

# INTERIMAGE: AN OPEN SOURCE KNOWLEDGE BASED FRAMEWORK FOR AUTOMATIC INTERPRETATION OF REMOTE SENSING DATA

D. A. B. Oliveira<sup>a</sup>, G. A. O. P. Costa<sup>a</sup>, R. Q. Feitosa<sup>a\*</sup>, E. F. Castejon<sup>b</sup>, L. M. G. Fonseca<sup>b</sup>,

<sup>a</sup> Department of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro.  
Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ – Brazil - (dario, gilson, raul)@ele.puc-rio.br

<sup>b</sup> Image Processing Division (DPI), National Institute for Spacial Research (INPE)  
Av. dos Astronautas, 1.758, Jd. Granja - São José dos Campos - SP – Brazil – (castejon, leila)@dpi.inpe.br

**KEY WORDS:** Interpretation, Classification, Knowledge Base, Open Systems, Inter-Operability

## ABSTRACT:

This paper introduces the InterIMAGE System, an open source knowledge based framework for automatic interpretation of remote sensing data being developed at the Computer Vision Lab (LVC) of the Catholic University of Rio de Janeiro, with the collaboration of the Brazilian National Institute for Spacial Research (INPE), among other scientific institutions. InterIMAGE is an extension of an earlier project called GeoAIDA, developed by the Insitut für Informationsverarbeitung (TNT) of the Leibniz Universität Hannover, and has inherited from that system the basic functional design and control mechanisms. The interpretation strategy implemented on the system is based on a knowledge model defined by the user, structured as a semantic network. The nodes of the network represent concepts, classes of image objects expected to be found in a scene. Specialized image processing operators can be attached to the system, which controls their execution. The interpretation process performed by InterIMAGE has two sequential and complementary phases. The first phase, called the top-down step, is model driven and generates, through the operators associated to the different nodes of the semantic network, hypotheses of the occurrence of image objects on the scene. The final instance network results from the bottom-up analysis, responsible from judging those hypotheses, validating or discarding them based on user defined rules, also associated to the nodes of the semantic network. In order to make a comprehensive introduction of the system, the methodology for the creation of an interpretation model in InterIMAGE will be presented, and the interpretation process will be discussed with the aid of an exemplary interpretation problem based on synthetic data.

## 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 (Ehlers, 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 modelling, 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.

The development of automatic, remote sensing data interpretation technology can benefit from its integration to GIS technological tools. GIS software is now ubiquitous; there are systems of different types and sizes, varying from the desktop to the corporate user and different solutions for Internet data access and distribution. In the vast majority of cases, such developments have been industry-driven, with few exceptions, such as the IDRISI (Eastman, 2003) and SPRING (Camara, 1993) systems. Fierce competition and growing user demand has resulted in a number of high-quality solutions, which are largely responsible for the vast increase in the GIS marketplace.

In this paper we present the architecture and features of a knowledge-based image interpretation system called InterIMAGE and comment on its integration with Terralib (Camara, 2000), a GIS classes and functions library developed in INPE. InterIMAGE is an open source software development initiative, leaded by the Computer Vision Lab of the Electrical Engineering Department, at the Catholic University of Rio de

---

\* Corresponding author.

Janeiro (PUC-Rio) and the by the Brazilian National Space Research Institute (INPE).

In order to make a comprehensive introduction of the system, we will also present an exemplary interpretation problem based on synthetic data, describing in detail aspects related to knowledge representation, and to the interpretation process carried out by InterIMAGE.

In the remainder of this paper we describe the basic characteristics of InterIMAGE (Section 2) and the interpretation strategy implemented by the system (Section 3). In Section 4 the synthetic data interpretation experiment is described and analysed. Finally, some conclusions and directions for future work are stated in Section 5.

## 2. INTERIMAGE 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 representation and processing functionality and image processing operators are being included in InterIMAGE.

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

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. In this context, an image processing operator is any operator that generates a labelled result image of a given image. They can carry out threshold operations, texture-based or model-based procedures. In those systems, the following grouping of the objects generated by such operators 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 the scene 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 operator results.

Moreover, 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 semantic concepts and edges represent the relations between concepts. During the analysis, guided by the semantic network, the system controls the execution of the operators and decision rules attached to the concept nodes, and generates a network of object instances, each instance defining a geographic region associated to a specific semantic concept.

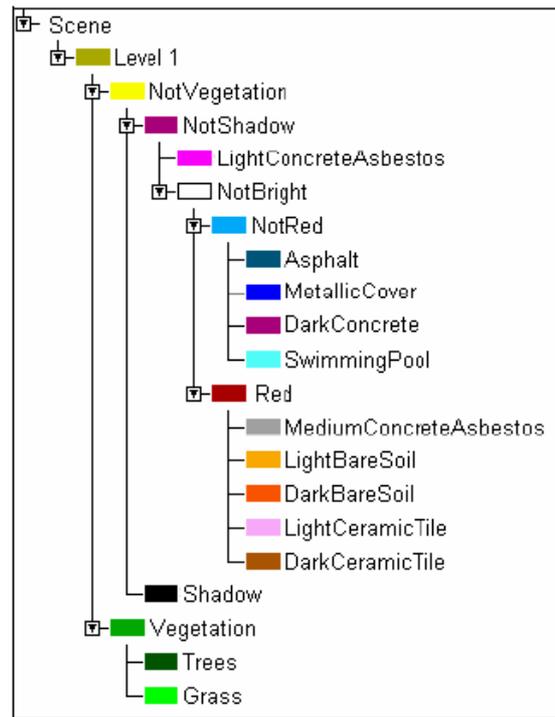


Figure 1: Semantic Network

Interpretation of remote sensing data means to transform input data into a structural and pictorial description that represents the result of the analysis. In InterIMAGE, the result of the interpretation is a structural description of the generated instance network and the corresponding label image.

The interpretation 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 object hypothesis – geographic regions associated to the occurrence of concepts that will be later subjected to validation. 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 analysis.

For each node of the semantic network the user defines the information used in the execution of each processing step, which includes the identity and respective parameters of the top-down and bottom-up operators associated to the node.

Top-down operators have the task of creating object hypotheses. For this purpose any external image processing classifying operator can be used, its output needs only to be organized in a predetermined structure. The object 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 object hypotheses is performed for that region.

In the top-down step the control process traverses the nodes of the semantic network, from the root to the leaf nodes, calling the top-down operators attached to each node. Top-down processing occurs in parallel with respect to the semantic network branches. Object hypotheses associated to the network nodes are created during this process and organized into a hypothesis network.

As processing reaches the leaf nodes, the bottom-up step initiates. The control process starts visiting nodes in the opposite direction, executing the bottom-up operators attached to the nodes recursively, until it reaches the root node. It is the task of the bottom-up operator to evaluate hypotheses associated to the child nodes of the node the operator is attached to. In this process object hypothesis are discarded or turned into object instances until the resulting instance network has been created.

#### 4. INTEGRATION WITH TERRALIB

Terralib library is an open source project, available from the Internet in a collaborative environment. TerraLib aims to provide a rich and powerful environment for the development of GIS science research, enabling the development of GIS prototypes, which include new concepts such as spatio-temporal data models, geographical ontologies and advanced spatial analysis techniques. TerraLib defines a geographical data model and provides support for this model over a range of different DBMS and is organized as a library of C++ classes and functions, written in ANSI-C++.

The code available in the Terralib can be used for the implementation of complex image processing operators suitable for InterIMAGE. An interface programming class has been added to the library for that purpose, and top-down operators have already been created from the library.

An interactive tool that implements an interface between Terralib and InteIMAGE, to build top-down operators using TerraLib image processing capabilities, is under development. This tool will allow the creation of an image processing pipeline, generating as output a stand-alone operator that can be used in InteIMAGE's interpretation process as an holistic operator, to be called by the system during the top-down step.

The main advantage of this kind of integration strategy is that it allows quick development of powerful top-down operators using the functions already available in Terralib, without requiring any adjustment of InteIMAGE's basic functionalities.

#### 5. INTERPRETATION EXAMPLE

The experiment described in this work was designed to help explaining the basic aspects of InteIMAGE's interpretation process and the role of top-down and bottom-up operators in such process.

The goal of the interpretation problem in this example is to classify predetermined regions on a scene into three different semantic classes: *Field*; *Residential*; and *Service*, based on

information about the occurrence of houses and large buildings (hereafter denoted as halls).

The general rules that define the interpretation problem, which will be structured as knowledge model in InterIMAGE's, are the following:

1. If a region does not contain any building – *house* or *hall* –, then it is classified as *Field*.
2. If a region contains only houses, then it is classified as *Residential*.
3. If a region contains at least one hall, then it is classified as *Service*.
4. A *house* is composed of a small *roof* and adjacent *shadow*.
5. A *hall* is also identified through its *roof* and *shadow*, but occupies an area larger than a certain threshold.

#### 5.1 Input Data

The input data is composed of three different images:

1. A synthetic *visual* – RGB – *image* (figure 2-right), simulating a geographic region in which it is possible to observe roofs of different sizes and the shadows of the respective buildings.
2. A *label image* (figure 2-center), dividing the geographic region into 4 different regions to be classified.
3. A *height image* (figure 2-left), simulating laserscan image data, where the pixels intensity value is proportional to the height of objects on the ground.



Figure 2: Input Data

#### 5.2 Knowledge Model

The interpretation process implemented by InterIMAGE is guided by a knowledge model, represented through a semantic network. So the first task in creating an interpretation application with InterIMAGE is the definition of a semantic network, whose nodes represent the semantic concepts involved in the interpretation problem.

In this application we expect to find *houses* and *halls*, which are characterized by *roofs* and *shadows*, to classify predefined *regions* into three different classes: *field*; *residential*; and *service*. The semantic network should comprise all those semantic concepts, respecting their intrinsic hierarchy. Figure 3 shows semantic network defined for this example.

Through the semantic network the underlying rules of the interpretation problem are made explicit. For example, residential regions are characterized by containing only houses, so the semantic network branch of the concept node *Residential* contains only a *House* node, and no *Hall* node.

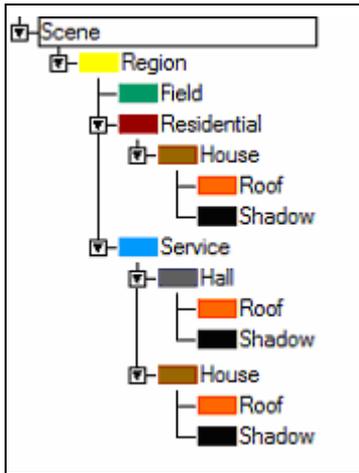


Figure 3: Semantic Network

### 5.3 Operators

As mentioned in Section 3, to each node two operators are attached: a top-down and a bottom-up operator. To the leaf nodes only top-down operators can be attached.

In this example the following operators were used:

1. *Regions from Raster* – a top-down operator that divides a scene into regions identified by different labels in an input label image.
2. *Segmentation of Roofs* – a top-down operator that segments roofs in an input height (laserscan) image.
3. *Segmentation of Shadows* – a top-down operator that segments shadows in a given visual (RGB) input image.
4. *Dummy Top-Down Operator* – a top-down operator that outputs hypotheses that are copies of the higher level node hypotheses.
5. *Generic Bottom-Up Operator* – a bottom-up operator that applies decision rules, defined by the user, in the evaluation of the hypotheses associated to the child semantic nodes.
6. *Dummy Bottom-Up Operator* – a bottom-up operator that only accepts the hypotheses associated to the child semantic nodes.

### 5.4 Top-Down Step

The top-down process is initialized at the *Scene* node. This node is a root node and is present in all applications. This node has a *Dummy Top-Down* operator that generates a single hypothesis, whose associated geographic region is that of the whole scene.

Then, the top-down process proceeds downward to *Region* node that is attached to the *Regions from Raster* top-down operator. The execution of this operator generates four different *Region* hypotheses, each one corresponding to a different region in the input label image (Figure 2-center).

The *Field* node is a leaf node with a *Dummy Top-Down* operator attached to. This means that all priorly generated *Region* hypotheses will become *Field* hypotheses. Top-down process of that branch ends after the execution of the operator, as this is a leaf node.

The *Dummy Top-Down Operator* is also attached to the *Residential* and *Service* nodes, so all *Region* hypotheses will also become *Residential* and *Service* hypotheses.

Similarly, the *House* and *Hall* nodes are also associated to the *Dummy Top-Down Operator*. Each *Residential* hypothesis will become a *House* hypothesis, and each *Service* hypothesis will become both *House* and *Hall* hypotheses.

The *Roof* nodes are associated to the *Segmentation of Roofs* top-down operator. It generates as many hypotheses as the quantity of roofs it finds in the regions that correspond to the hypotheses of its father nodes. Each time it is called, the operator only processes the corresponding geographic region in the input height image.

The *Segmentation of Shadows* top-down operator is attached to the *Shadow* nodes. It generates *Shadow* hypotheses by processing the input visual image. Each time it is called the operator only processes the geographic region associated to the hypothesis of the father of the *Shadow* nodes.

The top-down processing step ends when it reaches the leaves of the semantic network. At this point a network of hypotheses has been created. The interpretation process then enters the bottom-up step.

### 5.5 Bottom-Up Step

The bottom-up step evaluates all the hypotheses generated in top-down step and validates or discards them according to a particular decision rules. A bottom-up operator will evaluate the hypotheses associated to the children nodes of the node it is attached to. As the leaf nodes of the semantic network have no children, no bottom-up operators are attached to them.

The *Generic Bottom-Up Operator* has decision rules as parameters. These rules are defined in a specific stack-based language. So, at any node this operator is attached to, a specific rule to evaluate the hypotheses associated to its child nodes can be defined.

For example, the decision rule specified at the *House* node, to validate hypotheses of the child *Roof* and *Shadow* nodes, is shown below:

```
nodelist "Roof" selectClass
"area" calc 50 > select 500 < select
"compactness" calc 0.3 > select
"orthogonality" calc 0.3 > select

nodelist "Shadow" selectClass
"area" calc 5 > select 100 < select

"W" 70.0 neighbour
merge
```

Figure 4: Generic Bottom-Up Operator's decision rule

This specific decision rule is based on geometric aspects of *Roof* and *Shadow* hypotheses generated previously, in the top-down step.

Initially it selects the *Roof* hypotheses from the input list of hypotheses it receives from the children nodes (*Roof* and *Shadow*). Then it calculates the area of each *Roof* hypothesis, and selects those hypotheses that lie in the range of 50 and 500 m<sup>2</sup>. It selects, among the remaining hypotheses those with orthogonality and compactness lower than 0.3.

At this point the operator has a stored list of *Roof* hypotheses with area between 50 and 500 m<sup>2</sup>, and compactness and orthogonality lower than 0.3.

Then the complete list of input children hypotheses is reloaded and the ones associated to the *Shadow* node are selected. The operator selects from the *Shadow* hypotheses only those that have an area in the range of 5 and 100 m<sup>2</sup>, and compactness and orthogonality lower than 0.3.

At this point the operator has a stored list of *Shadow* hypotheses with area between 5 and 100 m<sup>2</sup>, and compactness and orthogonality lower than 0.3.

These two different lists are processed by the operation *neighbour*. This operation verifies if each *Roof* hypothesis has an adjacent *Shadow* hypothesis in East-West direction, at 70 degrees. Only *Roof* and *Shadow* hypotheses that present this topological relationship will be accepted, and turned into *putative* instances. The term *putative* is used to characterize the instances at this point because they can be later discarded, if hypotheses associated to the higher level nodes are discarded in subsequent bottom-up processing.

Then the validated child hypotheses are grouped and eventual spatial conflicts among them are resolved. At this point new *House* hypotheses are generated, and the correspondent original hypotheses, created through top-down processing, are substituted by the new ones. Each new hypothesis is associated to the geographical region covered by a group of child hypothesis.

A similar decision rule is defined for the *Hall* nodes, to which the *Generic Bottom-Up Operator* is also attached. The only difference is that the area threshold for roofs and shadows is larger than the ones defined for houses.

The *Dummy Bottom-Up Operator* is attached to the *Service* and *Residential* nodes. In this case all hypotheses associated to the respective child nodes will be accepted – turned into *putative* instances – after eventual spatial conflicts among them are resolved. Moreover, no new hypotheses are created from the grouping of child hypotheses. However, the original *Service* and *Residential* hypotheses, for which no instances of any of the respective child nodes have been found, are deleted. This means that an original *Residential* hypothesis is maintained only if at least one house is found inside it, and an original *Service* hypothesis is maintained if at least one house and one hall are found within the geographical region associated to the *Service* hypothesis.

It is interesting to notice that for the upper right region, one *Residential* one *Service*, and one *Field* hypothesis exist at this point. The decision rule associated to the *Generic Bottom-Up Operator* attached to the *Region* node resolves this spatial conflict by a simple rule that assigns different priority values for the respective concepts – highest priority for *Service* and lowest to *Field*, and chooses the hypothesis with higher priority.

So the *Service* hypothesis will be chosen for the upper-right region, and all other will be deleted.

When the bottom-up analysis reaches the root node *Scene*, all remaining putative instances are turned into final instances and the interpretation process finishes.

## 5.6 Results

Figure 4 shows the interpretation results in terms of thematic maps depicting the final classification, generated by the system's GUI. The colors assigned to the segments are the ones defined for the nodes of the semantic network. In Figure 4-top, roofs are shown in orange, shadows in black. Halls and houses are shown in Figure 4-center, in gray and brown respectively. Field regions are shown in green, Residential in red, and Service regions in blue (Figure 4-bottom).

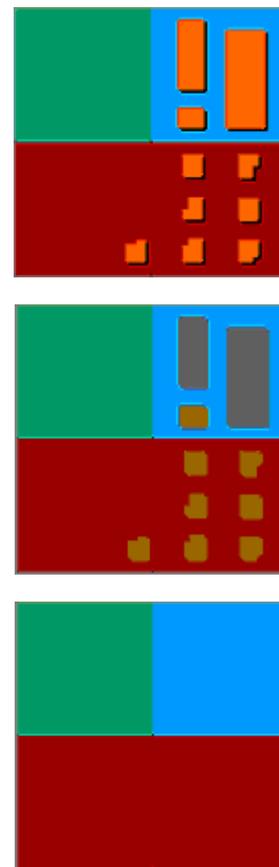


Figure 4. Classification Results

## 6. CONCLUSION

In this paper the basic characteristic and interpretation process implemented by the InterIMAGE System, an open source knowledge based framework for automatic interpretation of remote sensing data under developed at the Computer Vision Lab (LVC) of the Catholic University of Rio de Janeiro (PUC-Rio), have been presented.

The system has a number of other functionalities, which were not mentioned here due to the lack of space. Further details about the system, including technical and user documentation, its source code and executable versions compiled for Linux and

Windows are available at the InterIMAGE Project's website (<http://www.lvc.ele.puc-rio.br/projects/interimage/>).

The current version of the InterIMAGE System is still regarded as a prototype. Presently, however, a number of new functionalities are being implemented. Those include improvements directed towards augmenting user interactivity, such as debugging capabilities, visualization of partial results and feature spatialization.

A module for the creation of complex operators from primitive image processing operations is also under construction. The objective of this module is to enable users with no programming skills to create holistic operators from the functions available in Terralib (<http://www.terralib.org/>), and from the combinations of previously existing operators.

Finally, multitemporal interpretation capabilities related to the explicit representation and processing of temporal knowledge and multirate images are also currently being developed.

#### ACKNOWLEDGMENTS

The authors would like to acknowledge the support from FINEP (Financiadora de Projetos e Pesquisas) and FAPERJ (Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro).

#### REFERENCES

Ehlers, M., Janowsky, R., Gähler, M., 2002. New remote sensing concepts for environmental monitoring. In: Ehlers, M. ed. *Conf. on Remote Sensing for Environmental Monitoring, GIS Applications, and Geology*, 2002, Bellingham, WA. Proceedings SPIE, v. 4545, pp. 1-12.

Bückner, J., Pahl, M., Stahlhut, O., Liedtke, C.-E., 2001. 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).

Blaschke, T., Lang, S., Lorup, E., Strobl, J., Zei, P., 2000. Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. In: CREMERS, A., GREVE, K. ed. *Environmental Information for Planning, Politics and the Public*. Marburg: Metropolis-Verlag, v. 2, pp. 555-570.

Camara, G., Souza, R. M., Pedrosa, B.M., Vinhas, L., Monteiro, A.M., Paiva, J. C., Carvalho, M.T., Gattass, M., 2000. TerraLib: Technology in support of GIS innovation. In: *Workshop Brasileiro de GeoInformática*, São Paulo, Brasil, pp. 126-133.

Camara, G., Souza, R. M., Freitas, U. M., Paiva, J. C., 1993. SPRING: Concepcao, Evolucao e Perspectivas. In: *Simposio Brasileiro de Sensoriamento Remoto*, Curitiba, Brasil, v. 2, pp. 308-317.

Carrion, D., Gianinetta, M., Scaioni, M. 2002. 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, v.2, pp. 360.

Clément, V., Giraudon, G., Houzelle, S., Sandakly, F., 1993. 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, pp. 779-791.

Eastman, J. R., 2003. *IDRISI Kilimanjaro Guide to GIS and Image Processing*, Stuttgart, Germany.

Liedtke, C.-E., Bückner, J., Grau, O., Growe, S., Tönjes, R., 1999. 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.

Matsuyama, T., Hwang, V., 1990. SIGMA: A knowledge-base aerial image understanding system. *Advances in computer vision and machine intelligence*. New York: Plenum Publishing Corporation, 1990.

Mckeown, D. M., Harvey, W. A., Mcdermott, J., 1985. Rule-based interpretation of aerial imagery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 7, n. 5, pp. 570-585.

Rego, L. F. G., 2003. Automatic land-cover classification derived from high-resolution IKONOS satellite image in the urban Atlantic forest in Rio de Janeiro (Brazil) by means of an object-oriented approach. PhD Thesis, Forstwissenschaftlichen Fakultät, Albert-Ludwigs-Universität 193p.

Sagerer, G., Niemann, H. 1997. Semantic networks for understanding scenes. *Advances in Computer Vision and Machine Intelligence*. 1.ed. New York: Plenum Publishing Corporation, 500p.

Schiewe, J., Tufte, L., Ehlers, M., 2001. Potential and problems of multi-scale segmentation methods in remote sensing. *Geo-Information-Systeme*, v. 6, pp. 34-39.