A FRAMEWORK FOR AUTOMATIC LOW-RESOLUTION SATELLITE IMAGE INTERPRETATION BASED ON SPECTRAL, CONTEXTUAL AND MULTITEMPORAL KNOWLEDGE

G. L. A. Mota [a], K. Pakzad [b], S. Muller [c], M. S. P. Meirelles [d e], R. Q. Feitosa [a d], H. L. da C. Coutinho [e]
[a] Dept. of Electrical Engineering, Pontifical Catholic University, Rio de Janeiro, Brazil – (guimota, raul)@ele.puc-rio.br
[b] Institute of Photogrammetry and GeoInformation, University of Hannover, Germany – pakzad@ipi.uni-hannover.de
[c] Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, University of Hannover, Germany – mueller@tnt.uni-hannover.de
[d] University of the State of Rio de Janeiro, Rio de Janeiro, Brazil – maggie@eng.uerj.br
[e] EMBRAPA Soils, Rio de Janeiro, Brazil – heitor@cnps.embrapa.br

Commission IV , WG IV/7

KEY WORDS: Environment, Agriculture, Automation, Knowledge base, Fuzzy Logic, Multitemporal

ABSTRACT:

This work proposes a framework for increasing the degree of automation of low-resolution satellite images interpretation procedures. Basically, the method starts with an image segmentation which, according to a criterion of spectral response homogeneity, outlines the regions to be classified. The classification procedure is aided by expert knowledge. This procedure makes use of three types of knowledge: *spectral,* which relates the homogeneous classes of spectral response to the correspondent classes of interest; *contextual,* indicating the relevant contexts for the discrimination of classes with similar spectral responses; and *multitemporal reasoning*, considering both the former classification of the region and the plausible class transitions in that particular time interval. This strategy takes simultaneously into account spectral, contextual and multitemporal evaluations of the region which, are combined into a single membership value. Each membership value corresponds to a class of interest, and the highest indicates its classification. As a consequence, the proposed model requires as input: *satellite images of the region of interest acquired in different dates*; *the accurate classification of the former image* and *the previous mentioned categories of expert knowledge*. The prominent points of the proposed methodology are: its flexible structure, which allows for straightforward application of this model to low-resolution image interpretation cases; and the automatic learning/calibration of the spectral levels to the current time image. Experiments were performed in order to evaluate the potential of the proposed framework. The images employed in the experiments are situated in the Taquari Watershed, more exactly, in the County of Alcinópolis that belongs to the State of Mato Grosso do Sul, Brazil. The experimental results indicate that the use of knowledge can contribute to the increment of the degree of automation of the interpretation process

1. INTRODUCTION

Considering forests, ecosystems, productive areas and urban areas, remote sensing constitutes a significant tool for the monitoring of large regions of the surface of the Earth. Namely, such technology can play an important role for the management of the natural resources.

However, most software packages for remote sense data interpretation are based on visual analysis. Although such packages offer digital image processing and pattern recognition tools, they lack of a framework to automate the image interpretation procedure which could increase the productivity of the photo-interpreters. Additionally, despite software development allowing the automation of some tasks, high level structuring is limited. Consequently, in this field, the development of technology is essential.

The other motivation for this research derives from economical restrictions. The cost of aerial or high resolution satellite images – e.g. IKONOS –, considering the huge dimensions of threatened ecosystems currently under monitoring programs in countries like Brazil (tropical rainforests, savanna and wetlands) monitoring, is too high. This constitutes an obstacle to its utilization on a routine basis.

Visual (manual) interpretation and analysis dates back to the early beginnings of remote sensing for air photo interpretation. On the other hand, digital processing and analysis is more recent, it was brought by the advent of digital recording of remote sensing data and the development of computers. Both manual and digital techniques for interpretation of remote sensing data have their respective advantages and disadvantages. Consequently, by and large, the interpretation of remote sensing data encompasses a sequence of manual and automatic steps. Considering both knowledge background and additional data, in the manual steps, a photo-interpreter calibrates/trains the automatic algorithms and also solves the inconsistencies of their outcomes.

In low spatial resolution satellite images, distinct land cover classes may produce similar spectral responses turning more difficult their discrimination. Nonetheless, it does not represent a problem to experienced human photo-interpreters who take advantage of supplementary information (e.g. size, shape and texture) as well as their knowledge background in order to solve contradictory interpretations. The present proposal considers that by explicitly modeling the experience of a photo-interpreter about a specific site into a knowledge-basis, his reasoning can be computationally reproduced.

This research is part of an international cooperation project, so-called ECOWATCH, involving Brazilian and German institutions, aiming at setting up systems for the automatic interpretation of multitemporal low-resolution satellite images.

The visual (manual) interpretation process that this work aims at automating starts with an image segmentation algorithm which outlines contiguous segments with homogeneous spectral response; then, the photo interpreter selects some segments as a training set for the learning of spectral signatures. Finally, considering simultaneously the previous classification of one segment, its spectral pattern and the context where it is situated, the photo-interpreter, taking also into account his knowledge about the region under analysis, classifies the segment. The main objective of the present proposal is to improve the degree of automation of the process as a whole; therefore, the selection of the training set and knowledge-based classification must be automated.

The proposed framework starts with an image segmentation procedure, as explained above. A knowledge-based classification engine is employed, considering three types of knowledge: *spectral,* which relates the homogeneous classes of spectral response to the correspondent classes of interest; *contextual*, indicating the relevant contexts for the discrimination of classes with similar spectral responses; and *multitemporal reasoning*, considering both the former classification of the segment and the plausible class transitions in that particular time interval.

This paper is organized as follows: The next section discusses the state of the art of knowledge based image interpretation. Section 3 presents the proposed methodology, section 4 the experiments and section 5 the conclusions.

2. KNOWLEDGE BASED INTERPRETATION

The automatic interpretation of remotely sensed images has been intensively researched. By analyzing the existing systems available in the literature (Matsuyama, 1990; Clement, 1993; Niemann, 1990, 1997; Bueckner, 2001) their main components can be identified as follows:

a) Digital images from particular sensors,

b) GIS data of the focused region,

c) image processing algorithms,

d) prior knowledge on the focused region, and

e) a control logic to the interpretation process.

The control logic manages the interaction of the remaining components. This component triggers, accordingly to the scene semantic, the image processing algorithms. In this process, the prior knowledge, usually delivered by an expert, plays an important role, by supplying specific information about the expected objects.

In the literature, many approaches for image interpretation and sensor fusion have been presented; nevertheless, only some try to formalize the expert knowledge. Some cases, like SPAM (McKeown, 1985) and SIGMA (Matsuyama, 1990), implement a hierarchical control and construct the objects incrementally, considering multiple levels of detailing. MESSIE (Clement, 1993) models the objects explicity distinguishing four views: geometry, radiometry, spatial context and functionality. It employs frames and production rules. ERNEST (Niemann, 1990, 1997) applies semantic networks in order to represent the structure of the objects, serving as a knowledge base specific to the problem.

The increasing amount of regions represented in Geographic Information Systems (GIS) motivated the development of AIDA (Liedtke, 1997). It is able, in a single semantic network, to model the expert knowledge, GIS information and data from multiple sensors. Among other advantages, the incorporation of GIS information reduces the interpretation uncertainty. AIDA was applied to the analysis of aerial images, where the image components like buildings, houses, rivers, factories, and forests may be observed.

The semantic network into AIDA is organized in several layers. The highest layer provides the semantic of the objects foreseen, whilst the lowest level corresponds to the image primitives. As a whole, the several layers correspond to distinct abstraction levels, providing a structural description of the scene.

Employing a novel modality of scene description, GEOAIDA (Bueckner, 2001) incorporates a holistic approach to the main advantages of its predecessor. It considers an object as a whole, in a global way, in other words, without subdividing it into its subcomponents. In its implementation, holistic operators may be easily incorporated to the semantic network nodes. Rigorously, GEOAIDA provides a hybrid approach, in cases where the holistic operators are unable of acting a structural analysis proceeds. The main advantage obtained when the holistic operators perform properly is the reduction of the amount of time spent by the knowledge interpretation, which is such a time expansive process.

Considering the previously mentioned systems, their main divergence corresponds to the formalization of the expert knowledge and information acquisition. Even though historically formalisms of knowledge representation had been developed in order to process natural language, they are quite versatile.

Little is reported in the literature about the use of knowledge-based approaches to the interpretation of low-resolution satellite images. Zhang (1998) described a method which aims at detecting changes in the land use. The approach simultaneously employs SPOT and LANDSAT images to update the urban maps of Shanghai, China, and discriminates vegetation, water and urban areas. As a result, the urban maps are updated highlighting the new constructions.

Kunz et al. (1997) employed ERNEST to update the maps in a GIS database. The approach derives a semantic network from the contents of the GIS database. Beside the spectral response, the compactness, the mean curvature, the texture standard deviation and homogeneity are evaluated, and compared with the contents of the GIS database. Discrepancies are corrected, being the GIS updated.

Largouet et al. (2001), implemented a land cover analysis using a sequence of images of different satellites and a temporal model. The investigated scene corresponded to a rural area, and the analysis of the images required specific agricultural knowledge, modeled into a formalism, so-called timed automata, as a priori knowledge about the scene.

Suzuki et al. (2001), integrated structural knowledge to the image classification process. Basically, a fuzzy classifier generate a preliminary partition of the image and, then, the system tries to improve the initial classification.

In order to evaluate the potentiality of knowledge-based approaches for the interpretation of low-resolution satellite images, the following papers were produced as part of the ECOWATCH project: Müller (2003) and Mota (2003) investigated the usage of the GEOAIDA system to perform knowledge-based classification of a SPOT 3 XS and LANDSAT 7 TM, respectively. In such essays, only spectral and contextual knowledge were employed. The obtained results demonstrated the potential of knowledge-based approaches to the interpretation of low-resolution satellite images. Pakzad et al. (2003) presented a procedure for the multitemporal interpretation of LANDSAT 7 TM images of a region covered by different categories of vegetation. In such article, an object oriented classification was performed employing the commercial software eCognition simulating a multitemporal reasoning. Given the previous classification of an object, the procedure modeled possible temporal state transitions, taking into account either ecological, agronomic, or legal restrictions.

3. METHODOLOGY

3.1 General model description

The proposed framework considers data corresponding to different time instances. Hereafter, ***t*** represents the time in which the image to be interpreted was acquired, ***t-1*** a previous time instance and ***Δt*** the time interval between ***t-1*** and ***t***. By analogy, the time interval between ***t-2*** and ***t*** is equal to ***2Δt***.

In figure 1 the automatic low-resolution satellite image interpretation framework is presented. In order to perform the

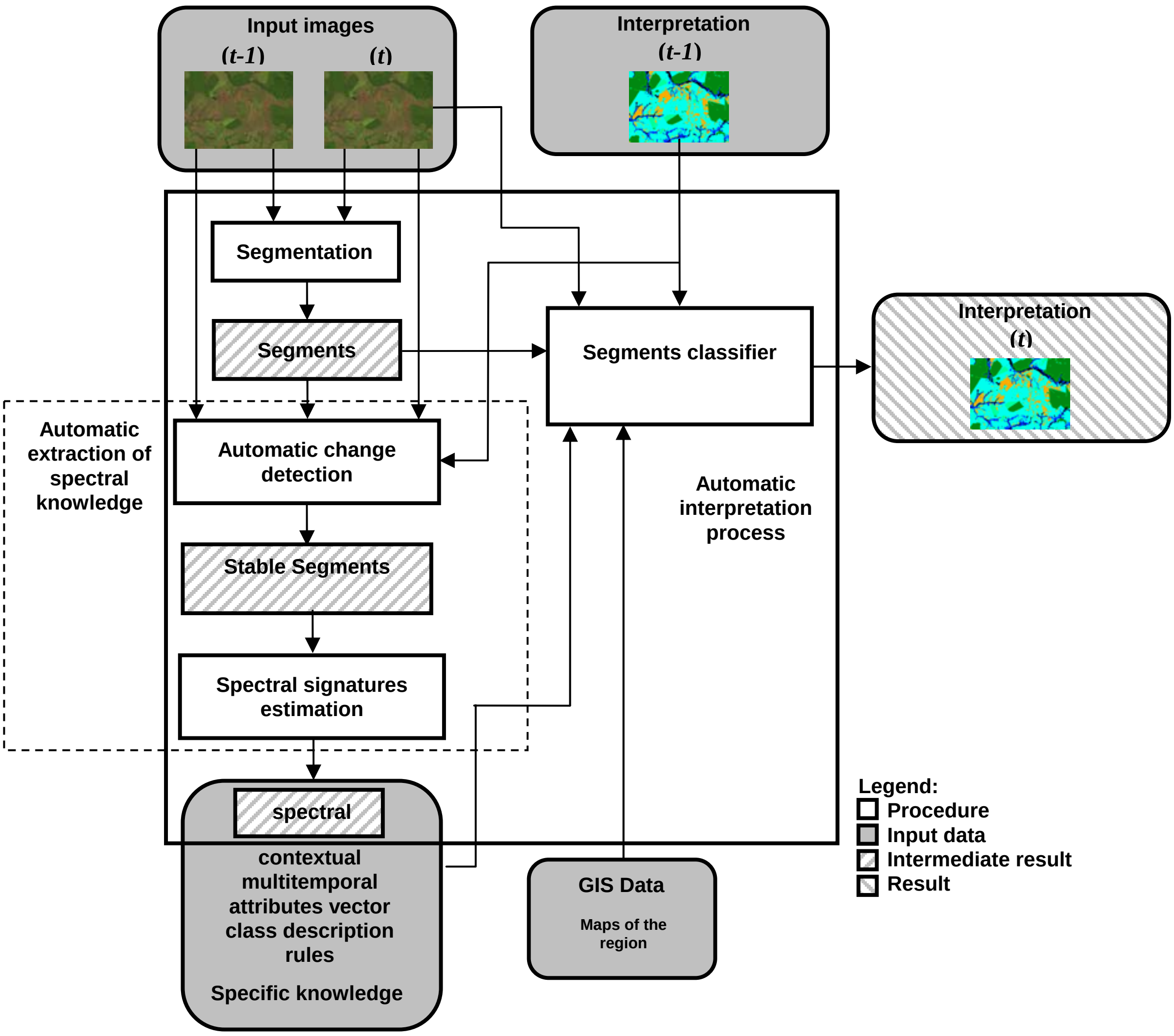

Figure 1. Framework to the automatic low-resolution satellite image interpretation

interpretation of the region of interest in the instant ***t***, the proposed procedure employs: ***1)*** Specific knowledge about the region of interest containing the classes of land use/cover and their respective spectral, contextual and multitemporal characterizations besides the rules that model the classification strategy; ***2)*** Two registered images acquired by the same satellite in the instances ***t*** and ***t-1***; ***3)*** The accurate interpretation of the image in the instant ***t-1***; ***4)*** GIS data about the region of interest.

The proposed interpretation procedure is based on segment classification. This allows the interpretation process to consider attributes such as shape, texture, and spatial relation between objects, in a more appropriate context for the representation of knowledge than the classic approach of pixel classification (Andrade, 2003; Darwish, 2003; Yan, 2003). Therefore, the procedure begins with an image segmentation which outlines the regions with homogeneous spectral response in both input images.

3.2 Specific knowledge

In the context of the present research, specific knowledge is the knowledge background gathered both theoretically and empirically which turns an individual able, or more competent, to perform a specific task. In this case, the task is interpretation of low-resolution satellite images.

Categories of knowledge

In this work, the following categories of specific knowledge are required:

1) Class description – Lists and describes the land use/cover (LULC) classes to be identified in the image.

2) Spectral knowledge – Spectral signatures of the LULC classes

3) Contextual Knowledge – Indicates the contextual information required for the discrimination of the LULC classes with similar spectral signatures.

4) Multitemporal knowledge – Given the segment classification in ***t-1***, relates its possible classifications in ***t*** and their respective possibilities. The possibility theory, introduced by Zadeh (1978), constitutes a context which allows treating the concepts of uncertainty in a non probabilistic way.

5) Vector of attributes – vector composed by the *n* attributes considered by a photo-interpreter while classifying a segment.

6) Rules of inference – Rules that describe the logic employed in the visual interpretation. Basically, it models the reasoning applied by the photo-interpreter in the interpretation process by combining different items of spectral, contextual and multitemporal knowledge.

The specific knowledge employed in the automatic interpretation of low-resolution satellite images can be acquired from experts on photo-interpretation, agronomy, ecology and, even, from the people of the region.

The Automatic Extraction of Spectral Knowledge

The contextual and multitemporal aspects of knowledge are, in general, dependent on the areas under analysis; however, they are independent of the image to be classified. On the other hand, the spectral knowledge is affected by the conditions in which the image was obtained. Climatic and atmospheric conditions, problems of sensor calibration, the level of soil humidity, etc., can lead to different spectral responses of the same LULC classes in images of the same regions, but acquired in different epochs.

Usually, this difficulty is solved by supervised classification methods in which the spectral signatures are estimated considering samples selected by the photo-interpreter in the image to be classified. Nonetheless, this selection is usually manually performed.

This section introduces a scheme designed to automate the procedure for estimation of spectral signatures (spectral knowledge). The dotted box in figure 1 encompasses the steps related to the automatic selection of training samples. Such procedure takes into account the input images, obtained in ***t*** and ***t-1***, and a reliable classification of the image of ***t-1***.

First, the automatic selection of training samples submits the objects produced by the segmentation to an automatic change detection procedure which discriminates the stable and changed segments, considering the spectral responses in the instants ***t-1*** and ***t***. Then, the stable segments are used as a training set to estimate the spectral signatures of the different classes.

The change detection approach assumes the following hypothesis: ***1)*** The amount of segments changed between ***t-1*** and ***t*** is small; ***2)*** The natural and man made events alters differently the spectral responses of the classes; ***3)*** The images in ***t-1*** and ***t*** are registered.

3.3 Segments Classifier

In this work, a fuzzy logic system is employed to model the reasoning of the photo-interpreter while performing the photo-interpretation, since this kind of system is a transparent model for modeling logical reasoning (Kuncheva, 2000).

Mendel (1995) warns that, in general, in order to build fuzzy logic systems, the following is required: ***1)*** employment of linguistic variables (Zadeh, 1965). ***2)*** Quantification of the linguistic labels associated to the linguistic variables. ***3)*** Definition of logical connectors, ***4)*** the combination of the rules.

In order to produce the interpretation of the scene in the instant ***t***, the segments classifier employs: ***1)*** Specific knowledge; ***2)*** low-resolution satellite images in ***t*** and ***t-1***; ***3)*** the interpretation of the image in ***t-1***; ***4)*** the segments; ***5)*** maps of the region of interest.

Initially, it is necessary to define the linguistic labels, their respective labels and the rules comprehended in the specific knowledge. The modeling of the linguistic variables, associated with contextual and multitemporal knowledge, is performed manually. On the other hand, the fuzzy sets, which define the spectral knowledge, are automatically modeled.

Then, segment classification is performed. Initially, the correspondent attributes vector is calculated. Then, the inference machine, based on the inference rules, the fuzzy sets, the rules and the attribute values, calculates the membership values of the segment to each one of the classes in the legend. The classification is given to the class with the highest membership value.

4. EXPERIMENTAL RESULTS

The experiments presented in this section aim at evaluating the potential of the proposed framework. The images employed in the experiments are situated in the Taquari Watershed, more exactly, in the County of Alcinópolis that belongs to the State of Mato Grosso do Sul, Brazil. The images were acquired on August 7, 2000; and August 10, 2001 by the satellite LANDSAT 7 (bands 3, 4 and 5).

The reference classification for this evaluation was produced by visual interpretation by a photo interpreter experienced in vegetal cover classification. In this procedure, it was considered, besides the images of 2000 and 2001, the classification of the image of 2000, the digital elevation model, the drainage map and the photo-interpreter's knowledge about the region.

In the region of interest, the following LULC classes can be found: Bare soil; Ancillary forest; Pasture; Water bodies; Dense savannah; and Dense savannah in regeneration. For a small segment of the input image, the figure 2 presents: a) the supervised multispectral classification result; b) the outcome produced by the proposed method; c) the reference classification.

The results were assessed in terms of percentage of segments that were correctly classified (classification ratio) and the percentage of segments wrongly classified (classification error), considering the previously mentioned reference classification.

Table 1 presents the evaluation of the results produced by a supervised multispectral classification and by the proposed method for the entire image. The purpose of this experiment is to evaluate the increment of the degree of automation provided by the proposed approach in comparison to a supervised multispectral classification.

	Supervised Classification	Knowledge based Classification
Classification ratio	69 %	90 %
Classification error	31 %	10 %

Table 1. Comparison of the results of the supervised multispectral classification and the proposed methodology

The results shown in Table 1 indicate the superiority of the results produced employing knowledge while compared to the outcomes produced by the spectral classification. If both results were given as basis to a photo-interpreter, he would post-edit the classification of 31 % of the segments previously classified by supervised classifier, but only 10 % of the segments classified by the knowledge based classifier. This fact indicates that the use of knowledge can contribute to incrementing productivity of the interpretation process.

5. CONCLUSION

This paper presented a framework for knowledge based interpretation of multitemporal low-resolution satellite images. The prominent points of the proposed methodology are: its flexible structure which allows for straightforward application of this model to low-resolution image interpretation problems; and the automatic learning/calibration of the spectral levels to the current time image.

The proposed methodology was preliminarily evaluated through experiments employing images of two regions inside the watershed of the Taquari river, northeast of the State of Mato Gosso do Sul. The evaluation showed that the knowledge based results were superior to the spectral classification results. This fact indicates that the use of knowledge can contribute to the increment of the degree of automation of interpretation process.

Multispectral classification using the manually selected training set provided a classification ratio of 69 % for the image of 2001. The remaining 31 % must be corrected during post-editing. On the other hand, for the same image, the proposed knowledge based classifier with the automatically selected training set provided a classification ratio of 90 %. In this case, only 10 % of the segments woul require post-editing. Therefore, the quantity of segments whose classification would be changed would decrease from 31 % to 10 %. Additionally, considering that the proposed procedure is capable of selecting automatically the training set, the work of the photo-interpreter would be, in this case, restricted to the post-edition of 10 % of the segments.

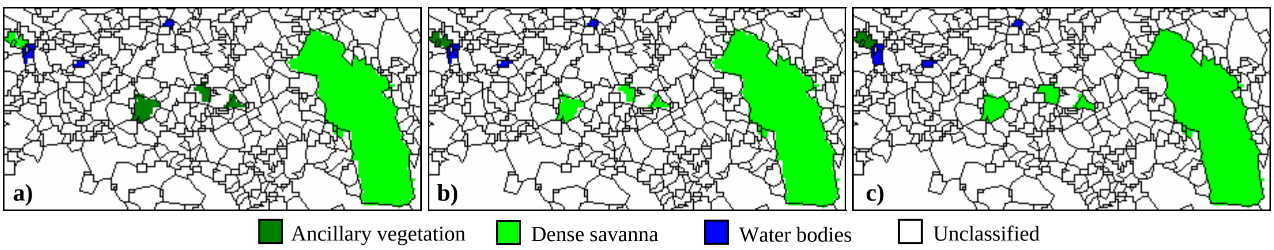

Figure 2. a) Supervised multispectral classification result; b) Proposed method outcome; c) reference classification

REFERENCES

Andrade, A., Botelho, M. F., Centeno, J., 2003. Classificação de imagens de alta resolução integrando variáveis espectrais e forma utilizando redes neurais artificiais. In: *XI Seminário Brasileiro de Sensoriamento Remoto***,** Belo Horizonte, Brazil, pp. 265-272.

Bückner, J., Stahlhut, O., Pahl, M., Liedtke, C.-E., 2001. GEOAIDA - A Knowledge Based Automatic Image Data Analyzer for Remote Sensing Data, In: *Computational Intelligence: Methods & Applications*, Bangor, Wales, United Kingdon, pp. 19-22.

Clément, V., Giraudon, G., Houzelle, S., Sandakly, F., 1993. Interpretation of Remotely Sensed Images in a Context of Mutisensor Fusion Using a Multispecialist Archteture. *IEEE Transactions on Geosciences and Remote Sensing*, 31(4), pp. 779-791.

Darvish, A., K. Leukert, W. Reinhardt, 2003. Image Segmentation for the Purpose of Object-Based Classification, In: *International Geoscience and Remote Sensing Symposium*, Toulouse, France.

Kuncheva, L. I., 2000. How good are fuzzy if-then classifiers?, *IEEE Transactions on Systems, Man, and Cybernetics*, Part B, 30(4), pp. 501-509.

Kunz, D., Schilling, K., Vögtle, T., 1997. A new approach for satellite image analysis by means of a semantic network. In: *Semantic Modeling*, W. Förstner and L. Plumer, Eds. Birkhäuser, Basel, pp. 20-36.

Largouet, C., Cordier, M.-O., 2001. Improving the landcover classification using domain knowledge, In: *AI Communications*, 14(1), pp. 35-43.

Liedtke, C.-E., Bückner, J., Grau O., Growe S., Tönjes, R., 1997. AIDA: A System for the Knowledge Based Interpretation of Remote Sensing Data. In: *Proceedings of the Third International Airborne Remote Sensing Conference and Exhibition*, Copenhagen, Denmark, Vol II, pp. 313-320.

Matsuyama, T., Hwang, V., 1990. SIGMA, a knowledge-base aerial image understanding system, New York: Plenum.

McKeown, D. M., Harvey, W. A., McDermott, J., 1985, Rule Based interpretation of aerial imagery, IEEE Transactions in Pattern Analysis and Machine Intelligence, 7(5), pp. 570-585.

Mendel, J. M., 1995, Fuzzy Logic Systems for Engineering: A tutorial, Proceedings of the IEEE, 83(3), pp.345-377.

Mota, G., Müller, S., Feitosa, R., Coutinho, H., Meireiles, M., Vieira, H. 2003. An Evaluation of Knowledge-Based Interpretation Applied to Low-Resolution Satellite Images. In: *IEEE International Geoscience and Remote Sensing Symposium - IGARSS 2003*, Toulouse, France.

Müller, S., Feitosa, R. Q., Mota, G. L. A., Costa, D., Silva, V. V., Tanisaki, K., 2003. GEOAIDA Applied to SPOT Satellite Image Interpretation, In: *URBAN 2003*, Berlim.

Niemann, H., Sagerer, G., Schröder, S., Kummert F., 1990. ERNEST: A Semantic Network System for Pattern Understanding. IEEE Transactions on Pattern Analysis and Machine Intelligence3, 12(9), pp. 883-905.

Niemann, H., Sagerer, G., 1997. Semantic Networks for Understanding Scenes, Advances in Computer Vision and Machine Intelligence, Book News, Portland, OR.

Pakzad, K., Mota, G., Meirelles, M., Coutinho, H., Feitosa, R. Q., 2003. Automatic Interpretation of Vegetation Areas in Brazil. In: Joint Workshop of ISPRS Working Groups I/2, I/5, IC WG II/IV and the EARSeL Special Interest Group 3D Remote Sensing "High Resolution Mapping from Space 2002"**,** Hannover, Germany.

Suzuki, H., Matsakis, P., Andrefouet, S., Desachy, J., 2001. Satellite Image Classification Using Expert Structural Knowledge: A Method Based on Fuzzy Partition Computation and Simulated Annealing. In: *Proceedings of IAMG'2001*, Cancún, Mexico.

Yan, G., 2003. Pixel based and object oriented image analysis for coal fire research. Master Thesis, ITC, Netherlands.

Zadeh, L. A., 1965, Fuzzy Sets, Journal of Information and Control, 8, pp. 338-353.

Zadeh, L. A., 1978. Fuzzy sets as a basis for a theory of possibility. *IEEE Trans. Systems, Man and Cybernetics***,** 3, pp. 28-44.

Zhang, Y, 1998. Detection of urban housing development using multisensor satellite data and maps, IAPRS 1998, Vol. 32, pp. 292-299, part 3/1.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to CAPES and DAAD, respectively a Brazilian and a German governmental research agencies, for supporting the ECOWATCH project, and Embrapa (World Bank funded Agtec Project), for providing the necessary data.