



## **Region similarity measures for automatic detection of building roofs from orthophotoplans**

Abdellatif Elidrissi, Youssef El Merabet, Yassine Ruichek, Cyril Meurie,  
Abderrahmane Sbihi, Ahmed Moussa

### **► To cite this version:**

Abdellatif Elidrissi, Youssef El Merabet, Yassine Ruichek, Cyril Meurie, Abderrahmane Sbihi, et al.. Region similarity measures for automatic detection of building roofs from orthophotoplans. SITIS 2014 - 10th International Conference on Signal Image Technology & Internet based Systems, Nov 2014, Marrakech, Morocco. 8p. <hal-01263487>

**HAL Id: hal-01263487**

**<https://hal.science/hal-01263487v1>**

Submitted on 27 Jan 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Region Similarity Measures for Automatic Detection of Building Roofs from Orthophotoplans

Abdellatif El Idrissi <sup>\*</sup>, Youssef El Merabet <sup>†</sup>, Yassine Ruichek <sup>‡</sup>, Cyril Meurie <sup>§</sup>, Abderrahmane Sbihi <sup>\*</sup> and Ahmed Moussa <sup>\*</sup>

<sup>\*</sup> National School of Applied Sciences

Abdelmalek Essaadi University, 90000 Tangier, Morocco

Email: elidrissi@ensat.ac.ma

<sup>†</sup> LASTID Laboratory, Department of Physics, Faculty of Sciences

Ibn Tofail University, B.P 133, 14000 Kénitra, Morocco

<sup>‡</sup> IRTES-SeT Laboratory, University of technology of Belfort Montbéliard,

13, Ernest Thierry-Mieg, 90010, Belfort, France

<sup>§</sup> University of Lille Nord de France, F-59000 Lille,

IFSTTAR, LEOST, F59650, Villeneuve d'Ascq, France.

**Abstract**—In this paper, an efficient and automatic method for detection of multiple-objects of interest from images is presented. This method is based on using region similarity measures. The method starts by constructing two knowledge databases in which significant and distinctive textures extracted from both objects of interest and background are respectively represented. The proposed procedure continues by an initialization step in which the processed image is segmented into homogeneous regions. In the purpose of separating objects of interest from image background, an evaluation of the similarity between regions of the segmented image and those of the constructed knowledge databases is then performed. The main advantages of this method are simplicity, applicability and suitability. Applying this method on building roof detection from orthophotoplans has shown its robustness and performance.

**Keywords**—Object detection; Region Similarity Measure; Feature extraction; Orthophotoplans.

## I. INTRODUCTION

In conception tasks of sophisticated and modern intelligent systems, object recognition component is often omnipresent. For this reason, automatic object recognition, has become a topic of growing interest for computer vision community. For instance, automatic extraction of man-made objects such as buildings and roads in urban areas has gained significant attention over the last decade. This problem is generally considered when we deal with high-level image processing in order to produce numerical or symbolic information. In this context, several methods have been proposed in the literature. Among the techniques most frequently used, one can cite semi-automatic methods that need user interaction to extract desired targets or objects of interest from images. Generally, this category of methods has been introduced to alleviate problems inherent to fully automatic segmentation that seems to never be perfect. These methods aim to produce a binary segmentation of the image: "object" and "background" regions. Interactivity consists in imposing

certain hard constraints for segmentation by pointing out certain pixels (seeds) that absolutely have to be part of the object and certain pixels that have to be part of the background.

To give a brief review of such methods, let's begin by citing the work of Boykov and Jolly [1]. In their paper, they developed an Interactive Graph Cuts (IGC) method for image segmentation. They performed this operation by using the min-cut/max-flow algorithm. User scribbles extract color information which will be used thereafter as hard constraints. An iterative algorithm called GrabCut by simplifying user interaction is also presented in [2] by Rother et al.. Their idea was to combine image segmentation, using graph cut, and Gaussian Mixture Models (GMMs) which are statistical powerful models for pattern recognition. Recently, a very useful segmentation benchmark with a platform implementing important algorithms has been proposed by McGuinness and Connor in [3]. Authors compared important algorithms such as IGC [1], Seeded Region Growing (SRG) [4], Simple Interactive Object Extraction (SIOX) [5] and Binary Partition Tree (BPT) [6]. This comparison provides a good coverage of the various techniques currently available for foreground extraction.

The SRG method, proposed by Adams and Bischof in [4], is very popular due to its simplicity and speed. It assumes that regions of interest are characterized by connected pixels with similar colors. The SIOX algorithm [5], based on color information, has recently been integrated into the popular imaging program GIMP as "Foreground Selection Tool". As for the BPT algorithm [6], it's based on hierarchical region segmentation and employs user interaction to split and merge regions in the tree. Fast kernel density estimation [8] for color statistics has been also used by Bai and Sapiro in [7] to improve the geodesic distance-based approach as described in [9].

A very recent algorithm that highly inspired us is the work done by Ning et al. [10]. They have proposed a novel maximal similarity based region merging (MSRM) mechanism for interactive image segmentation. The key idea of MSRM is to perform region merging between adjacent regions by exploiting an effective representation for color statistics based on (quantized) color histograms computed from the regions. First, the input image is segmented using the mean shift segmentation algorithm. User must then indicate location and regions of the object and background by using strokes as markers. Finally, a maximal-similarity based region merging mechanism is used in order to separate the object of interest from the background image while relying on the help of markers introduced by the user. A similar algorithm also based on maximal similarity based region merging has been proposed in [11]. The difference is that this algorithm uses regions as seeds and takes the regions as growth units for the region growing (i.e. adjacent regions merging).

Results obtained by all the methods mentioned above are in most cases accurate and good. However, they remain dependent on the degree of user interaction.

In the goal of avoiding this user interaction, automatic methods have been developed. Some of them tend to inspire from techniques used in pattern recognition and machine learning domains [22], [21] and [28]. Some others consist in joining prior information, like height data or Light Detection and Ranging (LIDAR) data, to classification algorithms as in [29], [30], [31]. Points of interest, that allow representing characteristics of targeted objects, have been also used to build automatic methods for object recognition as in [33], [34].

Generally, these methods are efficient, but finding the optimal building extraction / delineation requires a global optimization framework redhibitory over large scales. In addition they remain most of the time context-dependent and are often complexes leading to multiple parameters setting and significant computing times for large areas.

The method we propose is simple, but regarding the related literature, it presents numerous advantages. Firstly, it allows automatic extraction of interest objects and does so without any user interaction. Secondly, applying this method, it becomes possible to accurately detect multiple objects in the same time from a given image. In addition, one can achieve good results when considering complex images for which both foreground and background regions have similar colors. Moreover, the method can be applied in several fields like medical image processing (e.g cancer cell recognition) and remote sensing image processing (e.g vegetation and buildings detection). In this paper, we are interested in extracting building roofs from orthophotoplans.

The rest of this paper is organized as follows. In section II, we explain the proposed methodology and describe its main steps in details. Experiments and evaluations both

qualitative and quantitative are presented in section III in order to demonstrate the algorithm's potentials and efficiency. Section IV concludes the paper and addresses some perspectives.

## II. DESCRIPTION OF THE PROPOSED METHODOLOGY

### A. General presentation

The new proposed method incorporates two major stages: off-line and on-line stages. In the off-line stage, two knowledge databases have to be created in order to robustly avoid user interaction. One contains all distinctive textures of the objects to extract. The other is composed of textures picked up from the remaining objects that represent the background of the image. If we take the application of building roof extraction from aerial images as an illustrative example, the first knowledge database **B<sub>obj</sub>** will be constructed with  $m$  distinctive textures of building roofs while the second one **B<sub>back</sub>** will be constructed with  $n$  distinctive textures of other objects such as vegetation, road, forest,...etc. These two databases are provided considering some examples of images. Having these two knowledge databases **B<sub>obj</sub>** and **B<sub>back</sub>** as reference, it is now possible to automatically extract building roofs from any aerial image (orthophotoplan, in our case). Figure 1 illustrates an example of knowledge databases used in this work.

In the on-line stage, the object recognition process is performed. To do this, we begin by over-segment the initial image into many small and homogeneous regions. This is called a low-level processing step. In this paper, we have used SRM algorithm [12] (section II-B) as a tool of segmentation. Having a segmented image, the following task is a high-level processing step that consists of extracting features characterizing regions of both segmented image and constructed knowledge databases. We use RGB color histogram features (section II-C). The question that arises then is how can we measure the similarity between those regions. Several well-known goodness-of-fit statistical metrics using RGB color histogram features exist in the literature. In this work, the Bhattacharyya measure (section II-D) is adopted to accomplish this operation.

Once similarity measure is evaluated for all regions, each one of them can be classified as a part of an object of interest or rather as a part of the background of the image (section II-E). The procedure is finally, terminated by delineating object contours keeping only regions labelled as building roofs (objects of interest).

Figure 2 resumes the general flowchart of the proposed building roof detection method.

### B. Initial segmentation using Statistical Region Merging

The low-level processing step consists in over-segmenting the input image into many small and homogeneous regions with same properties. The goal of this initial segmentation, is to avoid the under-segmentation problem and thus correctly

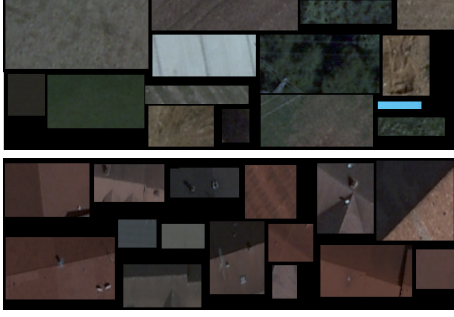


Figure 1. Example of knowledge databases used in this work. From top to bottom: knowledge database  $B_{back}$  of background (vegetation, road, forest, etc) and knowledge database  $B_{obj}$  of building roofs (red and non-red rooftop buildings).

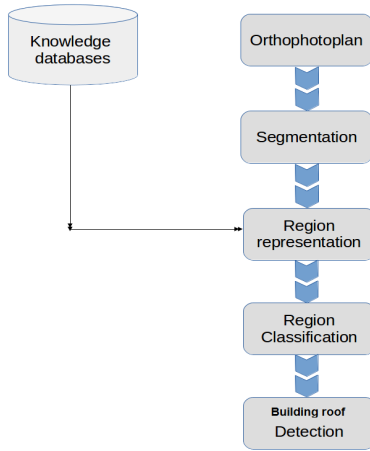


Figure 2. General flowchart of the proposed building roof detection method.

extract all significant regions where boundaries coincide as closely as possible with the significant edges characterizing the image. Of course, there are many low level segmentation methods in the literature. One can cite Mean-shift, Jseg unsupervised segmentation algorithm [13], watershed, Turbopixels [14], Statistical Region Merging (SRM) [12], etc. In this paper, we have chosen SRM algorithm to obtain this initial segmentation of the input image. Particular advantages of using this algorithm for dealing with large images are: it dispenses dynamical maintenance of a Region Adjacency Graph (RAG), allows defining a hierarchy of partitions and runs in linear-time by using bucket sorting algorithm while transversing the RAG. In addition, the SRM segmentation method does not only consider spectral, shape and scale information, but also has the ability to cope with significant noise corruption and handle occlusions. Figure 3 shows a segmentation result of the SRM method.

### C. Region feature extraction

In this stage of our method, we dispose of a segmented image obtained via the SRM algorithm. It is still a challeng-



Figure 3. Example of segmentation result using Statistical Region Merging (SRM) method. From left to right: Original image and its SRM segmentation result.

ing problem to accurately extract the object contours from this image because only segmented regions are calculated and no information estimation on their content, which is necessary for the extraction process, is yet done. Our main goal consists in classifying each segmented region as target object or background. For this purpose, we need first to characterize these regions using suitable descriptors.

It appears from the literature that there are several aspects that could be considered for representing a region such as edge [15], texture [16], shape, size or color. For our purpose, the most appropriate among them is the last one. In fact, the region texture which can be understood as repeatedly occurring local patterns in regions and their arrangement rules, is unfortunately difficult to describe; The same difficulties can be faced regarding shape and edge. Although it can be measured simply by computing the number of pixels, region size doesn't allow to distinct objects of interest since they can have different sizes from an image to an other or simply they can have the same size as other objects belonging to the background of the image. Hence, region color which can be tackled using simply its mean value or its histogram is an effective parameter to describe statistical information of object color distribution. Note that region histograms are local histograms and they reflect local features in images. Therefore, we exploit the color histogram to represent all regions of the segmented image and those of the constructed knowledge databases.

In this purpose, we first, uniformly quantize each color channel into  $l=16$  bins; afterwards, the color histogram of each region is calculated in the feature space of  $1 \times 1 \times 1 = 4096$  bins. Obviously, quantization reduces the information regarding the content of regions and it is used as trade off when one wants to reduce processing time. The RGB color space is used in order to perform these computations.

Now that we have defined the feature adopted for characterizing regions, the key issue is to determine the similarity between regions of the segmented image and those of the



constructed knowledge databases. In other words, we need to define the similarity measure rule  $\varrho(R, Q)$  between two regions R and Q basing on their color histograms.

#### D. Similarity measure rules

The most similarity measures commonly used are based on vector space model, i.e. taking image region features as points in the vector space, through the calculation of close degree of two points to measure the similarities between the image region features. Common similarity measures include Minkowski measure, histogram intersection method [17], second type distance [18], Bhattacharyya coefficient [19], and log-likelihood ratio statistic [20], etc. For two regions R and Q, using the notation  $\varrho(R, Q)$  for representing the similarity between them, the larger  $\varrho$  is, the larger similarity we will get. Denote by  $Hist_R^i$  the normalized histogram of a region R, the superscript  $i$  represents its  $i^{th}$  element and  $z = l \times l \times l = 4096$ .

In this work, we adopt Bhattacharyya coefficient which is given as:

$$\varrho(R, Q) = \sum_{i=1}^z \sqrt{Hist_R^i \cdot Hist_Q^i} \quad (1)$$

It represents the cosine of angle between the unit vectors  $(\sqrt{Hist_R^1}, \dots, \sqrt{Hist_R^z})^T$  and  $(\sqrt{Hist_Q^1}, \dots, \sqrt{Hist_Q^z})^T$ . This choice is due to its ability to simulate precisely the similarity value of vector shape. The higher the Bhattacharyya coefficient is, the higher the similarity is. That is to say their histograms are very similar and the angle between the two histogram vectors is very small. Certainly, two similar histograms do not necessarily involve that the two corresponding regions are perceptually similar. Nevertheless, coupling with the proposed classification process introduced in the next section (section II-E), the Bhattacharyya similarity works well in the proposed approach.

We mention that a histogram is a global descriptor of a local region and it is robust to noise and small variations. Given that the Bhattacharyya coefficient is the inner product of two histogram vectors, this coefficient is thus robust to noise and small variations too. It has been used in [10], [11] for image segmentation. Unlike these methods, our proposed technique aims for multiple extraction of objects of interest using the two constructed knowledge databases without any need for user to provide markers input usually necessary for region merging process.

#### E. Classification step

At this stage of our method, we aim to determine which of the two classes (objects of interest or background) will be affected to the regions composing the initial SRM segmentation result, which we denote  $\mathbf{M}_{SRM}$ . For this end, we identify candidate regions of  $\mathbf{M}_{SRM}$  that have a maximal

similarity with the knowledge database  $\mathbf{B}_{obj}$  of objects and those having a maximal similarity with the knowledge database  $\mathbf{B}_{back}$  of background. Once all regions of  $\mathbf{M}_{SRM}$  are classified, this leads directly to extracting the desired objects (building roofs in our case). The proposed object detection method can be summarized as in algorithm 1.

---

#### Algorithm 1 Multiple objects recognition algorithm

---

**Require:**  $I \leftarrow$  input image.

$\mathbf{B}_{obj} \leftarrow$  Knowledge database of objects of interest (building roofs).

$\mathbf{B}_{back} \leftarrow$  Knowledge database of background (vegetation, road, forest, etc)

- 1: (over)Segment  $I$  into regions through SRM algorithm in order to obtain the set  $\mathbf{M}_{SRM}$  of segmented regions.
  - 2: Calculate the RGB color histogram features for all regions of  $\mathbf{M}_{SRM}$  and for those composing the two constructed knowledge databases  $\mathbf{B}_{obj}$  and  $\mathbf{B}_{back}$ .
  - 3: **for** each candidate region  $R \in \mathbf{M}_{SRM}$  **do**
  - 4: Calculate the similarity vector  $V_{obj}^R = \{\varrho(R, Q_i); (Q_i)_{i=1..m} \in \mathbf{B}_{obj}\}$  between R and  $\mathbf{B}_{obj}$ .
  - 5: Calculate the similarity vector  $V_{back}^R = \{\varrho(R, Q_j); (Q_j)_{j=1..n} \in \mathbf{B}_{back}\}$  between R and  $\mathbf{B}_{back}$ .
  - 6: Calculate  $moy_{obj}^R = \frac{\sum_{i=1}^q \varrho(R, Q_i)}{q}$ , the mean of the  $q$  first elements of  $V_{obj}^R$  ( $q \leq m$ ).
  - 7: Calculate  $moy_{back}^R = \frac{\sum_{j=1}^p \varrho(R, Q_j)}{p}$ , the mean of the  $p$  first elements of  $V_{back}^R$  ( $p \leq n$ ).
  - 8: **if** ( $moy_{obj}^R \geq moy_{back}^R$ ) **then**
  - 9: The region R is classified as a part of building roof.
  - 10: **else**
  - 11: The region R is classified as a part of background.
  - 12: **end if**
  - 13: **end for**
  - 14: The classification process is finished.
  - 15: **return** The final detection map.
- 

As one can state, the similarity rule is very simple but it is efficient for the classification process. Its important advantage is that it avoids the presetting of similarity threshold for classification control. For the purpose of keeping a significant value of the mean value of similarity, avoid the dispersion phenomenon and hence obtain a good classification result, the two values  $moy_{obj}^R$  and  $moy_{back}^R$  are only calculated on the first  $q$  and  $p$  values of the sorted similarity vectors  $V_{obj}^R$  and  $V_{back}^R$  respectively. Although the mean values of similarity  $moy_{obj}^R$  and  $moy_{back}^R$  are sensitive to outlier values of  $q$  and  $p$ , we empirically found that there is a range of values where the classification result remains stable.

### III. EXPERIMENTAL RESULTS

In this section, we are interesting in assessing the ability of the proposed building-extraction method to deal with multiple detection of building roofs from orthophotoplans. This kind of images cover wide areas, where appear complex

and multiple objects of different classes, various shadows, occlusions and multiple colors and textures. As pointed out in the introduction, our building-extraction algorithm run automatically without any user interaction. To avoid each time calculating region features from the two constructed knowledge databases and thus reduce the computation time, we can calculate them once and save them in a binary file.

First, we evaluate the proposed method qualitatively by using several representative examples. Figure 4 illustrates results of roof detection over the set of processed images. In the upper row of this figure, we show the original images; in the midst row, the segmented images are given; the lower row shows the corresponding building roof extraction where the final detected building boundaries drawn with red color are superimposed upon the original images. Basing on visual evaluation of the results, one can state that the developed approach demonstrates excellent accuracy in terms of building boundary extraction, this means that the majority of the building roofs present in the image are detected with good boundary delineation. Indeed, our method gives reliable results across complex environment composed of buildings presenting red and non-red rooftop, road areas, vegetation, etc. The images of Figures 4.a, 4.b and 4.c include several rooftop buildings and road areas with same color and texture, the proposed approach is able to successfully distinguish between them.

However, as one can see from the experimental results of Figure 5, due to radiometric similarity between building roofs and image background, some false or imperfect detections can be generated. In fact, although we obtained notably accurate multiple detection of building roofs, our method missed some part of buildings when the contrast between their rooftop and the background is low. Also, some vegetation areas are extracted as part of buildings because of their radiometric characteristics which are similar. In Figure 5, some of building parts that have not been extracted are pointed out by yellow ellipses while some false detections are pointed out by green ellipses.

As for quantitative evaluation, we use the evaluation measures widely employed in measuring effectiveness and that constitute a useful and accepted tool in the object recognition field [35]. Within the orthophotoplans used in this work, 100 buildings were first manually delineated. Then, they are used as a reference building set to assess the accuracy of the automated building extraction. The extraction results and manually ones are compared pixel-by-pixel. Each pixel in the images is categorized as one of four types: 1/ True positive (TP), 2/ True negative (TN), 3/ False positive (FP), 4/ False negative (FN).

To examine detection performance, the number of pixels that fall into each of the four categories TP, TN, FP, FN are determined, and the following measures are computed: 1/ Branching Factor (B.F), 2/ Miss Factor (M.F), 3/ Detection Percentage (D.P), 4/ Quality Percentage (Q.P).

The results of the quality assessment of the method for the images illustrated in the figures 4 and 5 are given in table I. The last line gives the average values obtained on the full orthophotoplans used in this work. The values obtained on the set of processed images agree with the findings mentioned above regarding the performance of the proposed approach. Effectively, the results show that the building-extraction approach is quite successful for extracting the buildings from orthophotoplans with D.P and Q.P average values of 92.60% and 84.83%, respectively. In addition to this, the branching factor B.F and the miss factor M.F average values were found to be 0.105 and 0.083, respectively.

Table I  
THE QUALITY ASSESSMENT RESULTS OF THE BUILDING EXTRACTION.

<b>Images \ measures</b>	<b>B.F</b>	<b>M.F</b>	<b>D.P</b>	<b>Q.P</b>
Fig.4.a	0.1745	0.1116	90	77.75
Fig.4.b	0.14	0.0289	97.19	85.58
Fig.4.c	0.0871	0.1126	89.87	83.35
Fig.5	0.0184	0.0789	93.37	92.66
<b>Avr./100 building roofs</b>	<b>0.105</b>	<b>0.083</b>	<b>92.60</b>	<b>84.83</b>

Considering the complexity side of the algorithm, we note that it requires in average 5 seconds for extracting building roofs from images of about 1500 by 1000 pixels working on a machine with a processor of 2.75 MHZ(CPU) and a memory of 3 GO (RAM).

#### IV. CONCLUSION

In this paper, we have introduced an efficient method for automatic and accurate multiple objects detection from images. Unlike interactive methods, the proposed one does not require any user interaction or any setting of algorithm parameters (a threshold of similarity for example). The proposed method involves two knowledge databases where the first one is constructed with several distinctive textures of objects to be extracted and the second one is composed with textures representing background. After an over-segmentation of the original image, the segmented regions are classified as objects or background using a region similarity measure. The application of the proposed method on building roof extraction from orthophotoplans, that is very challenging regarding the complexity of scenes which contain a large number of buildings, has shown its performance.

In order to improve the method, there are some open questions that we still need to explore. For instances, the color histogram features are calculated using RGB color space. The orthophoto plan images in our possession contain a certain heterogeneity in terms of lights, illumination changes, shadows, etc, what constitutes a breeding ground for false detections. To overcome these drawbacks and hence reduce the effect of illumination and limit the artefacts of the acquired image, studying and evaluating different

color spaces and/or colorimetric invariants seem to be an interesting way forward [36].

#### ACKNOWLEDGMENT

The authors would like to thank the "Communauté d'Agglomération Belfortaine" for making available the "orthophotoplan" images on which the tests have been performed.

#### REFERENCES

- [1] Boykov, Y.; Jolly, M.: Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. *IEEE Intl. Conf. on Comput. Vision.* 1, 105-112 (2001)
- [2] Rother, C.; Kolmogorov, V.; Blake, A.: GrabCut -Interactive Foreground Extraction using Iterated Graph Cuts. *ACM Transactions on Graphics.* 23(3), 309-314(2004)
- [3] McGuinness, K.; O'Connor, N. E.: A comparative evaluation of interactive segmentation algorithms. *Pattern Rec.* 43, 434-444 (2010)
- [4] Adams, R.; Bischof, L.: Seeded region growing. *IEEE Trans. Pattern Anal. Mach. Intell.* 16, 641-647 (1994)
- [5] Friedland, G.; Jantz, K.; Rojas, R.: SIOX: simple interactive object extraction in still images. *IEEE Intl. Symposium on Multimedia.* 253-259 (2005)
- [6] Adamek, T.: Using contour information and segmentation for object registration, modeling and retrieval. PhD thesis, Dublin City Univ (2006)
- [7] Bai, X.; Sapiro, G.: Geodesic matting: a framework for fast interactive image and video segmentation and matting. *Intl. J. Comput.* 82, 113-132 (2009)
- [8] Yang, C.; Duraiswami, R.; Gumerov, N. A.; Davis, L.: Improved fast gauss transform and efficient kernel density estimation. *IEEE Intl. Conf. on Comput. Vision.* 1, 664-671(2003)
- [9] Protiere, A.; Sapiro, G.: Interactive image segmentation via adaptive weighted distances. *IEEE Transactions on Image Processing.* 16, 1046-1057 (2007)
- [10] Ninga, J.; Zhanga, L.; Wub, D. Z. C.: Interactive image segmentation by maximal similarity based region merging. *Pattern Recognition.* 43, n 2. 445-456 (2010)
- [11] Chen, M.; Hou, Y.; Zeng, X.; Lu, Y.: A multi-label interactive image segmentation method based on region merging. *IEEE International Conference on Mechatronic Science, Electric Engineering and Computer, Jilin, China.* 2336-2340 (2011)
- [12] Nock, R.; Nielsen, F.: Statistical region merging. *IEEE Trans. Pattern Anal. Machine Intelligence.* 26, n 11, 1452-1458 (2004)
- [13] Deng, Y.; Manjunath, B. S.: Unsupervised Segmentation of Color-Texture Regions in Images and video. *IEEE Trans. Pattern Anal. Mach. Intell.* 23, 800-810 (2000)
- [14] Levinshtein, A.; Stere, A.; Kutulakos, K. N.; Fleet, D. J.; Dickinson, S. J.; Siddiqi, K.: Turbo Pixels: Fast Super pixels Using Geometric Flows. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 31, 2290-2297 (2009)
- [15] Birchfield, S.: Elliptical head tracking using intensity gradients and color histograms. *IEEE Conference on Computer Vision and Pattern Recognition.* 232-237 (1998)
- [16] Ojala, T.; Pietikainen, M.; Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Transactions on Pattern Analysis and Machine Intelligence.* 24,971-987 (2002)
- [17] Swain, M. J.; Ballard, D. H.: Color indexing. *International Journal of Computer Vision.* 7, 11-32 (2002)
- [18] Stricker, M.; Orengo, M.: Similarity of color images. *SPIE Storage and Retrieval for Image and Video Databases.* 3, 381-392 (1995)
- [19] Kailath, T.: The divergence and Bhattacharyya distance measures in signal selection. *IEEE Transactions on Communications Technology.* 15,52-60 (1967)
- [20] Fukunaga, K.: Introduction to Statistical Pattern Recognition. Second ed., Academic Press (1990)
- [21] Teemu, K.; Joni-Kristian, K.; Lasse, L.; Heikki, K. : Unsupervised object discovery via self-organisation. *Pattern Recognition Letters* **2012**,33, 2102-2112
- [22] Tso, B.; Mather, P.M. : Classification Methods for Remotely Sensed Data, 2nd ed. ; *CRC Press: Boca Raton, FL, USA* **2009**
- [23] Salman, A.; , Valadan, M. J. Z.; Hamid, E.; Hamid, A. M.; Ali, M. : Automatic urban building boundary extraction from high resolution aerial images using an innovative model of active contours. *International Journal of Applied Earth Observation and Geoinformation* **2010**, 12, 150-157
- [24] Ying-Ho, L.; Anthony, J.T.L.; Fu, C. : Object recognition using discriminative parts.*Computer Vision and Image Understanding* **2012**, 116, 854-867
- [25] Giorgos, M.; Jungho, I.; Caesar, O. : Support vector machines in remote sensing: A review.*ISPRS Journal of Photogrammetry and Remote Sensing* **2011**, 66, 247-259
- [26] Knerr, S.; Personnaz, L.; Dreyfus, G. : Single-layer learning revisited: a stepwise procedure for building and training a neural network. *In: Neuro-computing: Algorithms, Architectures and Applications. In: NATO ASI Series , Springer, 1990*
- [27] Bischof, H.; Schneider, W.; Pinz, A.J. : Multispectral classification of Landsat-images using neural networks. *IEEE Trans. Geosci. Remote Sens.* **1992**, 30, 482-490
- [28] Soille, P.; Pesaresi, M. : Advances in mathematical morphology applied to geoscience and remote sensing. *IEEE Transactions on Geosciences and Remote Sensing* **2006**, 40, 2042-2055
- [29] Halla, Brenner, N.C. : Extraction of building and trees in urban environment. *ISPRS Journal of Photogrammetry Remote Sensing* **1999**, 54, 130-137
- [30] Zhao, B.; Trinder, J.C. : Integrated-approach-based automatic building extraction. *International Archives of Photogrammetry and Remote Sensing XXXIII* **2000**, Part B3

- [31] Lafarge, F.; Descombes, X.; Zerubia, J.; Pierrot-Deseilligny, M. : Automatic building extraction from DEMs using an object approach and application to the 3D-city modeling. *ISPRS Journal of Photogrammetry Remote Sensing* **2008**,63, 365-381
- [32] Miliaresis, G.; Kokkas, N. : Segmentation and object-based classification for the extraction of the building class from LIDARDEMs. *Computers Geosciences* **2007**, 33, 1076-1087
- [33] Lowe, D.G. : Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* **2004**, 60(2), 91-110
- [34] Berg,A.;Berg, T.; Malik, J. : Shape matching and object recognition using low distortion correspondence. *in: Proc. International Conference on Computer Vision and, Pattern Recognition* **2005**,26-33
- [35] Shufelt, J. A.: Performance evaluation and analysis of monocular building extraction from aerial imagery. *IEEE Trans. Pattern Anal. Machine Intell.*, 21, 311-326 (1999)
- [36] El-Merabet, Y.; Meurie, C.; Ruichek, Y.; Sbihi, A.; Touahni, R.: Orthophotoplan segmentation based on regions merging for roof detection. *S&T/SPIE Electronic Imaging 2012 - Image Processing: Machine Vision Applications V*, SPIE 8300, Burlingame USA (2012)

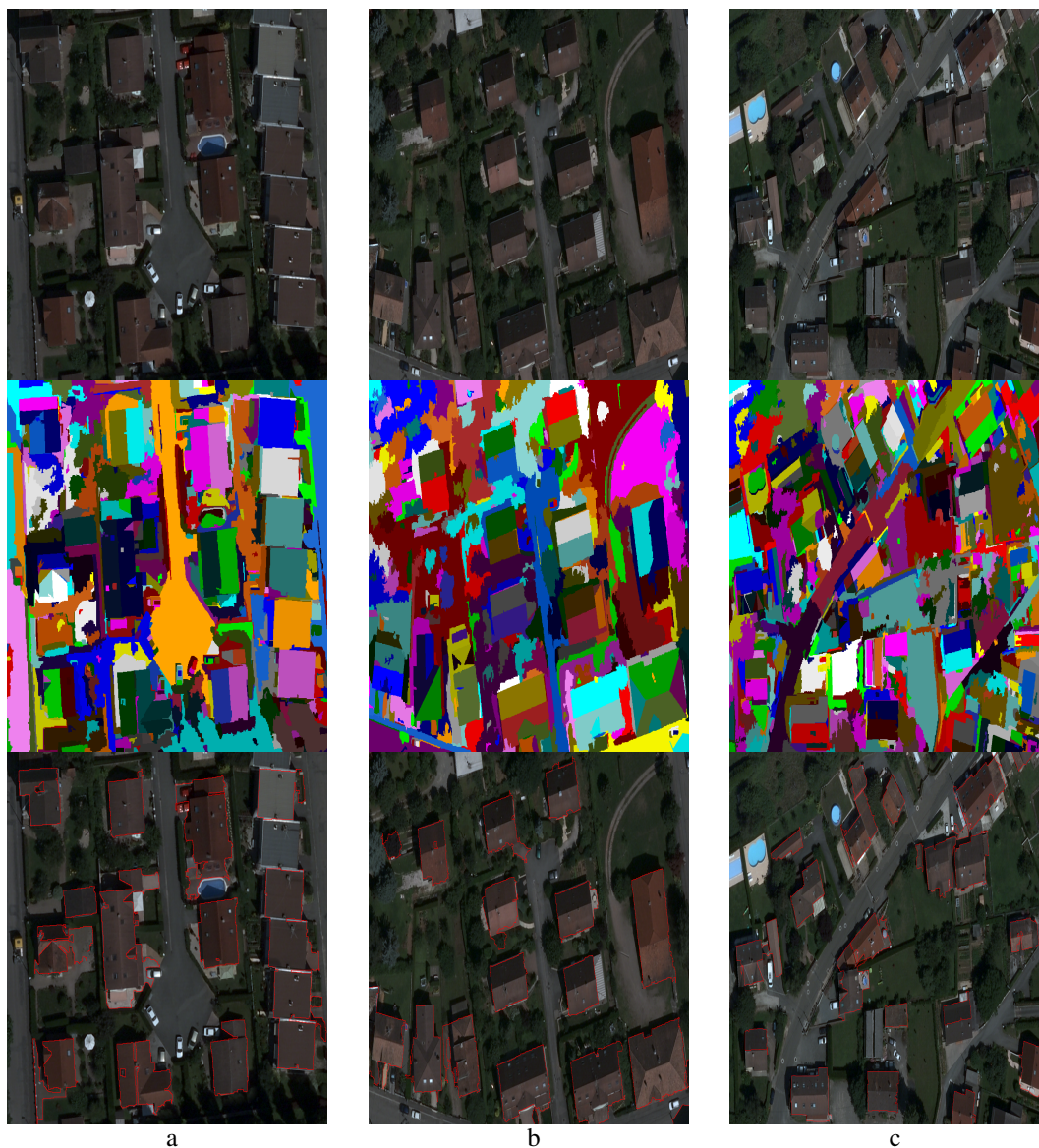


Figure 4. Automatic extraction of building roofs from the set of processed images. From top to the bottom: the initial image, the SRM segmentation result and the corresponding building roofs extraction



Figure 5. Examples of some building parts that have not been extracted (yellow ellipses) and some false detections (green ellipses)