

Earth Observation Images Information Mining for Humanitarian Crisis Management Related Applications

Daniela Faur¹, Inge Gavat¹, Mihai Datcu²

¹ University Politehnica Bucharest, 313 Splaiul Independentei 060042, Tel: +40214024623,
Fax: +4021 4024957 danielafaur@gmail.com, igavat@lpsv.pub.ro

² German Aerospace Center DLR, Oberpfaffenhofen, Germany, mihai.datcu@dlr.de

Abstract- Nowadays the Earth Observation sensors provide images containing detailed information relevant for applications related to hazard or security matters. Unfortunately, image information mining, its interpretation and transformation in products useful to the rescue or decision teams is still a laborious task, effectuated many times by visual inspection and manual annotations of the images, thus not appropriate to react in prerequisite time.

This paper presents a data analysis processing chain, which by interactive operations, enables analysis and interpretation of large volumes of images with high accuracy, flexibility and much faster than any existing methods.

I. Problem definition

Users in all domains require information to be focused, reliable, low cost and timely, and in a compatible format with the user's own activities. Earth observation data in general and images in particular are retrieved from archives based on attributes such as geographical location, acquisition time and type of the sensor, thus providing no insights into the image's actual information content.

In order to extract information experts interpret images using their own knowledge. In this scenario, the imagery or data sets does not match the customer's real need which is for information. The information extraction process is too complex, too expensive and too dependent on user conjecture to be applied systematically over an adequate number of scenes. There are domains that necessitate automatic and fast interpretation of the remote sensing data including here the field of humanitarian help.

In recent years, the number of complex humanitarian emergencies around the world has been steadily increasing [1]. Land mines left behind from wars worldwide are one of the century's main unsolved problems of war and remain the focus of humanitarian mine detection. Unlike the military demining, only requiring a safe corridor for the troops to move through, the humanitarian demining requires that the entire land area be free of mines, without any hazard to the environment. The detection of mined fields, compared with other scene classification tasks, has a very high complexity. The large number of models and the huge dimensionality of the feature space, mainly for hyperspectral data, make it difficult to select the appropriate data sets and the primitive features of the images by using classical methods.

We propose the use of a knowledge based information-mining system - KIM [3] and several related algorithms to demonstrate the effectiveness of a hierarchical data acquisition, interpretation and assessment services as well as to emphasize the benefits of remote sensing and geographical data in humanitarian crisis management. Two scenarios are envisaged: generation of a risk map for mined areas and refugee camp assessment.

II. Processing chain for remote sensing data analysis

Our goal, related to humanitarian applications is to find and delimited anomaly regions by combing the information extracted from the *knowledge based information mining system* - with the results of applying some information theory based algorithms [10][11] on the remote sensing images.

The goal of data analysis is to derive a symbolic description of image data, which is suitable for a given task and will often comprise objects detected in the scene along with their attributes and relations. Image analysis is the step allowing the extraction of intended feature indicators for risk assessments and provides useful layers for synthesized maps. Intended indicators are defined according to the considered problem. The specificity of the application calls for methods that can exploit context, spectral information, vegetation changes, shape and structural information for identifying direct and indirect indicators allowing the production of thematic maps for the considered Humanitarian Emergency applications. Fig.1 depicts the processing chain for data analysis.

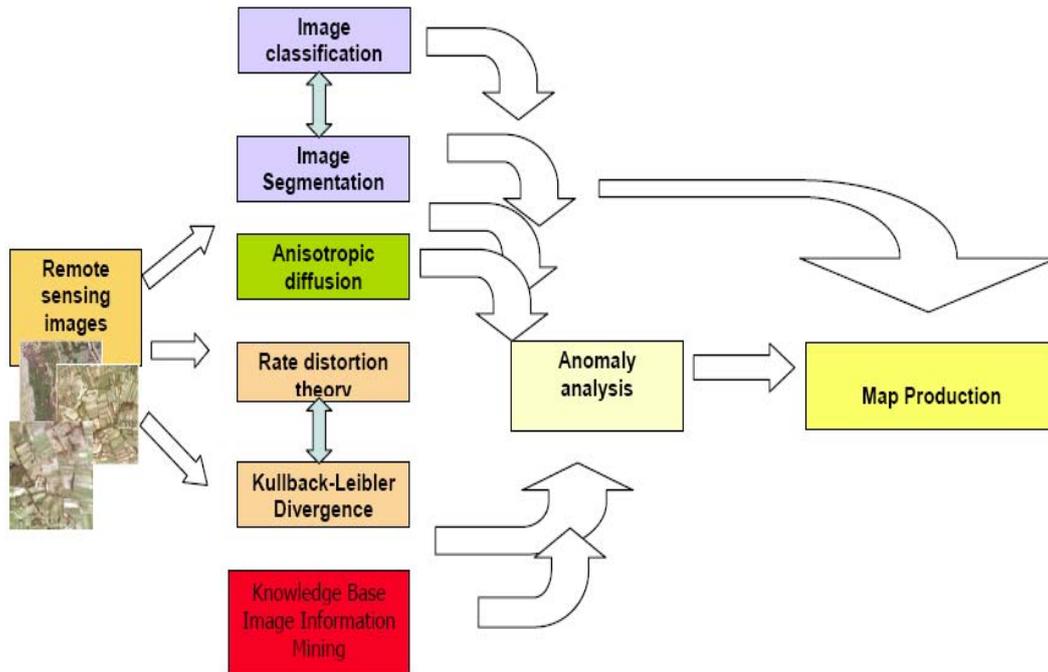


Figure.1 Processing chain for data analysis includes algorithms for image classification, segmentation and image information mining.

In order to exploit fully the synergy of human-computer, human centered concepts [3] are underlying the KIM system; the user guides the interactive learning process and the system continuously gives the operator relevance feedback about the performed training actions and searches the archive for relevant images.

A. Image classification and segmentation

Image segmentation or classification into a set of structures or objects is a fundamental aspect for understanding the nature of the observed scene [5]. Most of the work in the field is concentrating to solve the grouping of coherent or regular patterns in images. The objective is to merge connected components up to the whole scene represented by an image into more complex objects.

Diffusion process is useful for image processing and computer vision because it provides a convenient way of smoothing noisy data, of analyzing images at multiple scales. To overcome the limitations of linear methods leading to isotropic smoothing Perona and Malik proposed a nonlinear way to smooth an image. The idea is roughly to smooth out what has to be smoothed, the irrelevant, homogenous, regions and to enhance instead the boundaries.

B. Rate distortion theory for image complexity assessment

Rate distortion theory is one of the areas of information theory with important applications in multimodal signal processing. We propose an image characterization method based on rate distortion analysis in the feature space [8]; this space is coded using K-means clustering algorithm considering both spectral and texture features. Classification is done in view of an optimal digital representation of the signal, leading to a certain "compression" degree. Through compression, the processed signal is "distorted", rate distortion theory ensuring a balance between an acceptable distortion and a satisfactory

compression rate.

$$R(D) = \min_{p(y/x): E(X,Y)[d(x,y)] \leq D} I(X,Y) \quad (1)$$

Then, we analyse the rate distortion curve in order to detect anomaly regions. The point at which the distortion curve levels off gives the optimal number of classes. The number of clusters is proportional with image information content, so finding the number of clusters also gives us the possibility detect complexity regions.

C. Kullback Leibler divergence for anomaly detection

The association between image space and class space can further be used to measure the complexity of image's regions. We measure the information between image space and class space based on Kullback-Leibler divergence.

$$D(p,q) = \sum_{j,i} p(\omega_j^s / I_k) p(\omega_i^t / I_k) \log \frac{p(\omega_j^s / I_k) p(\omega_i^t / I_k)}{q(\omega_j^s) q(\omega_i^t)} \quad (2)$$

An unsupervised clustering algorithm like K-means was applied for each feature space, spectral and texture, across all image's interest regions. Then, we use posterior probabilities of signal classes ω_j^s and ω_i^t given a certain region I_k in order to compute $D(p,q)$, a "distance measure" between two probabilities distributions. The results denote how much a single region I_k is a mixture of the entire archive information content.

D. Knowledge based image information mining system KIM

Automatic interpretation of remote sensing images and the growing interest for image information mining and query by image content from large remote sensing images archives rely on the ability and robustness of information extraction from the observed data. One can focus on the modern Bayesian [3] way of thinking and introduce a pragmatic approach to extract structural information from remote sensing images by selecting those prior models which best explain the structure within an image.

The KIM system [6] is based on human centered concepts in order to fully exploit the synergy of human-computer: the user guides the interactive learning process and the system continuously gives the operator relevance feedback about the performed training actions and searches the archive for relevant images.

III. Anomaly detection- Case Studies

Anomaly detection in Earth Observation images refers to the detection of irregularity in the scene that appears unlikely according to a probabilistic or physical model of the scene, for example ammunition elements in a scene which is dominated by vegetation and soil. The image segmentation or classification into a set of structures or objects is a fundamental aspect for understanding the nature of the observed scene.

Image's anomaly regions can be regarded as very complex regions, different from the other regions or 'regular' patterns composing the image. An image can be partitioned into different objects, with different attributes depending on the observer interest and how carefully is observed. Similar techniques with the one used to locate words, regarded as molecules, formed by combining atoms, are proposed to describe images based on their information content.

We intent to present several case studies relevant to detection in classified images of the infrastructure and/or natural land cover in relation to humanitarian crisis management.

A. Environmental monitoring of a refugee camp.

Natural catastrophes or different conflicts continually force people to migrate against their will. A better understanding of the characteristics of refugee populations and of the population dynamics of these situations is vital.

Since for these areas only very little information about the existing infrastructure is available, the aim was to use remote sensing data for studying the situation. The proposed algorithms enabled us to select refugee camps and other information details in an image and to retrieve other image from the archive with similar content. Although the diversity of structures and object in the database did not always allow us to discriminate them clearly, KIM provides a gallery of the most relevant images and constraints the search. Fig. 2 presents this case study relevant to detection in classified images of the infrastructure and/or natural land cover in relation to displaced population.

Using classification our purpose is to identify and evaluate the water resources like rivers, lakes, the farming areas, energy sources like forests and finally the means to insure humanitarian help e.g. landing perspective for airplane or helicopters.

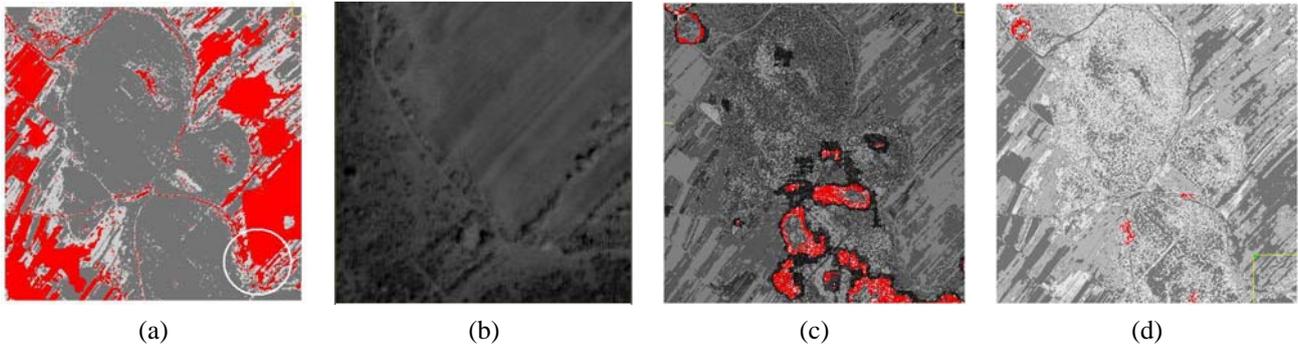


Figure 2 The result of user interactive and knowledge driven task solving the detection of areas with risk to be mined using an IKONOS image. (a) In red, area of high risk and in (b) image detail for confirmation. (c) Trees detection. (d) Refugee camps detection.

B. Generation of a risk map for mined areas

Land mines left behind from wars worldwide are one of the century's main unsolved problems of war and remain the focus of humanitarian mine detection. Unlike the military demining, only requiring a safe corridor for the troops to move through, the humanitarian demining requires that the entire land area be free of mines, without any hazard to the environment. The detection of mined fields, compared with other scene classification tasks, has a very high complexity. The large number of models and the huge dimensionality of the feature space, mainly for hyperspectral data, make it difficult to select the appropriate data sets and the primitive features of the images by using classical methods.

The expert in the demining action uses a special set of objects, structures, and relations among them to express a degree of belief in the fact that a field is mined or not. In this case, the following properties are relevant: bare terrain, agriculture fields in use or not, roads used and abandoned.

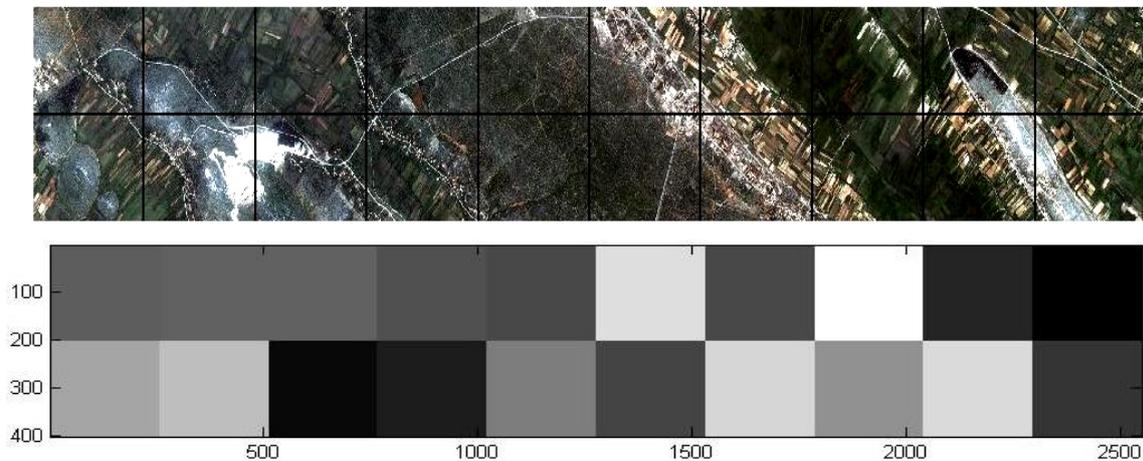


Figure. 3 Result of Kullback Leibler divergence on subimages (size 200x 200) IKONOS. White represents most different regions to entire image information content, black most similar.

To increase the degree of goodness of classification of mine fields, knowledge on placing mines in a field should also be used. Therefore the following categories could be relevant: roads, bridges, power lines, river site acces, etc.

Further, for every step of the processing chain we describe and present the results of the proposed algorithms. We have used an IKONS image, 5 Geotiff bands in visible, near infrared, and panchromatic, size of 400×2550 , and 4 m resolution, covering the land of Vrankovici, Croatia.

Figure 3 depicts the results of the Kullback Leibler algorithm denoting how much a n image region is similar to the entire image information content.

The results of the mean shift [7] and watershed segmentation are shown in Fig.4.

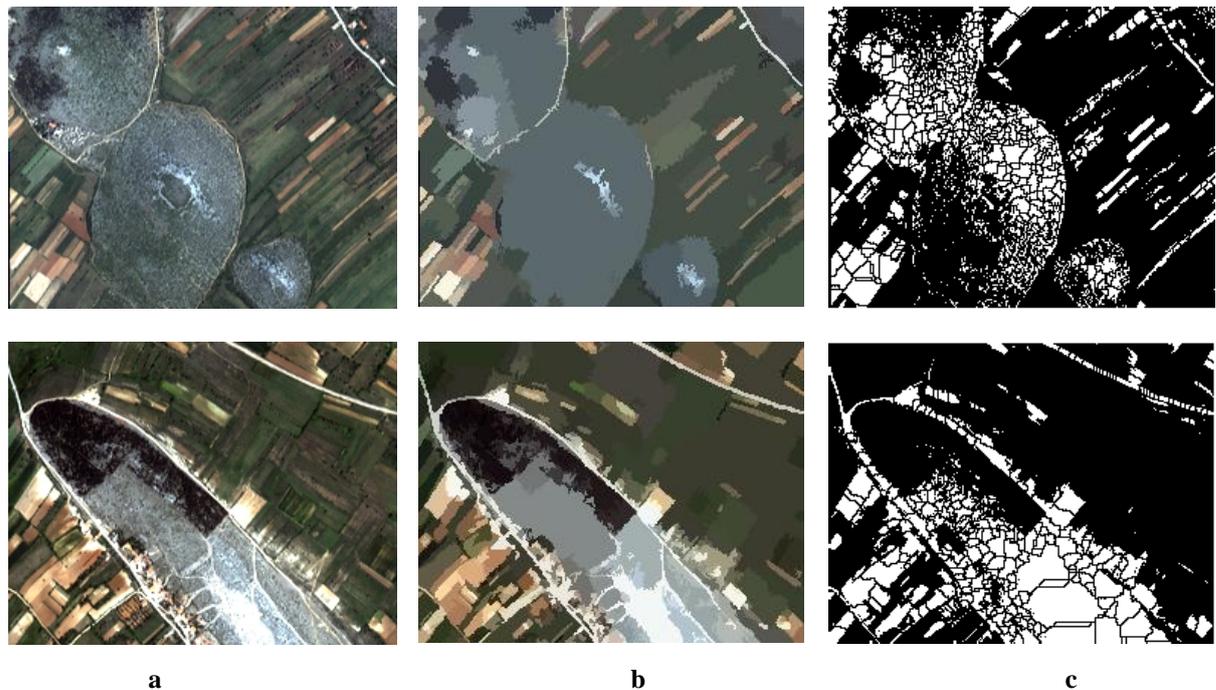


Figure 4. Image segmentation. a) Original image; b) Mean shift segmentation
c) Watershed segmentation

III. Final aspects and conclusions

Land cover classification from multi-sensor imagery that provides first-hand up-to-date information on land use represents the first step of scene analysis. For high resolution airborne imagery, temporary refugee camps can be identified. Multitemporal images acquired before and after the conflict can provide information on war activities, e.g. trenches, cradles created by explosives, destroyed houses and infrastructures, etc. Post-conflict multitemporal acquisition can update the situation of reconstruction process. Indicators can be derived to represent level of reconstruction: new constructions of road, power lines, housing, cultivated land, etc.

The detection of mined fields, compared with other scene classification tasks, has a very high complexity. Anomaly detection in Earth Observation images refers to the detection of irregularity in the scene that appears unlikely according to a probabilistic or physical model of the scene. Image's anomaly regions can be regarded as very complex regions, different from the other regions or "regular" patterns composing the image. An image can be partitioned into different objects, with different attributes depending on the observer interest and how carefully is observed. Similar techniques with the one used to locate words, regarded as molecules, formed by combining atoms, are proposed to describe images based on their information content.

Acknowledgments

This work was supported by the European Project FP6-STREAM in cooperation with IMEC.

References

- [1] <http://en.wikipedia.org/wiki/Demining>. Demining. Accessed on 27.09.2006.
- [2] A. Touzani and J.G. Postaire. "Clustering by mode boundary detection". *Pattern Recognition Letters*, 9:1-12, 1996.
- [3] M., Datcu, K., Seidel – "Human Centered Concepts for Exploration and Understanding of Earth Observation Images", *IEEE Trans. On Geoscience and Remote Sensing*, vol. 43, nr. 3, march 2005.
- [4] M. Herbin, N. Bonnet, and P. Vautrot. "A clustering method based on the estimation of the probability density function and on the skeleton by the influence zone". *Pattern Recognition Letters*, 17:1141-1150, 1996.
- [5] R.Wilson and M. Spann. "A new approach to clustering." *Pattern Recognition*, 23:1413-1425, 1990.
- [6] K. Fukunaga and L.D. Hostetler. "The estimation of the gradient of a density function, with application in pattern recognition". *IEEE Transaction on Information Theory*, 21:32-40, 1975.
- [7] D. Comaniciu, P. Meer, Mean shift: "A robust approach toward feature space analysis". *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol.24, no.5, May, 2002
- [8] Tasto, M., Wintz, P. "A bound on the rate-distortion function and application to images." *IEEE Transactions on Information Theory*, , vol.18, issue 1, Page(s): 150 -159, Jan 1972.
- [9] M. Datcu, H. Daschiel, A. Pelizzari, M. Quartulli, A. Galoppo, A. Colapocchioni, M. Pastori, K. Seidel, P.G. Marchetti, and S.D'Elia."Information mining in remote sensing image archive - part a: System concepts." *IEEE Transaction on Geoscience and Remote sensing*, 41(12):2923-2936, 2003.
- [10] D. Faur, I. Gavati, and M. Datcu. Mutual information based for image content characterization, vol. 4177, chapter *Current Topics in Artificial Intelligence*, pages 342{349. Springer Berlin/Heidelberg, 2006.
- [11] C. Iancu, I. Gavati, and M. Datcu. Image disorder characterization based on rate distortion, vol. 4177, chapter *Current Topics in Artificial Intelligence*, Springer Berlin/Heidelberg, 2006.