IMAGE INFORMATION MINING METHODS FOR EXPLORING AND UNDERSTANDING HIGH RESOLUTION IMAGES

Mihai Datcu[1,2], *Gottfried Schwarz*[1]

[1)] German Aerospace Center DLR

[2)] Télécom ParisTech

The users of Earth Observation (EO) images in all application domains require information to be focused, reliable, low cost and timely extracted, and also to be presented in a format compatible with their own needs. The EO information content extraction, particularly in the case of High Resolution satellite or airborne images, is mainly based on human expert knowledge and manual/visual interpretation. Thus, the information extraction process is too complex, too slow, too expensive and too dependent on interpreter subjectivity to be applied systematically over a large number of images.

Among the particularities of the High Resolution images (0.5-2m), and what makes their interpretation complicated, are: the 3D scene structures become visible, thus increasing the complexity; a 3D scene reconstruction requires multiple acquisitions, this is laborious and many times satisfactory only for natural landscapes; the image structure understanding is strongly dependent on the spatial contextual information and also on the interpretation context; understanding based on change analysis is hindered by the perspective and illumination effects.

Thus, a series of methods have been developed recently to help users to interpret images interactively, semi-automatically, and time effectively, based on 2D analysis methods. As an alternative to the 3D information, the methods are using semantic learning and coding based on 2D image analysis. These are interactive paradigms to decompose images into meaningful objects or structures using the user interaction for indicating a limited number of examples and generalizing the semantic annotation for large images, or image archives [3].

This is equivalent to link a whole image to a set of keywords. Among the methods used in [1], a Bayesian framework is used to assign satellite images to different classes according to the structures they contain. In [5], a training set of natural images with annotations is applied to produce a joint distribution linking images and words. The words which have a high posterior probability are associated with the image. However, this approach does not indicate which image structure gave rise to which word [6]. Thus, the co-occurrence of words with natural image tiles created on a regular grid is exploited in [7]. A segmentation based approach is applied in [8] to annotate multispectral satellite images. Another technique is developed in the KIM system [9], which is based on interactive learning and uses a 5-level hierarchical Bayesian model to semantically interpret remote sensing image data. An important and efficient class of methods and tools, initially developed for statistical text modeling in large document collections, have recently been used with success for natural image annotation, but also for object recognition, scene classification, and image retrieval [10], [11].

The use of a recent pure-spectral rule-based fully automatic classifier to define the basic 'vocabulary' provided a hybrid method to automatically understand and describe semantic rules that link existent data according to different specifications (e.g., CLC 2000) with end-results of unsupervised computer information mining methods. Following an agreement between the learning model and the cartographic scale implied, a Latent Dirichlet Allocation model (LDA) was exploited to map heterogeneous pixels with similar intermediate-level. The method supports automatic mapping and extracts the map generation rules [12].

Our new method is an image information mining concept, requiring both database and visual capabilities trying to bridge the gap between these systems. The theory of databases did not deal with multidimensional pictorial structures. Computer vision and pattern recognition systems do not provide database query capabilities. Many existing image databases have been created by using some extensions of a relational data model. Almost all queries of a database have a semantic nature since people use and construct discrete units of linguistic expressions. We also have to consider that the volume and multimodality of our data is growing; data and information is of spatio-temporal nature and unstructured; users want to have the knowledge, thus interactive operation is the only way to proceed; exploration is the predominant mode of interaction, context is critical and relevant, users are interested in information and

knowledge independent of conjecture, information must be obtained from the data, databases and search engines were not designed to provide contents, and visualization is very important.

REFERENCES

[1] S. Aksoy, K. Koperski, C. Tusk, G. Marchisio, J. C. Tilton, "Learning Bayesian Classifiers for Scene Classification with a Visual Grammar", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 43, No. 3, pp. 581-589, 2005

[2] K. Barnard, P. Duygulu, N. de Freitas, D. Forsyth, D. Blei, M. I. Jordan, "Matching Words and Pictures", *Journal of Machine Learning Research*, Vol. 3, pp 1107-1135, 2003

[3] M. Datcu, K. Seidel, "Human Centered Concepts for Exploration and Understanding of Earth Observation Images", IEEE *Transactions on Geoscience and Remote Sensing*, Vol. 43, No. 3, pp. 601-609, 2005

[4] L. Fei-Fei, R. Fergus and P. Perona , "One-Shot Learning of Object Categories", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, Issue 4, pp. 594-611, 2006

[5] J. Jeon, V. Lavrenko and R. Manmatha, "Automatic image annotation and retrieval using cross-media relevance models", in *Proc. of the Annual ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 119-126, 2003

[6] K. Barnard, P. Duygulu, D. Forsyth, N. de Freitas, D. M. Blei and M. I. Jordan, "Matching words and pictures", *Journal of Machine Learning Research*, Vol. 3, pp. 1107-1135, 2003

[7] Y. Mori, H. Takahashi and R. Oka, "Image-to-word transformation based on dividing and vector quantizing images with words", in *Proc. of the First International Workshop on Multimedia Intelligent Storage and Retrieval Management*, 1999

[8] T. Kusaka, H. Hegawa, and Y. Kawata, "Classification of the SPOT image using spectral and spatial features of primitive regions that have nearly uniform color", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 28, No. 4, 1990

[9] M. Datcu, H. Daschiel, A. Pelizzari, M. Quartulli, A. Galoppo, A. Colapicchioni, M. Pastori, K. Seidel, P.G. Marchetti and S. D'Elia, "Information mining in remote sensing image archives: system concepts", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 12, pp. 2923-2936, 2003

[10] G. Wilkinson, "Results and implications of a study of fifteen years of satellite image classification experiments", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 43, No. 3, 2005

[11] J. Sivic, B. Russell, A. A. Efros, A. Zisserman and B. Freeman, "Discovering objects and their location in images", in *Proc. of the 2005 International Conference on Computer Vision ICCV*, October 2005

[12] D. Bratasanu, I. Nedelcu, M. Datcu, “Knowledge acquisition and semantic rules discovery for CORINE LAND COVER mapping Latent Dirichlet Allocation”. *ESA-EUSC-JRC 2009* , 03.-05. Nov. 2009, Torrejon, Spain.