

# A Resegmentation Approach for Detecting Rectangular Objects in High-Resolution Imagery

Thales Sehn Korting, Luciano Vieira Dutra, and Leila Maria Garcia Fonseca

**Abstract**—Image segmentation covers techniques for splitting one image into its components as homogeneous regions. This letter presents a resegmentation approach applied to urban images. Resegmentation represents the set of adjustments from a previous segmentation in which the elements are small regions with a high degree of spectral similarity (a condition known as oversegmentation). The focus of this letter is the house roofs, which are assumed to have a rectangular shape. These regions are merged according to an objective function, which, in the technique presented here, maximizes the rectangularity. With oversegmentation, we create a graph known as a *region adjacency graph* (RAG) that relates border elements. The main contribution of this letter is a technique, which works with the RAG, to maximize the objective function in a relaxationlike approach that splits and merges oversegmented regions until they form a meaningful object. The results showed that the method was able to detect rectangles according to user-defined parameters, such as the maximum level of the graph depth and the minimum degree of rectangularity for objects of interest.

**Index Terms**—Graph theory, image classification, image segmentation, remote sensing.

## I. INTRODUCTION

IMAGE segmentation remains a great challenge in digital image processing. Techniques for segmentation, generally the first step in image analysis tasks, are expected to delineate the objects in the images, allowing further interpretation of such data. Several approaches have already been proposed in the literature, often covering specific areas of interest. Haralick and Shapiro [1] defined segmentation as a way to *separate the image into simple regions with homogeneous behavior*.

Segmentation is a broad area; it covers techniques for splitting one image into parts while considering a certain context. The context might include scale because image objects start with a single pixel. The image regions can be merged to produce elements with some meaning (the objects). The ultimate segmentation task is to extract the image objects, producing good results according to specific criteria. Moreover, the technique must be computationally efficient.

With regard to personal photographs, algorithms can distinguish faces or extract the background and stress the objects of interest, such as cars [2], [3], structures, and people [4], [5].



Fig. 1. Oversegmentation of roofs: Dark objects denote the background, yellow ones denote the foreground, and red ones are background objects that improve the rectangular segmentation of the foreground roof.

In remote sensing, the objects depend on the area of interest and generally include roofs [6], streets [7], and trees in urban images. In agricultural imagery, the algorithm should extract different crops [8] or deforested areas [9], which are used to distinguish land cover types by classification.

This letter presents a *resegmentation* approach applied to urban images. Resegmentation represents the set of adjustments from a previous segmentation, in which the elements are small regions with a high degree of spectral similarity, characterizing what is called an *oversegmentation*. These elements are merged according to one objective function, which maximizes the degree of rectangularity. We applied our technique to urban imagery in which the objects of interest (house roofs) are most often rectangular. Therefore, the resegmentation algorithm was programmed to detect rectangular objects.

Our approach was applied to optical multispectral images of urban areas. Such imagery has high spatial resolution, enough to distinguish the houses' roofs. An oversegmented image is first created by using traditional segmentation. With the oversegmentation, we create a graph known as a *region adjacency graph* (RAG) that relates border elements [10]. This graph is exploited to find graph cuts that maximize the objective function to detect rectangles. The main contribution of this letter is a technique to maximize the objective function in a relaxationlike approach that splits and merges oversegmented regions until they form a meaningful object.

The use of oversegmentation is important for generating segments with a smaller area than those of our target objects, i.e., roofs. The sun's illumination of a single roof, because of its distinct angulations, is different on every face of the roof. Such differences correspond to spectral differences, which, in traditional segmentation, result in more than one region for a single roof. It is expected that objects surrounding the roof will be small enough to fit the differences in the segmentation to match objects close to rectangles, as shown in Fig. 1.

Manuscript received May 28, 2010; revised October 20, 2010; accepted December 1, 2010. Date of publication January 19, 2011; date of current version June 24, 2011.

The authors are with the Image Processing Division (DPI), National Institute for Space Research (INPE), 12201-010 São José dos Campos, Brazil (e-mail: tkorting@dpi.inpe.br; dutra@dpi.inpe.br; leila@dpi.inpe.br).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LGRS.2010.2098389

In the next section, we discuss image segmentation in general and graph-based approaches specifically. In Section III, we present the resegmentation method, followed by Section IV, which contains the results and discussion. We provide our conclusions in Section V.

## II. IMAGE SEGMENTATION ANALYSIS

General image segmentation techniques are divided into two main classes. The first, called *pixel-oriented* segmentation, considers each pixel of the image as one graph node. Edges generally connect every pixel to its  $3 \times 3$  neighborhood. The second class considers regions as graph nodes. Such regions are pixel groups obtained by segmentation techniques. This class, called *region-oriented* segmentation, uses graph structures to represent the topology, connecting neighbor regions with graph edges. Neighboring regions apply the topological relation of *touch* [11].

Benediktsson *et al.* [12] presented a hybrid approach for segmenting urban images with high spatial resolution, applying morphological operations. Afterward, they applied a neural network to classify the extracted features. When dealing with panchromatic imagery, Gaetano *et al.* [13] segmented multispectral remote-sensing images, creating regions using panchromatic data. Subsequently, regions were merged using spectral features (i.e., *textural based*) from the remaining multispectral information.

The nature of urban areas is inherently complex. Although human operators may extract information from urban images easily, computer-based automated interpretation remains a challenging task. Cinque *et al.* [14] proposed the use of rectangles in image queries to express the properties of target regions. The authors used resegmentation to build the index of images in the database, which provided the features needed for determining the distance between the query and candidates [14].

### A. Oversegmentation

The term resegmentation defines a process of adjustments in a previously segmented image obtained by traditional methods, such as the watershed or region-growing methods [15]–[17]. The parameters of the original segmentation used by our technique were established in such a way that the result was composed of small regions with a high degree of spectral similarity, characterizing an oversegmentation.

Traditional approaches are tuned to match the objects as much as possible, losing several weak borders. The main reason for using oversegmentation is because it does not lose weak borders and, at the same time, establishes the object faces, such as those found in a roof object.

According to Comaniciu and Meer [18], oversegmentation involves breaking an image into many small regions from which any sought information can be assembled with some knowledge. This is recommended when the goal of the task is object recognition. This letter addresses the recognition of rectangles found in intraurban imagery. Oversegmentation is a necessary step in the process; by splitting or merging oversegmented elements, we can fit the rectangles with higher accuracy (see Fig. 1).

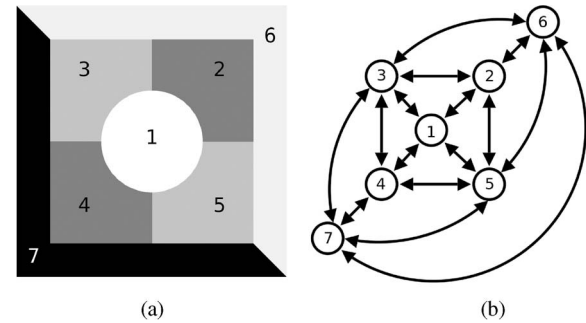


Fig. 2. RAG example. (a) Image with seven regions. (b) Produced graph.

### B. Graph Representation

The notation  $G = (V, E)$  represents one graph  $G$  with a set of nodes  $v_i \in V$ , and the set of connections is stored in  $E$  [16]. According to the employed segmentation, nodes will be pixels or segments. Borenstein *et al.* [19] combined a top-down segmentation and a bottom-up segmentation in a hierarchical graph structure. The top-down approach uses the representation learned from examples. The bottom-up approach uses image-based criteria to define coherent groups of pixels that are likely to belong together [19].

Zahn [20] first applied graph cuts to *minimum spanning trees* created from a pixel-based graph in which the edges were weighted as the differences among pixel intensities. Cuts were applied to the edges with larger weights. The maximum size of each edge is a user-defined threshold. However, depending on the threshold, simply making cuts at these edges may result in one region being split into multiple regions.

### C. RAG

A RAG is one data structure that provides a spatial view of an image. One description of a RAG involves the association of a vertex in each region with an edge in every pair of adjacent regions [17]. Fig. 2 shows an example image and the produced graph.

We created the RAG using oversegmented images, thus connecting adjacent regions in the graph representation. Subsets of the foreground regions connected to each other might define a rectangular object; therefore, they will be merged into a single segment. We propose a novel merging strategy. The oversegmented regions will be merged if they have similar spectral features (e.g., mean or variance) and if they maximize the objective function.

### D. Rectangularity Measure

High-resolution urban imagery contains rectangularly shaped objects, such as streets or buildings. However, because of image spectral properties and the atmospheric conditions under which they have been obtained, these objects present high spectral variance. Thus, by applying traditional segmentation algorithms to urban images, rectangular areas are not well segmented in most cases.

One approach to describe the rectangularity of an object, which relates the perimeters of a segment and its bounding box, was proposed by Benz *et al.* [21]. However, because of rotation, such a measure does not properly characterize the

rectangularity in many cases. Our approach of rectangular fitting is based on the work of Korting *et al.* [22]. The authors proposed one shape descriptor  $REC(p) \in [0, 1]$  called *rectangularity*, obtained by taking the ratio between the area and the bounding box area from element  $p$  [22]. To create one unbiased descriptor of rectangularity, the authors proposed using the main angle of the element, obtained from the first eigenvector of its internal coordinates. Let  $p$  be a polygonal object. By rotating  $p$ , we create a new object called  $p_r$ . Let  $AREA(s)$  be the area of object  $s$  and  $BOX(s)$  be the bounding box from object  $s$ . The rectangularity measure is then defined as follows:

$$REC(p) = \frac{AREA(p)}{AREA(BOX(p_r))} \quad (1)$$

with values falling within the interval  $[0, 1]$ . The closer to one the  $REC(p)$  is, the more rectangular the element  $p$  is. This measure is used as the objective function to detect rectangular roofs in urban imagery.

### III. METHODOLOGY

This section presents our approach to detecting rectangles in satellite images. As observed before, we focused on rectangles because several objects present in the urban scenario, mainly roofs, have such a shape. Roofs with particular shapes, e.g., “L” shaped, will not be focused on in this letter. Other rectangular objects found in intraurban images, like buildings and streets, will also not be addressed.

The main steps of the methodology, explained in detail later on, are as follows.

- 1) Obtain classified oversegmented regions.
- 2) Create a RAG description of the regions.
- 3) Filter the RAG to avoid unnecessary processing.
- 4) Perform resegmentation.
- 5) Return detected rectangles and background.

The first three steps denote the process initialization. The main step is the fourth one, and the fifth step relates to displaying the results.

#### A. Initialization

The methodology is applied to a classified set of regions from an oversegmented image, processed by a traditional method. Classification could be performed by any supervised approach and should divide the regions into two main groups, foreground and background. The foreground is composed of the elements of interest, i.e., house roofs. The rest of the image is considered background. Any classification technique may be employed in this step and should return only roofs in the foreground.

Oversegmentation generates several regions for the same object in the image, and the graph search considers all regions when searching for rectangles. In the initialization, the RAG is created; therefore, some prior manipulation of the RAG is performed in two steps, to reduce processing costs.

The first step merges the foreground regions that are surrounded only by other foreground regions. If a region has the same class as all of its first-order neighbors, this region is merged with a random neighbor. The second step aims to remove possible misclassification results. Because we are dealing with very small regions, it is assumed that one foreground

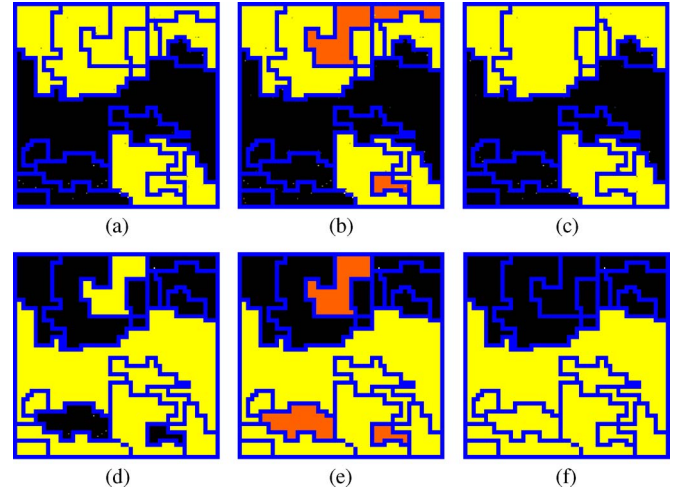


Fig. 3. Two steps of graph preprocessing. (a)–(c) show the first step, merging foreground regions that are only surrounded by foreground regions. (d)–(f) show the second step, removing possible misclassification results.

region without any foreground neighbor is misclassified. The same is applied to the background. When a region misclassification condition is detected, we reverse the region class. Fig. 3 shows both steps applied to synthetic data.

#### B. Resegmentation

The input parameters are defined as follows.

- 1) *Minimum rectangularity*: The minimum rectangularity needed to consider a merged set of regions as a rectangle.
- 2) *Maximum area for merging*: Neighboring regions can be merged only if they are smaller than this threshold, given as a percentage  $\alpha$  of the main foreground region.
- 3) *Graph levels  $M$* : The amount of  $M$ -order neighboring levels to be considered in a resegmentation.

The main parameter is the *minimum rectangularity*, and it must be tuned according to the needs of the user: Only well-defined rectangles<sup>1</sup> can be retrieved, or rough descriptions of roofs can be considered as well. The second parameter, *maximum area for merging*, avoids merging very large regions. The *graph level* is the connection order of the regions in the graph theory. First-order neighbors are the ones that touch a considered region. Second-order neighbors are the neighbors of the first-order ones, and so on. This parameter strongly depends on how oversegmented the input elements are and might be defined based on the visual inspection of the input.

Because there is a topological description of oversegmentation, the algorithm can pick one foreground region and start the graph search in a predefined level of neighbors. This level represents the number of regions that might satisfy the search for rectangles. An outline of the technique is provided as follows.

- 1) Randomly select one foreground region  $R$ .
- 2) Pick all  $N^{fg}$  foreground neighbors connected to  $R$  and, then, all  $M$ -order background neighbors  $N^{bg}$  using the RAG.
- 3) Let  $R_c$  be the merging of regions  $R$  and  $N^{fg}$ , and  $A \leftarrow \alpha \cdot AREA(R_c)$ .

<sup>1</sup>Practical exercises showed that a rectangularity higher than 0.70 returns well-defined ones.



- 4) Let  $N_s$  be the smallest available neighbor in  $N^{bg}$  or  $N^{fg}$ .
- 5) If  $AREA(N_s) \geq A$ , then go to step 13.
- 6) If  $N_s$  is in the foreground, then remove  $N_s$ .
- 7) Otherwise, insert region  $N_s$  in  $R_c$ .
- 8) Test the rectangularity  $REC(R_c)$ .
- 9) If the rectangularity has increased, then keep region  $N_s$ .
- 10) Otherwise, remove region  $N_s$ .
- 11) Let  $N_s$  be the next smallest available neighbor.
- 12) Return to step 5, or continue when no region is left.
- 13) If the final rectangularity is higher than the *minimum rectangularity*, then merge regions and store the result.
- 14) Randomly select another unprocessed foreground region, and return to step 2; otherwise, exit.

By inserting or removing regions in the process of rectangle detection, the interaction over regions characterizes a relaxationlike strategy. When no foreground element remains, the algorithm outputs the rectangles as foreground and the remaining regions as background.

#### IV. RESULTS AND DISCUSSION

In this section, we present some results from this algorithm. We show the oversegmentation and the final resegmentation with different parameters. All oversegmentations have been obtained using the region-growing approach based on the work of Bins *et al.* [23], with a ten-pixel threshold for the minimum area and pixel similarity to generate small segments. These thresholds were manually adjusted to be the largest value that did not lose roof edge details.

We used two different sections of a QuickBird image from São José dos Campos, Brazil, acquired in December 2005, with a spatial resolution of 1 m (fusion product). For the experiments, we used only the three bands corresponding to the visible wavelengths.

For the input parameters, we defined a minimum rectangularity value of 0.70 and a maximum area for merging of the foreground objects of 30%. The graph levels have been defined visually by selecting small image croppings and resegmenting them with different depths. Such a process can also be performed automatically by comparing the resultant objects with reference object samples or by testing different levels and storing the average rectangularity for further analysis.

The first result shows the differences obtained by changing the graph depths. Two roofs in the image contain a rectangular shape, and one is irregular because of the image crop. Fig. 4 shows the resegmentation results. We can see that, because of the large number of segments, only the result utilizing three graph levels could detect the rectangles.

In the second cropping, we exploited the differences between the input oversegmentation and the final resegmentation with proper parameters. Fig. 5 shows the input image, followed by the oversegmentation and the resegmentation using three levels of graph depth. Some rectangles were omitted to simplify the visualization.

To perform a quantitative assessment of the algorithm, we generated a manual segmentation of roofs in four proof images. The amount and the area of the manual segments were then compared with the ones generated by resegmentation and by one traditional approach (based on the work of Bins *et al.* [23]), whose parameters were optimized for the best visual

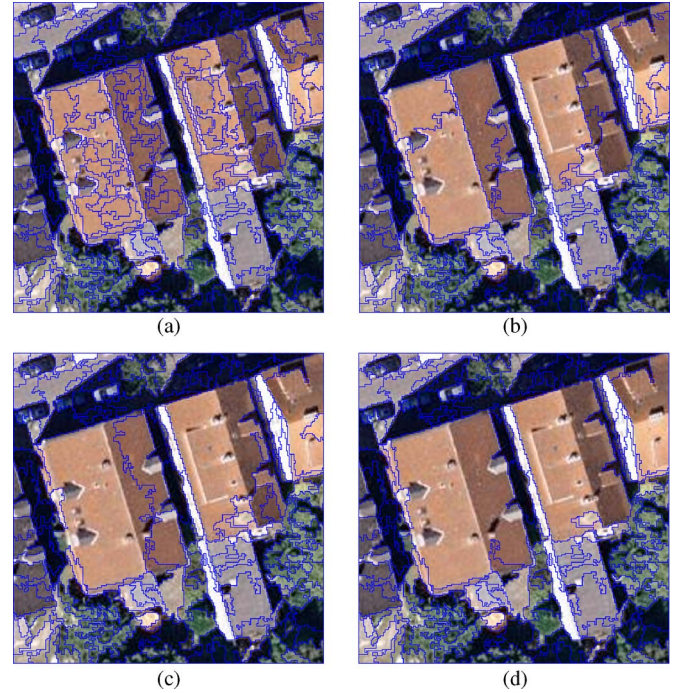


Fig. 4. Urban resegmentation. (a) Oversegmentation. (b)–(d) Resegmentation with one, two, and three levels, respectively.

segmentation. It was expected that the number of resegmented roofs would be similar to the number of manual segments. Moreover, the area of the output segments should match the manual segmentation.

For this purpose, the metrics  $RMA$  and  $RMSE$  were computed.  $RMA$  measures the *ratio* between the number of segments generated *manually* and that generated *automatically*. The metric  $RMSE$  stands for the *root mean squared error* obtained from the quadratic difference between the areas of the output elements obtained by the manual and automatic segmentations.

Let  $A$  be the set of automatically segmented elements and  $M$  be the set of manual segments. Let  $\#(\cdot)$  be an operator of cardinality, i.e., the number of elements inside a set. The metric  $RMA$  is defined by the following equation:

$$RMA = \frac{\#(M)}{\#(A)}. \quad (2)$$

$RMA$  is expected to be close to one, which means that the automatic segmentation is similar to the reference segmentation. The  $RMSE$  is expected to be close to zero; a smaller error means a more accurate segmentation. The obtained results are shown in Table I.

In the four experiments, the resegmentation approach matched the rectangles with a lower  $RMSE$  than the traditional approach of segmentation. It was observed that the traditional approach has somewhat oversegmented the regions corresponding to roofs, although it was visually tuned to the best match for rectangles. Moreover, the number of segments obtained by our approach is closer to that of the reference segmentation. This addresses our claim that traditional approaches lose weak borders, many of which are necessary to represent rectangles precisely.

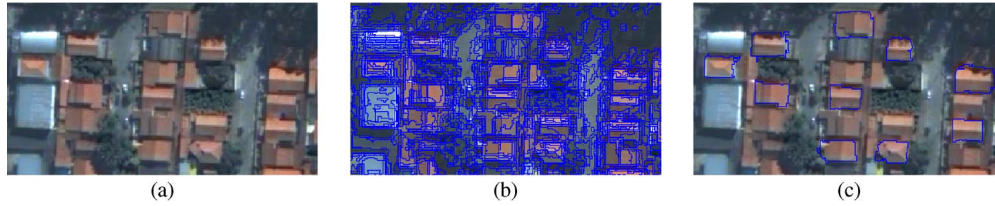


Fig. 5. (a) Input image. (b) Oversegmentation. (c) Resegmentation.

TABLE I  
QUANTITATIVE EVALUATION OF RESEGMENTATION COMPARED  
WITH A TRADITIONAL APPROACH

Experiment	Resegmentation		Traditional	
	<i>RMA</i>	<i>RMSE</i>	<i>RMA</i>	<i>RMSE</i>
1	0.834	0.251 pixels	0.416	0.715 pixels
2	0.625	0.354 pixels	0.385	0.631 pixels
3	0.769	0.574 pixels	0.455	1.357 pixels
4	0.667	0.215 pixels	0.527	0.446 pixels
$\mu$	0.724		0.446	

## V. CONCLUSION

One approach for the resegmentation of rectangular shapes was presented. It was used to detect roofs in remote-sensing high-resolution intraurban imagery, which often contain objects of high rectangularity. By replacing the objective function with any other shape detector, such as a circle or a triangle, for instance, our approach could be applied to other shapes in addition to rectangles.

The algorithm has certain drawbacks that should be considered. One of them is the processing time. Because the algorithm is executed in several steps, each one presents its own computational complexity. The RAG creation and filtering steps perform multiple comparisons with every element to create the topological representation. These operations have a computational complexity of  $O(n^2)$ . The remaining operations have a complexity of  $O(n)$ ; therefore, the overall complexity is  $O(n^2)$ . By storing a prior description of the oversegmentation topology, it is possible to reduce the amount of comparisons required to build the RAG.

One next step is to automate the search for the best graph depths, which, in this case, was performed by visual inspection. The algorithm was developed using the free library TerraLib, available at <http://www.terralib.org/>.

## REFERENCES

- [1] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Comput. Vis. Graph. Image Process.*, vol. 29, no. 1, pp. 100–132, Jan. 1985.
- [2] D. Koller, K. Danilidis, and H. Nagel, "Model-based object tracking in monocular image sequences of road traffic scenes," *Int. J. Comput. Vis.*, vol. 10, no. 3, pp. 257–281, Jun. 1993.
- [3] B. Leibe, A. Leonardis, and B. Schiele, "Combined object categorization and segmentation with an implicit shape model," in *Proc. ECCV Workshop Stat. Learn. Comput. Vis.*, 2004, pp. 17–32.
- [4] S. Z. Li and A. K. Jain, *Handbook of Face Recognition*. New York: Springer-Verlag, 2005.
- [5] R. Feris, V. Krueger, and R. Cesar, "A wavelet subspace method for real-time face tracking," *Real-Time Imaging*, vol. 10, no. 6, pp. 339–350, Dec. 2004.
- [6] A. L. Chesnel, R. Binet, and L. Wald, "Object oriented assessment of damage due to natural disaster using very high resolution images," in *Proc. IEEE IGARSS*, Jul. 2007, pp. 3736–3739.
- [7] Y. He, H. Wang, and B. Zhang, "Color-based road detection in urban traffic scenes," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 4, pp. 309–318, Dec. 2004.
- [8] A. Pérez, F. López, J. Benlloch, and S. Christensen, "Colour and shape analysis techniques for weed detection in cereal fields," *Comput. Electron. Agric.*, vol. 25, no. 3, pp. 197–212, Feb. 2000.
- [9] M. Silva, G. Câmara, R. Souza, D. Valeriano, and M. Escada, "Mining patterns of change in remote sensing image databases," in *Proc. 5th IEEE Int. Conf. Data Mining*, New Orleans, LA, 2005, pp. 362–369.
- [10] R. Schettini, "A segmentation algorithm for color images," *Pattern Recognit. Lett.*, vol. 14, no. 6, pp. 499–506, Jun. 1993.
- [11] M. Egenhofer and R. Franzosa, "Point-set topological spatial relations," *Int. J. Geographical Inf. Sci.*, vol. 5, no. 2, pp. 161–174, 1991.
- [12] J. Benediktsson, M. Pesaresi, and K. Amason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pt. 1, pp. 1940–1949, Sep. 2003.
- [13] R. Gaetano, G. Scarpa, and G. Poggi, "Hierarchical texture-based segmentation of multiresolution remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2129–2141, Jul. 2009.
- [14] L. Cinque, F. De Rosa, F. Lecca, and S. Levialdi, "Image retrieval using resegmentation driven by query rectangles," *Image Vis. Comput.*, vol. 22, no. 1, pp. 15–22, Jan. 2004.
- [15] A. Duarte, A. Sánchez, F. Fernández, and A. Montemayor, "Improving image segmentation quality through effective region merging using a hierarchical social metaheuristic," *Pattern Recognit. Lett.*, vol. 27, no. 11, pp. 1239–1251, Aug. 2006.
- [16] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Comput. Vis.*, vol. 59, no. 2, pp. 167–181, Sep. 2004.
- [17] A. Tremeau and P. Colantoni, "Regions adjacency graph applied to color image segmentation," *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 735–744, Apr. 2000.
- [18] D. Comaniciu and P. Meer, "Robust analysis of feature spaces: Color image segmentation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 1997, pp. 750–755.
- [19] E. Borenstein, E. Sharon, and S. Ullman, "Combining top-down and bottom-up segmentation," in *Proc. Conf. Comput. Vis. Pattern Recog. Workshop*, 2004, p. 46.
- [20] C. Zahn, "Graph-theoretical methods for detecting and describing gestalt clusters," *IEEE Trans. Comput.*, vol. C-20, no. 1, pp. 68–86, Jan. 1971.
- [21] U. Benz, M. Baatz, and G. Schreier, "Oscar-object oriented segmentation and classification of advanced radar allow automated information extraction," in *Proc. IEEE IGARSS*, 2001, vol. 4, pp. 1913–1915.
- [22] T. Korting, L. Fonseca, L. Dutra, and F. Silva, "Image re-segmentation—A new approach applied to urban imagery," in *Proc. 3rd Int. Conf. Comput. Vis. Theory Appl., VISAPP*, 2008, pp. 467–472.
- [23] L. Bins, L. Fonseca, G. Erthal, and F. Li, "Satellite imagery segmentation: A region growing approach," in *Proc. Simpósio Brasileiro de Sensoriamento Remoto*, 1996, vol. 8, pp. 677–680.