In: Stilla U et al (Eds) PIA07. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36 (3/W49B)

A MULTI-LEVEL IMAGE DESCRIPTION MODEL SCHEME BASED ON DIGITAL TOPOLOGY

Guobin Zhu [a, *], Xiaoli Liu [b], Zhige Jia [c] Qingquan Li [b]

[a] International School of Software, Wuhan University, 129 Luoyu Road, Wuhan 430079, China -zhuu_gb@china.com
[b] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China - liuxl.j@163.com
[c] Institute of Seismology, China Seismological Bureau, Xiaohongshan Road, Wuhan 430071, China - jia.zg@126.com

Commission III, WG III/4

KEY WORDS: Digital Image, Image engineering, Image Segmentation, Multi-level Model, Digital Topology, Object Extraction

ABSTRACT:

It's necessary to discuss the topology model of digital images for integrating remote sensing and geographic information system in higher levels. Based on complex theories, a multi-level hierarchical image representation is presented that preserves topological relation equivalency and a set of functional architectures that efficiently reflects this representation. In the proposed hierarchical framework, a novel progressive region growing method is proposed that incorporates spatial information related to adjacency between pixels. The particularity of this method is that connected regions and their topology generate objects in different scales, furthermore constructing a tree-object structure reflecting their spatial relationships.

1. INTRODUCTION

Image engineering consists of three layers: low-level image processing based on pixel, inter-level image analyzing based on object, and high-level image understanding based on semantics (Zhu, 2003). The course from image processing, analyzing to understanding is a progressive procedure reflecting information processing phases. In the processing, images are interpreted by description models using certain ways; while the transition from digital number within low-level origin images, objects within inter-level image extraction to knowledge retrieved from high-level architecture (e.g. spatial relationship) is progressively accomplished. It is amazed that such a hierarchical image information representation that preserves spatial relationships is constructed along with the progressive processing, which implies an information framework from abstract status to physical one and a hierarchical transition from discrete structure to continuous one. However, problems remain along with the evolvement. Digital images are discrete objects in nature, but they are usually representing continuous objects or, at least, they are perceived as continuous objects by visual perception of human beings. Therefore, having a discrete and continuous representation of one object at the same time may activate interests.

The field of digital topology grew out of this challenge, and its main purpose is to study the topology properties of digital images (Rosenfeld, 1979). Based on the theory, this paper presents a multi-level hierarchical image representation that preserves topological relation equivalency between discrete and continuous descriptions of the same object and a set of functional architectures that efficiently reflects this representation. This hierarchical framework involves a general method to associate each digital object, in an arbitrary digital space, with a Euclidean polyhedron (named as its continuous analogue), which naturally represents the "continuous perception" that an observer may take on that object. The multilevel architecture and, particularly, continuous analogues of objects can be applied to obtain new results in digital topology, by translating the corresponding continuous results through the levels of the architecture. So it may derive interests for integrating remote sensing and geographic information system in higher levels.

Image segmentation aim to subdivide the image into disjoint subsets of pixels, called regions, on the basis of some homogeneity criterion. Usually, this is the starting point in practical applications like content-based image retrieval or image compression (Chamorro-Martinez et al., 2003). Many types of segmentation techniques have been proposed in literatures (Chamorro-Martinez et al., 2003; Smith et al., 2000; Metternicht, 2003). Nevertheless, a drawback of most of these approaches is that they don't take into account that a region must be topologically connected. As a consequence, pixels belonging to separate and different regions could be assigned to the same cluster. The proposed hierarchical framework in this paper preserves spatial relationships and, so raises a suitable condition for image segmentation to incorporate spatial information related to adjacency between pixels. Based on the hierarchical representation, a novel progressive region growing method is proposed which incorporates spatial information related to adjacency between pixels. The particularity of this method is that connected regions and their topology generate objects in different scales, furthermore constructing a tree-object structure reflecting their spatial relationships.

* Corresponding author: Tel.: +86-27-68778916; fax: +86-27-68778910.
E-mail address: zhuu_gb@china.com (G. Zhu)

2. BASIC NOTIONS OF DIGITAL IMAGE

A topological model is used to explicitly specify adjacency and inclusion relationships between the different cells (vertices, edges, faces for dimensions 0, 1, 2) of a geometrical object. Information, called embeddings (labels), such as geometrical ones (for instance vertex coordinates) can be added to the model. The archetypical device model (R^2) is a standard decomposition of Euclidean plane ($\mathcal{R}^2$) into abstract cells. The kernel of our modeller is based on the topology of cellular complexes which is the only possible topology of finite sets (Kovalevsky, 1989; Kovalevsky, 2006). Under this topology no contradictions or paradoxes arise when defining connected subsets and their boundaries. In this section, we briefly summarize the basic notions of the multilevel architecture based on the topology of cellular complexes, as well as the notation that will be used throughout the paper.

Definition 1: An abstract cellular complex (AC complex) $C=(E, B, dim)$ is a set E of abstract elements (cells) provided with a natural reflexive, antisymmetric and transitive binary relation $B\subset E\times E$ called the bounding relation, and with a dimension function dim: $E\rightarrow I$ from E into the set I of non-negative integers such that $dim(e')\leq dim(e'')$ for all pairs $(e', e'')\in B$. An AC complex C, when regarded in this way as a topological space, is called an underlying polyhedron and written $|C| = \cup\{e, e\in C\}$. Moreover, C satisfies the following properties:

C1: $\leq$ is a partial order in E.
C2: If $e'\in C$ and e'' is a face of e' then $e''\in C$.
C3: If $e', e''\in C$ then $e'\cap e''$ is a face of both e' and e''.
C4: $\forall e\in C$ is a face of some n-dimensional cells in the C.

According to the definition of open star in the general topology, we introduce two types of "digital neighbourhoods": star and extended star of a cell $x\in C$ in a given digital object $O\subseteq cell_n(C)$.

Definition 2: The star of x in O is the set $st_n(x; O) = \{y\in O, x \leq y\}$ of n-cells (pixels) in O having x as a face. Similarly, the extended star of x in O is the set $st_n^*(x; O) = \{y\in O, x\cap y \neq \varnothing\}$ of n-cells (pixels) in O intersecting x.

Definition 3: Given an AC C and two cells x, $y\in C$. A centroid-map on C is a map $\Theta:C\rightarrow|C|$ such that $\Theta(x)$ belongs to the interior of x; that is, $\Theta(x) \in x^o = x-\partial x$, where $\partial x= \cup\{y, y<x\}$ stands for the boundary of x.

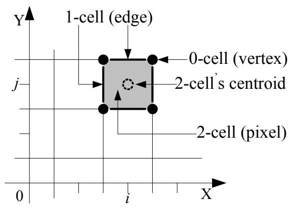

Fig. 1 An abstract cell in R^2

The maximum dimension of the cells of an AC complex is called its dimension. We shall mainly consider complexes of dimension 2. Their elements with dimension 0 (0-cells) are called points, elements of dimension 1 (1-cells) are called edges, elements of dimension 2 (2-cells) are called pixels (faces). As a consequence, the spatial layout of pixels in a digital image is represented by a device model, which is an n-dimensional locally finite AC complex determined by the collection of unit 2-squares in $\mathcal{R}^2$ whose edges are parallel to the coordinate axes and whose centres are in the set $\mathcal{Z}^2$. The centroid-map we will consider in R^2 associates to each square x its barycentre $\Theta(x)$. In particular, if $\dim(x)=2$ then $\Theta(x)\in\mathcal{Z}^2$. So that, every digital object O in R^2 can be identified with a subset of points in $\mathcal{Z}^2$.

Let the digital image (G) be a locally finite AC complex of dimension n (n=2). Each n-cell in G represents a pixel, and so the digital object ($O\in G$) displayed in a digital image is a subset of the set $cell_n(G)$ of n-cells in G, denoted by $O\subseteq cell_n(G)$; while the other lower dimensional cells in G are used to describe how the pixels could be linked to each other. If the full image is partition into m disjoint subsets (O_i, i=1, 2,…, m), then

$$\bigcup_{i=1}^{m} O_i = G \text{ and } \forall i\neq j : O_i\cap O_j=\varnothing \qquad (1).$$

The 2-cells (pixels) are the area elements. In image processing, 2-cells must be associated with the notion of pixels since a gray value assigned to a pixel originates from measuring the amount of energy radiated from an elementary area. Valid adjacencies are between adjacent pixels which are labelled by identical image values(Rosenfeld, 1984). In fact, pixels are usually a combination of materials, and frequently in multispectral and hyperspectral image (Bragato, 2004; Plaza et al., 2004). To solve this problem, region is regarded as a fuzzy subset of pixels (Bragato, 2004), in such a way that every pixel of the image has a membership degree to that region. It is a key that a fuzzy resemblance relation between neighbour pixels is obtained from a fuzzy resemblance relation between their corresponding feature vectors. So we characterize each pixel p by means of a vector of features $\vec{f}_p = [f_p^1, f_p^2, ..., f_p^k]$, where a feature $f_p^i \in \mathcal{R}$ with $i\in\{1, 2,…,k\}$, is a numerical measure of any relevant characteristic that may be obtained for p. In general, we define a fuzzy resemblance relation between feature vectors as a fuzzy subset FR of $\mathcal{R}^k\times\mathcal{R}^k$, with membership function FR:$\mathcal{R}^k\times\mathcal{R}^k\rightarrow[0, 1]$. For simplicity, we define a set of centroids and compute the membership value for all the pixels in the image to each centroid.

3. MULTI-LEVEL IMAGE REPRESENTATION

3.1 A Hierarchical Information Representation (HIR) for Images

In the processing from low-level to high-level, images are interpreted by description models using certain ways in different abstract levels; while the transition from digital number within low-level origin images, objects within inter-level image extraction to knowledge retrieved from high-level architecture is progressively accomplished. Amazing as it is, such a hierarchical information representation (HIR) for images that preserves spatial relationships is constructed along with the progressive processing, which shows an information framework from abstract to material and a hierarchical transition from discrete to continuous. Considering the gap between the discrete world of digital objects and the Euclidean world of their continuous interpretations, this paper proposes a progressively hierarchical image representation, in which there coexist discrete and continuous descriptions of the same object preserved topological relation equivalency.

This framework consists of four levels: two extreme levels: device and continuous levels, and additional two intermediate levels: logical and conceptual levels. The device level represents the physical problem (Fig.2a/a') whereas the models in the continuous level are topological spaces which allow us to use the well-known results of continuous topology (e.g. polyhedral topology) (Fig.2b/b'). The additional two intermediate levels are used to bridge the two extreme levels, which allow a progressive evolution from the discrete object to the Euclidean one, and vice versa. The logical level is closer to the device level and it is used for processing, for writing algorithms and to show their correctness (Fig.2c/c'). The conceptual level is the nearest to the continuous level and it is used to translate results and notions from the continuous level to the logical level (Fig.2d/d').

3.2 A Multi-Level Functional Architecture

Because a digital space fixes, among all the possible continuous interpretations, just one for each digital object, this continuous interpretation of a digital object is represented at each level of the architecture using a different model; in particular, the corresponding model at the continuous level is a Euclidean polyhedron, called the continuous analogue of the object. According to target applications, in each one of them we can use different models.

To illustrate our purpose, we introduce an empirical functional structure in dimension 2 that efficiently reflects the proposed representation. Inside the modelling structure, objects are represented with four different models: discrete, discrete contours, discrete analytical, and continuous. In the hierarchical structure, links between the different consecutive levels allow us to manipulate and propagate modifications locally. Updating the whole structure is thus not systematically needed. Of course such a framework comes with a prize. The complexity of the hierarchical structure is much more complex than classical topology based modelling or imaging softwares.

Let the full image G be subdivided into n disjoint objects $O_i \subseteq cell_2(G)$ (i=1,2,..,n), which is a set of 2-cells with domain f_p^* of homogeneous feature value (Fig. 2).

3.2.1 0-Device Level: corresponds to the classical discrete representation of image elements (pixels) in a computer screen. Each pixel is represented by a square topological face associated with a colour feature embedding. Moreover, integer coordinates are attached to each topological vertex.

A discrete model of O in this level is the subcomplex D_O^f ={x∈G; x ≤ y, y∈O} induced by the cells in O, and $\vec{f_p} \in f_p^*$ for any cells (Fig. 2a). This level has a very few degree of abstraction and we only represent the physical aspects of the objects.

3.2.2 1-logical level: corresponds to the contours obtained for each 4-connected region with homogeneous feature vector. The representation of these contours is based on the inter-pixel model (Kovalevsky, 1989). Each discrete point is represented by a 0-cell and two successive points are linked by a 1-cell (Fig. 2b). This representation simplifies the coverage of level 0 regions boundaries. There is no embedding at this level and geometrical information needed for the visualization of the level are located in level 0 and can be accessed from level 1.

A discrete border model of O in this level is an undirected graph L_O^f, whose vertices are 0-cells and linked by 1-cells in O, moreover two of those cells x, y are adjacent if there exists a common face $a \leq x \cap y$ such that $\vec{f}(x)$, $\vec{f}(y) \in f_p^*$.

In this level, we consider the proximity aspects of the objects and so, we can study some properties of topological nature. The main function of this level is to be the support for writing the algorithms and to prove their correctness. Because L_O^f is not planar, this level is far from the mathematical model. So we need the conceptual level as an interface between the level above and the continuous level.

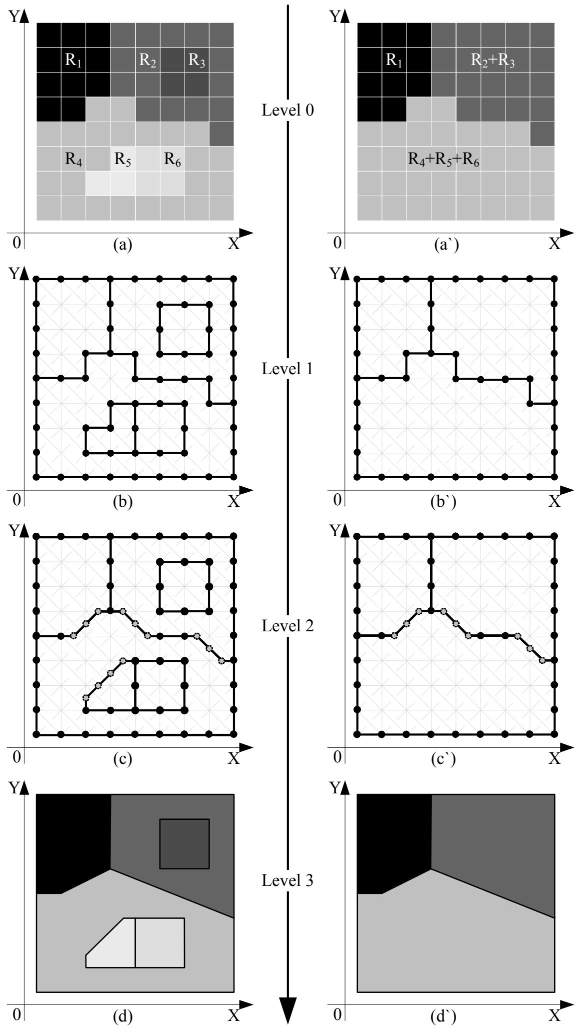

Fig. 2 Multi-level hierarchical representation for digital images: (a) discrete level; (b) logical level; (c) conceptual level; (d) continuous level.

3.2.3 2-Conceptual Level: is an implicit representation of the discrete border primitive. It corresponds to the discrete analytical description of the level 1 region contours. More precisely, each contour is described as a discrete analytical polygon computed according to the given models. For the sake of simplified solution, we flatten out the graph L_O^f in a natural way and we get the planar graph (Fig. 2c). In this graph there are two different kinds of vertices: 0-cells and centroids of 1-cells. Observe that this graph is a triangulation of the Euclidean plane, which makes up the conceptual level.

A discrete analytical model of O is an induced directed graph C_O^f, whose vertices are the centroids $\Theta(x)$ of all 0-cells and 1-cells $x \in G$ *with* $\vec{f}(x) \in f_p^*$, and its directed edges are pairs $(\Theta(x), \Theta(y))$ with $x<y$.

3.2.4 3-Continuous Level: is an explicit representation. In this level, each region is described as a Euclidean polygon with attribute features in the classical boundary representation form. The primitive of this level may be created using the tools available inside the modeller, or may be the result of the reconstruction process applied on level 0. 2D Euclidean vertex coordinates and face features are associated to the continuous model of O in this level (Fig. 2d).

A simplicial analogue S_O^f is an order complex associated to the directed graph C_O^f. This is, $<\Theta(\alpha_0), \Theta(\alpha_1), \ldots, \Theta(\alpha_m)>$ is m-simplex of S_O^f if $\Theta(\alpha_0), \Theta(\alpha_1), \ldots, \Theta(\alpha_m)$ is a directed path in C_O^f. This simplicial complex defines the simplicial level for the object O in the architecture and, finally, the continuous analogue of O is the underlying polyhedron $|S_O^f|$ of S_O^f.

When given a concrete problem, we must choose specific models in each level and functions which can support the functionality that we have described. Specifically, suppose that these chosen models are D, L, C and S for the device, logical, conceptual and continuous level, respectively. Let $\Omega(D)$, $\Omega(L)$, $\Omega(C)$ and $\Omega(S)$ be the sets of the objects of these models. Then here are several mappings: (1) a 1-1 mapping $i:\Omega(D)\rightarrow\Omega(L)$; (2) a natural mapping π: $\Omega(L)\rightarrow\Omega(C)$ and $\pi^*:\Omega(C)\rightarrow\Omega(L)$; (3) a suitable mapping $j:\Omega(C)$ and $\Omega(S)$. So we have the following functional architecture:

$$\begin{array}{l} \Omega(L) \xrightarrow{\ i\ } \Omega(D) \\ \pi \downarrow\uparrow \pi^* \\ \Omega(C) \xrightarrow{\ j\ } \Omega(S) \end{array}.$$

This architecture provides a link between the discrete world of digital pictures represented by a cellular complex, and a Euclidean space through several other intermediate levels and, embodies the transitions from low-level feature to high-level semantic. Further, this framework involves a general method to associate each digital object, in an arbitrary digital space, with a Euclidean polyhedron called its continuous analogue, which naturally represents the "continuous perception" that an observer may take on that object. The multilevel architecture and, particularly, continuous analogues of objects can be applied to obtain new results in digital topology, by translating the corresponding continuous results through the levels of the architecture. Thus it may be interesting for integrating geographic information system and remote sensing in higher levels.

4. A ROOT TREE STRUCTURE OF OBJECTS

The proposed hierarchical framework preserves spatial relationships, raising a suitable condition for image segmentation to incorporate spatial information related to adjacency between pixels. Thus this multi-level image representation allows manipulating and propagating modifications locally (Fig.2, Fig.3). We locally modify the segmentation of the digital image by filling the fragmental regions in order to obtaining desired-only regions. Thus, we can determine which cells must be removed in the other levels. Indeed, in the original digital image, these cells where surrounded by two faces are represented with different colors. After the edits, these cells are surrounded by two faces with the same color. Using the links between level 0 and level 1, we can easily find the cells of level 1 having been removed. And so on for level 2 and level 3. Here, the main interest of using links of the structure is that we can make local modification without re-computing the entire structure.

Suppose the full image and the discrete level correspond to two extreme levels of objects: the root and leafs of a tree, respectively. Starting with an arbitrary pixels, objects and features can be extracted easily through a down-top region merging cluster in different scales. While performing the down-top union of regions (children) at one level into a single larger region (parent) at the next higher level of the tree, regions are grouped together according to similarities between their feature vectors, which include such features as colour information, orientation, texture, size, energy, and neighbour information. Any other levels of segmented regions lie betwixt the two extreme levels. Hence a root tree structure is constructed in a simple and "natural" way the regions according to their topology and, where each level consists of two data structures, a weight graph $W_{level}(N_{level}, E_{level})$ and a disjoint-region set R_{level} (Corme, 1990).

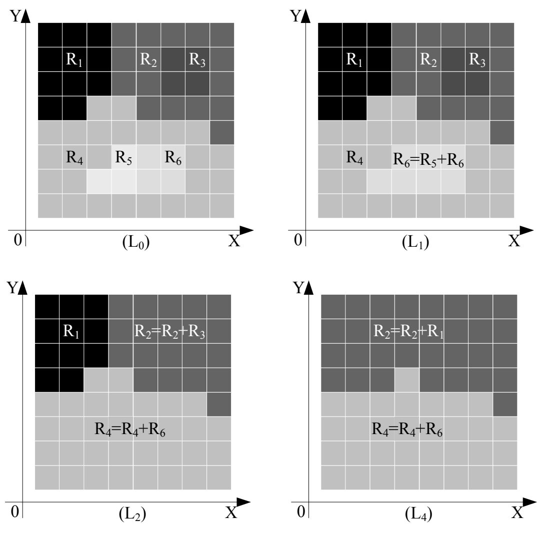

Fig. 3 Discrete segmentation operation based on the multi-level image representation

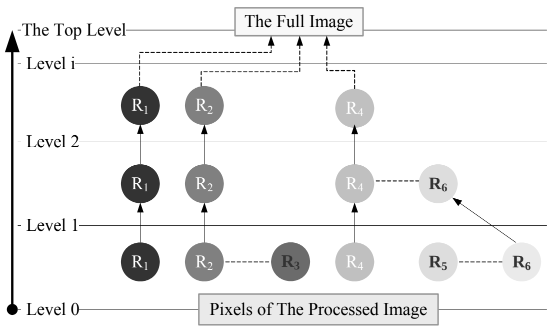

Fig.4. A root tree structure of objects

In the tree construction (Fig. 4), the root node denotes the border of the whole image where all the regions are included, the leaf nodes denote all pixels of the discrete level, other nodes of the tree denote regions N_{level} and has a feature vector characterizing those regions, each region in R_{level} is one node or a group of nodes from the next lower level. Two nodes are linked by an edge if and only if their corresponding regions are adjacent each other. The weighted graph in each level is composed of the nodes and edges within this level. The undirected dot-edges in each level define the spatial relationships between the nodes, and the weight of each edge defines the degree of the similarity between those two nodes. The relationship among the disjoint-set in each level is illustrated by the groupings. The directed solid-edges of regions in different levels reflect the inclusion relationships between them. Since regarded as a rooted "directed" tree, and in which from a given region stored, there is only one path to the root, this is, the map from the parent node to child nodes is 1: n ($n \in Z$). Given this directed notion of a rooted tree, a rooted sub-tree can be defined for each node of the tree. As a consequence, for each region which contains included regions, a sub-tree is recursively generated. The parent node shall directly inherit attribute features of the first child with smallest feature distance among all the edges connected to this node.

The tree-like organization embodies reflecting the relationships of regions and their attribute features. For each region stored in the tree, we can directly access to the included regions, the parent region, (e.g. vertical navigation), or adjacent regions (horizontal navigation) (Maire et al., 2005). Once the tree structure is stored, further complementary manipulating is possible for local modification without recomputing the entire structure.

5. DISCRETE IMAGE SEGEMETAION

As a practical instance, a region growing algorithm that efficiently constructs this root tree structure is proposed. The sequential process is described as four steps.

5.1.1 Step 1. Initialization: To reduce the computation complexity, the image is initially partitioned into small $K \times K$ (K varies by image size, usually equates to 4 or 8) blocks composed of topological connected pixels with similar image features. Each blob has an associated set of features measured from the original pixel spectral features of the image. Based on these extracted features and the neighboring relationships among blobs, the one-level is built.

5.1.2 Step 2. Region Merging Cluster: A recursive node clustering and region merging are performed at each level using a bottom-up strategy. At the end of each iteration, the algorithm has completed one level of the hierarchy, so a new level is constructed and the structure is updated. The procedure is repeated until the stopping condition has been attained, which is defined as either the desired final number of objects, or the maximum feature distance (threshold) below which clusters may be combined. Or if not specified, the algorithm will continue until a full tree structure of the original image is built with the root node of the tree being a single object corresponding to the whole image.

5.1.3 Step 3. Extraction of Geometry and the Adjacency Relation: Geometry and the topological relation in each level are extracted for each region detected during step 2, and stored in the database. When progressively recognizing, classifying and integrating image objects (nodes) from different regions in intra-level or inter-level, besides just colour and geometry information, multi-dimensional information, including the orientation, texture, size, energy, and neighbour relationships between objects are considered in the processing of region merging cluster.

5.1.4 Step 4. Building the tree structure: Based on segmentation results, and by analyzing topology relations and feature vectors, the regions are recursively created and stored in the tree structure and, so region inclusions are recursively deduced and propagated to the whole segmented image by adjacencies relations. Hence the tree-like organization of the hierarchical relations of spatial objects reflects hierarchical topological relations of spatial objects. From this tree structure, objects and features can be extracted easily through a top-down traversal of the final hierarchical structure of the image.

In our practical instance, to ensure low computation complexity, the second low computation complexity, the step 2 requires that each single-node region merge with at least one other region in this level. So, before a new level is generated, this iteration guarantees that all sets of the current S_{level} include at least two nodes. As a consequence, each new level of the hierarchy is guaranteed to have no more than half the number of nodes as the previous level, ensuring fast convergence of the algorithm.

6. CONCLUSION AND FUTURE WORK

In this paper we presented a multi-level hierarchical information representation (HIR) for images that preserves topological relation equivalency and a set functional architecture that efficiently reflects this representation. Each level corresponds to a particular representation of the same object: discrete, discrete border, discrete analytical of regions, and continuous representations. Each level is linked with the levels above and below itself by transition mapping. This ensures the topological relation equivalency between all the representations. In the proposed hierarchical framework, a progressive region growing method is used to subdivide image into regions and to construct a tree-object structure reflecting their spatial relationship. The particularity of this method is that it incorporates spatial information related to adjacency between pixels, while keeping connecting regions and their topology generated in different scales. As a short term goal, we plan to develop spatial relationship of discrete digital objects, while more studies in details for the propagation of local modifications along with the hierarchical multi-level will be pursued in the future.

7. REFERENCES

Zhu, G. 2003. Remote Sensing Image Analysis Based on Hierarchical Multi-resolution Structures. *Geomatics and Information Science of Wuhan University*, 28:315-320.

Rosenfeld, A., 1979. Digital Topology. *Am Math Month*, 86:621-630.

Chamorro-Martinez, J., Sanchez, D., 2003. A Fuzzy Color Image Segmentation Applied to Robot Vision. In Advances in Soft Computing - Engineering, Design and Manufacturing, J. M. Benitez, O. Cordon, F. Hoffman, and R. Roy, eds. (Springer), pp. 129-138.

Smith, J. R., Li, C.-S., 2000. Image Classification and Querying Using Composite Region Templates, Vol 77, Elsevier Science.

Metternicht, G. I., 2003. Categorical fuzziness: a comparison between crisp and fuzzy class boundary modelling for mapping salt-affected soils using Landsat TM data and a classification based on anion ratios, Vol 168, Elsevier Science.

Kovalevsky, V., 1989. Finite Topology as Applied to Image Analysis. *Computer Vision, Graphics, and Image Processing*, 46:141-161.

Kovalevsky, V., 2006. Axiomatic Digital Topology. *J Math Imaging*, 26:41-58.

Rosenfeld, A., 1984. The fuzzy geometry of image subsets. *Pattern Recognition Letters*, 2:311-317.

Bragato, G., 2004. Fuzzy continuous classification and spatial interpolation in conventional soil survey for soil mapping of the lower Piave plain, Vol 118, Elsevier Science.

Plaza, A., Martinez, P., 2004. A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles, Vol 37, Elsevier Science.

Cormen, T. H., Leiserson, C. E., and Rivest, R.L., 1990. Introduction to Algorithms, MIT Press, Cambridge, MA.

Maire, C., Datcu, M., 2005. Object and topology extraction from remote sensing images.Image Processing, 2005. ICIP 2005. IEEE International Conference on, Vol.2, Iss., pp: II-193-6.