entropy-maximizing term, we are able to

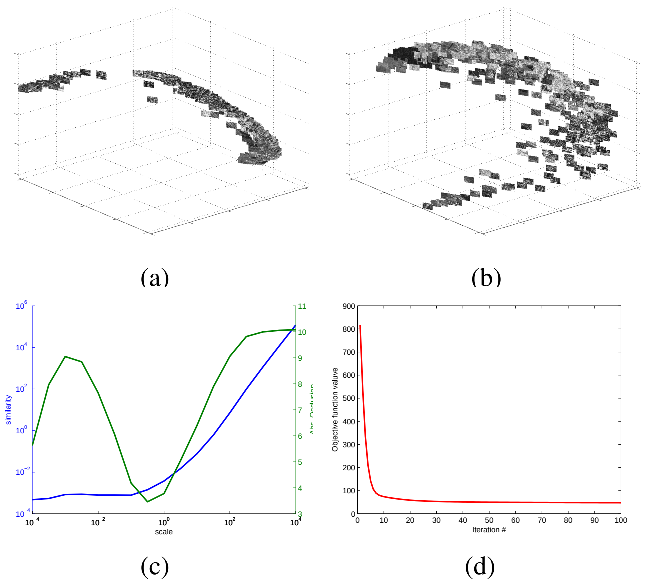

Fig. 2. SAR dataset: (a) Visualization of images for $\alpha = 100$ and $\lambda = 1$ (b) Visualization of images for $\alpha = 100$ and $\lambda = 0.2$ (c) absolute occlusion and similarity for different values of α and $\lambda = 0.2$ (d) Convergence rate.

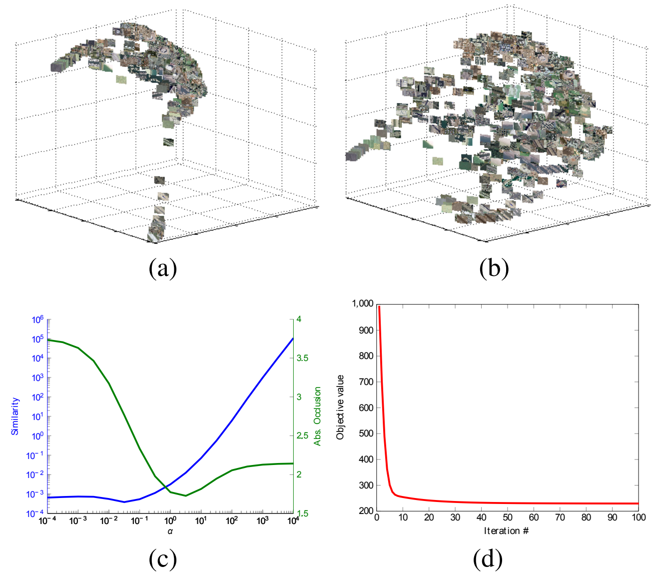

Fig. 3. Merced dataset: (a) Visualization of images for $\alpha = 100$ and $\lambda = 1$ (b) Visualization of images for $\alpha = 100$ and $\lambda = 0.2$ (c) absolute occlusion and similarity for different values of α and $\lambda = 0.2$ (d) Convergence rate.

achieve a similarity-preserving representation, which spreads out in the available space and minimizes the occlusion between images. Experimental results confirm this behavior.

One disadvantage of the proposed algorithm is the additional parameter introduced by the entropy term. Therefore, one possible direction for future work is to reduce the number of parameters by finding the optimal relationship between λ_1 and λ_2. Furthermore the definition of a similarity term, that is independent of the number of neighbors k, is another area open to further exploration.

7. ACKNOWLEDGMENTS

This work has been supported by Munich Aerospace.

8. REFERENCES

[1] Mohammadreza Babaee, Gerhard Rigoll, and Mihai Datcu, "Immersive interactive information mining with application to earth observation data retrieval," in *Availability, Reliability, and Security in Information Systems and HCI*, pp. 376–386. Springer, 2013.

[2] Mohammadreza Babaee, Mihai Datcu, and Gerhard Rigoll, "Assessment of dimensionality reduction based on communication channel model; application to immersive information visualization," in *Big Data, 2013 IEEE International Conference on*. IEEE, 2013, pp. 1–6.

[3] Deng Cai, Xiaofei He, Jiawei Han, and Thomas S. Huang, "Graph regularized nonnegative matrix factorization for data representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1548–1560, 2011.

[4] 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, October 1999.

[5] Sam T. Roweis and Lawrence K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *SCIENCE*, vol. 290, pp. 2323–2326, 2000.

[6] Giang P Nguyen and Marcel Worring, "Interactive access to large image collections using similarity-based visualization," *Journal of Visual Languages & Computing*, vol. 19, no. 2, pp. 203–224, 2008.

[7] Ruixuan Wang, Stephen J. McKenna, Junwei Han, and Annette A. Ward, "Visualizing image collections using high-entropy layout distributions.," *IEEE Transactions on Multimedia*, vol. 12, pp. 803–813, 2010.

[8] Daniel D. Lee and H. Sebastian Seung, "Algorithms for non-negative matrix factorization," in *In NIPS*. 2001, pp. 556–562, MIT Press.

[9] David G Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.