Snakes-based approach for extraction of building roof contours from digital aerial images

Aluir P. Dal Poz and Antonio J. Fazan

São Paulo State University – Dept. of Cartography, R. Roberto Simonsen 305 – 19060-900 Presidente Prudente, SP - Brazil
aluir@fct.unesp.br, ajfazan@gmail.com.cr

Abstract

This paper presents a method for building roof contour extraction from digital images taken over complex urban scenes. The proposed method is based on a snakes-based energy function that represents building roof contours in digital images, which is optimized by using the dynamic programming (DP) algorithm. As most of the building roof contours are characterized by rectilinear sides intercepting at right angles, appropriate geometrics constraints are enforced into the original snakes energy function. The main advantage of using the DP algorithm for optimizing the proposed snakes-based energy function is its better radius of convergence, when compared to the one that is usually obtained in the original solution based on variational approaches. Experimental evaluation, including visual inspection and numeric analysis, was performed by using real data and the obtained results showed the potentiality of the proposed method for extracting building roof contours from digital images.

Keywords: Snakes, dynamic programming, building extraction, image analysis.

1. Introduction

The concepts of snakes and dynamic programming (DP) have been widely exploited in applications involving the extraction of road network from digital images, and one may quote Gruen and Li (1995), Agouris et al. (2001), Dal Poz and Vale (2003) and Dal Poz et al. 2010). However, few approaches have been developed for building extraction from digital images by using snakes. Rüther et al. (2002) used snakes to model building contours associated with informal settlement areas. Guo and Yasuoka (2002) proposed an approach based on snakes for extracting buildings from the combination of IKONOS image and elevation data. Oriot (2003) presented a statistical model of snakes for the building extraction from stereoscopic pairs of aerial images. Taking into account radiometric and geometric properties of buildings, Peng et al. (2005) modified the traditional snakes' model to enable a more stable convergence to building contours.

This paper proposes a method for extracting building roof contours from digital images. This method uses snakes as a basis for developing a mathematical model of these objects, whose solution is obtained through the DP optimization technique. The basic assumption of our method is that buildings are projected onto the image space as rectilinear structures, along with their adjacent sides intercepting at approximately right angles. This article is organized as follows: Section 2 presents the proposed method. The experimental assessment and analysis of the results obtained is presented in Section 3. Finally, conclusions from the analysis of results obtained in our experiments are presented in Section 4.

2. Method

2.1. Mathematical model of building roof contours

A snake is described in an image by a curve that moves along directions x and y under the influence of internal and external forces (Xu and Prince, 1998). Originally, it was formulated based on a continuous parametric curve, but when involving computational implementation, it should be replaced by a discrete curve, given by a polyline or a polygon (v) defined by a sequence of n vertices, as follows:

$$v_i = \begin{bmatrix} x_i & y_i \end{bmatrix}, \qquad i = 0, \ldots, n-1 . \tag{1}$$

The snakes function can be put in the following way (Kass *et al.,* 1988),

$$E_2\left(v\right) = \sum_{i=1}^{n-1}\left(\alpha_i \left|v_{i+1} - v_i\right|^2 + \beta_i \left|v_{i-1} - 2v_i + v_{i+1}\right|^2\right) + E_e\left(v\right) \tag{2}$$

where: the expression under the sum is known as the internal energy of the snakes; the first and second terms of the internal energy are known as of first and second order; the constants α_i and β_i are weights that control the terms of first and second orders; and $E_e\left(v\right)$ is known as the external energy.

In terms of the object 'roof contour', the curve v is a polygon representing the roof contour. An important feature of the building roof contours in digital images is that they usually are delimited by step edges, which allows the definition of the external energy function as follows,

$$E_{edge}\left(v\right) = \sum_{i=0}^{n-1} \gamma_i \left|\nabla G\left(v_i\right)\right|^2 \tag{3}$$

where: γ_i is a positive constant and $\left|\nabla G\left(v_i\right)\right|$ is the image gradient magnitude at the vertex v_i of the contour.

In addition, in most cases, buildings are represented by rectilinear structures in the image, defined by consecutive edge segments forming right angles at their corners. Thus, an additional external energy term is proposed, as follows,

$$E_{corner}\left(v\right) = \sum_{i=0}^{k-1} \eta_i \left[\left(1 - \cos\delta_i\right) \cdot CS\left(v_i\right)\right]^2 \tag{4}$$

where: η_i is a positive constant; $CS\left(v_i\right)$ is a corner detection function that returns the corner strength at vertex v_i; and δ_i is the angle of incidence of edges that define a corner at vertex v_i. Please note that the term '$1 - \cos\delta_i$' is a penalty function that favors right-angle corners.

The mathematical model of building roof contour based on snakes may be finally expressed by incorporating Equations 3 and 4 in Equation 2, resulting in:

$$E_2(v) = \sum_{i=0}^{n-1} \left[\begin{array}{r} \alpha_i \left|v_{i+1} - v_i\right|^2 + \beta_i \left|v_{i-1} - 2v_i + v_{i+1}\right|^2 - \gamma_i \left|\nabla G(v_i)\right|^2 - \\ \eta_i \left[\left(1 - \cos(\delta_i)\right) CS(v_i) \right]^2 \end{array} \right] \quad (1)$$

The values of weights β_i, γ_i, and η_i depend on the type of discontinuity existing at the vertex v_i: step edge or corner. All weights (including α_i) are positive, but when it comes to a corner, there has to be an abrupt change in direction at vertex v_i of the polygon. This implies the need of nullifying β_i to allow a second-order discontinuity at the vertex v_i. Moreover, the edge energy term (Equation 3) does not have any discrimination power of the discontinuity at the vertex v_i, implying the need to have $\gamma_i = 0$. On the other hand, when at the vertex v_i there is a step-edge discontinuity, the corner energy term (Equation 4) will have no discriminatory power. In this case, it is necessary to nullify the weight η_i.

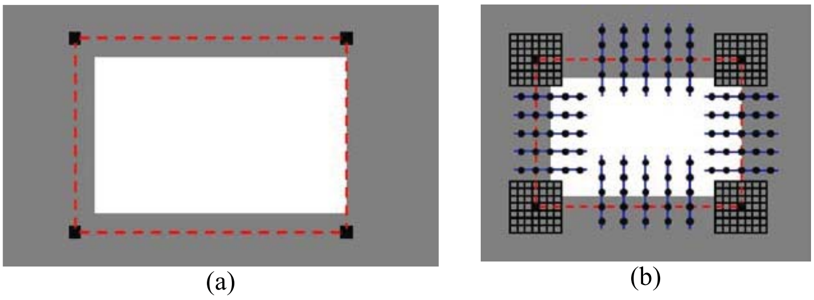
2.2. DP optimization of the building roof contour model

If the variables of an energy function in an optimization problem are not simultaneously interrelated, then an efficient way for solving this problem is through the DP technique (Ballard and Brown, 1982).

A simple analysis of the energy function given by Equation 5 shows that only three consecutive vertices (v_{i-1}, v_i, v_{i+1}) of the polygon *v* are simultaneously interrelated and, thus, it can be decomposed into a sum of $n-1$ sub-functions $E_i(v_{i-1}, v_i, v_{i+1})$, i.e.:

$$E_2(v) = \sum_{i=0}^{n-1} E_i(v_{i-1}, v_i, v_{i+1}) \quad (6)$$

In order to start the extraction process, a few seed points describing approximately the contour to be extracted should be provided by an operator. In general, the seed points should be positioned in the vicinity of the corners of the roof contour (Figure 1(a)). There is no need to provide seed points along the building contour sides, because these sides can be predicted considering that they are rectilinear.



(a) (b)

Figure 1: Illustrative example of the extraction process: (a) Initial polygon defined by seed points; (b) Sampling of the search space.

Two types of points generate the candidate polygons for the optimal polygon: points representing the sides and corresponding to the step-edge discontinuity; and points representing the corners and corresponding to the corner discontinuity. At a first glance, it seems simple to determine the corners and organize them to form the corresponding building roof contour polygon. However, false corners can be detected or even we can have missing true corners for a given roof contour. As a result, it is important to include building roof contour side points to provide better support to determine the correct corners.

In order to determine candidate points to represent a building corner, a reference point (seed point) must be provided by an operator (Figure 1(a)) and the remaining ones are determined by an algorithm to detect corners on a small sub-image around the reference point (Figure 1(b)). The dimensions of this sub-image should be sufficient so that it contains the correct corner of the building. The corner detector used is the one by Harris (Harris and Stephens, 1988) and the corners with the best response are stored together with the angles of incidence (δ_i) of edges that define the respective corners. The storage of multiple detected corners for each building corner is necessary for two reasons: first by the simple fact that there may be more than one corner in the sub-window; and the second by the possibility of the corner with the best response be a false positive.

Candidate points to represent building roof sides are sampled regularly along cross sections of sides of the initial polygon defined by pairs of seed points (Figure 1(b)). It is important to enhance that the cross sections are also sampled regularly along the sides of the initial polygon. As the cross sections are centered at points along the sides of the initial polygon, there is no preference on searching either side of the cross sections.

The strategy described above is applied to refine the initial polygon (defined by the seed points) provided at the beginning of the extraction process. Then, as points obtained by DP along the sides of the contour may be affected by local anomalies (e.g., an adjacent tree), causing local irregularities in the contour, a robust linear regression method is used to get polygon sides (straight lines) that better model the contour of the building sides. Finally, the final vertices of the refined polygon are determined by the intersection of straight lines obtained through the linear regression algorithm.

3. Experimental results

To preliminarily show the potential of the proposed method we present and analyze the results of one experiment using real data. We numerically compare building roof polygons extracted by the proposed method and by an operator. The polygons extracted manually are referred to as reference polygons. We used the followed quality parameters to evaluate the extracted building roof contours (Heipke et al., 1997): completeness, correctness, and RMSE (root mean square error).

Figure 2 shows an image segment containing several buildings. The initial, extracted and reference contours were overlaid on the image segment, as shown in Figures 2(a), 2(b), and 2(c), respectively. Please note that extracted contours 1, 2, 4, and 6 present some extraction problems. Although the extracted contour 4 is relatively compatible with the reference polygon, this one was accomplished manually by connecting its vertices, resulting in a rough approximation for the two curved sides. These sides are also roughly approximate by the proposed method, what was predictable because it was designed to handle rectilinear building roof contours.

The interaction of building 1, 2, and 6 with neighbor buildings affected the extraction process, as the corresponding extracted contours partially coincide with neighbor buildings. This is expected whenever a building roof boundary is very near and approximately parallel to a shadow or another building roof boundary. In the present context, the method can extract a neighbor building edge if its strength is higher than the correct edge. The performance of the method with the contours 3 and 5 can be considered satisfactory.

(a) (b) (c)

Figure 2: Experimental Results: (a) Initial contours; (b) Extracted contours; (c) Reference contours.

Table 1 shows the values of quality parameters. In general, the values show a good performance of the method, particularly in the extraction of building roof contours 3, 4, and 5. However, as already discussed above, the result obtained for the building roof contour 4 is in fact uncorrected for the two curved sides. As a result, the RMSE value is relatively high because the corner determination depends on the quality of extracted polygon sides. The method showed a partially satisfactory performance in the extraction of building roof contour 2, because one of the sides was incorrectly extracted. The RMSE and correctness parameters show that the worst result was obtained with the extraction of building roof contour 6.

Table 1: Quality parameters.

Quality parameter	**Building roof contour**					
	1	2	3	4	5	6
Completeness (%)	98	99	98	94	96	97
Correctness (%)	80	81	90	95	97	64
RMSE (pixels)	3.1	3.6	2.7	4.3	0.8	5.7

4. Conclusion and future work

The experimental results showed that the proposed method had a satisfactory performance in the task of extraction of different building roof contours from digital images. However, the experimental results also showed that the main disadvantage of the proposed method is that it cannot model local context. In fact, if a building roof boundary is very near and approximately parallel to a shadow or another building roof boundary, the extracted polygon may be affected to some degree.

The main direction for future work involves the development of strategies to circumvent the deficiency of the method in modeling local context. For example, the image gradient field can be exploited to separate correct building edge points from edge points representing shadow boundary or neighboring building roof boundary.

Acknowledgments

Authors thank FAPESP (São Paulo State Foundation for Scientific Research) for the financial support.

References

Agouris, P. *et al.* Dynamic node distribution in adaptive snakes for road extraction. *Proceedings of the 14th Annual Vision Interface Conference*, Ottawa, 2001, pp. 134-140.

Ballard, D., Brown, C. M. (1982), *Computer vision*, Prentice Hall, Inc., New Jersey, USA, 523 p.

Dal Poz, A. P., Gallis, R. B. A., Silva, J. F. C. (2010), "Three-dimensional semiautomatic road extraction from a high-resolution aerial image by dynamic-programming optimization in the object space". *IEEE Geoscience and Remote Sensing Letters*, Vol. 7(4): 796-800.

Dal Poz, A. P., Vale, G. M. (2003), "Dynamic programming approach for semi-automated road extraction from medium- and high-resolution images". *Proceedings of ISPRS*, V. 34: 87-91.

Gruen, A., Li, H. (1995), "Road extraction from aerial and satellite images by dynamic programming". *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 50(4): 11-20.

Guo, T., Yasuoka, Y. (2002), "Snake-based approach for building extraction from high-resolution satellite images and height data in urban areas". *Available online: http://www.gisdevelopment.net/aars/acrs/2002/vhr/018.pdf. (accessed in 30/June/2010)*.

Harris, C., Stephens, M. (1998), "A combined corner and edge detector". *Proceedings of the Alvey Vision Conferece*, 147-151.

Heipke, C., Mayer, H., Wiedemann, C., Jamet, O. (1997), "Evaluation of automatic road extraction". *Proceedings of ISPRS*, Vol. 32: 47- 56.

Kass, M., Witkin, A., Terzopoulous, D. (1998), "Snakes: Active contour models". *International Journal of Computer Vision*, Vol. 1(4): 321-331.

Oriot, H. (2003), "Statistical snakes for building extraction from stereoscopic aerial images". *Proceedings of ISPRS*, Vol. 34: 65-70.

Peng, J., Zhang, D., Liu, Y. (2005), "An improved snakes model for building detection from urban aerial images". *Pattern Recognition Letters*, Vol. 26: 587-595.

Rüther, H., Martine, H. M., Mtalo, E. G. (2002), "Application of snakes and dynamic programming optimisation technique in modeling of buildings in informal settlement areas". *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 56: 269-282.

Xu, C., Prince, J. L. (1998), "Snakes, shapes, and gradient vector flow". *IEEE Transactions on Image Processing*, Vol. 7(3): 359-369.