

A Region-based Level Set Segmentation for Automatic Detection of Man-made Objects from Aerial and Satellite Images

Konstantinos Karantzas and Demetre Argialas

Abstract

A region-based level set segmentation was developed for the automatic detection of man-made objects from aerial and satellite images. The essence of the approach is to optimize the position and the geometric form of an evolving curve, by measuring information within the regions that compose a particular image partition based on their statistical description. The present region-based variational model is fully automated without the need to manually specify the position of the initial contour. Furthermore, it converges after a small number of iterations, allowing real-time applications. The developed algorithm was tested for the detection of roads, buildings and other man-made objects in a number of aerial and satellite images. The effectiveness of the algorithm is demonstrated by the experimental results and the performed qualitative and quantitative evaluation.

Introduction

Automatic extraction of man-made objects, such as buildings, building blocks, or roads, is of major importance for supporting several government activities and various GIS applications like map generation and update (Gruen *et al.*, 1995; Gruen *et al.*, 1997; Baltsavias *et al.*, 2001; Mueller *et al.*, 2004; Baltsavias, 2004). Information extraction from images is not, however, a trivial task, due to the complexity of the information stored in images (Sowmya and Trinder, 2000) and requires the formulation of procedures and knowledge able to encapsulate their content. The structure of the terrain surface appears complex, being a combination of many different intensities representing natural features such as vegetation, geomorphological and hydrological features, man-made objects (buildings, roads, etc.), and artifacts caused by variations in terrain illumination, like shadows or occlusions. Curve propagation techniques (snakes, active contours, deformable models) have given promising results both for buildings and roads (Mayer, 1999; Mena, 2003).

Snakes (Kass *et al.*, 1987) were first employed by Cohen (1991) and by Gruen and Li (1997) for the semi-automatic extraction of linear objects and roads from remote sensing images. Later on, Zafiroopoulos and Schenk (1998) tackled the problem by embedding color information towards the

extraction and localization of road structures of small width. Jeon *et al.* (2002) proposed a snake model for roads which gave more accurate results, while Mayer *et al.* (1997 and 1998), Ferraro *et al.* (1999), Laptev *et al.* (2000) and Chiang *et al.*, (2001) used snakes for the detection of certain cartographic features. A modification of the classic snakes formulation was presented by Agouris *et al.*, (2001a and 2001b), where change detection and versioning were embedded in a single framework. Furthermore, the use of dynamic programming and snakes was also proposed (Gruen and Li, 1997; Agouris *et al.*, 2001a; Ruther *et al.*, 2002). Rochery *et al.* (2003) employed higher-order functionals with quadratic, instead of linear energies and applied such a formulation for line network extraction from satellite images. Recently, snakes were combined with a Hough transform voting and were tested for building detection (Bailloeuil *et al.*, 2005). Peng *et al.* (2005) proposed a postprocessing scheme for a more efficient detection of buildings in urban regions.

In the above efforts, the main limitation was snakes disability to automatically change their topology during propagation (Osher and Paragios, 2003). The initial contour(s) can not split or merge naturally; thus, hardly can an operational system be based on such formulations, even in cases that the number of desired (for extraction) objects is *a priori* known. The problem of snakes' restricted single topology was solved by Caselles *et al.* (1997), who introduced the geodesic active contours (GAC). GAC were further expanded and were formulated under the pioneer level set framework of Osher and Sethian (1998), and nowadays, variational geometric level set methods are an established technique for various computer vision applications (Osher and Paragios, 2003; Paragios *et al.*, 2005; Lee, 2005; Ayed *et al.*, 2006; Brox and Weickert, 2006; Lee and Seo, 2006; Martin *et al.*, 2006).

For photogrammetric and remote sensing applications, level sets have been used primarily for the task of image segmentation (Samson *et al.*, 2001; Ball and Bruce, 2005; Besbes *et al.*, 2006). Focusing on the detection of certain objects, Keaton and Brokish (2002), and more recently Niu (2006), used the GAC formulation in a semi-automatic framework. The position of the initial curve(s) was given manually before the curve evolution procedure. In addition, their energy functionals were based on image edges (gradients) inheriting, thus, the limitations of edge detection techniques, especially in

Konstantinos Karantzas is with the M.A.S, Ecole Centrale de Paris, Grande Voie des Vignes, 92 295, Chatenay-Malabry, France, (konstantinos.karantzas@ecp.fr).

Demetre Argialas is with the Remote Sensing Laboratory, National Technical University of Athens, 9 Heroon Polytechniou Str., 15780, Zografou Campus, Athens, Greece (argialas@central.ntua.gr).

Photogrammetric Engineering & Remote Sensing
Vol. 75, No. 6, June 2009, pp. 667–677.

0099-1112/09/7506-0667/\$3.00/0
© 2009 American Society for Photogrammetry
and Remote Sensing

complex scenes (Chan and Vese, 2001). Recently, Cao *et al.* (2005a and 2005b) proposed an energy functional which was based on a modified Mumford-Shah segmentation model for automatic man-made object detection in aerial images. Their model uses a coarse-to-fine strategy and a fractal error metric at the fast coarse evolution stage. Results, which were evaluated visually, show clearly that their algorithm can not extract the boundaries of specific objects, like buildings or roads, but instead extracts urban and semi-urban regions with insufficient spatial detection accuracy.

In this paper, a variational geometric level set functional is proposed for man-made object detection from aerial and satellite images. The novelty of the present approach lies in the development of a model-free, region-based, energy functional that is fully automated without the need for manually indicating the position of the initial contour. Furthermore, it converges after a small number of iterations, allowing for real-time applications. It was tested for the detection of different man-made objects, without any prior knowledge of their shape or position. The goal was to evaluate the algorithm for such a demanding task (Baltasvias, 2004), since low and medium level computer vision techniques (like edge detection and segmentation) do not incorporate high level procedures during their computations (in contrast to knowledge-based approaches) possessing, therefore, native limitations in order to overcome the complexity captured in images.

The paper is organized as follows. The next section is devoted to a short literature survey of variational level set segmentation techniques, followed by a detailed description of the algorithm developed in this paper. The next section discusses the efficiency of such a formulation which is demonstrated by the experimental results and their quantitative evaluation. Finally, concluding remarks are given, along with the perspectives for further research.

Region-based Variational Level Set Segmentation

Since the original work of Kass *et al.* (1987), extensive research has been performed, in the computer vision community on active contour models. The classical approach is based on deforming an initial contour towards the desired object boundaries. The deformation is obtained by trying to minimize a functional whose minimum is obtained at the boundary of the desired object. GAC (Caselles *et al.*, 1997) solved the problem of active contour's restricted topology, and with the use of the level set method (Osher and Sethian, 1988), such variational formulations became standard in computer vision (Osher and Paragios, 2003). With the advantage of being implicit, intrinsic, and parameter-free, level sets can detect object boundaries and track moving interfaces through either model-free (Paragios and Deriche, 2002) or model-based (Cremers, 2003) functionals.

Classical active contour models, GAC, and several similar functionals that have been proposed for computer vision (Paragios *et al.*, 2005) and geoscience applications (Keaton and Brokish, 2002; Ball and Bruce, 2005; Niu, 2006) rely on a specific edge-function, which depends on image gradients and stops the curve evolution procedure. These models can only detect objects whose edges can be well described by the gradient. In practice, though, since the discrete gradients are bounded, the stopping function is never zero at edges, and the curve may pass through the boundary (Chan and Vese, 2001). In addition, in complex scenes (which contain noise, shadows, occlusions, and numerous objects at various scales) the gradient has to be strongly smoothed and such an operation in most cases, blurs edges.

An active contour model, which is not based on image gradients but on the Mumford and Shah (1989) segmentation

model, has been proposed (Chan and Vese, 2001; Tsai *et al.* 2001). Let $I(x,y)$ be an image defined on a domain W without any particular geometrical structure. During segmentation, one aims at partitioning W into domains W_i , which are delimited by a system of crisp and regular boundaries K and within which the image I is homogeneous. In Bayesian models, two parts exist: the prior model and the data model. The prior model addresses the phenomenological evidence of what is qualitatively a segmentation, i.e., namely an approximation of image I by piecewise smooth functions u on $W-K$, which are discontinuous along a set of edges K (u is an optimal approximation of the observations, i.e, the data I). The aim is to introduce a way of selecting, from among all the allowed approximations (u, K) of I , the best possible one. To this end, Mumford and Shah (MS) proposed the following energy functional:

$$E(u,K) = \iint_{W-K} |\nabla u|^2 dx dy + \lambda \iint_W (u - I)^2 dx dy + \mu \int_K d\sigma. \quad (1)$$

The functional contains three terms: (a) the first term which measures the variation and controls the smoothness of u on the open connected components W_i of $W-K$, (b) the second term which controls the quality of the approximation of I by u , and (c) the third term which controls the length, the smoothness, the parsimony and the location of the boundaries K , and inhibits the spurious phenomenon of over-segmentation.

Due to the coefficients λ and μ , the MS-model (Equation 1) is a multi-scale one: if μ is small, the output is a "fine grained" segmentation, and if μ is large, the output is a "coarse grained" segmentation. The essence of such a region-driven functional is to use the evolving interface to define an image partition that is optimal in respect to the grouping criterion. Within the level set representation, such a partition is natural according to the sign of the embedding function. In addition, with such a curve evolution model, interior contours can be automatically detected, and the initial curve can be anywhere in the image (Chan and Vese, 2001). Such regional/global driven information can improve the performance of edge-based flows that suffer from being sensitive to the initial conditions (Paragios and Deriche, 2002).

The Man-Made Object Detection Algorithm

A region-based, level set model was developed and the relevant algorithm was implemented for the detection of man-made object boundaries. The developed algorithm consists of three steps. In the first step, an image simplification is implemented based on Anisotropic, Morphological Levelings (Karantzas *et al.*, 2007). The goal, of such a filtering was to enlarge the initial image flat zones by a small scale image simplification. A leveling of scale 3 was used, which could be compared, in terms of scale, with an isotropic Gaussian filtering with a standard deviation of $\sigma = 1.5$. A model-free, region-based, level set segmentation is applied, in the second step, resulting to the detection of man-made object boundaries. Towards this end, the approximations of Dirac (δ_a) and Heaviside (H_a) distributions were employed (Zhao *et al.*, 1996):

$$\delta_a(\phi) = \begin{cases} 0, & |\phi| < a \\ \frac{1}{2a} \left(1 + \cos\left(\frac{\pi\phi}{a}\right) \right), & |\phi| > a \end{cases},$$

$$H_a(\phi) = \begin{cases} 1, & \phi > a \\ 0, & \phi < -a, \\ \frac{1}{2a} \left(1 + \frac{\phi}{a} + \frac{1}{\pi} \sin\left(\frac{\pi\phi}{a}\right) \right), & |\phi| < a. \end{cases} \quad (2)$$

Based on the above distributions and letting w be the evolving interface, an image-partitioning objective function was introduced. The energy functional, that was implemented, is described by the following equation, defined in the image domain Ω :

$$E_{region}(\phi) = \iint_{\Omega} [H_a(\phi) \cdot r_{in}(I)] + [(1 - H_a(\phi)) \cdot r_{out}(I)] d\Omega + \beta \iint_{\Omega} |H_a(\phi)| d\Omega \quad (3)$$

where r_{in} and r_{out} are global region-driven functions which provide a statistical description of the inside and outside of the object area, and β a positive scalar. To effectively deal with the stated complexity captured on aerial and satellite imagery, the Chan and Vese's (2001) formulations were modified. The employed, here, normalized by the covariance matrix region descriptors aim to properly describe the intensity variation inside and outside object regions:

$$r_{in}(I) = \frac{(\mu_{in} - I)^2}{\sigma_{in}^2}, r_{in} : \mathbb{R}^+ \rightarrow [1, 0] \quad (4)$$

where μ_{in} is the mean and σ_{in} the covariance matrix of object appearance (similar formulations are defined for the background r_{out}). Such descriptors (r_{in} and r_{out}), being monotonically decreasing functions, measure the quality of matching between the observed image and the expected regional properties of the desired for extraction feature. Relying on a properly simplified image, such descriptors were expected to cope more efficiently with image complexity, than edge-based functionals. In Figure 1, the robustness of such a region descriptors formulation is demonstrated. The Chan-Vese formulation converged to a solution where both roads and other irrelevant objects were detected. The present formulation surpassed other features and resulted to the detection of the desired objects. Last in the processing scheme (third step) and following the level set propagation, segments with a small area (depending on the resolution of the data) were eliminated and the final detected boundaries were detected.

The calculus of variations and a gradient descent method were used to optimize the objective cost function. In order to optimize the objective functional, Euler-Lagrange

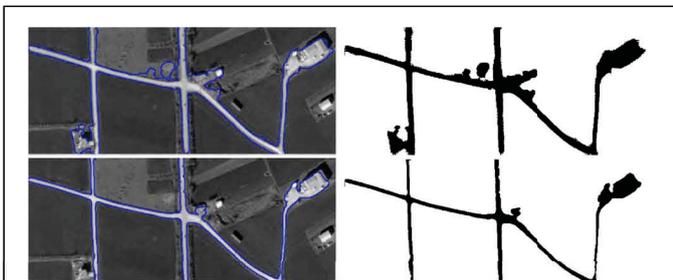


Figure 1. The Chan-Vese algorithm (top row) converged to a solution where both roads and other irrelevant objects were detected. The present formulation (bottom row) was more robust, surpassed, in most cases, other objects and resulted in the desired road boundaries. A color version of this figure is available at the ASPRS website: www.asprs.org.

equations in the implicit space ϕ were fulfilled through an iterative steady state process. The Euler-Lagrange equation for the presented functional (Equation 3) was implemented by the following gradient descent:

$$\frac{\partial \phi}{\partial t} = \delta_a(\phi) \left[-\frac{\partial r_{in}(I)}{\partial t} + \frac{\partial r_{out}(I)}{\partial t} + \beta \left(\text{div} \frac{\nabla \phi}{|\nabla \phi|} \right) \right] \quad (5)$$

where div is the two dimensional divergence, and β a constant.

A flowchart describing the implementation of the developed algorithm is presented in Figure 2. For all color images, the algorithm was applied in the luminance 3D polar coordinate of the color map, as proposed by Hanbury (2003). An ellipse, inside the image limits (covering approximately 80 percent in both its axis), was chosen as the initial contour. However, it can be any arbitrary one allowing automation, in terms of avoiding manually indicating the shape and the position of the initial condition. Even without code optimization and in the Mathworks Matlab® environment, the algorithm is relatively fast and converges after approximately three minutes in an ordinary Pentium-V computer and for an image of approximately two million pixels. Absolute real-time results can be achieved by considering multigrid implementations and other relevant approaches (Shi and Karl, 2005; Paragios *et al.*, 2005).

Results and Discussion

The developed algorithm was tested for the detection of man-made objects like roads, buildings, and airports. A quantitative evaluation of the results was performed based on ground truth data which were derived from a manual digitization procedure. In the case of roads the reference data comprised major and secondary roads while all short driveways or pathways were excluded. The standard measures of Completeness (detection rate), Correctness (under-detection rate), and Quality (a normalization between the previous two) were employed in the same way as in Wiedemann *et al.* (1998) and Doucette *et al.* (2004). To this end, the measures of True Positives (total length/area of correctly detected segments), False Positives (total length/area of incorrectly detected segments), and False Negatives (total length/area of missing segments), were computed.

Last, but not least, and unlike Mayer *et al.* (2006) where algorithm parameters had been appropriately tuned for each image, the goal here was to use a constant setting for every feature. Towards this end, the parameter α (Equation 2), which controls the width of distributions, was tuned by a trial and error investigation and left stable for each feature. Thus, for

```

Given an initial image  $I_1$ ,
if  $I_1$  is colored then compute its luminance 3D polar coordinate (Hanbury, 2003)
Compute its simplified version  $I_2$  (Karantzalos et al., 2007)
Compute an elliptical initial curve and calculate initial values for  $\phi_0, r_{in}, r_{out}$ 
for every iteration  $t$ 
    calculate a new value for  $\phi_t$  (Equation 5)
    update  $r_{in}, r_{out}$ 
end
Erase detected segments which cover a small area
Get resulted binary image  $I_3$ 

```

Figure 2. A flowchart describing the implementation of the developed algorithm. A color version of this figure is available at the ASPRS website: www.asprs.org.

the road network detection parameter α was set to 1.5, allowing a broader variation in region descriptors and was set to 1.0 for the building and airport detection tasks. In all cases, β was set to 0.1 sufficiently smoothing object boundaries.

Road Network Detection

The algorithm was applied to an aerial image (Aerial_1, RGB, approximately 1 m ground resolution) of a rural scene with low complexity. Roads were mostly homogeneous and not disturbed by shadows or occlusions. The sequence of active contour propagation resulting to the final solution for road boundaries is shown in Figure 3. The final result overlaid on the initial image and the ground truth data are shown in Figure 4. The algorithm did detect the whole road network and failed in only, two cases where an over-detection took place. In Table 1, the quantitative evaluation supports visual observations, since the completeness measure was over 99 percent and the correctness and quality measures at 96 percent.

In Figure 5, results from the application of the algorithm to another aerial image (Aerial_2, RGB, approximately 1.5 m ground resolution) of a more complex scene are presented. The algorithm detected the road network and only failed to extract the smaller part of the secondary road at the top right of the image. Just a couple of over-detection cases have, also, taken place, where image descriptors failed to follow strictly the variation of the road network's intensity values and were influenced by the intensity of adjacent objects. These visual observations are confirmed by the quantitative evaluation (Table 1), where the completeness measure is 96 percent, the correctness at 90 percent and the overall quality at 87 percent.

The quantitative results can be compared with the lower rates reported by other automatic algorithms (Doucette *et al.*, 2004; Mayer *et al.*, 2006) but not directly since different data were used. However, the efficiency of the developed algorithm should be emphasized. The quantitative rates, reported here, are expected to decrease in

a more extensive evaluation that will be concentrated strictly to roads and will be performed on several types of complexities and over larger areas (i.e., images of at least $4,000 \times 4,000$ pixels as was, similarly, the initial intention in Mayer *et al.* (2006)).

Finally, in all the previous cases a morphological skeletonization was performed in order to extract road network center lines from the resulted road boundaries. Optimizing the skeletonization procedure was beyond the goals of this work.

Building Boundaries Detection

The algorithm was, moreover, tested for the detection of buildings. Results from its application to an aerial image (Aerial_3, panchromatic, approximately 0.4 m ground resolution) of an urban scene are presented in Figures 6 and 7. The different steps of the level set propagation, until algorithm convergence, are shown in Figure 6. The final result overlaid on the initial image, and the ground truth data are shown in Figure 7. The algorithm did detect all building boundaries and successfully managed to surpass the intensity variation of the rooftops, caused by different objects. A failure (under-detection) that can be observed in the bottom left building is due to a significant and wide shadow. In Table 1, the quantitative evaluation supports qualitative observations, since the completeness measure was reported at 91 percent and the correctness at 98 percent. The overall quality was at 90 percent.

Furthermore, the developed algorithm was applied for the detection of buildings to an aerial (RGB, approximately 0.7 m ground resolution) and a satellite (QuickBird© fused RGB, 0.6 m ground resolution) image. Each of these two images covers a wide area, of a complex scene where multiple objects of different object classes, shadows, occlusions, different texture patterns, and some terrain height variability exist. In Figures 8 and 10 (respectively, for the aerial and satellite test image), the different steps of the level set propagation, until algorithm's convergence,

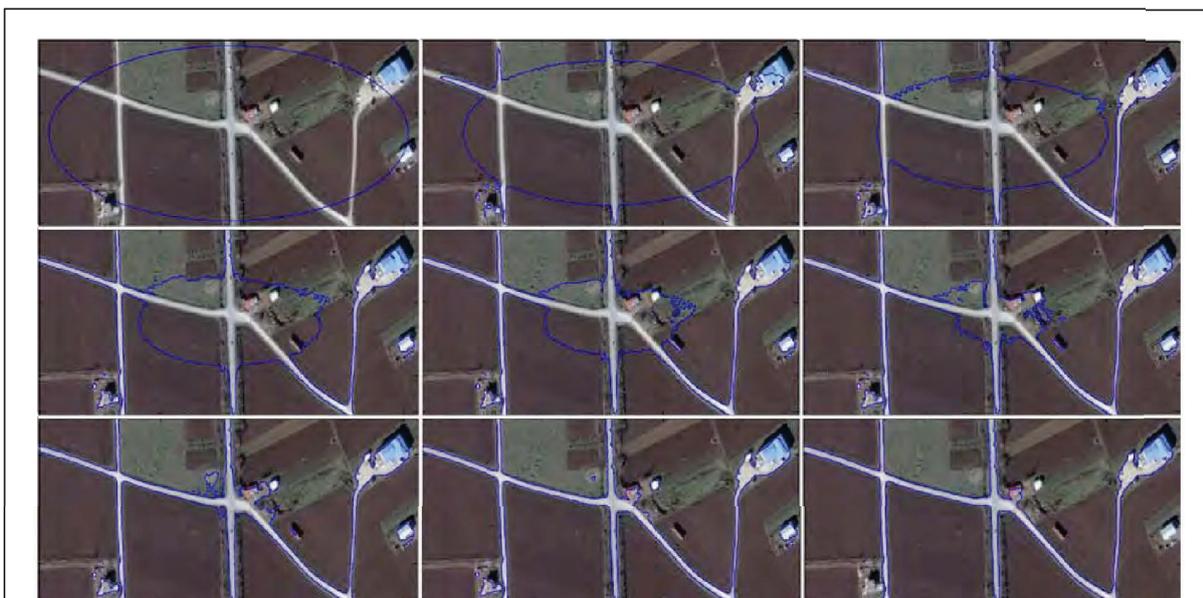


Figure 3. Road network detection from the Aerial_1 test image using the present region-based level set functional. The different steps of contour propagation are shown, starting from an initial, arbitrary, elliptical curve leading to the final detection. View this figure starting with each row from the left. A color version of this figure is available at the ASPRS website: www.asprs.org.

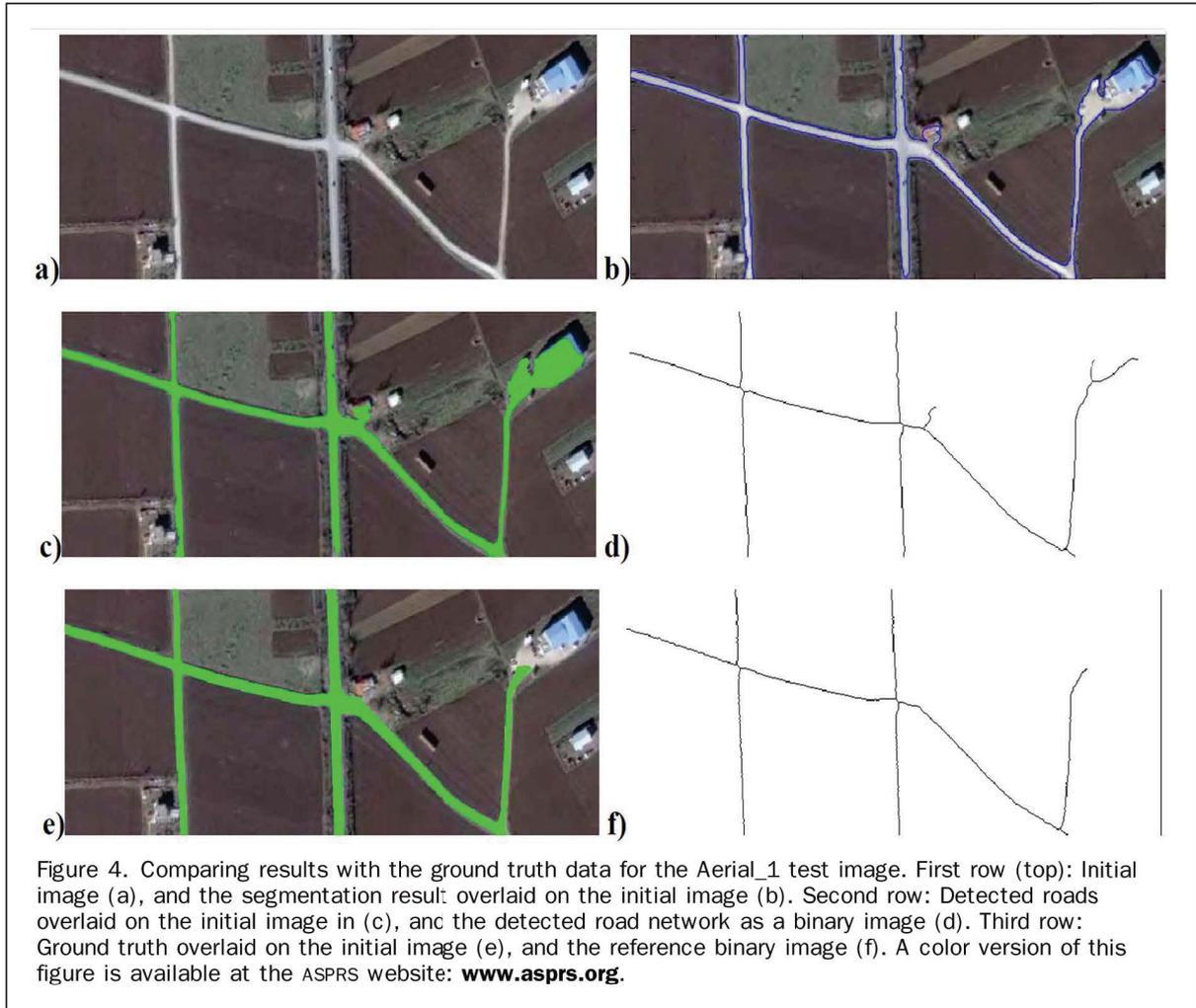


TABLE 1. QUANTITATIVE EVALUATION AFTER THE APPLICATION OF THE DEVELOPED ALGORITHM TO DIFFERENT TEST IMAGES. FOR EACH CASE, THE STANDARD MEASURES OF COMPLETENESS, CORRECTNESS, AND OVERALL QUALITY, ARE SHOWN. THE ALGORITHM SCORES HIGH IN ALL CASES

Test data	Quantitative Measures		
	Completeness	Correctness	Quality
Aerial_1 (roads)	0.9966	0.9764	0.9732
Aerial_2 (roads)	0.9626	0.8978	0.8675
Aerial_3 (buildings)	0.9114	0.9846	0.8987
Aerial_4 (buildings)	0.8113	0.9188	0.7588
Satellite (buildings)	0.8266	0.8471	0.7193

are shown. The final detected building segments superimposed on the initial aerial image are shown in Figure 9. All buildings, except one, were fully or partly detected. Most of them have been recognized as different identities (are labeled and numbered uniquely) apart from the three-building segment in the top right of the image. The later was poorly detected but appears as one segment in the ground truth data, as well. The correctness of the detection was high at approximately 92 percent with a lower completeness at 81 percent (Table 1) due to above described errors. For the satellite test case, the final

detected building segments superimposed on the initial image are shown in Figure 11. All buildings, curved or rectangular shaped ones, were fully or partially detected. Due to the misleading low-level information, caused by shadows, occlusions, and the intensity variation on and around building rooftops, some over and under detection cases took place. The quantitative evaluation, comparing results with the reference image (Figure 11d), indicated that the algorithm scored approximately 83 percent regarding its completeness and approximately 85 percent its correctness (Table 1).

The above three experimental results can highlight algorithm's efficiency for building detection. A more extensive evaluation can set the basis for the construction of an operational system that will be based on the present research.

Airport Detection

Finally, the algorithm was applied to a SPOT (HRV, panchromatic, 10 m ground resolution) satellite image for the detection of an airport. Figure 12 presents the different steps of the level sets propagation up to the point where the algorithm converged and the final result overlaid on the initial image. The algorithm resulted in a single segment, which describes efficiently the boundaries of airports pavement. All other irrelevant objects with approximately similar intensities were robustly surpassed.



Figure 5. Comparing results with the ground truth data for the Aerial_2 test image. First row: Initial image (a), and the segmentation result overlaid on the initial image (b). Second row: Detected roads overlaid on the initial image (c), and the detected road network as a binary image (d). Third row: Ground truth overlaid on the initial image in (e), and the reference binary image (f). A color version of this figure is available at the ASPRS website: www.asprs.org.

Conclusions and Future Perspectives

A variational framework for the automatic detection of man-made objects in aerial and satellite images has been presented. Contrary to efforts reported in Baltasvias (2004) and Meyer (2006), here, the problem of detecting man-made objects has been addressed under a unique framework, covering all object classes (buildings, roads, etc). The novelty of the present approach lies, also, in the development of a model-free region-based energy functional that is fully automated, without the need for manually indicating the initial contour, as in Keaton and Brokish (2002) and Niu (2006). The algorithm converges after a small number of iterations (approximately six), allowing real-time applications. In addition, the present approach detects objects whose boundaries are not necessarily defined by the gradient or an edge-map like most previous active contour, GAC, or similar formulations. The energy functional for minimization can be seen as a particular case of a minimal partition task, i.e, a particular segmentation of the image following the Mumford and Shah (1989) ideas.

The algorithm was evaluated qualitatively and quantitatively and it was found efficient. By extending region descriptors to utilize color and hyperspectral information, its effectiveness is expected to increase. By further extensive evaluation through its application on several datasets, a promising operational system could be introduced. Moreover, by using available information from cadastral maps or any other vector database and forcing the initial contour to take the form of the existing in the map geometric shapes (like building or road boundaries), the algorithm can be used for map revision and updating.

References

- Agouris, P., S. Gyftakis, and A. Stefanidis, 2001a. Dynamic node distribution in adaptive snakes for road extraction, *Vision Interface 2001*, Ottawa, Canada, pp. 134–140.
- Agouris, P., A. Stefanidis, and S. Gyftakis, 2001b. Differential snakes for change detection in road segments, *Photogrammetric Engineering & Remote Sensing*, 67(12):1391–1399.

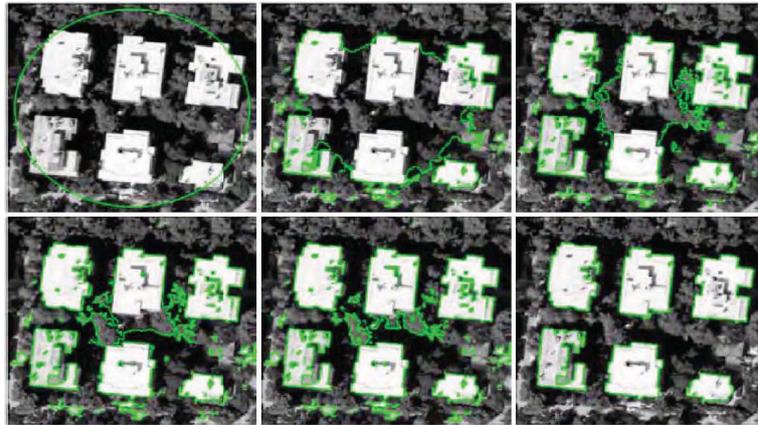


Figure 6. Building detection from the Aerial_3 test image, using the present region-based level set method. The different steps of contour propagation are shown, starting from an initial arbitrary elliptical curve leading to the final detection. View this figure starting each row from the left. A color version of this figure is available at the ASPRS website: www.asprs.org.

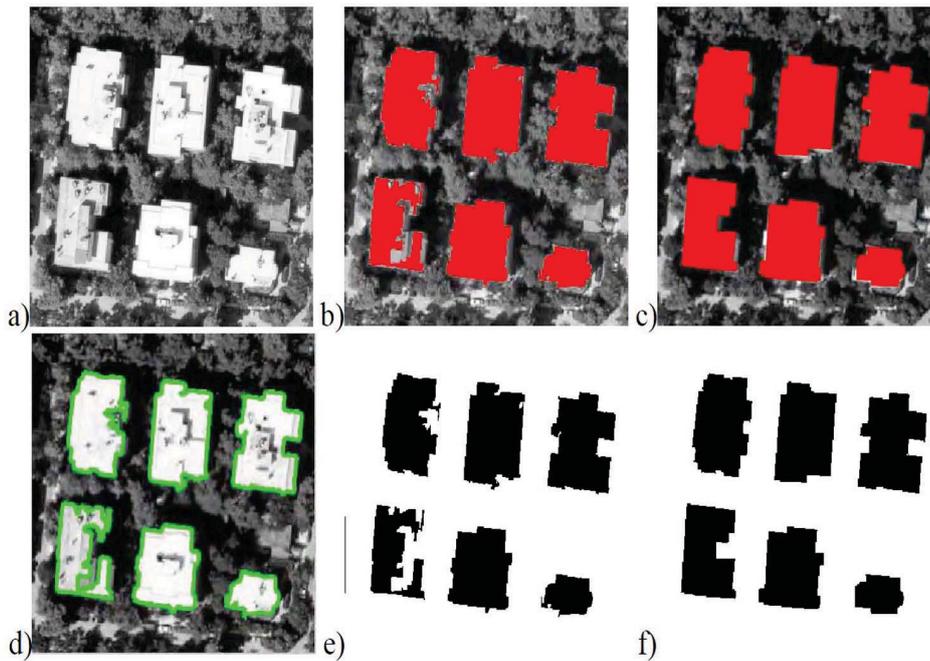


Figure 7. Comparing result with the ground truth data for the Aerial_3 test image. Top row: Initial image (a), the detected output segments (b), and the ground truth (c) overlaid on the initial image in red. Bottom row: Detected building boundaries overlaid on the initial image (d), the detected buildings as a binary image (e), and the reference binary image (f). A color version of this figure is available at the ASPRS website: www.asprs.org.

Ayed, B.I., A. Mitiche, and Z. Belhadj, 2006. Polarimetric image segmentation via maximum likelihood approximation and efficient multiphase level sets, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(9):1493–1500.

Bailloleul, T., V. Prinet, B. Serra, P. Marthon, P. Chen, and H. Zhang, 2005. Urban building land use change mapping from high resolution satellite imagery, active contours and Hough voting, *Proceedings of the International Symposium on Physical*

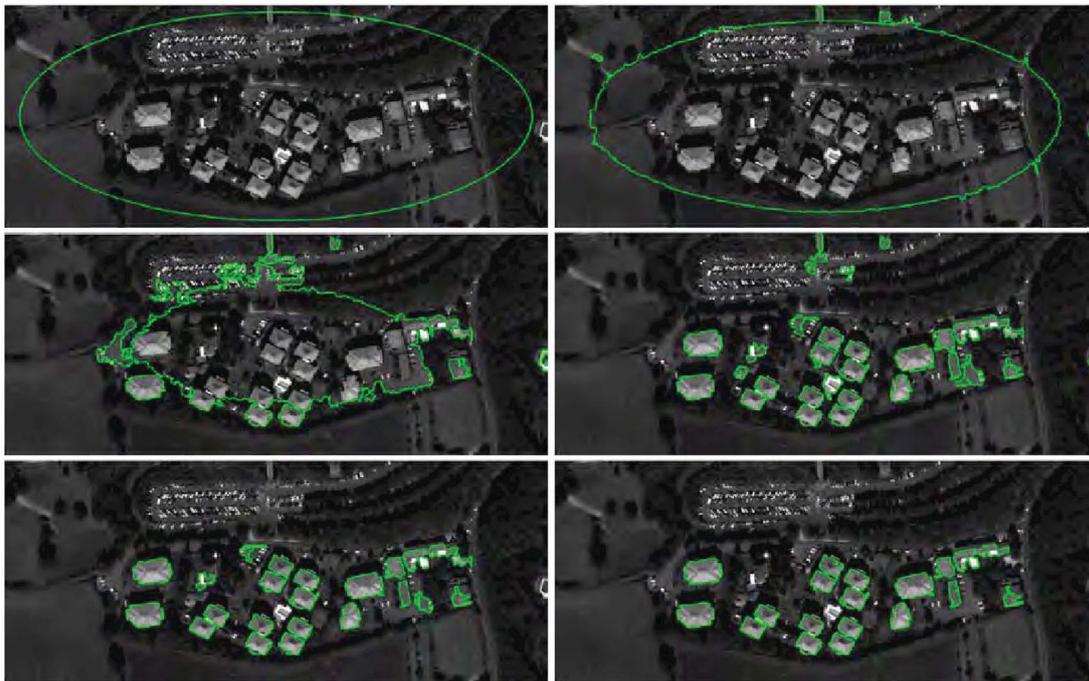


Figure 8. Building detection to the Aerial_4 test image, using the developed region-based level set method. The different steps of contour propagation are shown, starting from an initial arbitrary elliptical curve leading to the final detection. View this figure starting each row from the left. A color version of this figure is available at the ASPRS website: www.asprs.org.

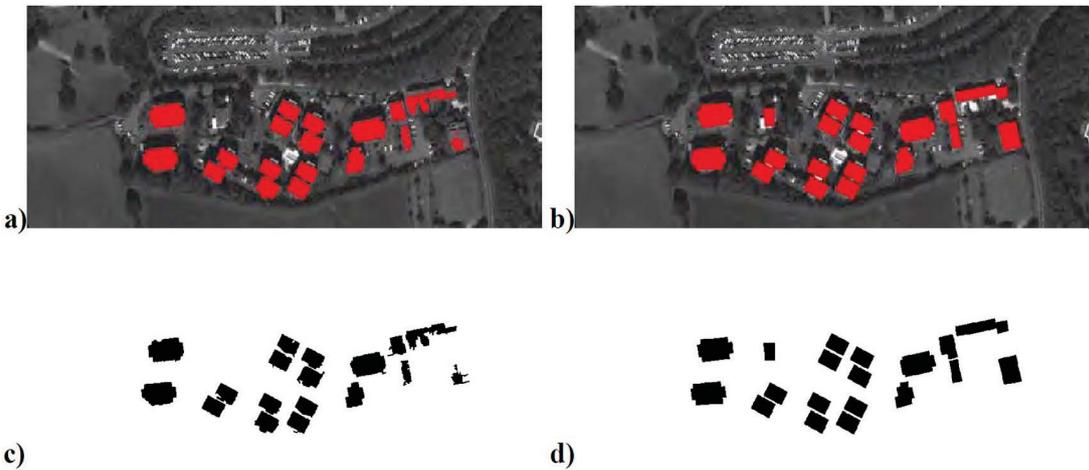


Figure 9. Comparing results with the ground truth data for the Aerial_4 test image. Top row: Detected buildings (a), and the ground truth overlaid on the initial image (b). Bottom row: The detected buildings as a binary image (c), and the reference binary image (d). A color version of this figure is available at the ASPRS website: www.asprs.org.

Measurements and Signatures in Remote Sensing, Beijing, China.

Ball, J.E., and L.M. Bruce, 2005. Level set segmentation of remotely sensed hyperspectral images, *Proceedings of the IEEE Geoscience and Remote Sensing Symposium*, (8):5638–5642.

Baltsavias, E.P., 2004. Object extraction and revision by image analysis using existing geodata and knowledge: Current status and steps

towards operational systems, *Journal of Photogrammetry and Remote Sensing*, 58(3–4):129–151.

Baltsavias, E.P., A. Gruen, and L. VanGool, 2001. *Automatic Extraction of Man-made Objects from Aerial and Satellite Images III*, Taylor & Francis, 425 p.

Besbes, O., Z. Belhadj, and N. Boujemaa, 2006. *Adaptive Satellite Images Segmentation by Level Set Multiregion Competition*, INRIA, Research Report 5855.

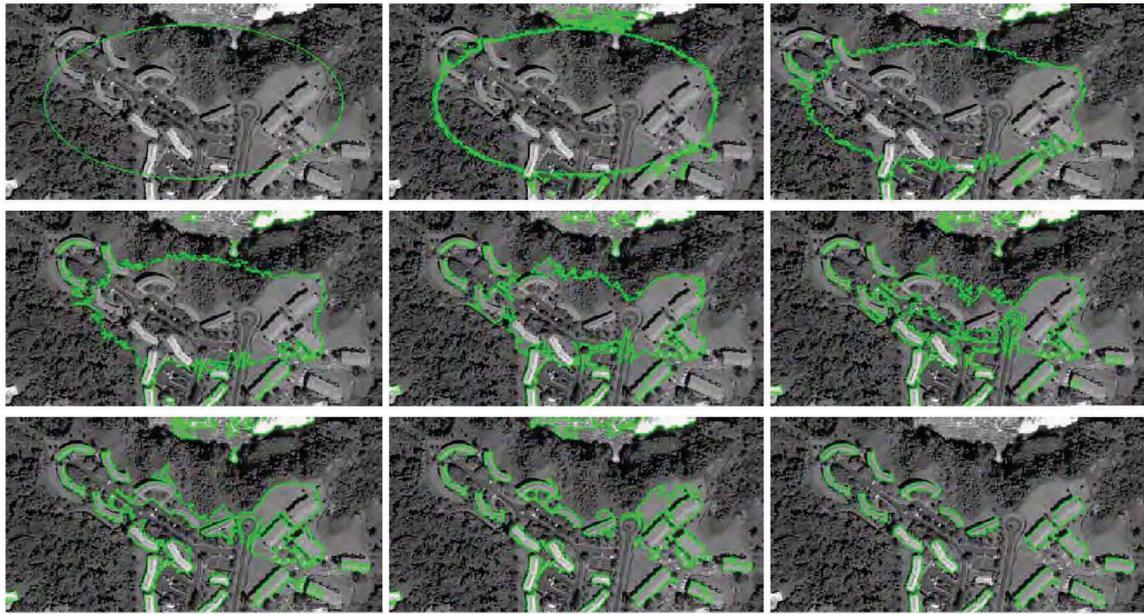


Figure 10. Building detection to the satellite (QuickBird) test image, using the present region-based level set method. The different steps of contour propagation are shown, starting from an initial arbitrary elliptical curve leading to the final detection. View this figure starting each row from the left. A color version of this figure is available at the ASPRS website: www.asprs.org.

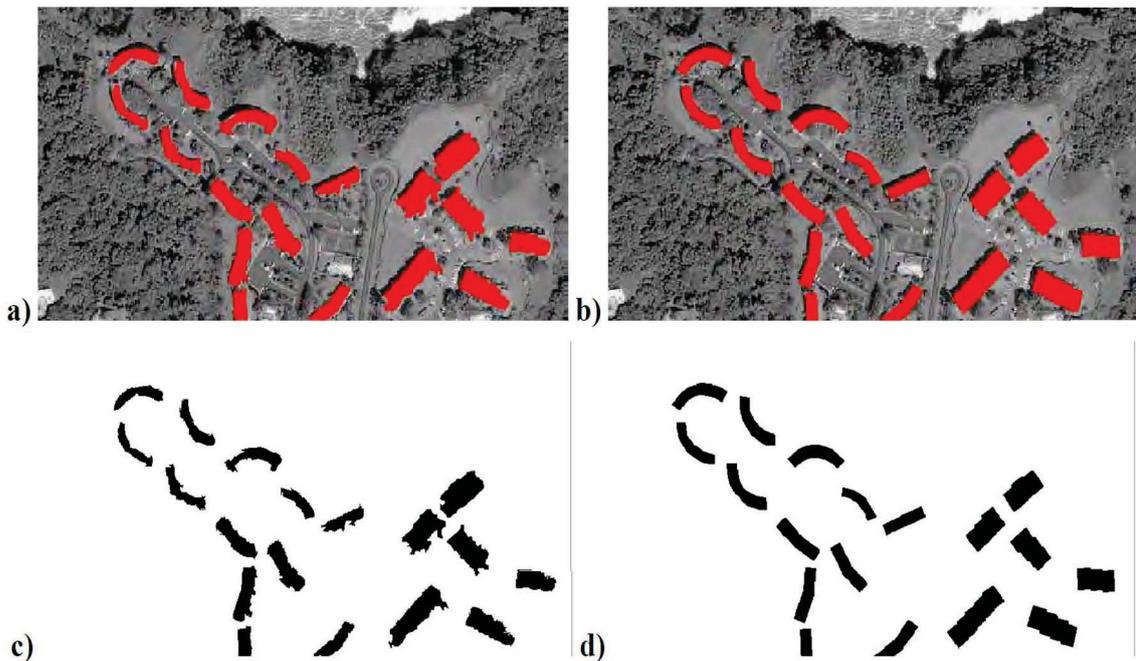
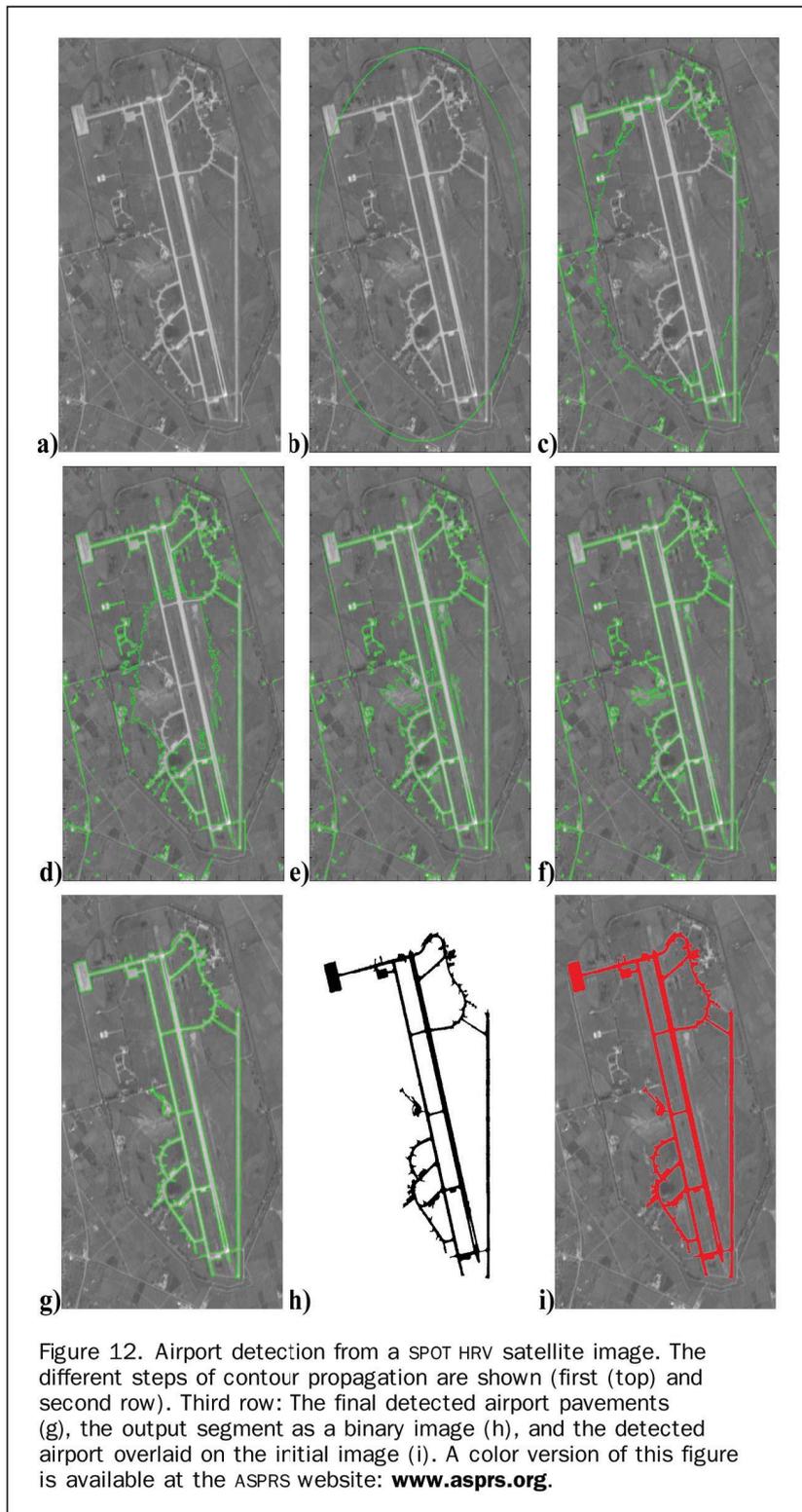


Figure 11. Comparing algorithm's result with the ground truth data for the satellite (QuickBird) test image. Top row: Detected segments (a), and the ground truth (b) overlaid on the initial image. Bottom row: The detected buildings as a binary image (c), and the reference binary image (d). A color version of this figure is available at the ASPRS website: www.asprs.org.



Brox, T., and J. Weickert, 2006. Level set segmentation with multiple regions, *IEEE Transactions on Image Processing*, 15(10):3213–3218.

Cao, G., X. Yang, and Z. Mao, 2005a. A two-stage level set evolution scheme for man-made objects detection in aerial images, *Proceedings of the International Conference of Computer Vision and Pattern Recognition*, (1):474–479.

Cao, G., X. Yang, and D. Zhou, 2005b, Mumford-Shah model based man-made objects detection from aerial images, *Proceedings of the International Conference on Scale-Space Theories in Computer Vision*, pp. 386–395.

Caselles, V., R. Kimmel, and G. Sapiro, 1997. Geodesic active contours, *International Journal of Computer Vision*, 22:61–79.

- Chan, T., and L. Vese, 2001. Active contours without edges, *IEEE Transactions on Image Processing*, 10(2):266–277.
- Chiang, T.Y., Y.H. Hsieh, and W. Lau, 2001. *Automatic Road Extraction from Aerial Images*, Stanford Education.
- Cohen, L.D., 1991. On active contour models and balloons, *Computer Vision, Graphics, and Image Processing: Image Understanding*, 53(2):211–218.
- Cremers, D., 2003. A variational framework for image segmentation combining motion estimation and shape regularization, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2003*, pp. 53–58.
- Doucette, P., P. Agouris, and A. Stefanidis, 2004. Automation and digital photogrammetric workstations, *Manual of Photogrammetry*, Fifth edition (C. McGlone, editor), American Society for Photogrammetry and Remote Sensing, pp. 949–981.
- Ferraro, M., G. Boccignone, and T. Caelli, 1999. On the representation of image structures via scale space entropy conditions, *IEEE Transactions Pattern Analysis Machine Intelligence*, (21):1199–1203.
- Gruen, A., E. Baltsavias, and O. Henricsson, 1997. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhaeuser, Basel.
- Gruen, A., O. Kuebler, and P. Agouris, 1995. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhaeuser, Basel, 321 p.
- Gruen, A., and H. Li, 1997. Semi-automatic linear feature extraction by dynamic programming and LSB-snakes, *Photogrammetric Engineering & Remote Sensing*, 63(9):985–995.
- Hanbury, A., 2003. A 3D-polar coordinate colour representation well adapted to image analysis, *Proceedings of the Scandinavian Conference on Image Analysis*, Sweden, Springer-Verlag, pp. 804–811.
- Jeon, B.K., J.H. Jang, and K.S. Hong, 2002. Road detection in spaceborne SAR images using a genetic algorithm, *IEEE Transactions on Geoscience and Remote Sensing*, 40(1):22–29.
- Karantzalos, K., D. Argialas, and N. Paragios, 2007. Comparing morphological levelings constrained by different markers, *Proceedings of the International Symposium on Mathematical Morphology*, Rio de Janeiro, Brasil. Published in *Mathematical Morphology and its Applications to Signal and Image Processing*, edited by G.J.F. Banon, J. Barrera and U. Braga-Neto, MCT/INPE, 2007, pp. 113–124.
- Kass, M., A. Witkin, and D. Terzopoulos, 1987. Snakes: Active contour models, *International Journal of Computational Vision*, pp. 321–331.
- Keaton, T., and J. Brokish, 2002. A level set method for the extraction of roads from multispectral imagery, *Proceedings of the Workshop on Applied Imagery Pattern Recognition*, pp.141–150.
- Laptev, I., H. Mayer, T. Lindeberg, W. Eckstein, C. Steger, and A. Baumgartner, 2000. Automatic extraction of roads from aerial images based on scale space and snakes, *Machine Vision and Applications*, 12(1):23–31.
- Lee, C.P., 2005. *Robust Image Segmentation Using Active Contours: Level Set Approaches*, Ph.D. dissertation, North Carolina State University, Raleigh, North Carolina, 146 p.
- Lee, S.H., and J.K. Seo, 2006. Level set-based bimodal segmentation with stationary global minimum, *IEEE Transactions on Image Processing*, 15(9):2843–2852.
- Martin, P., P. Refregier, F. Galland, and F. Guerauld, 2006. Nonparametric statistical snake based on the minimum stochastic complexity, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(9):1493–1500.
- Mayer, H., 1999. Automatic object extraction from aerial imagery - A survey focusing on buildings, *Computer Vision and Image Understanding*, 74(2):138–149.
- Mayer, H., S. Hinz, U. Bacher, and E. Baltsavias, 2006. A test of automatic road extraction approaches, *International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences*, 36(3):209–214.
- Mayer, H., I. Laptev, and A. Baumgartner, 1998. Multi-scale and snakes for automatic road extraction, *Proceedings of the Fifth European Conference on Computer Vision*, 2:720–733.
- Mayer, H., I. Laptev, A. Baumgartner, and C. Steger, 1997. Automatic road extraction based on multi-scale modeling, context and snakes, *International Archives of Photogrammetry and Remote Sensing*, 32(3-2W3):106–113.
- Mena, J.B., 2003. State of the art on automatic road extraction for GIS update: A novel classification, *Pattern Recognition Letters*, 24(16):3037–3058.
- Mueller, M., K. Segl, and H. Kaufmann, 2004. Edge- and region-based segmentation technique for the extraction of large, man-made objects in high-resolution satellite imagery, *Pattern Recognition*, 37(8):1619–1628.
- Mumford, D., and J. Shah, 1989. Optimal approximation by piecewise smooth functions and associated variational problems, *Communications on Pure and Applied Mathematics*, 42:577–685.
- Niu, X. 2006. A semi-automatic framework for highway extraction and vehicle detection based on a geometric deformable model, *ISPRS Journal of Photogrammetry and Remote Sensing*, 61:170–186.
- Osher, S, and J. Sethian, 1988. Fronts propagation with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations, *Journal of Computational Physics*, (79):12–49.
- Osher, S., and N. Paragios, 2003. *Geometric Level Set Methods in Imaging Vision and Graphics*, Springer Verlag, ISBN 0387954880.
- Paragios, N., and R. Deriche, 2002. Geodesic active regions: A new framework to deal with frame partition problems in computer vision, *Journal of Visual Communication and Image Representation*, Vol. 13, pp. 249–268.
- Paragios, N., Y. Chen, and O. Faugeras, 2005. *Handbook of Mathematical Models of Computer Vision*, Springer, ISBN 0387263713.
- Peng, J., D. Zhang, and Y., Liu, 2005. An improved snake model for building detection from urban aerial images, *Pattern Recognition Letters*, 26(5):587–595
- Rochery, M., I.H. Jermyn, and J. Zerubia, 2003. Higher order active contours and their application to the detection of line networks in satellite imagery, *Proceedings of the IEEE Workshop on Variational, Geometric and Level Set Methods in Computer Vision*, Nice, France.
- Rüther, H., H.M. Martine, and E.G. Mitalo, 2002. Application of snakes and dynamic programming optimisation technique in modeling of buildings in informal settlement areas, *ISPRS Journal of Photogrammetry and Remote Sensing*, 56(4):269–282.
- Samson, C., L. Blanc-Feraud, G. Aubert, and J. Zerubia, 2001. Two variational models for multispectral image classification, *Proceedings of the Third International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, Lecture Notes In Computer Science, (2134):344–358.
- Shi, Y., and W.C. Karl, 2005. Real-time tracking using level sets, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2(20–25):34–41.
- Sowmya, A., and J. Trinder, 2000. Modelling and representation issues in automated feature extraction from aerial and satellite images, *ISPRS Journal of Photogrammetry and Remote Sensing*, 55(1):34–47.
- Tsai, A., A. Yezzi, and A. Willsky, 2001. Curve evolution implementation of the Mumford-Shah Functional for image segmentation, de-noising, interpolation, and magnification, *IEEE Transactions on Image Processing*, 10(8):1169–1186.
- Wiedemann, C., C. Heipke, H. Mayer and S. Hinz, 1998. Automatic extraction and evaluation of road networks from MOMS-2P imagery, *International Archives of Photogrammetry and Remote Sensing*, 30(3/1):285–291.
- Zafiroopoulos, P., and T.F. Schenk, 1998. Extraction of road structures with color energy models, *Proceedings of SPIE, Videometrics VI* (F. Sabry, El-Hakim, and A. Gruen, editors), 3641:276–290.