COGNITIVE APPROACHES AND OPTICAL MULTISPECTRAL DATA FOR SEMI-AUTOMATED CLASSIFICATION OF LANDFORMS IN A RUGGED MOUNTAINOUS AREA

F. F. Camargo [a, *], C. M. Almeida [a], G. A. O. P. Costa [b], R. Q. Feitosa [b], D. A. B. Oliveira [b], R. S. Ferreira [b], C. Heipke [c]

[a] INPE, Division for Remote Sensing, 12227-010 São José dos Campos – SP, Brazil - (fortes, almeida)@dsr.inpe.br
[b] Dept. of Electric Engineering, Catholic University of Rio, R. Marquês de São Vicente, 225, 22451-900,
Rio de Janeiro - RJ, Brazil - (gilson, raul, dario)@ele.puc-rio.br; webcsm@gmail.com
[c] IPI, Leibniz University of Hannover, Nienburger Str. 1, D-30167, Hannover, Germany – heipke@ipi.uni-hannover.de

Commission WG IV/4

KEY WORDS: InterIMAGE, Remote Sensing; Semantic Network; Fuzzy Logic; Semi-Automated Classification

ABSTRACT:

This paper introduces a new open source, knowledge-based framework for automatic interpretation of remote sensing images, called InterIMAGE. This framework owns a flexible modular architecture, in which image processing operators can be associated to both root and leaf nodes of the semantic network, which constitutes a differential strategy in comparison to other object-based image analysis platforms currently available. The architecture, main features as well as an overview on the interpretation strategy implemented in InterIMAGE is presented. The paper also reports an experiment on the classification of landforms. Different geomorphometric and textural attributes obtained from ASTER/Terra images were combined with fuzzy logic and drove the interpretation semantic network. Object-based statistical agreement indices, estimated from a comparison between the classified scene and a reference map, were used to assess the classification accuracy. The InterIMAGE interpretation strategy yielded a classification result with strong agreement and proved to be effective for the extraction of landforms.

1. INTRODUCTION

Remote sensing technology delivers the most important subsidies for the identification and monitoring of land cover changes and physiographic features on the earth surface, effectively supporting the investigation of the interactions between the environment and agricultural, urban and environmental planning activities.

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

In this sense, some commercial software packages for the automatic interpretation of images have been launched, aiming to overcome the drawbacks imposed by conventional classifiers. Although this new generation of programs represents a considerable advance in relation to the conventional classifiers, some important challenges remain in the domain of automatic interpretation of images, so as to assure a greater accuracy and detailing capacity in feature extraction and in classification. There is consequently a strong demand for the development of robust techniques for automatic information extraction and interpretation of remote sensing data (Blaschke et al., 2000).

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 (Liedtke et

al., 1997; Bückner et al., 2001). In this approach human experts' knowledge is organized in a knowledge base (Graham and Jones, 1997) to be used as input for automated interpretation processes, enhancing the productivity and accuracy and reducing at the same time the subjectivity of the interpretation process.

In this paper we introduce the architecture and main features of a knowledge-based image interpretation system called InterIMAGE (Section 1.1), 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 by the Brazilian National Institute for Space Research (INPE).

In the remainder of this paper we describe the interpretation strategy implemented by the system (Sections 1.1 and 1.2), and a brief overview on the study area is provided in Section 2. Sections 3 and 4 present an actual image interpretation experiment related to the classification of landforms in the municipality of São José dos Campos, located in the southeastern State of São Paulo, Brazil. Section 5 describes the results of this experiment, and a critical evaluation is made on the potential and drawbacks of the input data and methodological procedures. Finally, some conclusive remarks and directions for future work are drawn in Section 6.

1.1 System Description

InterIMAGE is based on the software GeoAIDA (Bückner et al., 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

---

* Corresponding author.

control mechanisms. A new graphical user interface, a knowledge model debugging tool, multitemporal interpretation capability and some image processing operators are also available at InterIMAGE.

In short, InterIMAGE implements a specific image interpretation strategy. Such strategy is based upon and guided by a hierarchical description of the interpretation problem, structured in a semantic network.

The bases for interpretation of digital image data are the results generated by image processing operators. In this context, an image processing operator is any operator that generates a labeled result image of a given image. Such image processing operators are denoted here as 'classifying operators'. They can fulfill 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 et al., 1997) can be used to reduce the combinatory 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.

1.2 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 contains nodes and edges, whereat nodes represent concepts and edges represent the relations between the concepts. The network is actually a connected graph with no cycles, i.e. a tree. In each concept node, information necessary for the analysis, such as the image processing operator specialized 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 and a top-down one. The top-down step is model-driven and generates a network of hypotheses based on the semantic network. The grouping of hypotheses and their acceptance or refusal 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) as well as the decision rules to be used in the bottom-up step.

The top-down operators are entrusted with separating regions into sub regions and with building hypotheses for the concepts of the semantic network, regions of the image associated to the concepts. This task is realized recursively from the root to the leaf nodes. For this purpose any (external) image processing operator, which creates hypotheses for the sub region, can be used in the analysis process.

When the top-down analysis reaches the leaf nodes, the interpretation turns from model-driven to data-driven (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.

2. STUDY AREA

The municipality of São José dos Campos, with a total area of 1098.6 km², is located in the Paraíba do Sul Valley, in São Paulo State, southeastern Brazil (Fig. 1). This valley is situated in a Pre-Cambrian mountain range, which is the most outstanding relief feature in the east of the South American continent. The lithology of this region is composed by gneisses (Arquean/Mid-Proterozoic), sin- and post-tectonic granite suites (Upper Proterozoic), sedimentary rocks (Late Tertiary), and quaternary deposits. The geomorphological features consist of ridges, mountains, rivers plains, terraces and sedimentary hills.

3. DATA ACQUISITION, DATA PREPROCESSING, AND REFERENCE MAPS

3.1 Data Acquisition and Pre-processing

The data used comprised ASTER/Terra VNIR images (bands 3N and 3B of 08/31/2004 with a nominal spatial resolution of 15 m, processing level L1B; WRS-2 Path: 219; WRS-2 Row: 076; ID: AST_L1B.003:2071432223); 109 GPS points (with orthometric heights); 1:10 000 vector files of the streets network, drainage, 10 m interval contour lines, and scattered elevation data points. Initially, the images were oriented using Toutin´s model, available in the software Geomatica, and ground control points extracted from the vector files. Epipolar images were then generated and the DEM (with 15 m spatial resolution) was obtained by means of stereo-correlation and parallax estimation. The 109 GPS points were supplied by FUNCATE foundation, INPE and IEAv (Institute for Advanced Studies) research institutes and these points were used as the elevation reference in the DEM accuracy validation. Two kinds of statistics were obtained based on elevation discrepancies (D) between the DEM and the GPS elevations: RMSE and t-Student test (H0: $\bar{\Delta} = 0$ or H1: $\bar{\Delta} \neq 0$) with a 90% confidence interval.

The validated DEM was not subject to any kind of post-processing algorithm or filtering, since it was generated with the same resolution as the stereo pair images (15 m). This DEM
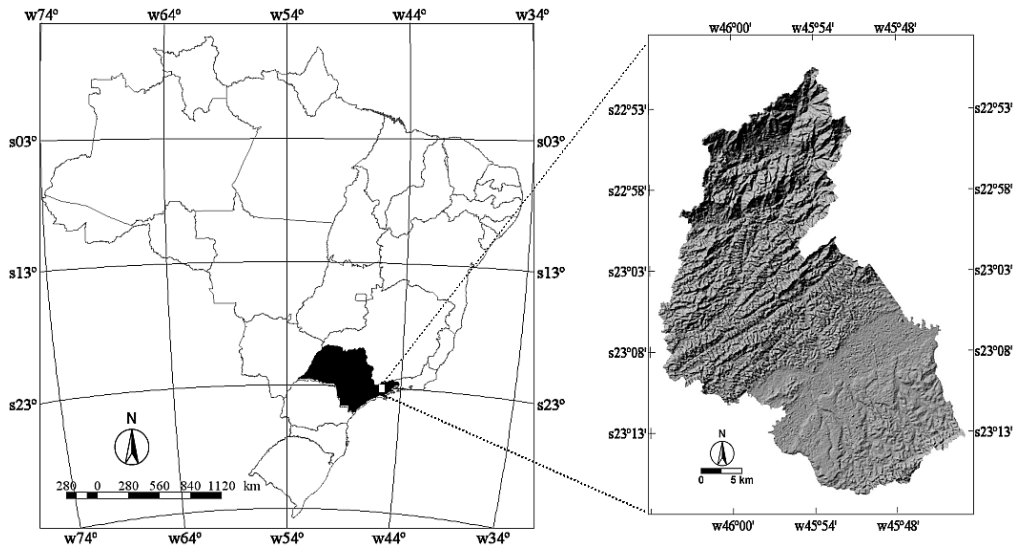
Fig. 1. Study area: Brazil and São Paulo State (in black) on the left side, and on the right side the municipality of São José dos Campos (view of its shaded relief).

was then used to extract the geomorphometric and textural variables, which drove the expert classification system employed in this work. Table 1 presents the input variables of the classification model. All these variables entered the model as images with different formats (float, binary, 8 bits, 11 bits) and the extension '.pfm'.

Type of Variable	Description
Geomorphometric	Elevation (ASTER/Terra DEM)
	Slope
	Vertical curvature
	Horizontal curvature
Textural	Entropy
	Second Angular Moment
Others	Accumulated Flow
	Drainage density
	Lakes and dams
	Shaded Relief

Table 1. Input variables of the classification model

Drainage density was previously generated using the vector layer of water streams and a coarse segmentation level. For each segmentation region, we calculated the ratio of the surface area of water streams to the area of the segment. The final map was sliced into two classes: 'high drainage density' and 'medium and low drainage density'. This binary layer was then converted into an image with the '.pfm' format. And finally, the input variable 'lakes and dams' was obtained by a region classification of the band 3N, which was orthorectified based on the generated ASTER/Terra DEM. This classification was converted into a shape file, and this file was then reformatted as a thematic raster layer, with extension '.pfm'.

3.2 Landform Classes and Reference Maps

The landform classes were based on the legend of the São Paulo State geomorphological map, issued at a 1:1 000 000

scale. For the final classes selection, two procedures were adopted: (i) a visual interpretation of the ASTER/Terra epipolar images stereo pair, and (ii) a field work, which was precisely important for the discrimination of the different river terraces levels.

The legend followed the International Institute for Geo-Information Science and Earth Observation (ITC) system of geomorphological survey, which emphasizes the main terrain aspects (morphography, morphometry, morphogenesis, and morphocronology) and is employed in geomorphological mapping and natural resources and hazards evaluations. The landforms were then classified into groups of prevailing morphogenetic processes: (i) structural-denudational landforms (in purple); (ii) denudational landforms (in brown), and (iii) aggradational landforms (in green).

Due to the unavailability of detailed geomorphological maps for the area (1:100 000 or greater scales), reference maps of landforms was specially produced in a 3D digital station by an experienced geomorphologyst, adopting the same set of classes used in the semi-automated classification. This task was accomplished by means of a restitution of the stereo pair of epipolar images, and these reference maps were used for validating the classification result.

4. EXPERIMENT DESIGN: OBJECT-BASED CLASSIFICATION

The experiment implemented in this work was designed to evaluate the performance of InterIMAGE for the specific purpose of classifying landforms (geomorphological features). The knowledge model for the automatic interpretation experiment is based on a semantic network comprising five classes of morphogenetic landforms at a macro level (Fig 2), and eleven classes of morphographic landforms at a detailed level (Fig. 3). These two levels of classification correspond to two distinct InterIMAGE projects.
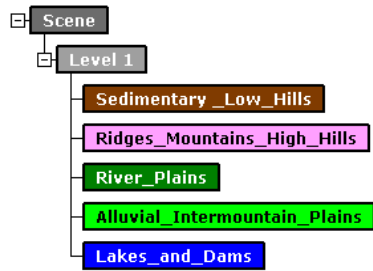
Fig. 2. Semantic network for the classification of morphogenetic landforms in São José dos Campos.

In InterIMAGE, the semantic network is structured so as to contain an upper node, known as 'scene node', to which all other nodes are subjugated. In the first network, the node 'Level 1' contains the bottom-up operator. No top-down operator is attached to the nodes 'Scene' and 'Level 1'. A top-down operator was attached to each of the nodes 'Sedimentary Low Hills', 'Ridges/Mountains/High Hills', 'River Plains', and 'Alluvial Intermountain Plains'. This operator performs a segmentation of the ASTER/Terra DEM, using the region-growing algorithm (Baatz and Schäpe, 2000), generating for each segment one hypothesis of each of those concepts – geographically coincident hypotheses. A shape file related to the study area borders was inserted in the node 'Level 1', so as to prevent segmentation in the scene background. The parameters used in the segmentation were set as: 50, for the scale parameter; 0.7 for the color factor; 0.3 for the shape factor; 0.3 for compactness; and 0.7 for smoothness.

As stated in Kressler and Steinnocher (2006), the scale parameter controls the maximum allowed heterogeneity per segment, and thus, larger scales will lead to larger segments. The color and shape factors are complementary, i.e., they sum to one and indicate how color and shape information is used in the segmentation process. The shape factor is further divided into compactness and smoothness. A high value of compactness leads to smaller and very compact segments, and hence, is more suitable for man-made objects, while a high value of smoothness leads to segments optimized to have smooth borders, which are on their turn more suitable for natural objects (Kressler and Steinnocher, 2006).

The bottom-up rules for the classification of morphogenetic landforms – placed at node 'Level 1' of the semantic network – as they are represented in the graphical user interface of InterIMAGE. The column to the left contains colored boxes associated with the so-called elements, which correspond to the basic structures of a decision rule. In our particular case, the elements comprise: Class, Logic, Expression, Membership, and Classify. The Class element is used for selecting hypotheses of a specific class. The element Logic selects hypotheses that fulfill a given criterion. It enables the user to choose attributes and specify crisp selection criteria based on logical expressions. Expression is used to set an attribute value. Membership is an element that uses fuzzy logic to calculate and aggregate class membership values. And finally, the element Classify solves spatial conflicts and classifies the selected hypotheses according to their membership values.

Initially all hypotheses associated to the children of 'Level 1' have their membership to the respective concepts set to null, so as to set equal conditions as to the competition among hypotheses in relation to the existing classes of landforms.

Membership to 'Lakes and Dams' is exclusively defined through an attribute derived from the thematic raster layer containing such water bodies, coded as 'media8'. The attribute value indicates the percentage of pixels in the segment associated with a hypothesis that fall into such water bodies, according to the thematic layer.
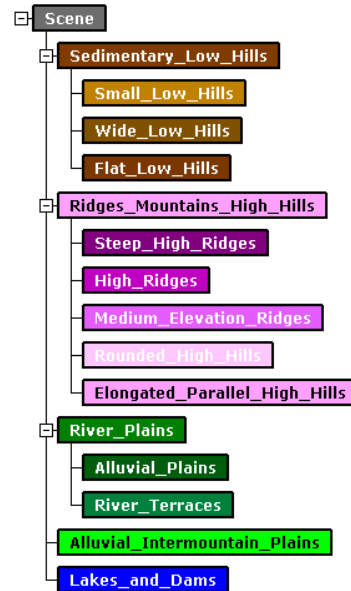

Fig. 3. Semantic network for the classification of morphographic landforms in São José dos Campos.

The membership to the class 'Sedimentary Low Hills' is based on three attributes, calculated for all segments generated by the top-down operator: mean of the DEM (i.e. average height), mean of entropy, and mean of slope (calculated for all pixels contained in each segment). The element 'Min' refers to the fuzzy minimum operator and was employed for the aggregation of membership values.

The membership to 'Ridges/Mountains/High Hills' is also based on three attributes. The first one is the drainage density. This first attribute is used to select only the hypotheses (segments) that comply to crisp thresholds , corresponding to the existence of the class 'high drainage density' on the drainage density image (Section 3.1). The non-complying hypotheses will remain with membership to 'Ridges/Mountains/High Hills' equal to zero. The second attribute corresponds to the mean of the DEM, and the third one to the mean of slope. The final membership to 'Ridges/Mountains/High Hills' is defined by an aggregation (minimum) of the values obtained through the membership functions related to the mean of the DEM and of slope.

The membership to 'Alluvial Intermountain Plains' is once more based on three attributes: amplitude of the DEM, mean of the DEM, and mean of slope. The amplitude refers to the difference between the maximum and minimum height values observed within a hypothesis (segment). On its turn, the membership to 'River Plains' is based on two attributes: mean of entropy and mean of slope.

Membership to 'Lakes and Dams' is exclusively defined through an attribute derived from the thematic raster layer containing such water bodies. The attribute value indicates the percentage of pixels in the segment associated with a hypothesis that fall into such water bodies, according to the thematic layer.
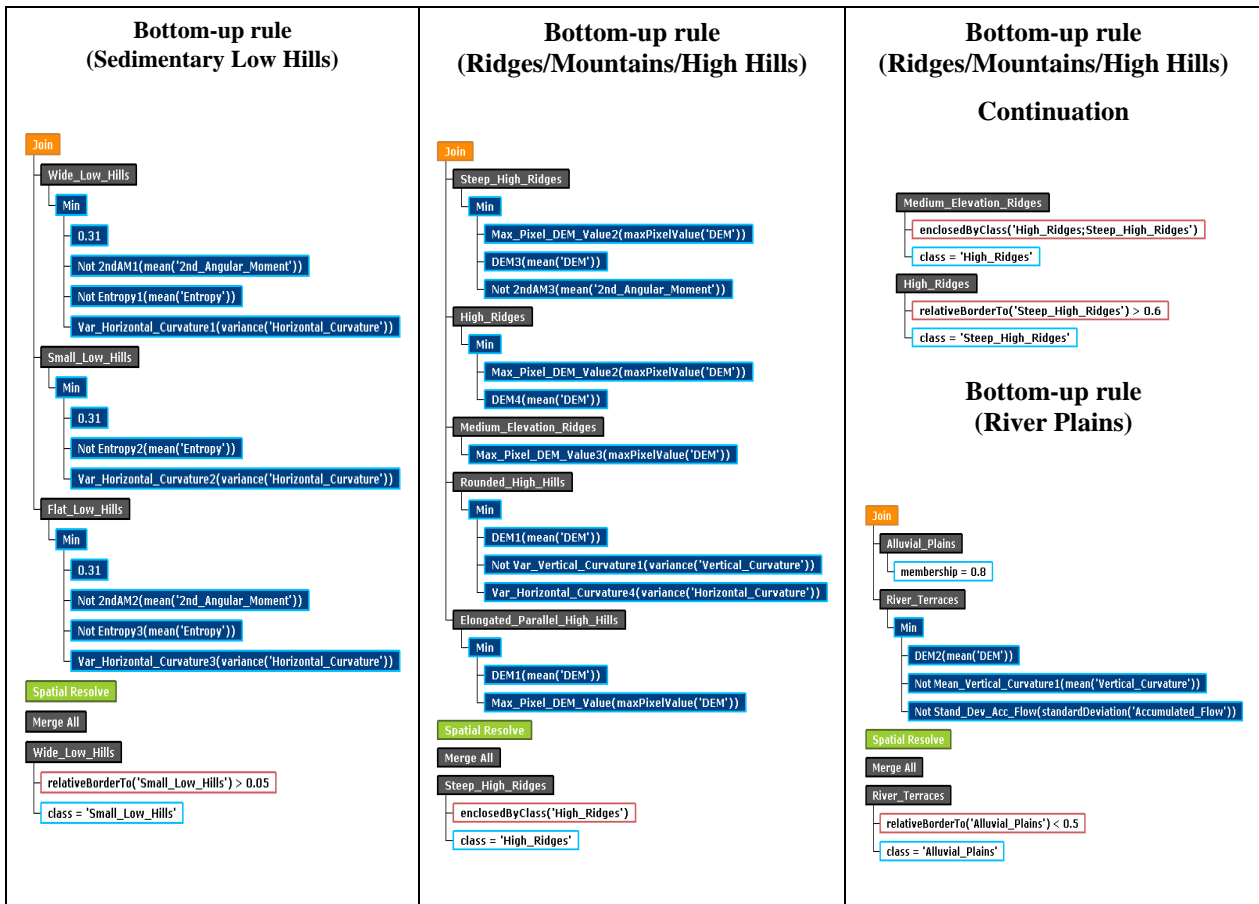
Fig. 4. Bottom-up rules for the morphographic landforms classification.

Finally, the 'Global merge' operation synthesizes the classes assignment and concludes the classification process. This operation is responsible for selecting the hypotheses with the highest membership values among all geographically concurrent hypotheses, and for discarding the other competing hypotheses.

Fig. 4 illustrates the bottom-up rules for the classification of morphographic landforms. The class 'Sedimentary Low Hills' was subdivided into 'Wide Low Hills', 'Small Low Hills', and 'Flat Low Hills'. For all of them, an upper threshold of 0.31 was established for the membership value. The former one was described by negative functions of the mean of both the second angular moment and entropy as well as by the variance of the horizontal curvature. 'Small Low Hills' comprised two attributes: the negative function of the mean of entropy and the variance of the horizontal curvature. On its turn, the membership to 'Flat Low Hills' was defined by the same attributes used to describe the class 'Wide Low Hills', but with different thresholds.

A topological operator − relative border to − was employed to refine the classification. This operator reclassified objects initially belonging to the class 'Wide Low Hills' as 'Small Low Hills', in the cases where the ratio of their adjacent boundaries to the class 'Small Low Hills' to their total perimeter surpassed 0.05. Before using topological operators, all spatial conflicts among objects must be previously solved, what is accomplished by the element 'Spatial Resolve'. After that, all neighbouring objects belonging to the same class were merged into one single object. 'Ridges/Mountains/High Hills' were detailed into five

subclasses. 'Steep High Ridges' was described by the mean and the maximum pixel value of the DEM and by the negative function of the mean of the second angular moment. 'High Ridges' and 'Elongated Parallel Hills' were discriminated by the mean and the maximum pixel value of the DEM. The class 'Rounded High Hills', on its turn, was identified by the mean of the DEM, the variance of the horizontal curvature and the negative function of the variance of the vertical curvature. In contrast, the membership to the class 'Medium Elevation Ridges' was solely defined by the maximum pixel value of the DEM. Topological operators were as well employed to refine the classification. Some objects of 'Steep High Ridges' and 'Medium Elevation Ridges' were reassigned to the class 'High Ridges' by the operator 'enclosedByClass', which seeks objects that are totally enclosed by a specific class and reclassifies them into another class (generally, the enclosing class).

The class 'River Plains' was subdivided into two morphographic subclasses: 'Alluvial Plains' and 'River Terraces'. The latter one was described by the mean of the DEM, the negative function of the mean of the vertical curvature and the negative function of the standard deviation of the accumulated flow. When the final membership value was equal or greater than 0.8, the objects were classified as 'River Plains', and as 'Alluvial Plains' otherwise.

5. RESULTS AND DISCUSSION

The validation results for the ASTER/Terra DEM obtained an RMSE of 9.38 m, and hence is consistent with similar

works. On the other hand, the t-Student test null hypothesis (H_0) of no difference between the DEM and GPS elevations was rejected. This is in some sense expected, because the DEM is influenced by the height of objects that extend above the surface, such as buildings and trees.

Fig. 5 presents the reference map and the classification of the morphographic landforms, respectively.

A contingency table was generated for the final classification (morphographic landforms) − showing errors of commission rows) and omission (columns) and the respective statistical validation indices (global accuracy and Kappa index).

As to the morphogenetic landforms classification, there were ambiguity errors between some classes. The class 'Sedimentary Low Hills' was confused with both 'Ridges/Mountains/High Hills' and 'River Plains'. This can be ascribed to the diversity of hill types, ranging from mildly flat to moderately dissected surfaces. On the other hand, the confusion between 'Ridges/Mountains/High Hills' and 'River Plains' was of reduced extent and limited to the contacts between these landforms
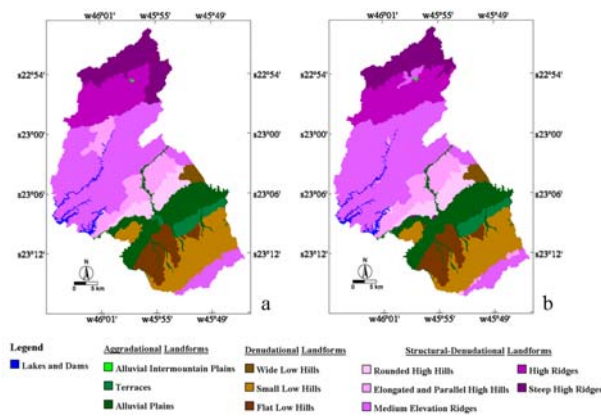

Fig. 5. a) Detailed morphographic landforms obtained from visual interpretation. b) Detailed morphographic landforms obtained from semi-automated classification.

In Fig. 5, it is noticeable that a few errors occurred between classes of different morphogenetic domains (for instance, between Quaternary sedimentary terrains that comprise the 'River Plains' and Precambrian crystalline terrains, which refer to the class 'Ridges/Mountains/High Hills'). This is due to the appropriate distinction of these domains in the stage of the morphogenetic landforms classification.

The calculated global accuracy was 0.89, very close to the Kappa index value (0.86). This small difference between these two indices can be explained by the low values of both commission and omission errors and also by the employed multilevel classification strategy, which is able to substantially reduce confusion among classes. The Kappa index value is regarded as indicating 'strong agreement'.

Analyzing the errors of commission (rows) and omission (columns), we realize that the class 'River Plains' was confused with the class 'Terraces'. This is mainly due to the similar geomorphometric characteristics (slope and elevation) between both classes. In the case of the structural-denudational landforms, considerable errors were observed between the following pairs of classes: 'Rounded High Hills' and 'Medium Elevation Ridges', 'Elongated and Parallel High Hills' and 'Medium Elevation Ridges' as well as 'Steep High Ridges' and 'High Ridges'. Although morphographically different, these classes present

morphometric similarities regarding particular attributes, such as elevation.

6. CONCLUSIONS

This work introduced InterIMAGE, a new knowledge-based image interpretation platform developed in compliance with the open source philosophy. As a practical application, this paper reported an experiment dealing with landforms classification.

The presented semi-automated classification approach allows the user to get acquainted with and to explore the behaviour of morphographic and morphometric characteristics of the concerned landforms classes. It also enables further investigations related to the connections between such characteristics and the landforms genesis (tectonics, basin sedimentation, etc). Conventional surveying and mapping methods do not render these inquiries available in such a short span of time.

In terms of interpretation strategy, InterIMAGE demonstrated to own a flexible architecture. The combination of a model-driven followed by a data-driven analysis, as performed by InterIMAGE, has the potential of an improved computational efficiency.

As the InterIMAGE Project evolves, the task of implementing more sophisticated knowledge-based models with this system will certainly become much easier. Further development of InterIMAGE is under way. Multi-temporal interpretation functionalities, automatic knowledge extraction functions, as well as built-in image processing operators are some of the developments under way.

REFERENCES

Baatz, M., Schäpe, A., 2000. Multiresolution segmentation—an optimization approach for high quality multi-scale image segmentation. In: Strobl, J., Blaschke, T., Griesebner, G. (Eds.), *Angewandte Geographische Informations-Verarbeitung XII*, Wichmann Verlag, Karlsruhe, pp. 12-23.

Blaschke, T., Burnett, C., Pekkarinen, A., 2004. New contextual approaches using image segmentation for object-based classification. In: De Meer, F., de Jong, S. (Eds.), *Remote Sensing Image Analysis: Including the spatial domain*, Kluver Academic Publishers, Dordrecht, pp. 211-236.

Bückner, J., Pahl, M., Stahlhut, O., Liedtke, E.-C., 2001. geoAIDA - A knowledge-based automatic image data analyser for remote sensing data. In: *The Second International ICSC Symposium on Advances in Intelligent Data Analysis - CIMA 2001*, Bangor, Wales, United Kingdom.

Graham, J., Jones, P., 1997. *Knowledge-Based Image Processing Systems*. Springer-Verlag, London.

Kressler, F.P., Steinnocher, K., 2006. Image data nd LIDAR – An ideal combination matched by object-oriented analysis. In Lang, S., Blaschke, T., Schöpfer, E. (Eds.), *The First International Conference on Object-based Image Analysis – OBIA 2006*, ISPRS, Viena.

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: *Third International Airborne Remote Sensing Conference and Exhibition*, Copenhagen, Denmark.