

# Simultaneous Segmentation-Recognition-Vectorization of Meaningful Geographical Objects in Geo-Images

Serguei Levachkine, Miguel Torres, Marco Moreno, and Rolando Quintero

Geoprocessing Laboratory (GEOLAB)

Centre for Computing Research (CIC) - National Polytechnic Institute (IPN)  
{palych, mtorres, marcomoreno, quintero}@cic.ipn.mx

**Abstract.** We present an approach to color image segmentation by applying it to recognition and vectorization of geo-images (satellite, cartographic). This is a simultaneous segmentation-recognition system when segmented geographical objects of interest (alphanumeric, punctual, linear, and area) are labeled by the system in same, but are different for each type of objects, gray-level values. We exchange the source image by a number of simplified images. These images are called composites. Every composite image is associated with certain image feature. Some of the composite images that contain the objects of interest are used in the following object detection-recognition by means of association to the segmented objects corresponding “names” from the user-defined subject domain. The specification of features and object names associated with perspective composite representations is regarded as a type of knowledge domain, which allows automatic or interactive system’s learning. The results of gray-level and color image segmentation-recognition and vectorization are shown.

## 1 Introduction

Segmentation is fundamental to the field of image processing because it is used to provide the basic representation on which understanding algorithms operate. The ability to build up a representation from individual pixels of an image, which exploits relationships such as local proximity and highlights the structures of the underlying components, is important for the extraction of features during interpretation and recognition [1]. In general, the nature of this representation is application dependent. In the present work, we developed an application independent segmentation.

Up to the now a great variety of segmentation algorithms for gray-level images has been proposed. The majority of color segmentation approaches are based on monochrome segmentation approaches operating in different color spaces [4]. Gray-level segmentation methods can be directly applied to each component of a color space; thus, the results can be combined in some way to obtain a final segmentation result. However, one of the problems is *how to employ the color information as a whole for each pixel*. When the color is projected onto three *RGB* color components, the color information is so scattered that the color image becomes simply multispectral image

and the color information that humans can perceive is lost [2]. Another problem is *how to choose the color representation for segmentation* [3], [4]. There is no single color representation that can surpass others for segmenting all kinds of color images. The use of nonlinear spaces, such as *HSI* and the normalized color space can solve the problem to certain approximation. However, the nonlinear spaces have essential, non-removable singularities and there are spurious modes in the distribution of values [5].

An alternative solution presented in this work is invariant image representation (composite images, or simply *composites*) that does not depend on the choice of particular color space. The processing of a color image is individual segmentation by each color component into image meaningful (or invariant with respect to a given, unnecessary color feature) regions, first and, then – image’s joint segmentation-recognition (or “objects of interest designing”). Moreover, the prescribed set of features is regarded as a type of *knowledge domain*. The *composite image technique* includes object-fitting compact hierarchical segmentation, binarization of segmented images, and synthesis of binary representations. The main goal of image synthesis consists of the object linking by its associated names. In the following sections, we build up composite image representations based on object-fitting compact hierarchical segmentation. See also [6], [7], [8], and [9].

## 2 Object-Fitting Compact Hierarchical Segmentation, Recognition, and Vectorization

In our method, the image segments obtained as the result of the iterative procedure of successive increasing of the admitted gray-level and color thresholds in the segment merging form subsequently increasing compact hierarchical structure of the flat segment networks. Each segment of this structure can have the ancestor or a descendant. Thus obtained structure is called *the adaptive dynamic data structure*. The segment of image is a node of the spatial structure, which attributes are primary numbers defined by the averages of color/gray-level segment’s features and by a set of pixels that represent the area and the shape of the segment (Section 6). This allows organizing the object-oriented identification of semantically meaningful image’s regions. Our system has the interactive procedure of *compulsory restructuration* of the segment relationships as a tool of the semantic analysis of visual data. In other words, the system’s learning and self-learning with the prescribed set of associative identifiers are possible in the interactive regime.

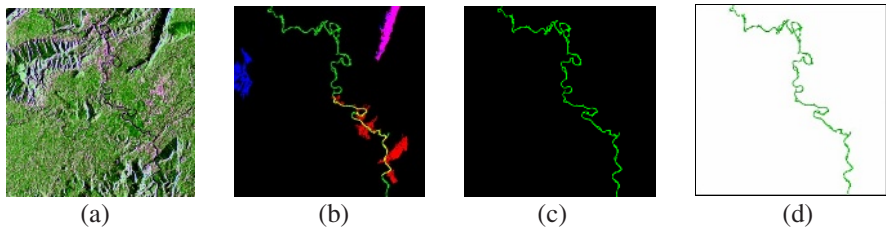
Successive segment merging by some criteria leads to the segment structuring in a *multi-level hierarchy* that represents by the dynamic trees [7]. This hierarchy or *the multi-level image partition* is an efficient method of semantic identification of the image’s objects. The relationships between the dynamic tree nodes indicate *the neighboring semantically meaningful regions*. Because the image’s regions are identified by the corresponding tree nodes, *the neighbor relation* between them can be completely defined by a table of adjacency. A modification (elimination of some edges, i.e. segment relationships) of the dynamic tree allows modifying the resulting region and thus more exact object detection is reached. Each level of the tree of segments can

be considered as alternative image interpretation in different semantics (see Fig. 3, Section 3 and Fig. 5, Section 5).

Adaptive dynamic tree structure regards the search for meaningful objects as the combination of the object features that fit to the corresponding ranges and the following analysis of all admitted areas. This makes possible to use *the automatic learning* algorithms when the set of searched objects is given and it is necessary to define only the corresponding feature ranges (this is natural supposition for geo-images [9]). The learning process can be organized as follows. The user selects the appropriate level of segment hierarchy and points out the set of the suitable areas. These areas can be defined by combining the corresponding segments. Then the program computes the characteristics of the located segments and relationships between them and establishes the formal criterion of the search for the similar objects.

*Object-fitting compact hierarchical segmentation* is a sequence of embedded partitions without repetition of composed segments in different partitions. A partition is obtained by iterative segment splitting or merging. In the merging mode, any segment in each iteration merges into the nearest adjacent segment. The number  $2^i$ , where  $i$  is the number of iteration, is bound total number of segments  $N$ , generated at each iteration [6]. The number  $N$  has to be taken into account for automatic color image analysis. Indeed, the merging of segments into objects defines the image semantics. The image's semantics in this context corresponds to the association of segment fields of different hierarchical levels being identified with identifying conceptions from the subject domain. For example, detection of a segment identifying a coastline or highway becomes semantically meaningful. Further, this set of segments is renamed as "coastline", "highway", etc. (Fig. 1).

Segmented and recognized objects are subsequently vectored by applying a method described in [9] to be finally included into GIS. These three stages (segmentation-recognition-vectorization) are Objected Oriented Data Integration for GIS [9], [12].

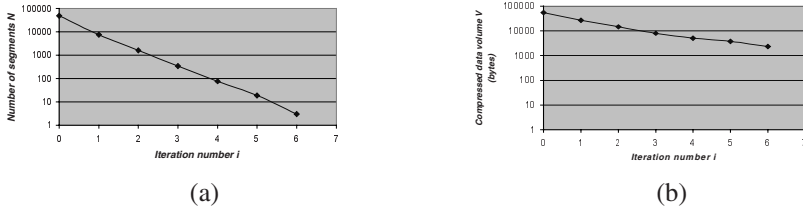


**Fig. 1.** (a) Source image, (b) Segmented image, (c) Recognized image, and (d) Vector image of the river

Precisely all this is what we mean under simultaneous segmentation-recognition-vectorization process. Fig. 1 shows segmentation-recognition-vectorization of the river in a SAR image of Kalimantan Island.

The number of segments  $N$  decreases approximately as  $(4+5)^i$ , where  $i$  is the iteration number (Fig. 2a). From our point of view, deviation from this exponential de-

pendence leads to image's semantics violation. Disclosed regularity can be useful for automatic analysis of gray-level and color images.



**Fig. 2.** Linear dependences: (a) Number of segments  $N$  on iteration number  $i$ , (b) Compressed image volume  $V$  on iteration number  $i$

Our experiments have shown (Fig. 2b) that the compressed volume of an image is decreased in the same exponential mode (Cf. image compression with information-lossless standard algorithms: RAR, LZH, etc.).

Eliminating the dependence on iteration number  $i$ , we can obtain the exponent-mode relation of compressed data volume  $V$  on number of segments  $N$  as follows:

$$N/N_0 = (V/V_0)^\alpha \quad (1)$$

In equation (1),  $N_0$ ,  $V_0$  denote number of segments and compressed volume of the source image respectively;  $\alpha$  is some real coefficient. We obtained that in the case of object-fitting compact hierarchical segmentation the exponent  $\alpha$  is approximately 2.9. Note that for non-adaptive pyramidal segmentation [3]  $\alpha$  is approximately 1.4. It is known that the volume of compressed data is closely related to the *amount of information into data*. Thus, a theoretical explanation of the obtained experimental dependencies (Fig. 2) represents an interesting research topic of Pattern Recognition.

### 3 Composite Image Representation

We have found that in addition to natural decomposition (e.g.,  $R$ ,  $G$  or  $B$  – component splitting) of color images, artificial representations can also be useful for objects of interest detection-recognition [6], [7], [8], and [9]. Our approach provides composite representations of the source image by means of reduced number of color or tone components and segments. Composite image representation is a sequence of binary representations, which are packed into different bit planes. These binary images are the result of two-valued classification of source image by some feature (intensity, area, invariant moments, etc.; Section 6).

A bit component of composite image (Fig. 3) computes by means of global dynamic thresholding of the current segmented image. The threshold is equal to the average all over the image intensity, geometric or other feature, denoted by  $I^{(i)}$ . To threshold the image,  $I^{(i)}$  is compared with its average over the pixels of each segment, denoted by  $I^{(s)}$ , as follows:  $I^{(i)} \geq (<) \xi I^{(s)}$ , where  $\xi$  is a tuning parameter.



**Fig. 3.** Bit components of composite images obtained by means of dynamic adaptive thresholding of Lena's source and segmented images

To compose these images, we also used the geometric features from the feature set (Section 6) in addition to intensity feature. The bit components are packed in the resulting representation, where the extrema of intensity indicate the pixels associated with unchanged binary feature. Essentially, the composite images form a “book” in which the objects of interest can be found on appropriated page(s). Thus, a “page number” defines the method of thresholding and the tuning parameter  $\xi$ .

## 4 Color Composites

Compact hierarchical segmentation of a color image is performed by each independent color components ( $R$ ,  $G$ , and  $B$ ) considering these as semi-tone images. In this way, coinciding intensities of resulting  $R$ ,  $G$ , and  $B$  composite images indicate the segments of equal color with respect to using feature. This can be used for invariant color image description. As a rule, compact hierarchical image segmentation implies that color segments are enlarged simultaneously in accordance to regularities presented in Section 2. Due to the self-consistence of  $RGB$ -segmentation behavior, visual quality improvement in composite intensities becomes available [6]. The method requires significant operative memory space. To overcome this disadvantage, we used special data organization in the form of irregular dynamic trees (Section 2 and [6], [7], [8], [9]) that provides optimal in memory space computing for the successive scanning of image scales. Due to data organization, a practical use of our program package does not require further algorithmic development. The user needs only to make adequate choice to carry out task features from prescribed feature set (*knowledge domain*).

## 5 Applications of Composites

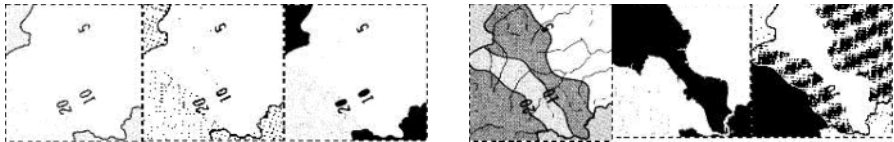
Image two-valued classification (binarization) is one of the most important tasks in modern recognition methods. In the frameworks of the composite image technique, we obtained a few solutions for this task [6], [7], [8], and [9]. By applying composites, we

are able to extract cartographic data using  $R$ ,  $G$ , and  $B$ -components of full-size color raster-scanned image (Fig. 4).



**Fig. 4.** Cartographic pattern retrieval from a color map image 1082 x 1406 pixels (extreme left)

Fig. 5 shows how our method insures object detection in the task of recognition of inclined digits embedded in graphics (note that this is old and very difficult problem that has been attracted much attention by image processing specialists [1], [4], [9]). This illustrates that each composite image contains machine-treatable bit-planes for target object detection and also purposeless bit-planes. Indeed, to effectively recognize the objects of interest, it is better to search these objects on appropriated bit-planes. Although, our system can generate some errors in interpretation, it is much more useful for the following understanding algorithms because its output is nearly recognized objects of interest.



**Fig. 5.** Bit-planes suitable for digit recognition (left side) and other purposes (right side)

## 6 Comments

The presented approach exploits the user's experience providing the knowledge domain in the form of the prescribed feature-attribute set. This set contains a number of attributes and numerous features. The attributes are a primary set of segment characteristics estimated and dynamically stored for all image segments at any level of the composite image representation. This provides a full-value use of object-fitting hierarchical segmentation. The features are numerical segment characteristics, which are obtained as output of data conversion, and are selected in function of the processing stage and the problem context. Thus, prescribed segment attributes are the following - 1) *Extrema of numerical characteristics*: a) *global* (for the whole image), b) *local* (for a neighborhood of the segment); 2) *Additive*: a) *integral intensity* (the sum of pixel intensities), b) *number of pixels*, c) *integral first and second moments computed with*

respect to the origin; 3) *Non-additive perimeter*; 4) *Description of the adjacent segments in terms of binary relationships*.

These data provide an estimation of the intensity and geometric segment features: pixel intensity range, average intensity, invariant moments, parameters of linear sizes and shapes, etc. In this manner, the features used in generation of object-fitting hierarchy of the segments can be different from the features used in object recognition [6], [7], [8], and [9]. As a rule, at first step of image processing, only intensity features are useful, because the source image pixels do not form geometrically meaningful segments and objects. Consequently, up to reaching image invariant representation other features are used for object designing and recognition. To our knowledge, this is one of the first attempts to design a segmentation-recognition computer system for complex color images of arbitrary type (Cf. [10] and [12]).

## 7 Conclusion

The problem of *how and to what degree the semantic information should be employed in image segmentation* has led us to the conception of composite image representation for mutual object detection-recognition at low level processing. We conjecture that modern segmentation systems must support mutual object detection-recognition-interpretation, starting at low level, memorizing results at the intermediate level, and effectively communicating these results to the high level. The approach proceeding from this conjecture is called *composite image technique*. The idea is to prepare the source image as much as possible for subsequent high-level processing of image regions. In most of the existing color image segmentation approaches, definition of a region is based on similarity of color. This assumption often makes it difficult for any algorithms to separate the objects with highlights, shadows, shadings or texture, which cause inhomogeneity of colors of the object's surface. Using *HSI* can solve this problem to some extent, except that hue is unstable at low saturation. Some physics-based models have been proposed to solve this problem [4], [11].

We saw an alternative solution of the problem defining image regions by *quantitative, qualitative, and nominal features* (in addition to color feature), which on the whole render the *user's knowledge domain*. We believe that this is a kind of advanced simulation of the human's visual perception. However, it is necessary to emphasize that for optimization of labor-intensive program training a strong formalization of composite image technique is now required. We are under way to solve this problem.

At the same time, automatic interpretation of color images presents certain difficulties for state-of-the art in image processing and also artificial intelligence. To date, it appears unrealistic to obtain fully automatic computer-based interpretation system free of errors [4], [9], and [12].

We believe that only a system approach to the problem can be fruitful. In the context of the present work, this means first, decomposition of source image by multiple hierarchical components to achieve a stable, accurate representation in the presence of degraded images. Second is the segmentation with mutual recognition of appropriate primitives (compression stage) and, if required, their vectorization to be directly in-



cluded into application-oriented database, e.g., GIS (application-dependent stage). Finally, there is the development of a unified knowledge-based trainable and self-trainable system with optimal human-machine interaction for color image treatment.

## References

1. Doermann, D.S.: An Introduction to Vectorization and Segmentation. Lecture Notes in Computer Science, Vol. 1389. Springer-Verlag, Berlin Heidelberg New York (1998) 1–8
2. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. 3<sup>rd</sup> edn. PTR, NJ: Prentice Hall, New York (2002)
3. Alexandrov, V.V., Gorsky, N.D.: Image Representation and Processing: A Recursive Approach. Mathematics and Its Applications, Vol. 261. Kluwer Academic Publishers, Dordrecht Boston London (1993)
4. Cheng, H.D., Jiang, X.H., Sun, Y., Wang, J.: Color Image Segmentation: Advances and Prospects. Pattern Recognition, Vol. 34, No. 12. (2001) 2259–2281
5. Kender, J.: Saturation, Hue, and Normalized Color: Calculation, Digitization Effects, and Use. Computer Science Technical Report, Carnegie Mellon University (1976)
6. Alexandrov, V.V., Kharinov, M.V., Velázquez, A., Levachkine, S.P.: Object-oriented Color Image Segmentation, In: Hamza, M.N. (ed.): Proc. IASTED International Conference on Signal Processing, Pattern Recognition, and Applications (SPPRA 2002) 25–28 June 2002, Crete, Greece, ACTA Press, Anaheim Calgary Zurich (2002) 493–498 (ISBN: 0-88986-338-5; ISSN: 1482–7921)
7. Alexandrov, V.V., Kharinov, M.V., Levachkine, S.P.: Conception of Hierarchical Dynamic Structure in Application to Audio and Video Data Recognition. In: Hamza, M.N. (ed.): Proc. IASTED International Conference on Artificial Intelligence and Soft Computing (ASC 2001) 21–24 May 2001, Cancun, Mexico, ACTA Press, Anaheim Calgary Zurich (2001) 348–353 (ISBN 0-88986-283-4; ISSN 1482-7913)
8. Levachkine, S.P., Velázquez, A., Alexandrov, V.V.: Color Image Segmentation using False Colors and its Applications to Geo-images Treatment: Alphanumeric Character Recognition. Proc. IEEE International Geosciences and Remote Sensing Symposium (IGARSS 2001) 9–13 July 2001, Sydney, Australia (2001) (IEEE Catalog Number (CD-ROM): 01CH37217C; Library of Congress Number: 01-087978; ISBN CD-ROM: 0-7803-7033-3)
9. Levachkine, S.P., Velázquez, A., Alexandrov, V.V., Kharinov, M.V.: Semantic Analysis and Recognition of Raster-scanned Color Cartographic Images. Lecture Notes in Computer Science, Vol. 2390. Springer-Verlag, Berlin Heidelberg New York (2002) 178–189
10. Kumar, K.S., Desai, U.B.: Joint Segmentation and Image Interpretation. Pattern Recognition, Vol. 32, No. 4. (1999) 577–589
11. Levachkine, S.P., Sossa, J.H.: Image Segmentation as an Optimization Problem. Computation and Systems, Vol. 3, No. 4. (2000) 245–263 (ISSN 1405-5546)
12. Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M.: Multi-resolution, Object-oriented Fuzzy Analysis of Remote Sensing Data for GIS ready Information, In: Levachkine, S., Ruas, A., Bodansky, E. (eds.), Proc. International Workshop on Semantic Processing of Spatial Data (GEOPRO 2002) 3–4 December 2002, Mexico City, Mexico (2002) (ISBN CD-ROM: 970-18-8521-X)