VISUALIZATION OF IMAGE COLLECTION IN 3D: APPLICATION TO IMMERSIVE INFORMATION MINING

Mohammadreza Babaee, Gerhard Rigoll

Munich Aerospace Faculty
Institute for Human-Machine Communication
TU München, Germany

Mihai Datcu

Munich Aerospace Faculty
German Aerospace Center (DLR)
Weßling, Germany

ABSTRACT

Dimensionality reduction for visualization is widely used in visual data mining where the data is represented by high dimensional features. However, this leads to have an unbalanced and occluded distribution of visual data in display space giving rise to difficulties in browsing images. In this paper, we propose an approach to the visualization of images in a 3D display space in such a way that: (1) images are not occluded and the provided space is used efficiently; (2) similar images are positioned close together. An immersive virtual environment is employed as a 3D display space. Experiments are performed on an optical image dataset represented by color features. A library of dimensionality reduction is employed to reduce the dimensionality to 3D. The results confirm that the proposed technique can be used in immersive visual data mining for exploring and browsing large-scale datasets.

***Index Terms*—** Immersive data mining, visualization, entropy, image layout, dimension reduction

1. INTRODUCTION

Browsing and visualizing image collections is a key part in Visual Data Mining (VDM) systems. In such systems, the content of each image (e.g., color, texture, shape) is represented by a high-dimensional feature point [1, 2]. Furthermore, the similarity relationship between images is measured on the basis of the distance between feature points. In VDM, a query image might be loaded into the system and the resulting similar images are visualized as thumbnails in a 2D or 3D display space. For visualization dimensionality reduction is widely used to determine the position of images [3, 4]. However, the images are mostly occluded and much of the display space is not used. Therefore, it is essential to arrange the images in the display space in such a way that: 1) the similar images are positioned close together; 2) the display space is used as efficiently as possible by uniformly distributing the images.

The related work in the area of visualization of images is mainly well suited for 2D display space, where the positions of images are determined by minimizing a cost function that measures the overlap between images [5, 6, 7].

In this paper, we propose a technique of arranging image collections in an immersive 3D virtual environment for the task of image retrieval. More specifically, we arrange the images in a 3D space so that similar images are near to each other with an acceptable distance from one another so that there is no overlap between images. In this visualization, the user can watch the images without having the problem of cluttered images.

The contribution of the paper is as follows: Section 2 describes our proposed methodology for the visualization of image collections. This section is organized in two parts. In the first part, we provide the information about immersive visualization system and in the second part we formulate the problem of arranging image collections in 3D display space. Finally, Section 3 presents the experiments and results.

2. METHOD

2.1. Image positioning

Let us assume that the initial positions of a set of images $\{I_i\}$ are represented by three dimensional points X_i, where $i = 1..N$, and N states the number of images. The goal is to find the optimal position of each image, Y_i. For each image, i, we define asymmetric probabilities, P_{ij} and Q_{ij}, from the initial and optimal positions of images respectively, showing the probability that i is a neighbor of j.

$$P_{ij} = \frac{exp(\|x_i - x_j\|^2)}{\sum_{i \neq k} exp(\|x_i - x_k\|^2)} \tag{1}$$

$$Q_{ij} = \frac{exp(\|y_i - y_j\|^2)}{\sum_{i \neq k} exp(\|y_i - y_k\|^2)} \tag{2}$$

The sum of Kullback-Leibler divergences between distributions P_{ij} and Q_{ij} for each point gives us the dissimilarity of these two distributions:

$$C_{dis} = \sum_i \sum_j P_{ij} \log \frac{P_{ij}}{Q_{ij}} \tag{3}$$

We assume that each image occupies a sphere with radius $r = max\{w, h\}$, where w and h represent the width and height of each image, respectively and the layout of 3D display space is a sphere with radius R. For each image, we define a Gaussian, G_i, with $\mu = Y_i$ and $\sigma = r\sqrt{\frac{R^3}{N*r^3}}$. Then the Gaussian of all optimal image positions can be combined to define a Gaussian mixture. More specifically,

$$P(Y) = \sum_{i=1}^{N} G_i(Y; Y_i, \sigma) \quad (4)$$

In order to have less overlap between images, we need to increase the entropy of $P(Y)$. Therefore, we use quadratic Renyi entropy due to the fact that it can be simply estimated for a Gaussian mixture model by:

$$H = -\log \int P(Y)^2 dy = -\log\{\frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} G_i(Y_j; Y_i, \sigma)\} \quad (5)$$

Obviously, to increase the entropy, H, we should decrease $-H$. Now we have two cost functions such that their combination leads to a single cost function controlled by the parameter λ. This parameter controls the trade-off between preserving image similarity and uniform positioning of images in the 3D display space. Therefore, the total cost function is defined by

$$C_{tot} = (1 - \lambda)C_{dis} - \lambda H \quad (6)$$

The minimization of C_{tot} gives rise to an optimal positioning of images. Practically, this minimization can be solved by the gradient descent method, where the gradient of C_{dis}, and H are

$$\frac{\delta C_{dis}}{\delta Y_i} = 2\sum_{j=1}^{N}(Y_i - Y_j)(P_{ij} - Q_{ij} + P_{ji} - Q_{ji}) \quad (7)$$

and

$$\frac{\delta H}{\delta Y_i} = \frac{1}{2\sigma\alpha} \sum_{j=1}^{N}\{G_j(Y_i; Y_j, 2\sigma) * (Y_i - Y_j)\} \quad (8)$$

where

$$\alpha = \sum_{i=1}^{N} \sum_{j=1}^{N} G_j(Y_i; Y_j, 2\sigma) \quad (9)$$

2.2. Immersive visualization

In order to have 3D visualization of image collections, we employ immersive virtual environment playing the role of 3D display interface. This interface is composed of a CAVE Automated Virtual Environment (CAVE) as display space and a cluster of PCs which work together to render and control the scene. The CAVE consists of four room-sized walls as display surfaces, mounted on top of which there are optical cameras for tracking the user. There are also projectors behind the walls projecting the scene on the walls. A schematic of CAVE is depicted in Fig. 1.

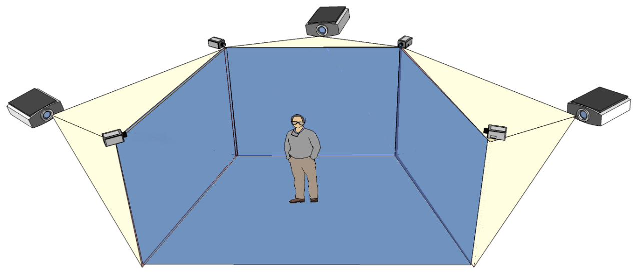


Fig. 1. Immersive virtual environment for 3D visualization of image collections. The user goes inside the CAVE and can get insight into data and navigate to explore and browse the images.

3. EXPERIMENTAL RESULTS

An experiment was performed on a dataset of optical images composing of 1500 images. These images are categorized into 15 groups so that each group contains 100 images. We use a bag-of-words model of the color-histograms [2] as a feature vector of length 200 to represent each image. For visualization we employed both linear and nonlinear dimensionality reduction techniques. Multidimensional Scaling (MDS)[8] and Principal Component Analysis (PCA)[9] are employed as linear techniques and Laplacian Eigenmaps (LE) [3] and Locally Linear Embedding (LLE) [10] are used as nonlinear ones to reduce the dimensionality of features from 200 to 3. We also applied our visualization algorithm to the output of the employed dimensionality reduction techniques to arrange images and the results are shown in Fig 2.

By comparing the visualization results we see that the proposed method arranges the images with an acceptable distance from each other so that there is less overlap between images and also images are not cluttered. Meanwhile, the similarity of images is almost preserved. In conclusion, the proposed algorithm for arranging the images is a trade-off between preserving image similarity and the efficient use of display space which is controlled by a parameter.

4. CONCLUSION

In this paper we have proposed an algorithm to position an image collection in 3D space such that similar images are close together and also the display space is used as much as possible. To this end a cost function is defined that measures the dissimilarity of images and also the entropy of image positions. We have shown that the minimization of this cost function leads to a non-occluded images visualization to make it easier to the user to understand the structure of data. Several examples confirm the effectiveness of proposed methods for the visualization of optical image collection.

Acknowledgement

This work is supported by PhD scholarship Award by Munich Aerospace Faculty.

5. REFERENCES

[1] D.G. Lowe, "Object recognition from local scale-invariant features," in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*. Ieee, 1999, vol. 2, pp. 1150–1157.

[2] K.E van de Sande, T. Gevers, and C.G. Snoek, "Evaluating color descriptors for object and scene recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 9, pp. 1582–1596, 2010.

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

[4] G. Hinton and S. Roweis, "Stochastic neighbor embedding," *Advances in neural information processing systems*, vol. 15, pp. 833–840, 2002.

[5] R. Wang, S.J. McKenna, and J. Han, "High-entropy layouts for content-based browsing and retrieval," in *Proceedings of the ACM International Conference on Image and Video Retrieval*. ACM, 2009, p. 16.

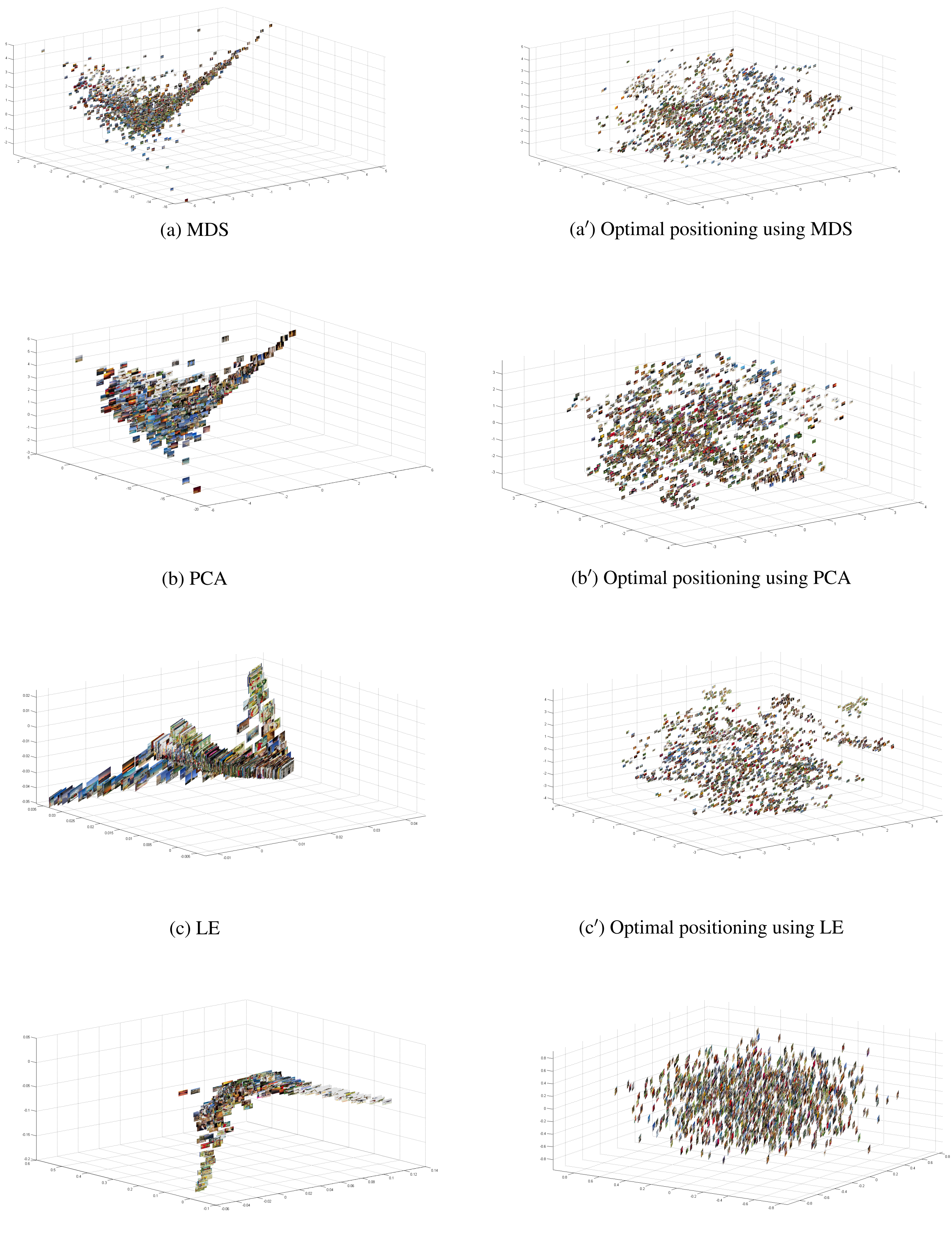
[6] B. Moghaddam, Q. Tian, N. Lesh, C. Shen, and T.S. Huang, "Visualization and user-modeling for browsing personal photo libraries," *International Journal of Computer Vision*, vol. 56, no. 1-2, pp. 109–130, 2004.

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

[8] I. Borg, *Modern multidimensional scaling: Theory and applications*, Springer, 2005.

[9] I. Jolliffe, *Principal component analysis*, Wiley Online Library, 2005.

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



(a) MDS

(a′) Optimal positioning using MDS

(b) PCA

(b′) Optimal positioning using PCA

(c) LE

(c′) Optimal positioning using LE

(d) LLE

(d′) Optimal positioning using LLE

Fig. 2. Visualization of optical image dataset. The left column depicts the visualization utilizing MDS, PCA, LE, and LLE. The right column depicts the visualization utilizing the proposed method. In (a′)-(d′), the initial position of images come from PCA , MDS, LE, and LLE respectively.