

INTERIMAGE: AN OPEN SOURCE PLATFORM FOR AUTOMATIC IMAGE INTERPRETATION

¹GILSON ALEXANDRE OSTWALD PEDRO DA COSTA

²CAROLINA MOUTINHO DUQUE DE PINHO

¹RAUL QUEIROZ FEITOSA

²CLÁUDIA MARIA DE ALMEIDA

²HERMANN JOHANN HEINRICH KUX

²LEILA MARIA GARCIA FONSECA

¹DARIO AUGUSTO BORGES OLIVEIRA

¹Pontifícia Universidade Católica do Rio de Janeiro - PUC-RIO

{gilson,raul,dario}@ele.puc-rio.br

²Instituto Nacional de Pesquisas Espaciais – INPE

{carolina, almeida, hermann@dsr.inpe.br, leila@dpi.inpe.br }

ABSTRACT – This paper introduces a new open source knowledge-based platform for automatic image interpretation, called InterImage. The architecture, main features as well as an overview of the interpretation strategy implemented in InterImage is presented. The paper also reports an experiment in which an application built and tested upon a commercial software (eCognition) is translated into InterImage. No change was introduced in the original knowledge model itself. Only the knowledge formulation has been changed in order to comply with InterImage's knowledge representation structures. The results showed practically identical classifications obtained with both platforms.

1 INTRODUCTION

Remote sensing technology delivers the most important subsidies for the identification and monitoring of land cover changes on the surface of the Earth, effectively supporting the investigation of the interactions between the environment and agricultural, environmental and urban planning activities (Ehles, 2002).

Presently, however, the lack of efficient automatic image interpretation tools makes it difficult to achieve the goals of many land cover monitoring applications. The large amount of time spent from the acquisition of an image to its classification results in insufficient time to substantiate critical decisions that may avoid or diminish the effects of environmental degradation or unplanned urban expansion (Rego, 2003).

Currently most remote sensing data analysis techniques require intense human intervention. The commercially available softwares for image interpretation usually deliver incomplete and fragmented results, which require careful scrutiny by a human specialist for the identification and rectification of the inconsistencies produced by the conventional image analysis algorithms (Bükner, 2001). There is, consequently, a strong demand for the development of robust techniques for automatic information extraction and interpretation of remote sensing data (Blaschke, 2000) (Carrion, 2002).

A rather successful approach for automatic image interpretation is based on the explicit modeling, on a high level computational environment, of the human interpreter's knowledge concerning the interpretation problem (McKeown, 1985; Matsuyama, 1990; Clement, 1993; Sagerer, 1997; Liedtke, 1997; Bükner, 2001; and Schiewe, 2001). In this approach human expert's knowledge is organized in a knowledge base (Graham, 97), to be used as an input of automated interpretation processes, enhancing the productivity and accuracy and reducing the subjectivity of the interpretation process.

In this paper we present the architecture and features of a knowledge-based image interpretation system called InterImage, an open source software development initiative, leaded by the Computer Vision Lab of the Electrical Engineering Department, at the Catholic University of Rio de Janeiro (PUC-Rio) and the by the Brazilian National Space Research Institute (INPE).

We also present the results of an actual image interpretation experiment, the land cover classification of an area of the city of São José dos Campos. This interpretation application was based on the work presented in (Pinho, 2007), implemented in the software eCognition (Schiewe, 2001). The specific knowledge model designed in that research was translated into the knowledge structures of InterImage, and the interpretation

executed over a subset of the same input data. A comparison of the results of the interpretation with both softwares was then performed, showing almost identical results.

In the remainder of this paper we describe the basic characteristics of the InterImage (Section 2) and the interpretation strategy implemented by the system (Section 3). In Section 4 the land cover classification interpretation experiment is described, and in Section 5 the results of the experiment are presented. Finally, some conclusions and directions for future work are stated in Section 6.

2 SYSTEM DESCRIPTION

InterImage is based on the software GeoAIDA (Bükner, 2001), developed at the TNT Institute of the Leibniz Hannover University, Germany, and it inherited from that system the basic functional design, knowledge structures and control mechanisms. As a work in progress, a new graphical user interface, knowledge extraction functionality and image processing operators are planned to be included in InterImage in the near future.

In short, InterImage implements a specific image interpretation strategy. A strategy based and guided by a hierarchical description of the interpretation problem, structured in a semantic network.

The basis for interpretation of digital image data are results generated with image processing operators. In this context, an image processing operator is any operator that generates a labelled result image of a given image. Such image processing operators are denoted here as 'classifying operators'. They can fulfil threshold operations, texture-based or model-based methods and build the basis for the interpretation of a scene.

In most of the systems that use semantic networks for knowledge representation, only the leaf nodes of the network can be associated to image processing operators. The following grouping of the objects often produces a very high combinational diversity, because all objects extracted from the image have to be taken into account at the same time.

In InterImage, holistic operators (Liedtke, 1997) can be used to reduce the combinational diversity problem. Holistic operators aim at identifying specific types of objects independently of the identification of their structural components. They can be connected to any node of the semantic network, and their basic task is to divide a region into sub-regions, reducing the need of processing alternative interpretations. The structural interpretation of the sub-regions that follows can verify or disprove the holistic results.

Moreover, InterImage permits the integration of any of such classifying operators in the interpretation process. The problem that different operators can generate different information for the same region in the image is solved by the use of additional knowledge regarding the judgment of the competing interpretations. Furthermore, as different operators can process different types of data, the system permits the integrated analysis of image and GIS data from multiple sources.

3 INTERPRETATION STRATEGY

In InterImage explicit knowledge about the objects expected to be found in a scene is structured in a semantic network, defined by the user through the system's graphic user interface (GUI).

A semantic network, such as the one in Figure 1, contains nodes and edges, whereat nodes represent concepts and edges represent the relations between the concepts. In each concept node, information necessary for the analysis, such as the image processing operator specialised in the search of occurrences of the concept, is defined. During the analysis, guided by the semantic network, the system controls the execution of the operators and generates a network of instances, each instance defining a geographic region associated to a specific concept.

Interpretation of remote sensing data means to transform input data into a structural and pictorial description of the input data that represents the result of the analysis. In InterImage, the result of the interpretation contains a structural description of the result (an instance network) and thematic maps. The final and all intermediate results, in terms of region descriptions, are stored in XML format, and can be used for further external examinations.

The analysis process performed by InterImage has two steps: a bottom-up step and a top-down step. The top-down step is model driven and generates a network of hypothesis based on the semantic network. The grouping of hypothesis and their verification or falsification is a task of the data driven bottom-up analysis. The final instance network results from the bottom-up, data driven analysis.

In each node of the network the user defines the information necessary for the execution of each processing step, that is, the image processing (classifying) operator and respective parameters to be used in the top-down step (top-down operator), and the decision rules to be used in the bottom-up step.

The top-down operators have the task of separating regions into sub regions and of building hypotheses for the concepts of the semantic network, regions of the image associated to the concepts. The task is realized recursively from the upper to the lower nodes. For this purpose any (external) image processing operator, which

creates hypotheses for the sub regions can be used in the analysis process. The sub regions hypotheses can be defined by means of consistency measurements. If the contemplation of texture, for instance, allows only a few possible hypotheses for a particular region, no further investigation of other concept hypotheses is performed for that region.

When the top-down-analysis reaches the leaf nodes, the analysis turns from model driven interpretation to data driven interpretation (bottom-up). The decision rules for the bottom-up step are defined in a particular stack based language that provides functions for deciding between spatially concurrent hypotheses generated in the top-down step.

4 EXPERIMENT DESIGN

The experiment implemented in this work was designed to evaluate the performance of InterImage in a specific interpretation application, already implemented with another knowledge-based system, in this case, the system eCognition, distributed commercially by the company Definiens AG, that is virtually the only knowledge-based image interpretation system available. In doing so, we also wanted to investigate the possibility of transporting to InterImage a knowledge model implemented with the aid of eCognition.

The selected application was the one described at (Pinho, 2007), which produced an automatic land cover classification of an area of the city of São José dos Campos. The specific knowledge model designed for the application, described in terms of the knowledge structures defined in the eCognition system, was translated into the knowledge structures of InterImage. The same set of input data (a Quickbird image and GIS vector data with street limits) was used. In this experiment, however, only a 129,600 m² subset of the original Quickbird scene was interpreted.

As the focus of the experiment was on the translation of the eCognition knowledge model to InterImage, we decided to purge eventual differences in the initial region hypothesis and in the measurements of those regions from the analysis. That was done by using the same segments generated by the eCognition segmentation algorithm (for the first segmentation level of that particular application), and also same feature values for the segments. A special top-down operator was implemented for that purpose. The operator imports the segmentation (image objects) and respective feature values exported by eCognition.

Also implemented for the application was a function that calculates fuzzy membership values, to be used in the decision rules of the bottom-up interpretation step. Figure 1 shows the semantic network defined in InterImage. It should be noted that the network designed in InterImage is very similar to the Class Hierarchy defined in the

eCognition project, the marginal differences are due to particularities of the different interpretation strategies of the two systems.

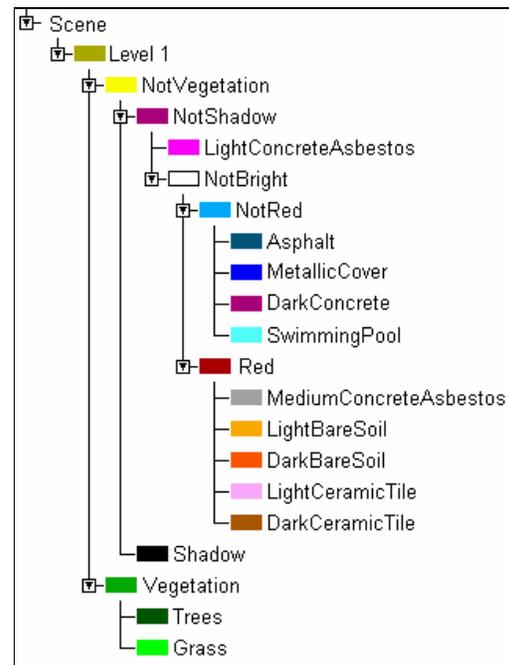


Figure 1 – Semantic Network

5 EXPERIMENT RESULTS

Figure 2 shows the interpretation results in terms of a thematic map depicting the final classification performed with InterImage.

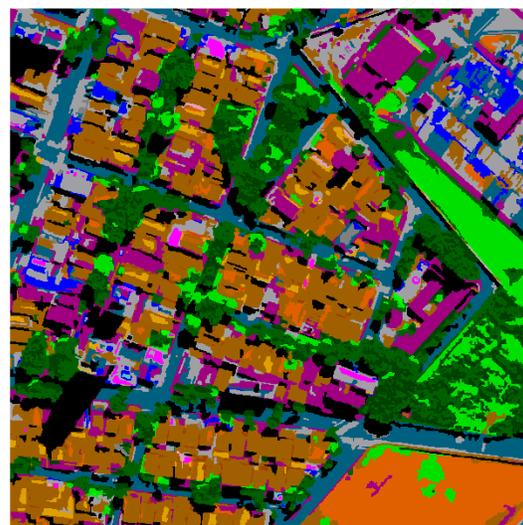


Figure 2 – Classification Results

Table 1 shows the confusion matrix between the two classifications. The rows of the matrix show the classification results obtained with InterImage and the columns show the results obtained by eCognition.

Table 1 – Confusion Matrix

| InterImage \ eCognition | Asphalt | Metallic Cover | Light Concrete Asbestos | Medium Concrete Asbestos | Dark Concrete | Swimming Pool | Bare Soil | Shadow | Ceramic Tile | Trees | Grass | Total |
|--------------------------|---------|----------------|-------------------------|--------------------------|---------------|---------------|-----------|--------|--------------|-------|-------|-------|
| Asphalt | 380 | | | | 1 | | | | | | | 381 |
| Metallic Cover | | 109 | | | | | | | | | | 109 |
| Light Concrete Asbestos | | | 65 | | | | | | | | | 65 |
| Medium Concrete Asbestos | | | | 632 | | | 11 | | 4 | | | 647 |
| Dark Concrete | 49 | | | | 446 | | | | | | | 495 |
| Swimming Pool | | | | | | 2 | | | | | | 2 |
| Bare Soil | | | | | | | 99 | | 21 | | | 120 |
| Shadow | | | | | | | | 450 | | | | 450 |
| Ceramic Tile | | | | 4 | | | 93 | | 1090 | | | 1187 |
| Trees | | | | | | | | | | 1130 | | 1130 |
| Grass | | | | | | | | | | 4 | 323 | 327 |
| Total | 429 | 109 | 65 | 636 | 447 | 2 | 203 | 450 | 1115 | 1134 | 323 | 4913 |

The global result in terms of the overall coincidence in the two classifications is 96.2%. The small difference in the classification can be explained by a precision problem, related to the fuzzy sets defined in the eCognition knowledge model. As there is no way to extract precise formulation of the fuzzy sets, defined graphically by the user through eCognition's GUI, the shape of the fuzzy sets in InterImage was an approximation of the original ones. That conclusion is supported by the fact that the confusion occurred basically between classes that have similar characteristics, such as *Asphalt* and *Dark Concrete*, and *Ceramic Tile* and *Exposed Soil*.

6 CONCLUSION

This work introduced InterImage, a new knowledge-based image interpretation platform being developed in accordance to the open source philosophy. The paper also reported an experiment version in which an interpretation model built in eCognition was translated into the current InterImage version. The classification results achieved by both software platforms were very similar.

In terms of interpretation strategy InterImage has a more flexible architecture than eCognition. The combination of a model-driven followed by a data-driven analysis, as performed by InterImage, has the potential of an improved computational efficiency in comparison to eCognition that follows a pure data-driven strategy. Thus InterImage offers knowledge modelling features not available in eCognition.

The experiment described in this paper provides evidence that most interpretation models built on eCognition may be transposed to InterImage. For the particular application taken as example in this paper, time was mostly spent in reformulating the fuzzy rules in the InterImage environment. As the InterImage project evolves, the task of translating the eCognition models into InterImage will certainly become much easier.

Further development of InterImage has been already planned. Implementation of multi-temporal capabilities, automatic knowledge extraction functions, as well as built-in image processing operators are some of the developments envisaged for the near future.

AKNOWLEDGMENTS

The authors acknowledge the support provided by FAPERJ (*Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro*) and FINEP (*Financiadora de Estudos e Projetos*).

REFERENCES

BLASCHKE, T.; LANG, S.; LORUP, E.; STROBL, J.; ZEI, P. Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. In: CREMERS, A., GREVE, K. (Eds.). **Environmental Information for Planning, Politics and the Public**. Marburg: Metropolis-Verlag, 2000. 2 v. p.555-570.

BÜCKNER, J.; PAHL, M.; STAHLHUT, O.; LIEDTKE, C.-E. GEOAIDA - A Knowledge Based Automatic Image Data Analyzer for Remote Sensing Data. In: ICSC CONGRESS ON COMPUTATIONAL INTELLIGENCE METHODS AND APPLICATIONS 2001 - CIMA 2001, Bangor, Wales, UK. **Proceedings of The Congress on Computational Intelligence Methods and Applications 2001 - CIMA 2001 (CD-ROM)**.

CARRION, D.; GIANINETTO, M.; SCAIONI, M. GEOREF: A Software for Improving the Use of Remote Sensing Images in Environmental Applications. In: IEMSS 2002 - INTEGRATED ASSESSMENT AND DECISION SUPPORT, 2002, Lugano, Switzerland. **Proceedings of IEMSS 2002**, 2 v., pp. 360.

CLÉMENT, V.; GIRAUDON, G.; HOUZELLE, S.; SANDAKLY, F. Interpretation of Remotely Sensed Images in a Context of Multisensor Fusion Using a Multispecialist Architecture. **IEEE Transactions on Geoscience and Remote Sensing**, v. 31, n. 4, p. 779-791, jul. 1993.

EHLERS, M.; JANOWSKY, R.; GÄHLER, M. New Remote Sensing Concepts for Environmental Monitoring. In: Ehlers, M. (Ed.), Conf. Remote Sensing for Environmental Monitoring, GIS Applications, and Geology, 2002, Bellingham, WA. **Proceedings SPIE**, Vol. 4545, p.1-12,

GRAHAM, J.; JONES, P.: **Knowledge-Based Image Processing Systems**. London: Springer-Verlag, 1997. 178 p.

LIEDTKE, C.-E.; BÜCKNER, J.; GRAU, O.; GROWE, S.; TÖNJES, R. AIDA: A System for the Knowledge Based Interpretation of Remote Sensing Data, In: THIRD INTERNATIONAL AIRBORNE REMOTE SENSING CONFERENCE AND EXHIBITION, 1999, Copenhagen, Denmark,

Proceedings of the Third International Airborne Remote Sensing Conference and Exhibition. Environmental Research Institute of Michigan, AGU, 1997.

MATSUYAMA, T.; HWANG, V. Advances in computer vision and machine intelligence. **SIGMA: A knowledge-base aerial image understanding system**. New York: Plenum, 1990. ISBN: 030643301X.

MCKEOWN, D. M.; HARVEY, W. A.; MCDERMOTT, J. Rule Based interpretation of aerial imagery. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 7, n. 5, p. 570-585, 1985.

PINHO, C.M.D., KUX, H., ALMEIDA, C.M. Elaboração de rede semântica para classificação de cobertura do solo de ambientes intra-urbanos: o Caso de São José dos Campos - SP, Brasil. In: XIV SIMPÓSIO BRASILEIRO DE SENSOREAMENTO REMOTO - SBSR, 2007, Florianópolis, Brasil. **Proceedings of the XIV Simpósio Brasileiro de Sensoriamento Remoto - SBSR 2007**. In press.

REGO, L. F. G.; **Automatic land-cover classification derived from high-resolution IKONOS satellite image in the urban atlantic forest in Rio de Janeiro, Brasil by means of an objects-oriented approach**. 2003. 193p. Tese (Forstwissenschaftlichen Fakultät). Albert-Ludwigs-Universität.2003

PINHO, C.M.D. **Análise orientada a objetos de imagens de satélite de alta resolução espacial aplicada à classificação de cobertura do solo no espaço intra-urbano: o caso de São José dos Campos-SP, São José dos Campos**, 181 p. Dissertação (Mestrado em Sensoriamento Remoto, INPE, 2005 (INPE-14183-TDI/1095).

SAGERER, G.; NIEMANN, H. Semantic Networks for Understanding Scenes. **Advances in Computer Vision and Machine Intelligence**. 1.ed. New York: Plenum Publishing Corporation, 1997. 500p. ISBN: 0306457040.

SCHIEWE, J., TUFTE, L., EHLERS, M.,. Potential and Problems of multi-scale segmentation methods in remote sensing. **GIS-Informationssysteme**, v. 6, pp. 34-39, 2001.