Interactive clustering for SAR image understanding

Mohammadreza Babaee, Munich Aerospace Faculty, Institute for Human-Machine Communication, Technische Universität München, reza.babaee@tum.de, Germany
Reza Bahmanyar, Munich Aerospace Faculty, German Aerospace Center (DLR), gholamreza.bahmanyar@dlr.de, Germany
Gerhard Rigoll, Munich Aerospace Faculty, Institute for Human-Machine Communication, Technische Universität München, rigoll@tum.de, Germany
Mihai Datcu, Munich Aerospace Faculty, German Aerospace Center (DLR), mihai.datcu@dlr.de, Germany

Abstract

The increasing amount of high resolution Earth Observation (EO) data during recent years, has brought the content analysis of the provided data into the spotlight. Most of the current content analysis is based on unsupervised methods (e.g., clustering). However, the structure discovered by these methods is not necessarily human understandable. Moreover, they require some prior knowledge of the structure of the data for initialization.
In this paper, we propose an interactive method to discover the semantic structure behind SAR image collections. Thus, we use a modified version of k-means, namely weight-balanced k-means, to perform clustering on the given images. The interaction mechanism allows users to provide the clustering method with relevant knowledge about the structure of the data. Experimental results demonstrate that the structure discovered by the proposed interactive method is closer to human understanding of the data.

1 Introduction

High resolution *Synthetic Aperture Radar* (SAR) images provide a large volume of detailed information of land coverage. To understand the contents of the images, the semantic structure of the data should be known. There has been much research done in recent years, using supervised and unsupervised techniques, to deal with this problem. While supervised methods suffer from the lack of enough prior annotated data (esp. in SAR data scenario), unsupervised techniques require prior knowledge based on the structure of the data to set some parameters (e.g., number of clusters in clustering methods) [1]. Moreover, the data structure discovered by unsupervised methods is not necessarily understandable for human user due to the well-known *semantic gap* [2].
In this paper, we propose an interactive technique to discover the semantic structure of SAR image collections. The interaction mechanism in this method allows users to provide unsupervised methods (e.g., clustering) with relevant prior knowledge. While the relevant prior knowledge leads to a shortening of the semantic gap, the unsupervised nature of clustering helps to avoid the need for prior annotated data.
In this work, we first represent each SAR image by a vector of its primitive features, the so-called *feature descriptors* (e.g., Gabor and WLD). This representation maps all the images into a high dimensional vector space, so called feature space, as feature points. Then a clustering technique is applied to the feature space to discover the structure behind the feature points.

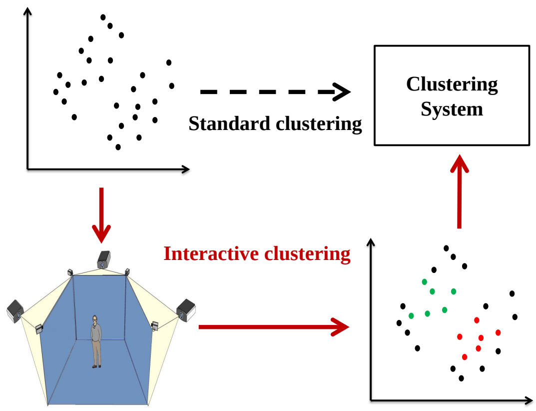


Figure 1: Our proposed interactive clustering method allows users to provide more relevant prior knowledge by exploring the entire dataset.

In our work, we use *weight-balanced k-means* [3], a modified version of the widely used *k-means* clustering technique [4]. This clustering method allows to weight different feature points individually. In our method, these weights are assigned based on the feedback provided by users. To provide an interface for users to interact with the data, we visualize the feature space using a 3D virtual environment, the so-called CAVE [1].
In our proposed method, images are visualized in 3D space based on the position of their corresponding feature descriptor in the feature space. CAVE allows users to explore the entire image collection in the feature space and group samples of related images. In the clustering step, the centers of these groups are taken as the initial-

[1]CAVE is installed in Institute for Human-Machine Communication, Technische Universitaẗ Muünchen, Germany.

ization for k-means. Moreover, the points touched by the user are assigned higher weight than the others in the optimization procedure. In other words, these points play a more important role in determining the positions of the clusters, Figure 1.

The rest of the paper is organized as follows. Section 2 introduces the weight-balanced k-means. In Section 3 we briefly describe our CAVE. In Section 4, we propose our interactive clustering method. Section 5 shows some experimental results. We conclude this paper in Section 6.

2 Weight-balanced k-means

In data mining, k-means is a widely used clustering technique to partition a set of d-dimensional real valued points, $X =: \{x_1, x_2, ..., x_n\}$, into k clusters, $C := (C_1, C_2, ..., C_k)$, in such a way that the sum of within-cluster distances is minimized,

$$\underset{C}{\operatorname{argmin}} \sum_{i=1}^{k} \sum_{x_j \in C_i} \|x_j - \mu_i\|^2, \tag{1}$$

where μ_i is the mean of the cluster C_i.

In standard k-means, all of the points have similar weight in determining the clusters. However, recently modified k-means, namely *weight-balanced k-means* [3], allows to weight each data point individually as well as defining upper and lower bounds for the cluster sizes. While weighting is applied by vector $\Omega := (\omega_1, \omega_2, ..., \omega_n) \in \mathbb{R}^n$, each element y_{ij} in assignment vector $y := (y_{11}, ..., y_{1n}, ..., y_{k1}, ..., y_{kn}) \in [0, 1]^{kn}$ with $\sum_{i=1}^{k} y_{ij} = 1$ assigns a fraction of the weight ω_j of the point x_j to the cluster C_i. Thus, the final clustering is achieved by,

$$\sum_{i=1}^{k} \sum_{j=1}^{n} y_{ij} \omega_j \|x_j - s_i\|^2, \tag{2}$$

where s_i is the center of gravity of the cluster C_i,

$$s_i = \frac{1}{\sum_{j=1}^{n} y_{ij} \omega_j} \sum_{j=1}^{n} y_{ij} \omega_j x_j. \tag{3}$$

3 Cave Automatic Virtual Environment

Cave Automated Virtual Environment (CAVE) is a a tool designed for visualization and simulation of real scenes. In our research, we adapt this tool for visualization of feature spaces generated from collections of EO images. CAVE allows users not only to explore the entire dataset in a 3D environment , but also to interact with the data in the feature space. Figure 2 shows a schematic view of the CAVE.

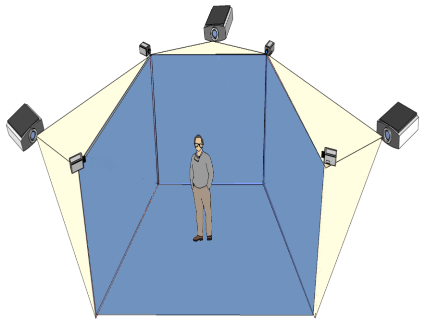


Figure 2: Schematic view of the CAVE. Four room-sized walls make a cube such that the user with stereoscopic glasses sees the 3D graphical elements.

CAVE is composed of four room-sized walls connected together generating a cube. There are four projectors directed to the walls in such a way that each wall performs as a display screen. The user wears stereoscopic LCD shutter glasses inside the CAVE to see the 3D scene. In order to provide capability of tracking the position of the user inside the CAVE, a motion capture system is performed. This system is composed of several infrared cameras mounted on top of the walls. They capture the motion of the user's head by tracking the markers mounted on top of the glasses using the infrared cameras.

4 Interactive k-means clustering

In this section we describe our proposed interactive k-means clustering. Although k-means is a widely used unsupervised method in data mining scenarios, the discovered clusters by k-means are not necessarily semantically understandable for users. Moreover, it requires some prior knowledge of data such as the numbers and the initial position of the clusters.

In this paper, we deal with these problems by introducing an interactive method which allows users to explore the entire dataset to provide k-means with more relevant prior knowledge. In this method, images are represented by primitive feature descriptors (e.g., Gabor and WLD). Then, using our CAVE, we visualize all the images in 3D virtual environment based on the position of their corresponding feature descriptors in feature space. Figure 3 shows examples of visualization of a SAR image collection in our CAVE.

The idea is that the user explore the images in feature space; then, he selects and groups some of representative related images together. These groups are considered as the initial clusters in k-means and also indicate the number of clusters. Since the initial clusters are selected by the user based on their semantic contents, the resulted clusters are more semantically understandable for users, as experimental results indicate. In the next step, we use a modified version of k-means, namely weight-balanced k-means, which allows to apply the interaction of the user by weighting each individual data point using vector $\Omega := (\omega_1, \omega_2, ..., \omega_n)$. In this way, we assign higher weights to the points touched by the user. It means, these

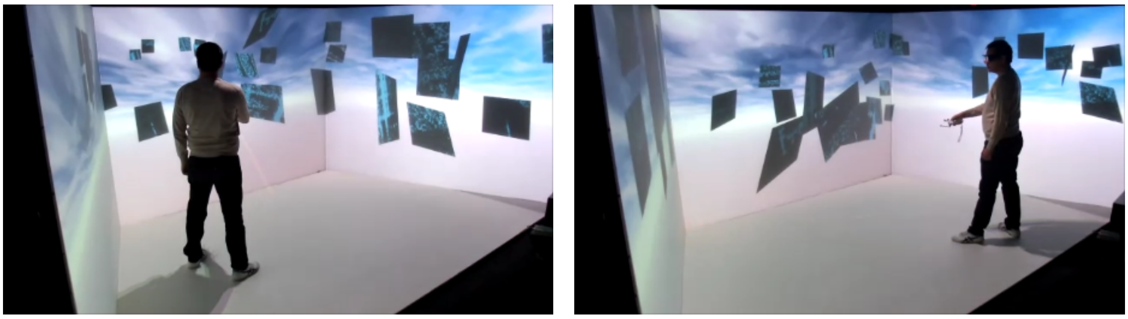


Figure 3: Visualization of a collection of SAR images in our CAVE. User can navigate through the image cloud and explore the entire data set.

points play a more important role than the others in determining the clusters. Since in the current paper, we do not assess the influence of the sizes of the clusters, the assignment vector $y := (y_{11}, ..., y_{1n}, ..., y_{k1}, ..., y_{kn})$ is defined in such a way that all of the points are uniformly probable to be assigned to all of the clusters.

5 Experiments and discussions

5.1 Experimental setups

In this section we compare the performance of the standard k-means with the performance of our proposed interactive method. In our experiments, we use a collection of 1230 SAR images of size 160×160, Figure 4. The images are grouped in seven non-equal size classes (forest (198 images) , water (210 images), medium density urban area (204 images), forest + water (114 images), roads (67 images), high density urban area (279 images), and urban area + roads (158 images))[2].

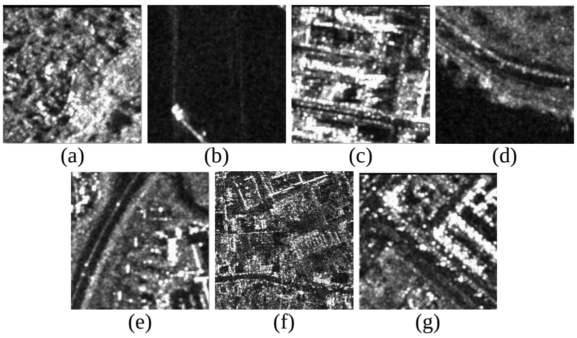


Figure 4: EO data set of 1230 SAR images grouped in seven non-equal size classes. (a) forest, (b) water, (c) medium density urban area, (d) forest + water, (e) roads, (f) high density urban area, and (g) urban area + roads).

In order to perform clustering on the data, we represent the images by a discrete set of their primitive features. In this paper, we study the structure of the feature space generated by two different feature descriptors, e.g., Gabor [5] and WLD [6]. The feature descriptors are extracted globally for each image, where the vector sizes are 36 for Gabor and 144 for WLD.

5.2 Evaluation methods

In our work, we used two evaluation metrics, namely *Accuracy* (AC) and *normalized Mutual Information* (nMI) [7]. The accuracy is the percentage of data that is labeled correctly based on a priorly provided annotation.

Mutual Information is computed between two sets of clusters. Assume C and C' are the ground truth and the resulted clusters by our method, respectively, the Mutual Information is defined as:

$$MI(C, C') = \sum_{c_i \in C, c'_j \in C'} p(c_i, c'_j).log_2 \frac{p(c_i, c'_j)}{p(c_i)p(c'_j)}. \quad (4)$$

Where $p(c_i)$ and $p(c'_j)$ denote the probability of a data point belonging to cluster c_i and c'_j, respectively. $p(c_i, c'_j)$ is the joint probability which denotes the probability of the selected point belonging to the both clusters c_i and c'_j at the same time. Thus, the normalized MI is defined as:

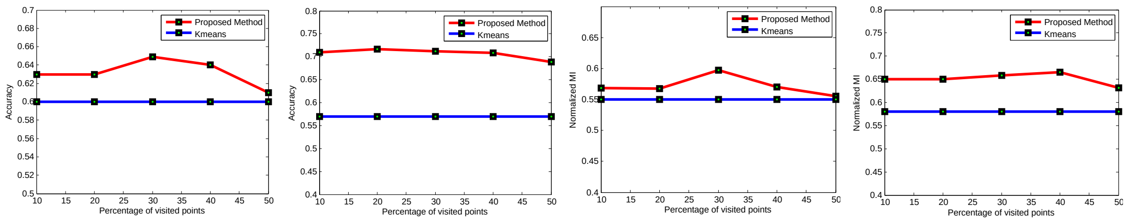
$$nMI(C, C') = \frac{MI(C, C')}{max(H(C), H(C'))}, \quad (5)$$

where $H(C)$ and $H(C')$ represents the entropies of C and C', respectively. The normalized mutual information is bounded between 0 and 1. It takes 1 when the two clusterings are identical, whereas it is 0 when the clusterings are independent.

5.3 Experimental results

In our completed experiments, we have assessed the influence of the percentage of the data touched by the user, in clustering results. In other words, the user only selected images which have been distinguished as informative regardless of their groupings. These images have been assigned weight value $\omega = 1$, whereas the rest have taken weight value $\omega = 0.5$.

[2]The images are collected from TeraSAR-X data by Shiyong Cui, Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Germany, shiyong.cui@dlr.de.



(a) MI for Gabor **(b)** AC for WLD **(c)** MI for Gabor **(d)** MI for WLD

Figure 5: Evaluation of the performance of the proposed interactive clustering in comparison to the performance of standard k-means. In the figures, the blue line is the result of standard k-means, where the red line indicate the results by our proposed method.

Experimental results demonstrate that using interactive clustering generally outperforms standard k-means (e.g., the nMI is $\simeq 0.65$ by interactive clustering, whereas k-means achieved $\simeq 0.55$). However, increasing the number of selected points does not necessarily increase the performance of clustering. As is shown in Figure 5, the performance of clustering increases until a certain point, then it decreases as the number of selected points increases. Because the more points are selected, the more points are assigned a high weight value 1. This makes our method to perform similar to standard k-means.

6 Conclusion

In this paper, we deal with SAR data analysis by proposing an interactive clustering method, where the interaction mechanism allows users to initialize the clustering method (e.g., k-means) with relevant knowledge about the structure of the data. As a clustering technique, weight-balanced k-means is performed, which allows to weight each data points individually in the feature space as well as controlling the size of the clusters. In the current paper, we use the users' feedback to determine the informative points and weight them higher than the others in clustering procedure. Experimental results demonstrate that the resulted clusters are closer to human understanding of the data.

References

[1] Junhua Li and Wenjun Chen, "Clustering synthetic aperture radar (sar) imagery using an automatic approach," *Canadian Journal of Remote Sensing*, vol. 33, no. 4, pp. 303–311, 2007.

[2] Reza Bahmanyar and Mihai Datcu, "Measuring the semantic gap based on a communication channel model," in *IEEE International Conference on Image Processing, ICIP 2013*, 2013.

[3] Steffen Borgwardt, Andreas Brieden, and Peter Gritzmann, "A balanced k-means algorithm for weighted point sets," *arXiv:1308.4004v1*, 2013.

[4] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, L. M. Le Cam and J. Neyman, Eds. 1967, vol. 1, pp. 281–297, University of California Press.

[5] B.S. Manjunath and W.Y. Ma, "Texture features for browsing and retrieval of image data," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 18, no. 8, pp. 837–842, 1996.

[6] Jie Chen, Shiguang Shan, Chu He, Guoying Zhao, M. Pietikainen, Xilin Chen, and Wen Gao, "Wld: A robust local image descriptor," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 9, pp. 1705–1720, 2010.

[7] Wei Xu, Xin Liu, and Yihong Gong, "Document clustering based on non-negative matrix factorization," in *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, New York, NY, USA, 2003, SIGIR '03, pp. 267–273, ACM.