Expert Systems with Applications 39 (2012) 541–554

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

An open source object-based framework to extract landform classes

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

[a] National Institute for Space Research (INPE), Remote Sensing Division, PO Box 515, São José dos Campos, SP, Brazil
[b] Catholic University of Rio de Janeiro (PUC-Rio), Electrical Engineering Department, PO Box 38097, Rio de Janeiro, RJ, Brazil
[c] Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover, Nienburger Str. 1, D-30167 Hannover, Germany

ARTICLE INFO

Keywords:
Cognitive approaches
Object-based image analysis
Semantic network
InterIMAGE
Geomorphology

ABSTRACT

This paper introduces a new open source, knowledge-based framework for automatic interpretation of remote sensing images, called InterIMAGE. This framework exhibits a flexible modular architecture, in which image processing operators can be associated to both root and leaf nodes of a semantic network, which accounts for 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 are 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 to drive 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 information 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 (Ehlers, Janowsky, & Gähler, 2002).

Presently, however, the lack of efficient automatic image interpretation tools renders 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 turns out to be inappropriate to support critical decisions that may avoid or mitigate the effects of environmental degradation or unplanned urban expansion (Rego, 2003).

Currently, most remote sensing data analysis techniques require intense human intervention. The conventional digital images analysis platforms, which exclusively operate with statistical methods, have proved to be constrained for detecting targets of greater complexity. Their results usually require careful inspection by a human specialist for the identification and correction of interpretation errors (Bückner, Pahl, Stahlhut, & Liedtke, 2001).

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, Lang, Lorup, Strobl, & Zei, 2000; Carrion, Gianinetto, & Scaioni, 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 (Bückner et al., 2001; Clément, Giraudon, Houzelle, & Sandakly, 1993; Liedtke, Bückner, Grau, Growe, & Tönjes, 1997; Matsuyama & Hwang, 1990; McKeown, Harvey, & McDermott, 1985; Sagerer & Niemann, 1997; Schiewe, Tufte, & Ehlers, 2001; Witlox, 2005). In this approach human experts' knowledge is organized in a knowledge base (Graham & 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, led 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 article we describe the interpretation strategy implemented in the system (Section 1.2) and a brief

* Corresponding author. Tel.: +55 12 3208 6428; fax: +55 12 3208 6488.
E-mail address: almeida@dsr.inpe.br (C.M. Almeida).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.07.044

overview on the study area is provided in Section 1.3. Section 2 basically describes the data acquisition and pre-processing. Section 3 presents 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 4 describes the results of this experiment, and a critical evaluation is conducted on the potential and drawbacks of the input data and methodological procedures. Finally, some conclusions and directions for future work are drawn in Section 5.

1.1. System description

InterIMAGE is based on the software GeoAIDA (Bückner et al., 2001), developed at the TNT Institute of the Leibniz Universität Hannover, Germany, and it inherited from that system the basic functional design, knowledge structures, and control mechanisms (Fig. 1). A new graphical user interface, a knowledge model debugging tool, and some image processing operators are also available in InterIMAGE (Laboratório de Visão Computacional – LVC, 2009).

In short, InterIMAGE implements a specific image interpretation strategy, which is based on 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 implement virtually any decision rule based on spectral, texture or structural features and form 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 with 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 combinatorial 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 enables the integration of any 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 allows 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 graphical user interface (GUI).

A semantic network contains nodes and edges; nodes represent concepts and edges represent the relations between the concepts. The network is actually a connected graph without 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 such data that represents the result of the concerned analysis. In InterIMAGE, the result contains a structural description (an instance network) and thematic maps (Fig. 2). The final and all intermediate results,

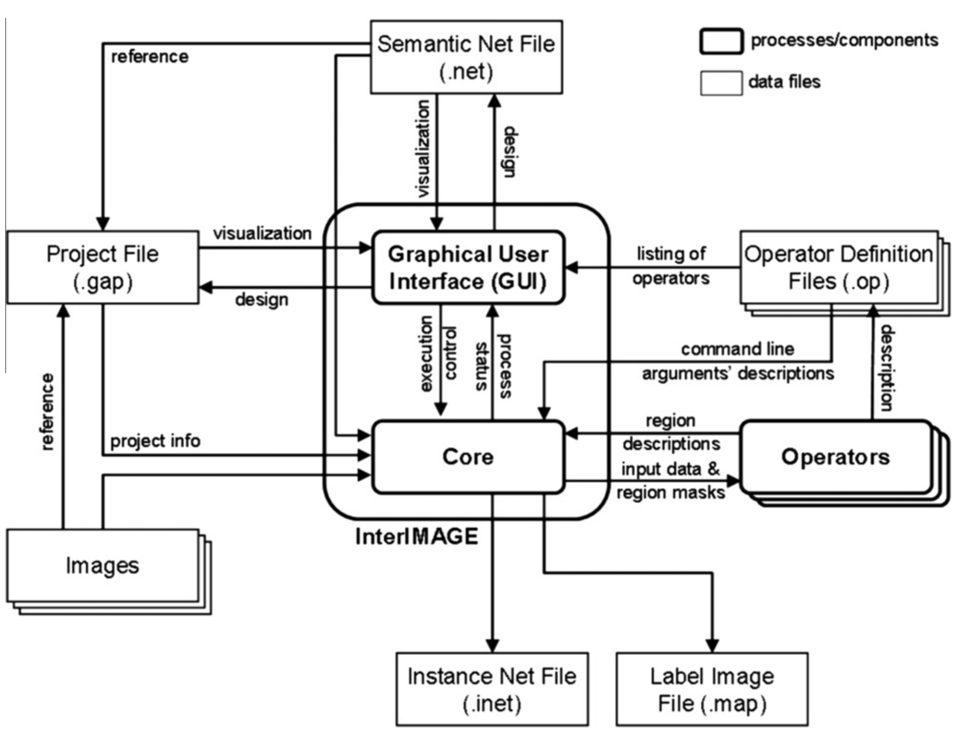


Fig. 1. InterImage system architecture.

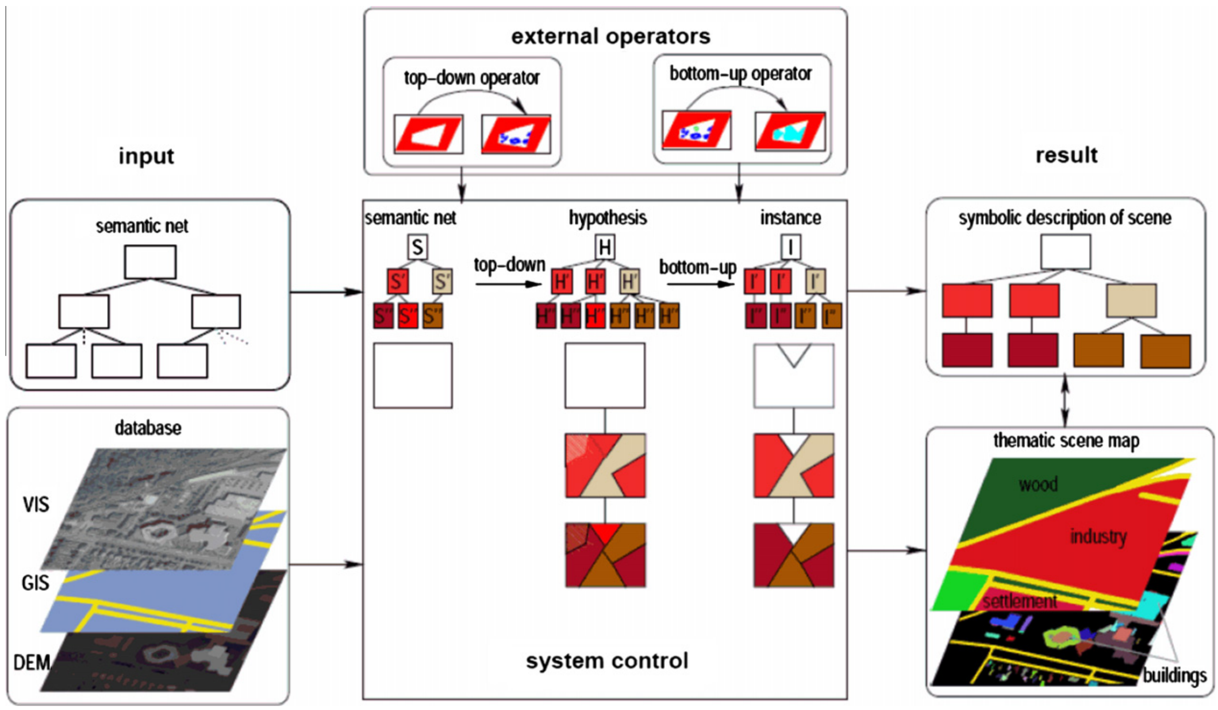


Fig. 2. Components of the interpretation strategy. Source: Pahl (2003).

in terms of region descriptions, are stored in XML format, and can be used for further external investigations.

The analysis process performed by InterIMAGE has two steps: a bottom-up and a top-down one. The top-down step is model-driven (Fig. 3) 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 (classification) 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.

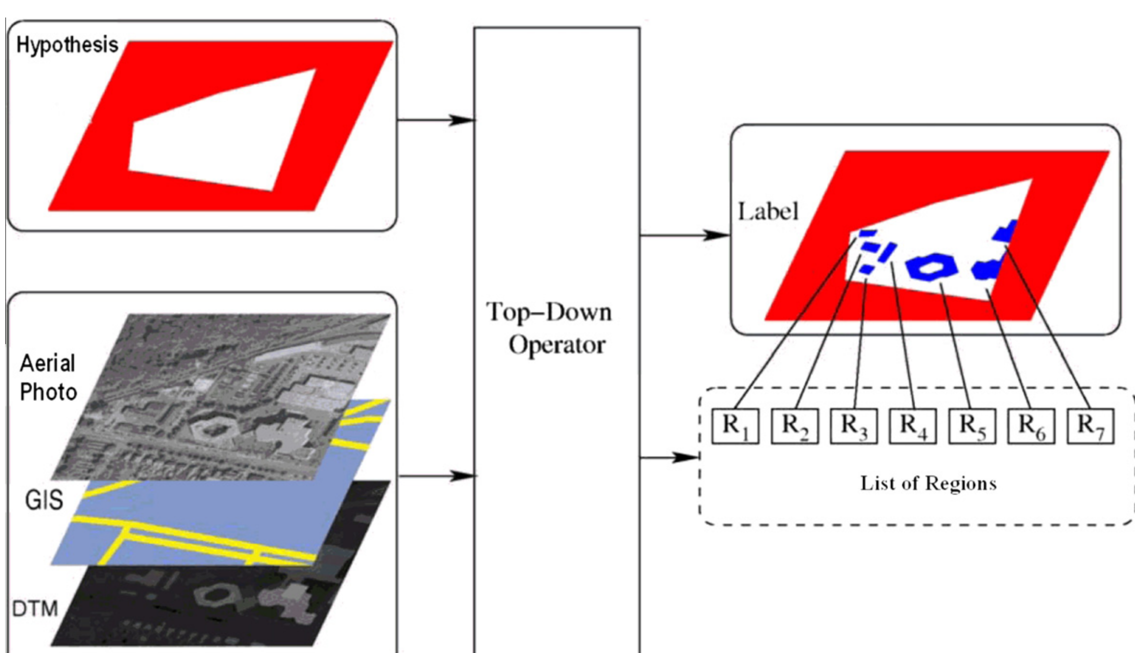


Fig. 3. Process flow of the top-down operator: this operator receives as input an image segment to be processed. An optional mask can reduce the area to be processed by the operator. The top-down operator generates as output a labeled image and the corresponding list of classified regions, which thoroughly describes the generated segments (objects). Source: Pahl (2003).

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. The sub-region hypotheses can be defined by means of consistency measurements. If certain textural metrics, for instance, allow 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 interpretation turns from model-driven into data-driven (bottom-up). The decision rules for the bottom-up step (Fig. 4) are defined in a user friendly graphical language that provides functions for deciding between spatially concurrent hypotheses generated in the top-down step.

1.3. Study area

The municipality of São José dos Campos, with a total area of 1098.6 km^2, is located in the east of São Paulo State, southeastern Brazil (Fig. 5). The area of interest was selected for analysis due to its landform diversity, the availability of previous local studies (Florenzano & Csordas, 1993; Verdade & Hungria, 1966), and the good accessibility for field work.

The study area is embedded in a mountain range, comprising the Mar and Mantiqueira Ridges. This range is the most prominent landform in eastern South America and it dates back to the Precambrian era (Almeida & Carneiro, 1998). Three events were responsible for its physiographic settings and hence defined the present lithological and geomorphological characteristics: (i) continuous interactions between continental plates (in the Proterozoic era), which formed and successively reworked a series of accretionary, collisional or transpressional mobile belts, (ii) the erosion and leveling of a preexisting large region, with mainly igneous and metamorphic rocks, informally known as 'Japi Surface', during the Upper Cretaceous to Lower Eo-Tertiary ages; and (iii) taphrogenesis (in the early Paleogene period), sedimentation, and half-grabens filling (Almeida, 2000; Almeida & Carneiro, 1998).

Regarding its lithology, the area of investigation is composed of crystalline rocks (igneous and metamorphic) belonging to the following complexes: (i) Amparo, (ii) Embu, (iii) Paraíba do Sul, and (iv) Paraisópolis. Sedimentary rocks are also found in the Taubaté Formation (Late Tertiary) and Quaternary deposits (alluvia). The above mentioned complexes mostly consist of gneisses (Archaen/mid-Proterozoic), sin- and post-tectonic granite suites (Late Proterozoic), both derived from crust movements of the Brasiliana Orogenesis (Precambrian) (DNPM, 1983).

In the crystalline rocks, the geomorphological features comprise high hills and mountains (Florenzano & Csordas, 1993). According to Almeida and Carneiro (1998), plateaus and steep slopes lie on more resistant rocks, while rocky lineaments and the drainage network are influenced by faults, fractures, and shear zones. On sedimentary terrains, Alluvial Plains (Florenzano & Csordas, 1993), terraces (Verdade & Hungria, 1966), and Tertiary hills (Florenzano & Csordas, 1993) are found. It is worth highlighting the presence of several patterns of alluvial intermountain plains. Verdade and Hungria (1966) acknowledged the existence of two levels of river terraces in the right margin (southwest) of the Paraíba do Sul river, that crosses the study area in the NW–SE direction. This can be ascribed to the river trend in shifting to the northwest direction (Verdade & Hungria, 1966), caused by the north-northwest inclination of the half-graben (Almeida, 2000).

2. Data acquisition, data preprocessing, and reference data

2.1. Data acquisition

The data used comprised: (i) ASTER/Terra VNIR images (bands 3N and 3B, within the range 0.78–0.86 µm, with a nominal spatial

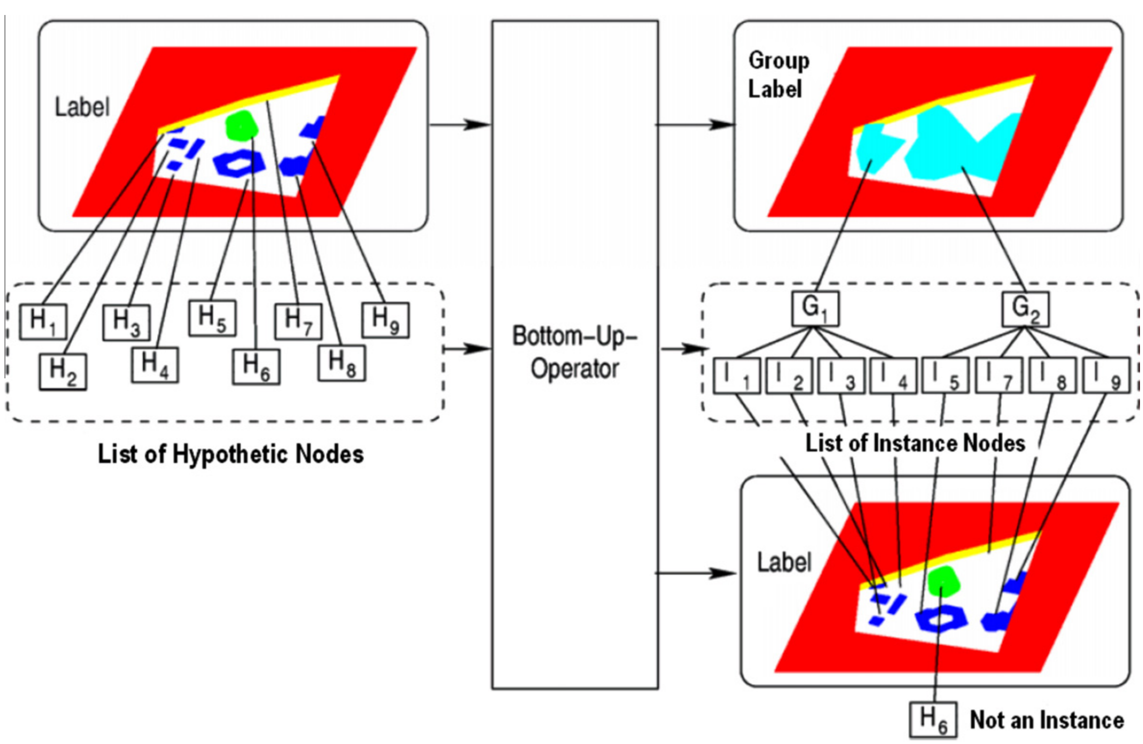


Fig. 4. Process flow of the bottom-up operator: this operator receives a list of hypothetic nodes with the corresponding labeled image. These nodes are evaluated, which may change the corresponding labeled images, and they can be grouped into classes of objects of a higher level. A labeled image is obtained with the generated groups. Source: Pahl (2003).

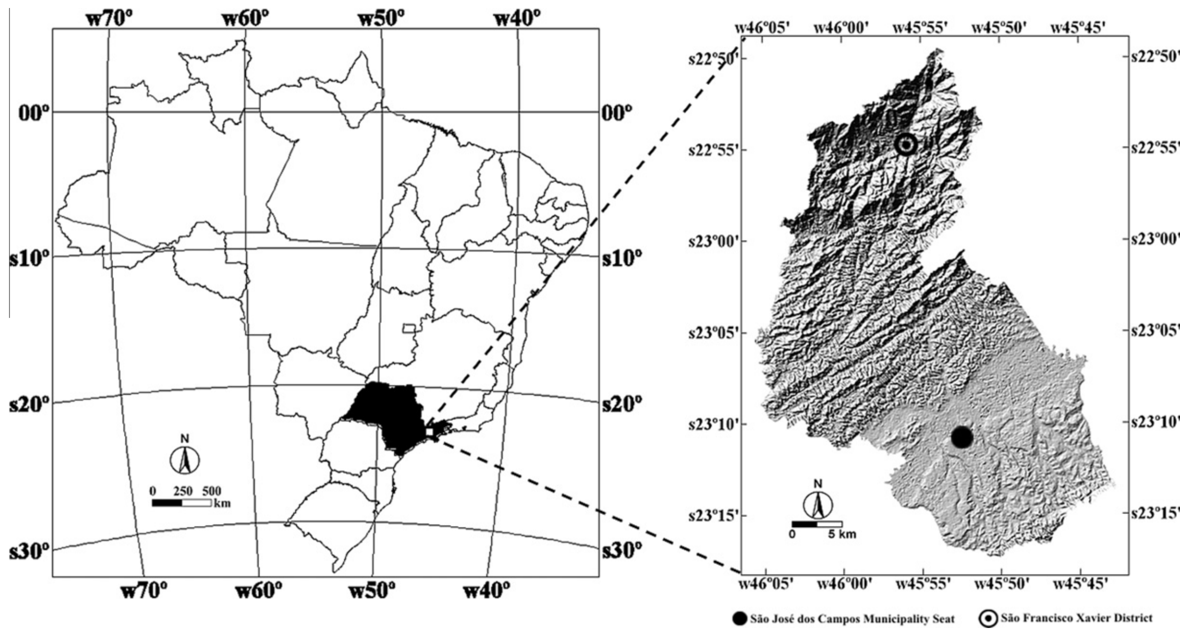


Fig. 5. Study area: Brazil and São Paulo State (in black) on the left side, and on the right side the municipality of São José dos Campos (view of its shaded relief). Source: Adapted from INPE – Instituto Nacional de Pesquisas Espaciais (2006).

resolution of 15 m); (ii) a 1:10 000 vector file of the street network; (iii) a 1:10 000 vector file of water streams; (iv) a 1:10 000 vector file of 10 m interval contour lines; (v) scattered elevation data points; (vi) a 1:50 000 vector file of geological units, and (vii) 109 GPS points with orthometric heights.

The ASTER/Terra images, acquired on 08/31/2004, are of processing level L1B (geometric and radiometric correction) and they present a base/height ratio of 0.6 (Abrams, Hook, & Ramachandran, 1999). The vector files belong to the geographic database ''Cidade Viva'', issued by the local government of São José dos Campos (PMSJC). The GPS coordinates were jointly obtained by the Aeronautics Institute for Advanced Studies (IEAv), the National Institute for Space Research (INPE), and the Foundation for Space Science, Applications and Technology (FUNCATE).

The Zscreen 2000 was used for stereoscopic visualization, and the softwares PCI Geomatica 10.0.3, ENVI 4.3, and SPRING 4.3.3 were used for image processing as well as for the generation of the digital elevation model (DEM) and the textural and geomorphometric variables.

2.2. Data preprocessing

Data preprocessing comprised the DEM generation, its validation, and the extraction of textural and geomorphometric variables, which were further used for the multilevel segmentation and the object-based classification. For generating the DEM, the software Geomatica 10.0.3 (module OrthoEngine) was used, in which the following processing operations were executed: (1) collection of ground control points (GCPs) and tie points (TPs); (2) estimation of the mathematical model parameters and the stereo pair orientation; (3) generation of epipolar images; (4) calculation of parallaxes through stereo-correlation, and (5) DEM generation. In steps 1 and 2, the ASTER/Terra sensor attitude and ephemeris data, Toutin's Model (Toutin, 2004a, 2006), 43 three-dimensional GCPs, and 90 bi-dimensional TPs were used.

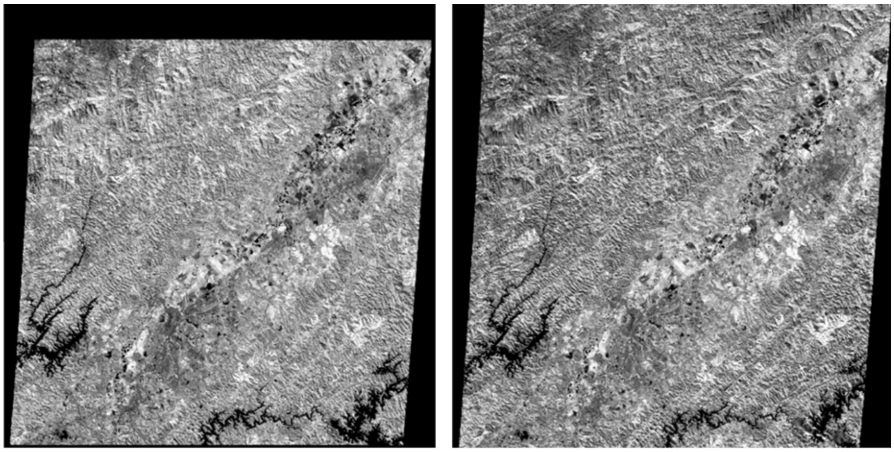
The employed model considers the collinearity (single images) and coplanarity (stereo pair) equations, and is thus based on a rigorous photogrammetric model. Since the parameters for correcting distortions (associated with the platform/sensor combination, Earth rotation, and cartographic projection) are correlated, Toutin's model reduces such parameters to a decorrelated set, upon which all the above mentioned corrections are simultaneously made (Toutin, 2004a). The model positional accuracy can be improved with the use of GCPs, employed in an iterative refinement procedure based on the least square method (Toutin, 2002, 2004a, 2004b, 2006).

The GCPs planimetric coordinates were collected directly on the streets network file, and the respective elevation coordinates were obtained in an ancillary DEM, generated as well in Geomatica 10.0.3 for this specific purpose, using the contour lines and the scattered elevation data points by means of bilinear interpolation. The TPs are simultaneously collected on both images of the stereo pair, and hence provide more robust models.

After the stereo pair orientation and the generation of epipolar images (Fig. 6a and b), the parallaxes were automatically calculated. This task was accomplished by means of search and correlation windows, which locate homologous pixels on both images (Ehlers & Welch, 1987). The correlation measure employed by the OrthoEngine algorithm is the normalized cross-correlation coefficient (PCI Geomatics, 2006). Further information on this calculation can be found in Russ (1998).

The DEM validation used the 109 GPS points, previously mentioned in Section 2.1. The ASTER/Terra DEM and these GPS points were referenced to the same horizontal datum (SAD69/Brazil) and to the same cartographic projection system (UTM). The discrepancies (residuals) between the ASTER/Terra DEM and the GPS points elevations were estimated. Further on, a descriptive analysis of residuals was executed followed by a *t-Student* test, meant to evaluate if the residuals mean showed a trend. The obtained results were cross-checked with other scientific works, which also evaluated the elevation accuracy of ASTER/Terra data.

The validated DEM, generated with the same resolution as the stereo pair images (15 m), was not subject to any kind of post-processing. This DEM 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



a b

Fig. 6. (a) Left epipolar image – Band 3N. (b) Right epipolar image – Band 3B.

Table 1
Input variables of the classification model.

Type of variable	Description
Geomorphometric	Elevation (ASTER/Terra DEM) Slope Vertical curvature Horizontal curvature
Textural	Entropy
Others	Drainage density Lakes and dams

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'.

Slope is one of the components of an isotropic two-dimensional vector of elevation gradient, which is obtained from the first DEM derivative. The length of this component indicates slope steepness. The second derivative provides the horizontal and vertical curvatures (Valeriano, 2003; Valeriano & Carvalho, 2003). The vertical curvature indicates areas of gravitational acceleration and disacceleration and characterizes slopes in different types: concave (negative values), convex (positive values) or straight (zero values) (Shary, 2007). The horizontal curvature highlights the convergence (negative values) and divergence (positive values) lines of surface runoff (Shary, 2007).

In order to extract the textural variables, the gray level co-occurrence matrix (GLCM) was used. A GLCM is a two-dimensional histogram of gray levels for a pair of pixels which are separated by a fixed spatial relationship, defined by the interpixel distance (δ) and orientation (θ) observed in the spatial window of interest. Entropy measures the deviation from the uniform texture or lack of 'organization' of orientations. It is high when the elements of GLCM have relatively equal values, and low when the elements are close to either 0 or 1 (i.e. when the image is uniform in the window) (Clausi, 2002; PCI Geomatics, 2006).

Drainage density was generated using a coarse segmentation level of the shaded relief and the vector layer of water streams (Section 2.1) as input data for classification. For each segmented 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 in '.pfm' format. 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'.

2.3. Landform classes and reference map

The landform classes were based on the legend of the São Paulo State geomorphological map (IPT, 1981), issued at scale 1:1 000 000. For the final classes selection (Table 2), two procedures were adopted: (i) a visual interpretation of the ASTER/Terra epipolar stereo pair, and (ii) field work, which was important for the discrimination of the different river terraces levels, as reported in Verdade and Hungria (1966).

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 morphochronology) and is employed in geomorphological mapping and natural resources and hazards evaluations (Verstappen & van Zuidam, 1991). 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). Further information on the lithology of the study area obtained from the vector file of geological units (Section 2.1) can be found in Table 2.

Due to the unavailability of detailed geomorphological maps for the area (1:100 000 or greater scales), a reference map of landforms was specially produced in a 3D digital station by an experienced geomorphologist, 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 this reference map was used for validating the classification result.

3. 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: (i) 'Sedimentary Low Hills'; (ii) 'Ridges/Mountains/High Hills'; (iii) 'Alluvial Plains'; (iv) 'Alluvial Intermountain Plains', and (v) 'Lakes and Dams' (Fig. 7).

Table 2
Landforms classes, their main lithologic characteristics and prevailing morphogenetic processes.

Landforms0	Lithology	Prevailing morphogenetic processes
Alluvial Plains	Sandy clay sediments	Aggradation
Alluvial Intermountain Plains	Sandy sediments	
Sedimentary Low Hills	Sandy sediments and sandy clay sediments	Denudation
Ridges/Mountains/High Hills	Migmatites, gneisses, schistes, philites, and granite suites	Tectonism-Denudation

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. 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', 'Alluvial Plains', and 'Alluvial Intermountain Plains'. This operator performs a segmentation of the ASTER/Terra DEM, using the region-growing algorithm (Baatz & Schäpe, 2000), generating for each segment one hypothesis of each of those concepts – geographically coincident hypotheses. 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 & Steinnocher, 2006).

Fig. 8 shows the bottom-up rules – 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 (LVC, 2009).

As shown in Fig. 8, 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.

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). In Fig. 8, these attributes were coded respectively as 'media0', 'media5', and 'media1'. The final membership to 'Sedimentary Low Hills' was defined as an aggregation of the values obtained through the membership functions (or fuzzy sets) 'dem_col', 'entro_col' and 'decli_col' (Fig. 9). The element 'Min' refers to the fuzzy minimum operator (Bonham-Carter, 1994) and was employed for the aggregation of membership values. The fuzzy membership functions associated to each of the attributes describing the class 'Sedimentary Low Hills' can be seen in Fig. 9.

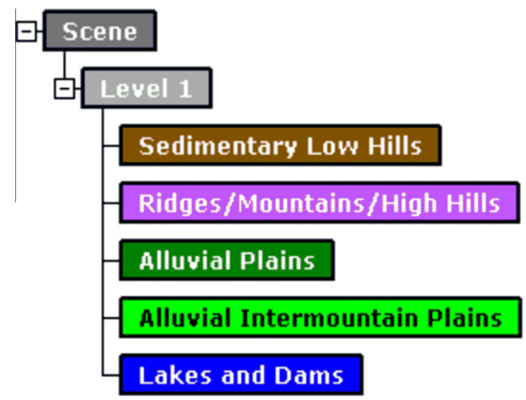


Fig. 7. Semantic network for the classification of landforms in São José dos Campos.

The membership to 'Ridges/Mountains/High Hills' is also based on three attributes. The first one is the drainage density, coded as 'media4'. 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 2.2). 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 ('media0'), and the third one to the mean of slope ('media1'). The final membership to 'Ridges/Mountains/High Hills' is defined by an aggregation (minimum) of the values obtained through the membership functions 'dem_ser_mor' and 'decli_ser_mor' (Fig. 10).

The membership to 'Alluvial Intermountain Plains' is once more based on three attributes: amplitude of the DEM ('amp0'), mean of the DEM ('media0'), and mean of slope ('media1'). The amplitude refers to the difference between the maximum and minimum height values observed within a hypothesis (segment). Fig. 11 presents the membership functions associated to the attributes describing the class 'Alluvial Intermountain Plains'. On its turn, the membership to 'Alluvial Plains' is based on two attributes: mean of entropy ('media5') and mean of slope ('media1'). The membership functions associated to these attributes are shown in Fig. 12.

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. In a similar way, the scene background, acknowledged as a class, was also identified through a single attribute. This attribute considered the lower threshold of the DEM values in terms of its background, thus creating a mask over the study area.

Finally, the 'Spatial Resolve' operator synthesizes the class assignment and concludes the classification process. This operator is responsible for selecting the hypotheses with the highest membership values among all geographically concurrent hypotheses, and for discarding the other competing hypotheses. This classification operator is followed by the element 'Merge All', in which all neighboring objects assigned to the same class are merged into one single object.

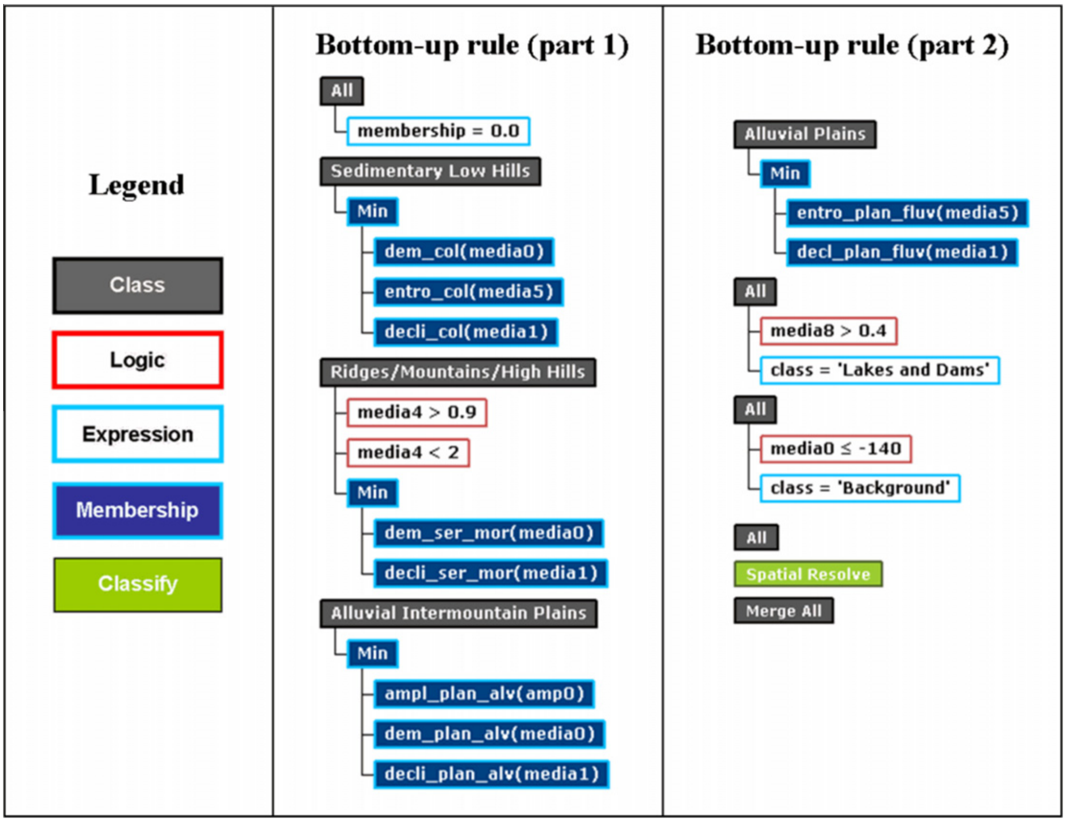


Fig. 8. Bottom-up rules for the landforms classification.

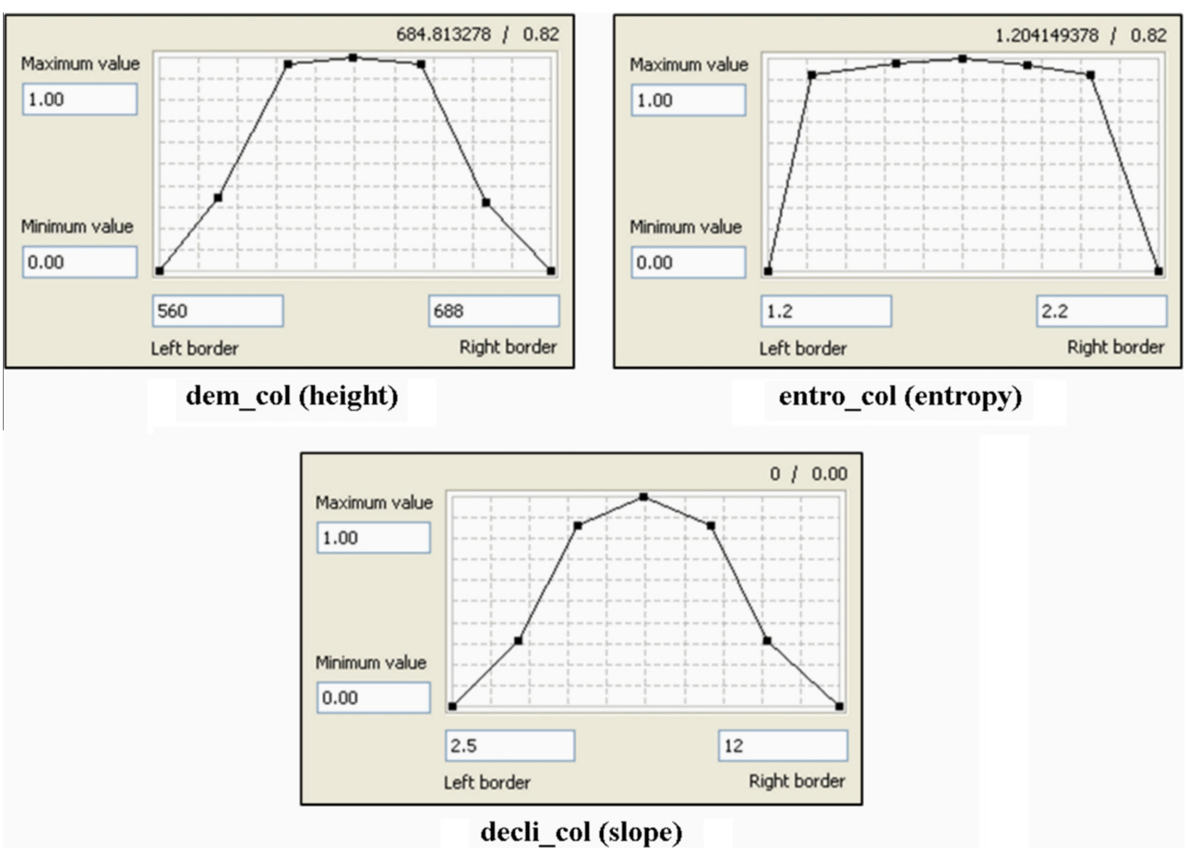


Fig. 9. Membership functions employed for identifying the class ‘Sedimentary Low Hills’.

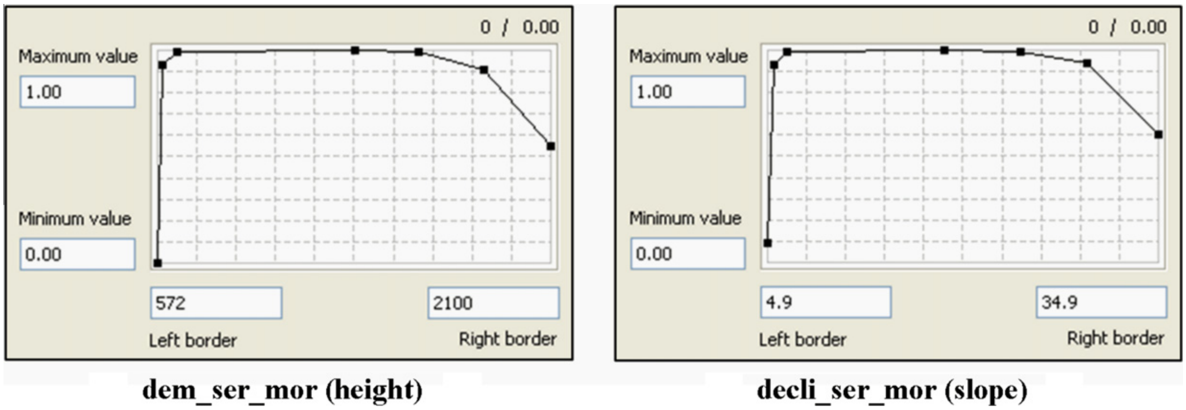


Fig. 10. Membership functions employed for identifying the class 'Ridges/Mountains/High Hills'.

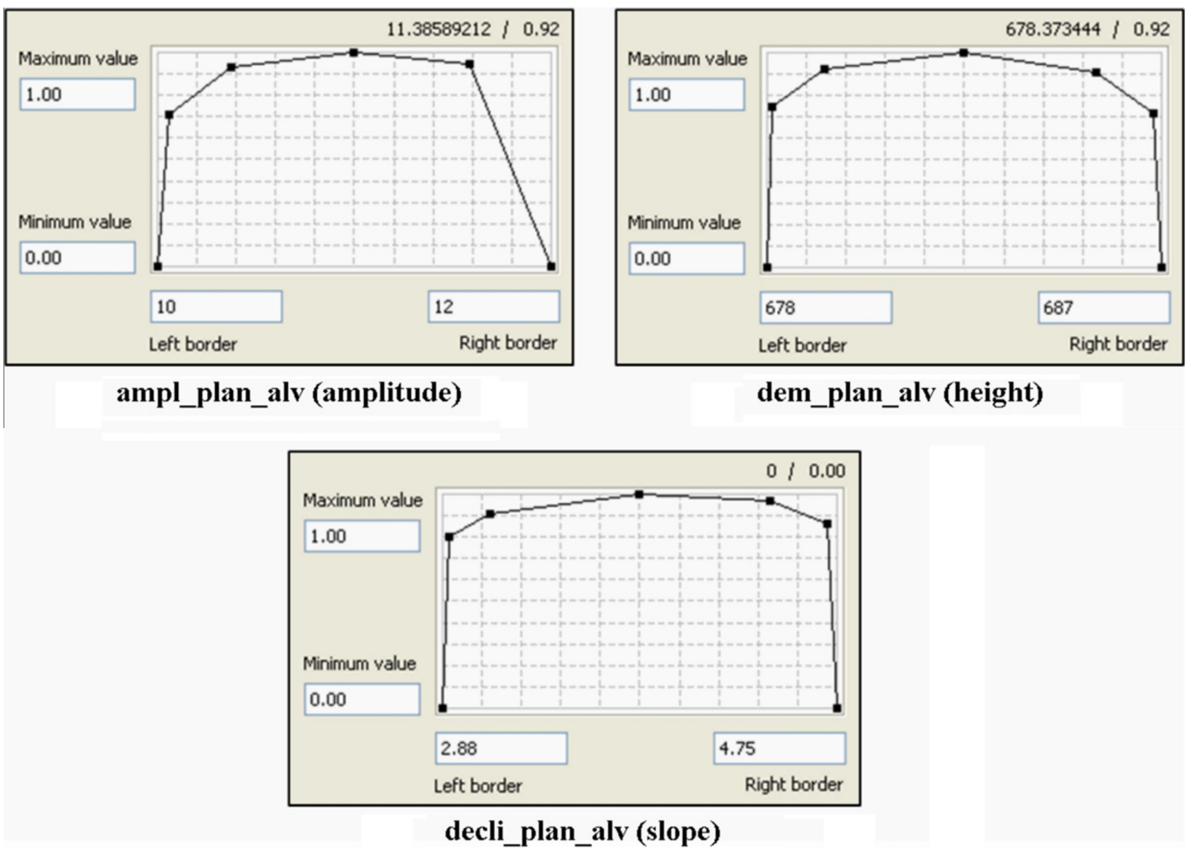


Fig. 11. Fuzzy membership functions employed for identifying the class 'Alluvial Intermountain Plains'.

In order to refine the classification initially produced, a sequence of topological operators was added at the end of the semantic network. They basically consist of post-processing routines designed to reclassify wrongly classified objects according to contextual rules, by taking into consideration the classes of neighboring objects. Only one type of topological operator was employed, namely 'enclosed by class', which seeks objects that are totally enclosed by one or more specific classes and reclassifies them into another class (generally, the enclosing class). Please refer to LVC (2009) for additional information. Fig. 13 shows the topological operators used in this experiment, as they are seen in the interface of the bottom-up decision rule.

4. Results and discussion

The validation results for the ASTER/Terra DEM obtained a root mean square error (RMSE) of 9.38 m, and hence is consistent with similar works (Oliveira & Paradella, 2008, 2009; Toutin, 2008). This level of accuracy is up to the elevation standards required for mapping products at a 1:100 000 scale. Since the geomorphological maps presented in this work consist of thematic products that are not concerned with elevation accuracy, it is reasonable to assume that these maps are up to the standards of regional mapping, with scales ranging from 1:50 000 to 1:100 000, what complies with the works of Bolten and Bubenzer (2006) and Bubenzer and

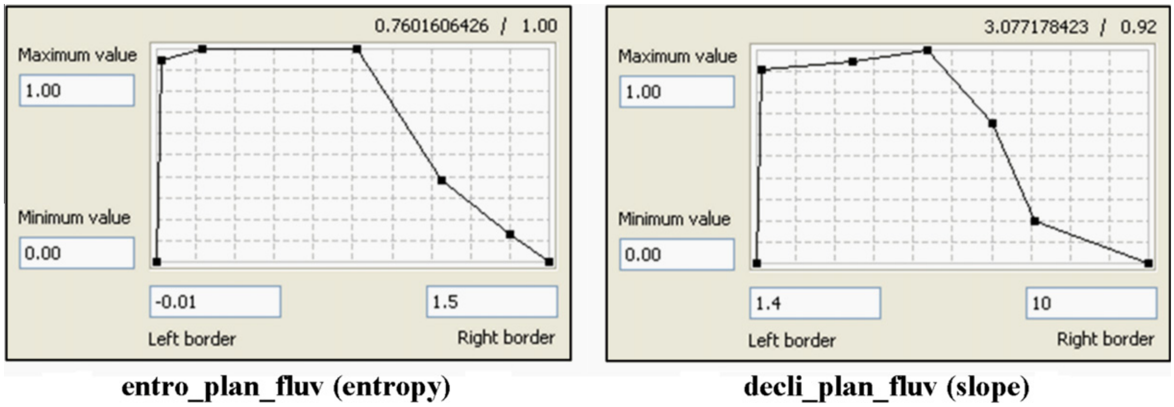


Fig. 12. Fuzzy membership functions employed for identifying the class 'Alluvial Plains'.

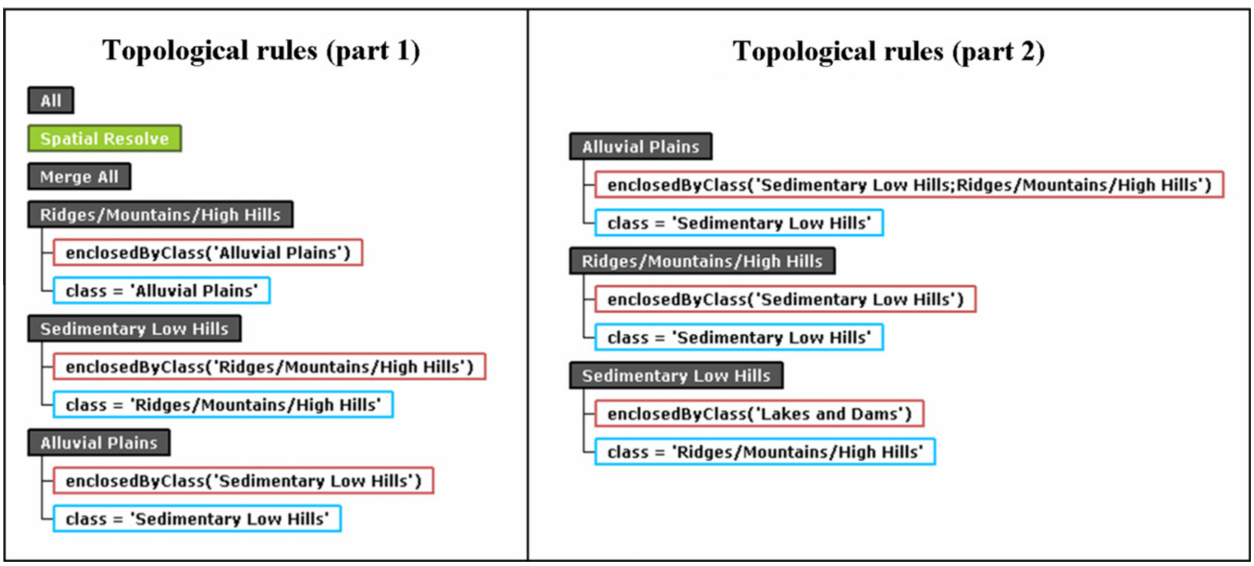


Fig. 13. Topological rules for the landforms classification.

Bolten (2008). According to Siart, Bubenzer, and Eitel (2009), the intended scale always governs quality thresholds, and, in reverse, data quality predetermines the type of detectable landforms.

The statistical indices derived from the discrepancies between the DEM and the GPS points and employed for the DEM accuracy assessment are presented in Table 3. Additionally, a two-tailed *t-Student* test ($\alpha = 10\%$) was also applied in order to check whether a systematic trend could be detected. The *t-Student* test null hypothesis (H_0) of no difference between the DEM and GPS elevations was nevertheless rejected. This is in some sense to be expected, because the DEM is influenced by the height of objects that extend above the surface, such as buildings and trees. Fig. 14 shows the DEM with the study area boundary superimposed as a black line, the 109 GPS points (black dots), and areas with no data (black circles), which mainly correspond to water bodies and shadow. Since the variance between pixel gray values in such targets is close to zero, the stereo-correlation and the automatic extraction of parallaxes becomes unfeasible.

In Fig. 15a the reference map for the study area obtained from a 3D visual interpretation of the ASTER stereo pair (epipolar images of bands 3N and 3B) using a digital photogrammetric station is depicted, whereas Fig. 15b shows the thematic map generated from the final classification performed in InterIMAGE. Illustrative photos of the morphogenetic landforms, taken during the field work, are presented in Fig. 16 together with their respective inset boxes with the shaded relief image. These photos are detached from the final classification of the landforms draped over shaded relief.

Table 3
Statistical indices used for the ASTER/Terra DEM validation.

ASTER/Terra DEM	
Number of samples (GPS points)	109
Minimum error – $\Delta_{minimum}$ (m)	−18.40
Maximum error – $\Delta_{maximum}$ (m)	30.60
Mean error – Δ_{mean} (m)	4.14
Standard deviation – S_{Δ}	8.40
Root mean square error – RMSE (m)	9.38
$\lvert t_{sample} \rvert$	5.146
$t_{(n-1,\ 5\%)}$	1.659

Table 4 shows the confusion matrix between the ground truth and the classification. All the classified objects (segmented regions) excluding the ones used for training were taken into account for assessing the accuracy indices (global accuracy and Kappa index). The rows of the matrix show the classification results obtained with InterIMAGE and the columns show the data from the reference map. The global result in terms of the overall coincidence (global accuracy) between the reference map and the classification is 95%, and the Kappa index attained 86%. Considering the

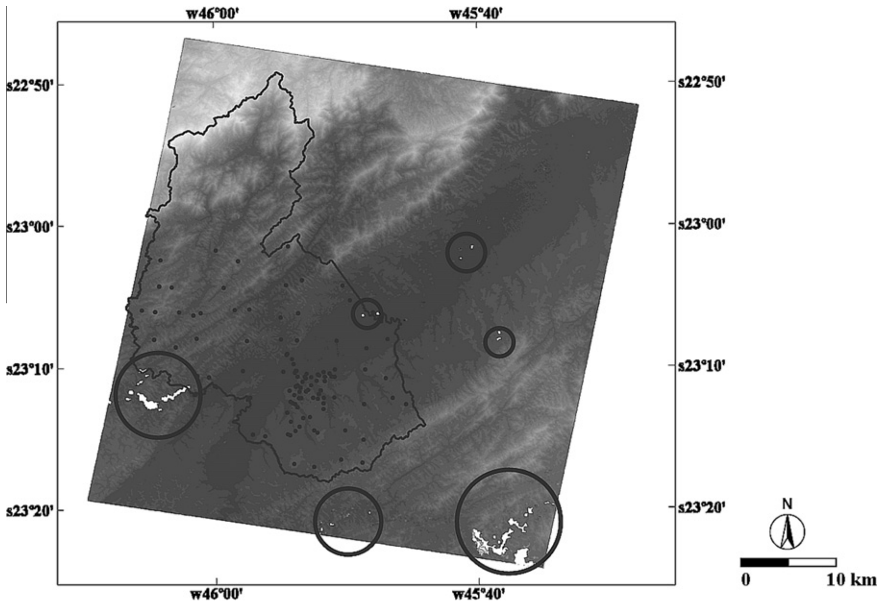


Fig. 14. DEM, study area boundary (black line), no data areas (black circles), and GPS points (black dots).

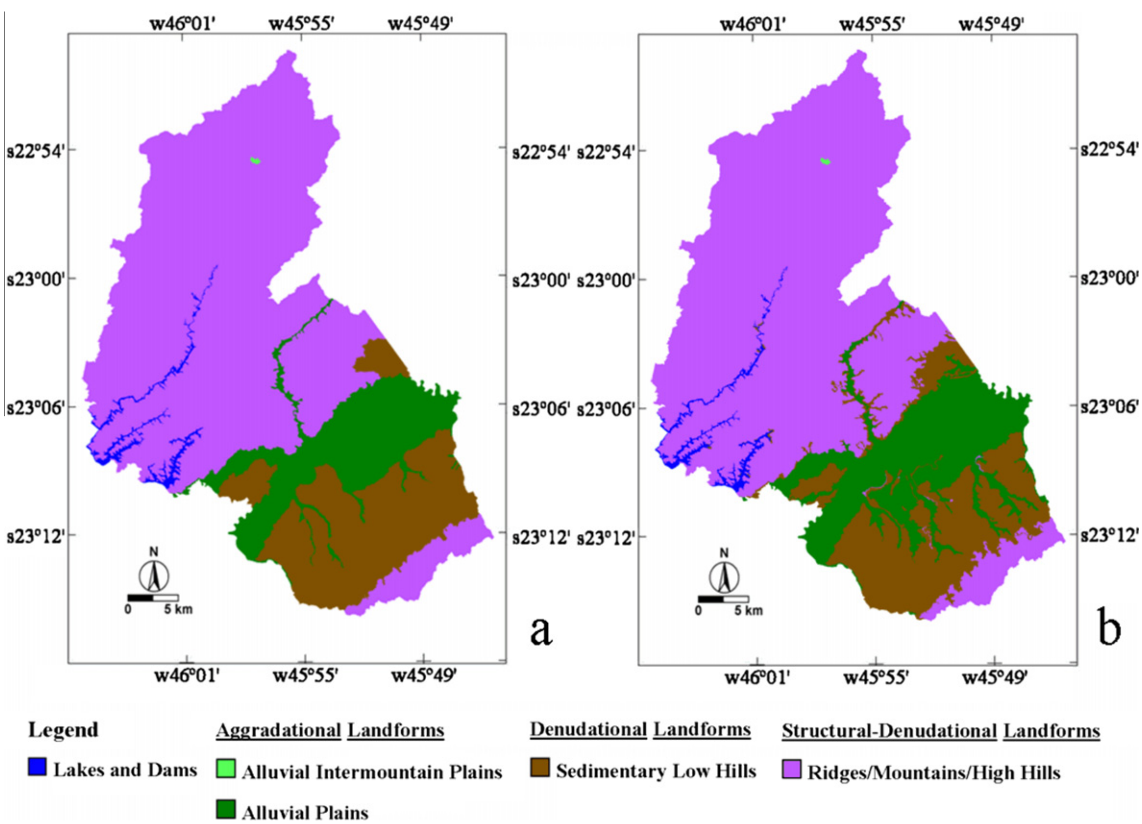


Fig. 15. (a) Morphogenetic landforms obtained from visual interpretation. (b) Morphogenetic landforms obtained from semi-automated classification.

commission and omission errors, it is possible to observe that the class 'Sedimentary Low Hills' presented confusion with both 'Ridges/Mountains/High Hills' and 'Alluvial Plains'. This can be ascribed to the diversity of hill typologies, ranging from mildly flat to moderately dissected surfaces. On the other hand, the confusion between 'Ridges/Mountains/High Hills' and 'Alluvial Plains' was of reduced extent and limited to the contacts between these units. The Kappa index is regarded as of 'strong agreement' (Landis & Koch, 1977), and the producer's and user's accuracies (Table 4) suggest a substantial agreement between the reference map and the classification result.

5. Conclusions

This article introduced InterIMAGE, a new knowledge-based image interpretation platform developed in compliance with the

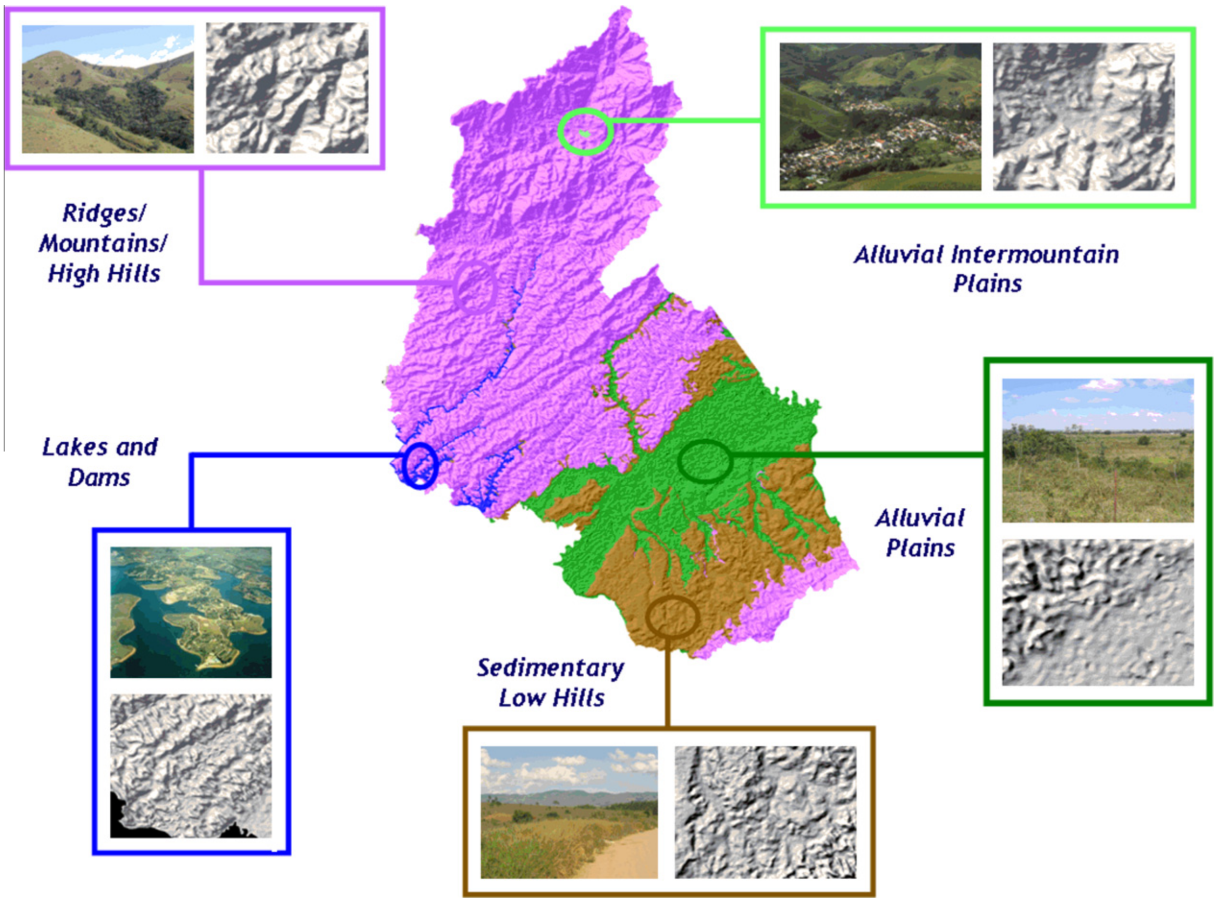


Fig. 16. Final classification of the morphogenetic landforms draped over shaded relief of the study area, illustrative photos of the landforms and respective inset boxes with the shaded relief image.

Table 4
Error matrix, global accuracy, producer's and user's accuracies, and Kappa index for the landforms classification.

Classification	Reference map (visual interpretation)					
	Ridges/Mountains/ High Hills	Sedimentary Low Hills	Alluvial Plains	Alluvial Intermountain Plains	Producer's accuracy	User's accuracy
Ridges/Mountains/High Hills	4192	13	1	0	**0.97**	**0.99**
Sedimentary Low Hills	99	595	26	0	**0.84**	**0.83**
Alluvial Plains	9	100	317	0	**0.92**	**0.74**
Alluvial Intermountain Plains	0	0	0	2	**1.00**	**1.00**
Global accuracy	**0.95**					
Kappa index	**0.86**					

open source philosophy and with 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 in comparison to commercially available softwares for object-based image analysis. In this way, InterIMAGE offers innovative knowledge modeling possibilities.

The purpose of this work was to develop a semi-automated method for geomorphological mapping in InterIMAGE, exploring new ways of data fusion, integrating an ASTER/Terra spectral band (3N), a DEM generated from a stereo pair of bands 3N and 3B, geomorphometric and textural variables extracted from this DEM, and GIS vector layers. In face of the obtained results, the following conclusions can be drawn.

In spite of its medium spatial resolution (15 m), the ASTER/Terra DEM was suitable for the identification of subtle landforms, like terraces (contained within the Alluvial Plains) and a tiny alluvial intermountain plain. In addition to their high discrimination power, the ASTER/Terra images are price worthy, enable stereoscopic vision, and also an appropriate integration between multispectral and elevation data. This integration is able to eliminate errors prone to arise when handling data from different sources (DEMs, multispectral images, etc.) and of different spatial resolution. Future works dealing with semi-automated identification of landforms can now rely on ASTER GDEM, which is globally available and will enable time-saving classification procedures.

The hierarchical semantic network could embody rules representing the topological, contextual, geomorphometric, and textural characteristics of the corresponding landforms. The cognitive approach employed in this work enabled the expert knowledge modeling with a very good degree of fidelity, what led to a high performance in the semi-automated classification result, confirmed by the strong agreement of the Kappa index.

The presented semi-automated classification approach allows the user to get acquainted with and to explore the behavior of morphographic and morphometric characteristics of the concerned landform 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 deliver these inquiries in such a short span of time.

Another major contribution of this approach is that it generates detailed digital data on terrain features that can be easily retrieved and disseminated for practical applications, like hydrological modeling, mass movement simulations, cut and fill earth works, and alike, which usually demand fast and precise information. And finally, an outstanding advantage of this method is the possibility of replicating the hierarchical semantic network to other areas with similar geomorphological characteristics, once the InterIMAGE platform allows the fine tuning of fuzzy membership functions and their respective thresholds, so as to fit the model to the peculiarities of particular landforms. In this way, it is possible to save modeling efforts from the photointerpreter point of view as well as computational processing time in terms of selecting the optimal attributes. This is especially useful in geomorphological mapping of large areas, as in the case of Brazil.

As the InterIMAGE project evolves, the task of implementing more sophisticated knowledge-based models in InterIMAGE 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 envisaged for the near future.

Acknowledgments

The work reported in this paper has been jointly supported by the German Aerospace Agency (DLR) and the Brazilian National Council for Scientific Research (CNPq) under grant number 491084/2005-6. The authors acknowledge the financial support provided by the Rio de Janeiro Foundation for Scientific Research (FAPERJ) and the Brazilian Federal Agency for Scientific Projects and Research (FINEP).

References

Abrams, M., Hook, S., & Ramachandran, B. (1999). *ASTER user handbook* (Version 2). Jet Propulsion Laboratory – EROS Data Center, Sioux Falls/South Dakota (135 pp.). <http://www.asterweb.jpl.nasa.gov/content/03_data/04_Documents/aster_user_guide_v2.pdf> Accessed 02.02.07.

Almeida, F. F. M. (2000). The origin and evolution of the South American Platform. *Earth-Science Reviews, 70*, 77–111.

Almeida, F. F. M., & Carneiro, C. D. (1998). Origem e evolução da Serra do Mar. *Revista Brasileira de Geociências, 28*, 135–150.

Baatz, M., & Schäpe, A. (2000). Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation. In *Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium, Salzburg* (pp. 12–23). Karlsruhe, Germany: Herbert Wichmann Verlag.

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 A. Cremers & K. Greve (Eds.), *Environmental information for planning. Politics and the public* (pp. 555–570). Marburg, Germany: Metropolis-Verlag.

Bolten, A., & Bubenzer, O. (2006). New elevation data (SRTM/ASTER) for geomorphological and geoarchaeological research in arid regions. *Zeitschrift für Geomorphologie Supplementband, 142*, 265–279.

Bonham-Carter, G. F. (1994). Tools for map analysis: Multiple maps. In G. F. Bonham-Carter (Ed.), *Geographic information systems for geoscientists: Modelling with GIS* (pp. 267–338). Kidlington, UK: Pergamon.

Bubenzer, O., & Bolten, A. (2008). The use of new elevation data (SRTM/ASTER) for the detection and morphometric quantification of Pleistocene megadunes (draa) in the eastern Sahara and the southern Namib. *Geomorphology, 102*, 221–231.

Bückner, J., Pahl, M., Stahlhut, O., & Liedtke, C.-E. (2001). GEOAIDA – A knowledge-based automatic image data analyzer for remote sensing data. In *Proceedings of the ICSC congress on computational intelligence methods and applications – CIMA, Bangor, Wales, UK (CD-ROM)*.

Carrion, D., Gianinetto, M., & Scaioni, M. (2002). GEOREF: A software for improving the use of remote sensing images in environmental applications. In *Proceedings of the IEMSS – integrated assessment and decision support, Lugano, Switzerland*.

Clausi, D. A. (2002). An analysis of co-occurrence texture statistics as a function of gray level quantization. *Canadian Journal of Remote Sensing, 28*, 45–62.

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, 31*, 779–791.

DNPM (1983). *Folhas SF 23/24 (Rio de Janeiro/Vitória): Geologia, geomorfologia, pedologia, vegetação e uso potencial da terra*. Rio de Janeiro, Brazil: DNPM, 775p.

Ehlers, M., Janowsky, R., & Gähler, M. (2002). New remote sensing concepts for environmental monitoring. In *Proceedings of the conference on remote sensing for environmental monitoring, GIS applications, and geology, Toulouse, France* (pp. 1–12).

Ehlers, M., & Welch, R. (1987). Stereocorrelation of Landsat TM images. *Photogrammetric Engineering and Remote Sensing, 53*, 1231–1237.

Florenzano, T. G., & Csordas, S. M. (1993). Mapa geomorfológico da região do Vale do Paraíba e Litoral Norte do Estado de São Paulo (escala 1:250 000). *Technical Report*. São José dos Campos (SP), Brazil: INPE.

Graham, J., & Jones, P. (1997). *Knowledge-based image processing systems*. London, UK: Springer-Verlag.

INPE – Instituto Nacional de Pesquisas Espaciais (2006). SPRING – *Sistema de Processamento de Informações Georreferenciadas: Banco Demo*. São José dos Campos, SP, Brazil. <http://www.dpi.inpe.br/spring/portugues/download.php> Accessed 09.05.06.

IPT (1981). *Mapa Geomorfológico do Estado de São Paulo (escala 1:1 000 000)*. São Paulo (SP), Brazil: IPT.

Kressler, F. P., & Steinnocher, K. (2006). Image data and LIDAR – An ideal combination matched by object-oriented analysis. In *Proceedings of the first international conference on object-based image analysis (OBIA 2006), Vienna, Austria*. Salzburg, Austria: Universität Salzburg, ISPRS.

Laboratório de Visão Computacional – LVC, Catholic University of Rio de Janeiro (2009). *InterImage: Interpreting images freely*. <http://www.lvc.ele.puc-rio.br/projects/interimage/documentation/>, Accessed 12.01.09.

Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.

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*. Michigan, USA: Environmental Research Institute of Michigan, AGU.

Matsuyama, T., & Hwang, V. (1990). *SIGMA: A knowledge-based aerial image understanding system*. New York, USA: Plenum Publishing Corporation.

McKeown, D. M., Harvey, W. A., & McDermott, J. (1985). Rule-based interpretation of aerial imagery. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 7*, 570–585.

Oliveira, C. G., & Paradella, W. R. (2008). An assessment of the altimetric information derived from spaceborne SAR (RADARSAT-1, SRTM3) and optical (ASTER) data for cartographic application in the Amazon region. *Sensors, 8*, 3819–3829.

Oliveira, C. G., & Paradella, W. R. (2009). Evaluating the quality of digital elevation models from ASTER stereoscopy for topographic mapping in the Brazilian Amazon region. *Annals of the Brazilian Academy of Sciences, 81*, 217–225.

Pahl, M. (2003). *Architektur eines wissensbasierten Systems zur Interpretation multisensorieller Fernerkundungsdaten*. PhD thesis. Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, Universität Hannover, Hannover, Germany.

PCI Geomatics (2006). *Geomatica v. 10.0.3 user guide*. Ontario, Canada: Richmond Hill.

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, Freiburg, Germany.

Russ, J. C. (1998). *The image processing handbook*. Boca Raton, FL: CRC Press.

Sagerer, G., & Niemann, H. (1997). *Semantic networks for understanding scenes* (first ed.). New York, USA: Plenum Publishing Corporation.

Schiewe, J., Tufte, L., & Ehlers, M. (2001). Potential and problems of multi-scale segmentation methods in remote sensing. *GIS-Informationsysteme, 6*, 34–39.

Shary, P. A. (2007). *Local morphometric variables*. http://www.giseco.info/pb/wp_d4052550/wp_d4052550.html?0.7718744701805855 Accessed 17.03.08.

Siart, C., Bubenzer, O., & Eitel, B. (2009). Combining digital elevation data (SRTM/ASTER), high resolution satellite imagery (Quickbird) and GIS for geomorphological mapping: A multicomponent case study on Mediterranean karst in Central Crete. *Geomorphology, 112*, 106–121.

Toutin, T. (2002). Three-dimensional topographic mapping with ASTER stereo data in rugged topography. *IEEE Transactions on Geoscience and Remote Sensing, 40*, 2241–2247.

Toutin, T. (2004a). Review article: Geometric processing of remote sensing images: Models, algorithms and methods. *International Journal of Remote Sensing, 25*, 1893–1924.

Toutin, T. (2004b). Comparison of stereo-extracted DTM from different high-resolution sensors: SPOT-5, EROS-A, IKONOS-II, and QuickBird. *IEEE Transactions on Geoscience and Remote Sensing, 42*, 2121–2129.
Toutin, T. (2006). Generation of DSMs from SPOT-5 in-track HRS and across-track HRG stereo data using spatiotriangulation and autocalibration. *ISPRS Journal of Photogrammetry and Remote Sensing, 60*, 170–181.
Toutin, T. (2008). ASTER DEMs for geomatic and geoscientific applications: A review. *International Journal of Remote Sensing, 7*, 1855–1875.
Valeriano, M. M. (2003). Curvatura vertical de vertentes em microbacias pela análise de modelos digitais de elevação. *Revista Brasileira de Engenharia Agrícola e Ambiental, 7*, 539–546.
Valeriano, M. M., & Carvalho, O. A. de. Jr., (2003). Geoprocessamento de modelos digitais de elevação para mapeamento da curvatura horizontal em microbacias. *Revista Brasileira de Geomorfologia, 4*, 17–29.
Verdade, F. C., & Hungria, L. S. (1966). Estudo genético de bacia orgânica do Vale do Paraíba. *Bragantia, 25*, 189–202.
Verstappen, H. T., & van Zuidam, R. A. (1991). *The ITC system of geomorphologic survey: A basis for the evaluation of natural resources and hazards*. Enschede, The Netherlands: ITC.
Witlox, F. (2005). Expert systems in land-use planning: An overview. *Expert Systems with Applications, 29*, 437–445.