IMMERSIVE VISUAL INFORMATION MINING FOR EXPLORING THE CONTENT OF TERRASAR-X ARCHIVES

Mohammadreza Babaee[1], Rreza Bahmanyar[2], Gerhard Rigoll[1], Mihai Datcu[2]

1) Munich Aerospace Faculty, Institute for Human-Machine Communication, Technische Universität München,
Theresienstr. 90, 80333 Munich, Germany - {reza.babaee,rigoll}@tum.de
2) Munich Aerospace Faculty, German Aerospace Center (DLR), Oberpfaffenhofen, 82234 Wessling, Germany
{gholamreza.bahmanyar, mihai.datcu}@dlr.de

ABSTRACT:

The amount of collected Earth Observation (EO) data is increasing intensively in order of several Terabytes a day. Consequently, new trends to explore and retrieve information from the available data is highly demanded. Recent proposed methods for exploring EO data are mainly based on Image Information Mining (IIM). Because the main steps of this approach are image features extraction, data reduction, and labeling, developing a new process chain, mainly based on human interaction, might be a promising solution. More precisely, this chain provides an active learning system by interacting human users with features descriptors in a virtual environment.
The focus of this article is an Immersive Visual Information Mining in which feature descriptors/images are visualized in a 3-D virtual environment, so called CAVE. This environment also allows the users to manipulate the feature space to increase the performance of learning process in an interactive manner. In our experiments, we use a dataset of Synthetic Aperture Radar (SAR) images. To process the data, the contents of images are represented by three different feature descriptors comprising Gabor, adapted WLD, and Bag-of-Words.

1 INTRODUCTION

The Earth is facing unprecedented climatic, geomorphological, environmental and anthropogenic changes which require global scale observations and monitoring. The collected EO data volumes are increasing immensely with a rate of several Terabytes a day. The current EO technologies will be soon amplify these figures to exceed beyond Zettabytes of data. This increases the challenge in exploring the available data to provide focused information and knowledge in a simple and understandable format. Because, as it is well known in EO research area, the data repositories are too large to be scanned or analyzed thoroughly by humans. Moreover, advent of very high resolution EO adds another big challenge, so called explosion of the information content of data.
The EO data content comprises many facets. They carry quantitative information about physical, geometrical, or other types of attributes of the observed scene. The instrument (sensor) and image formation parameters lend understandability to the data. They also refer to geographical names as well as to geomorphologic, tectonic, and political categories. They have cartographic symbols, and last but not least, have ubiquitous names. Therefore, any proposed approaches to use this data must necessarily automates the retrieval of relevant images from the repository.

The focus of this article is on Immersive Visual Information Mining in which the structure of a given image collection is visualized and explored in a 3D virtual environment (namely, CAVE). As the first step, the contents of images are extracted and represented by feature descriptors. A library of feature descriptors for SAR is used, comprises Gabor, adapted WLD, and Bag-of-Words.
Since the extracted features are usually high-dimensional, in order to visualize them in 3D environment, the data is mapped from high-dimensional space into the 3D space. Because the Dimensionality Reduction (DR) are lossy techniques, using a suitable DR method is one of the challenges in visualization-based techniques. In our experiments we use two different approaches (e.g., Laplacian Eigenmaps and Stochastic Neighbor Embedding). Rest of the paper is organized as follows. In Section 2, the methodology used in our work is explained. A brief review of the used feature descriptors and dimensionality reductions are also included.

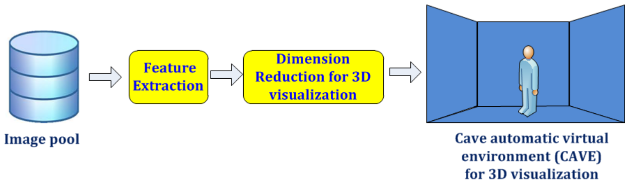


Figure 1: An immersive visualization system provides the user with a visual representa-tion of data. Here, high-dimensional features are extracted from a database of Earth Observation images and are fed into a dimension reduction technique to be visualized in an immersive 3D virtual environment

Section 3 introduces the immersive visualization technique. Experimental results in Section 4 demonstrate the applicability of our approach. Finally, in Section 5, we conclude our work and state what one can do in future based on this approach.

2 METHODOLOGY

The main idea behind our approach is to visualize the images based a set of their representative features (i.e., feature descriptors) in an immersive virtual 3D environment, so called CAVE. This visualization allows the users to explore the feature space for better understanding of the structure of the given collection of images. Further CAVE make it possible for the users to interact with the feature space. In other words, users can manipulate the feature space to better representation of data as well as providing feedback. The feedback can be used by learning systems in data learning process (i.e., active learning).
One of the main steps in any visualization chain is Dimensionality Reduction (DR). Further, DR is highly used in data mining area to reduce the dimensionality of the feature space not only to decrease the computation complexity, but also to ignore the redundant and irrelevant features. Since DR is a lossy technique, improving the performance of DR techniques is a big challenge in data mining.

As it is illustrated in Figure. 1, in our approach, we extract high-dimensional feature descriptors to represent the content of images. Then, we reduce the dimensionality of the feature descriptors using state-of-the-art techniques. Finally results are visualized by projecting them in an immersive 3D virtual environment.

2.1 Feature Extraction and Assessment

In image retrieval and information mining, feature extraction and indexing are still challenging problems. The main objective of our research is to investigate the behavior of different feature descriptors in extracting the content of images. These descriptors represent the content of images based on their low-level features (e.g., shape, texture, color). In this research, a set of feature descriptors, representing image from various aspects, for multi-spectral and SAR data are used (e.g., spectral-SIFT, spectral-WLD, color-histogram).
Spectral SIFT and WLD are local feature descriptors (Lowe, 1999), (Chen et al., 2008). While SIFT is known for extracting lines and corner structures, WLD is mostly used for describing textural structures in images. Since human vision system is able to capture knowledge about the content of image data based on the frequency of different colors, in optical EO images color-histogram is extracted as a kind of feature descriptor. Color-histogram is produced by concatenating the local histograms of pixel intensity values for the three, RGB, channels (van de Sande et al., 2010). Exploring the feature spaces produced by different feature descriptors provides useful knowledge about not only the content of the data but also the discriminability of the features. This helps to develop more sophisticated feature descriptors which can be tuned to recognize a particular human-semantic concept. Moreover, they can be more general to group a collection of images into human-understandable classes.

2.2 Dimensionality Reduction

In order to visualize the features in the CAVE we have to reduce their dimensionality from ND to 3D. Laplacian Eigenmaps (LE) (Belkin and Niyogi, 2003), and Stochastic Neighbor Embedding (SNE) (Hinton and Roweis, 2002) are the two techniques that have been used in this work. LE is a graph based algorithm that reduces the dimension of data and represent it as a curve in a high dimensional space. This curve is represented as a graph whose nodes are data points and edges are determined based on K-NN algorithm. SNE is a probabilistic approach that aims to preserve the neighborhood identities in the low dimensional representation of data. The same construction of neighborhood is also done in a low-dimensional space and the goal is to match these two constructions as much as possible.

3 IMMERSIVE VISUALIZATION

A collection of SAR images, provided by German Aerospace Center's EO Digital Library, Figure 2, as well as a multi-spectral dataset, Figure 3, UCMerced-LandUse (Yang and Newsam, 2010), are used in our experiments. In the first step the feature descriptors are extracted from the given images (e.g., SIFT and WLD for SAR data and SIFT, WLD, and color-histogram for multi-spectral data). Then the feature space is analyzed and projected adaptively in 3D space, i.e., the CAVE, jointly with multi-modal rendering of the images and their content. The users then are able to interact with the content of the images to explore, analyze, manipulate, and provide feedback for the information learning algorithms and navigation systems.

3.1 CAVE

Images and features are represented in an Immersive environment, so called, CAVE. Physically, the CAVE consists of four

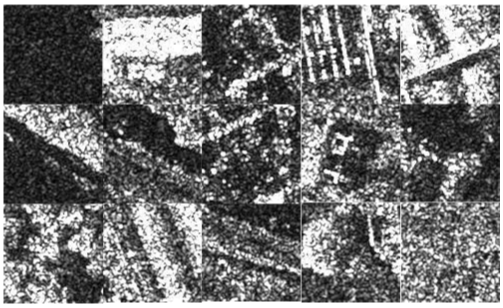


Figure 2: Collection of Synthetic Aperture Radar (SAR) images.

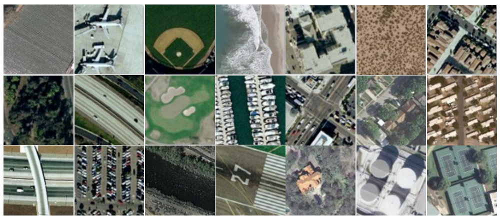


Figure 3: UCMerced-LandUse dataset of multi-spectral images.

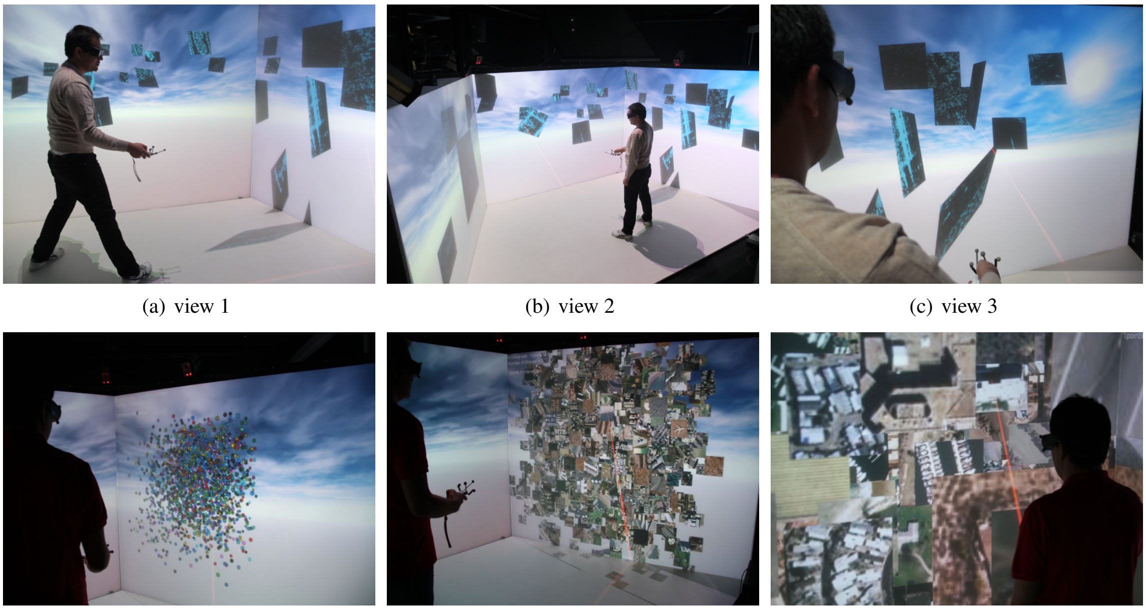
walls as displays and a three-layer cluster of PCs, Figure 4. The first layer collects the motion and control signals, then sends them to the middle layer. Middle layer is responsible for transmitting the control signals to the other PCs. The third layer comprises four PCs which render and display the scene on the walls. In this virtual environment, users are able to move and observe the data in different views. Furthermore, users can manually cluster the data and change the position of features. In this context, the synergy of HMC and information retrieval becomes an interdisciplinary approach in automating EO data analysis.

4 EXPERIMENTS

For demonstration of our approach, we visualize the feature space, built by different feature descriptors (e.g., SIFT, WLD, color-histogram), for the SAR and the multi-spectral images. In order to provide better perception of the structure of the data, the corresponding images to the feature points are shown in the feature space.
Observations shows that navigation and exploring the feature space in 3D virtual environment as well as manipulating the structure of the feature points, facilitate data understanding. There are some images from different classes that are represented very close to each other. Looking to the data from different views make it possible to find out the relations and sense of similarity of these images.
Figure 5 shows some photos of the CAVE provided from different views. The user wear a special glasses to percept the 3D view of the scene. There are some small balls mounted on the glasses which used for tracing the user's motion into the CAVE, (i.e., marker-based tracing). Moreover, the user hold a Wii controller in his hand which is used to navigate into the feature space and manipulate the structure of the points.

5 CONCLUSIONS AND FUTURE WORK

In this paper we focus on Immersive Visual Information Mining which is a visualization-based technique to understand the structure of the information provided by collection of images. In our approach we introduce an immersive 3D virtual environment, so



(a) view 1 (b) view 2 (c) view 3

(d) view 4 (e) view 5 (f) view 6

Figure 5: Sample views of visualization of SAR and multi-spectral images in the CAVE. The user inside the CAVE interacts with images for better understanding of the data.

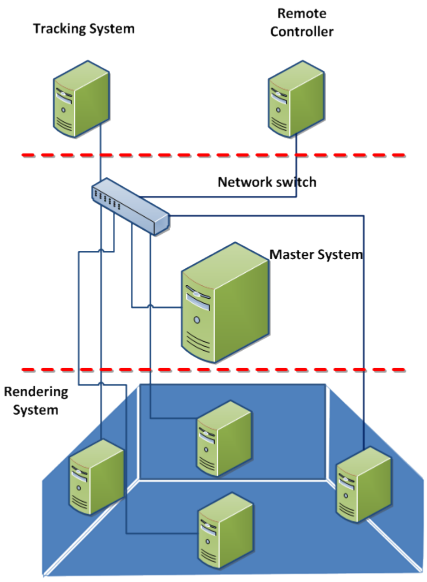


Figure 4: The physical diagram of immersive visualization. The visualization system is composed of three layers with different responsibility. The first layer comprises two PCs for motion capturing (tracking) and controlling. A master PC in the middle layer for the synchronization, and finally four PCs render the scene for each wall of the CAVE. All PCs are connected together via an Ethernet network

called CAVE. Visualizing the feature space in CAVE allows the users to explore, and manipulate the structure of feature points. Moreover, they can provide feedback which can be used by learning algorithms to improve the learning process (e.g, active learning). To demonstrate the applicability of our system, we visualize the feature space, built by the extracted feature descriptors form a SAR and a multi-spectral datasets for different feature descriptors.

ACKNOWLEDGEMENTS

The authors would like to thank Munich Aerospace Faculty for their support and German Aerospace Center (DLR) for providing the SAR image dataset.

REFERENCES

Belkin, M. and Niyogi, P., 2003. Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput. 15(6), pp. 1373–1396.

Chen, J., Shan, S., Zhao, G., Chen, X., Gao, W. and Pietikainen, M., 2008. A robust descriptor based on weber's law. In: Computer Vision and Pattern Recognition (CVPR), pp. 1 –7.

Hinton, G. and Roweis, S., 2002. Stochastic neighbor embedding. In: Advances in Neural Information Processing Systems 15, MIT Press, pp. 833–840.

Lowe, D., 1999. Object recognition from local scale-invariant features. In: Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, Vol. 2, pp. 1150 – 1157.

van de Sande, K., Gevers, T. and Snoek, C., 2010. Evaluating color descriptors for object and scene recognition. Pattern Analysis and Machine Intelligence, IEEE Transactions on 32, pp. 1582 –1596.

Yang, Y. and Newsam, S., 2010. Bag-of-visual-words and spatial extensions for land-use classification. In: Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS '10, ACM, pp. 270–279.