An Automatic Approach for Building Top Silhouette Extraction Using a PGVF Snake Model

ZHEN LIU
Center of Information & Network Technology, 
Beijing Normal University, Beijing 100875, China; 
and
YANHUI XIE
School of Information Science and Technology, 
Beijing Normal University, Beijing 100875, China; 
and
JIN CHEN
Department of Statistics, Indiana University, 
Bloomington, IN 47408; 
and
JUN WU
Program in Public Health, College of Health Sciences, 
University of California, Irvine, CA 92697-3957, USA

ABSTRACT:
The recent availability of high-resolution remote sensing technology provides a new data source for urban geospatial data acquisition, which has made it possible to detect buildings quickly and automatically. The building data are used for a variety of applications, including urban planning, land-use monitoring, map updating, and 3D city modeling. However, conventional approaches for building extraction are inefficient due to the high spatial heterogeneity between objects, as well as the increased texture and details from high-resolution images. This paper presents a new, automated method to detect the tops of buildings from high spatial resolution imagery using a Poisson Gradient Vector Flow (PGVF) Snake model. Based on the methodology of overall outline extraction for the tops of buildings from partitioned images, improved Canny edge detection was employed to produce a Boolean edge image, which is the initial condition necessary to calculate the PGVF Snake model for building detection. The methods are illustrated with a QuickBird image over the city of Los Angeles, USA. The results indicated that the proposed approach results in 92.3% completeness and 89.5% correctness compared with 61.5% and 65.5% for the traditional model and 84.6% and 85.5% for the GVF model, respectively.

Keywords: PGVF Snake Model, Building, Silhouette Extraction

1 INTRODUCTION

The development of current remote sensing technology has provided a new method for updating urban information. Because buildings are one of the most important land cover types in cities, the extraction of building information has extensive applicative values in fields of 3D urban reconstruction, urban planning, urban map charting, traffic control, military surveillance, etc. Furthermore, due to the structural complexity of urban buildings, rich interior texture, and the problems of shadow and shelter in the remote sensing images, extracting buildings in an automatic, quick, and accurate manner is still a difficult and avid topic of scientific research.

Thus far, methods for building extraction are mainly divided into two categories: edge-driven and region-driven methods. In the edge-driven method, lines are first detected by means of an edge detection algorithm, and assumptions are made on building polygons, with abstraction and verification done gradually. In the 1980s, by combining with the perception theory, Nevatia et al. [1] (1998) made assumptions on the building target shapes first and then proved their validity. By detecting lines, corner points, and other characteristics, they used a single image to generate a rectangular silhouette (which was assumed to be the building top silhouette), and then the height, shadow, and walls were used to verify the building top silhouette to extract the building. Liu et al. [2] (2006) proposed a method that applied the perception group to building top silhouettes with parallelogram shapes and performed extraction step by step. Based on the extraction of edge points, this method searches for lines and corner points, then builds U-type structures, and gradually forms parallelograms.

The region-driven method is mainly based on region segmentation, region verification, and extraction of a master curve. Edge detection and connection are performed last. Stassopoulou et al. [3] (2000) combined edge detection with region segmentation, using a method that combined Canny edge detection with multi-scale segmentation to segment images and extract regional characteristics. The building characteristics are recognized and extracted by combing other imaging conditions under the support of a Bayesian network. Guo et al. [4] (2001) generated buildings’ initial silhouette by combining gray images with range images, then used rectangle templates with proper orientation and scale parameters to approximate the silhouette.
The rectangle templates were deformed further to incorporate more edges to generate the true silhouette with concave-convex shapes. However, the aforementioned methods are mostly based on assumptions, verification, building model bases, or extracting the edge line characteristics of buildings with relatively regular shapes. Thus, the degree of automation and accuracy are relatively low.

This paper combined the edge detection with the energy field, and adopts a Poisson Gradient Vector Flow (PGVF) Snake model that uses optimized binary edge images to perform building top silhouette extraction. On the basis of Canny operator edge detection, optimized binary edge images and the Laplacian operator were adopted to solve the Poisson equation in order to obtain the external energy field of the active silhouette, on which updating and iterating were performed. Finally, methods with gradual convergence were used to approximate the building top silhouette.

2 PGVF SNAKE ALGORITHM MODEL

2.1 Snake algorithm

The active silhouette model, which is also called the Snake model, is a deformable curve model with minimum energy. This model was first proposed by Kass et al.[5] in 1987, and was called the “conventional active silhouette model” to distinguish it from other active silhouette models after improvement. The main idea of the Snake model is to evolve the curve toward the target image silhouette through continuous evolution by defining an initial energy function curve and initializing it around the target silhouette to be segmented, under the constraints of the minimized value of the energy function.

This model obtains the main axis of the active curve based on the problem of “energy minimization”. The energy of the active silhouette model is influenced by two types of energies: internal energy and external energy. The parameter $C(x(s), y(s))$ is set as the active silhouette curve, and the mathematical model of the active silhouette curve’s energy $E_{snake}$ can be expressed as:

$$E_{snake} = \int_0^1 [E_{int} (C(s)) + E_{image} (C(s))] ds \quad (1)$$

specifically,

$$E_{int} = \frac{1}{2} (a(s) |C'(s)|^2 + \beta(s) |C''(s)|^2) \quad (2)$$

The conventional active silhouette model unifies the constraints of image data, initial estimate, target silhouette, and basic knowledge in one characteristic extraction process and minimizes the energy from coarseness to fineness in the scale space, which has greatly extended the capture regions and reduced the computational complexity. However, it is sensitive to initial positions and noises, and it is hard to converge to edge silhouettes with concave-convex characteristics. In order to solve these problems, subsequent researchers have attempted to improve its application results in aspects of the initial silhouette, the model, etc. Improved models include the Balloon model[6], convolution model[7], etc. Improvements on the computational methods for energy minimization include the dynamic programming of Animi[8], the greedy algorithm of Williams[9]. In the improved model of the active silhouette algorithm, the Gradient Vector Flow (GVF) Snake model, proposed by Xu et al.[10], successfully solved the problem that the target silhouette edge cannot be approximated completely. GVF is an external force field that captures the edges of the images from both directions simultaneously, and it is a gradient vector diffusion of the image function’s edge graphs. It replaces the image external force $\nabla E_{ext}$ of the conventional models with two-dimension gradient vector flow

$$V(x, y) = [u(x, y), v(x, y)] \quad (3)$$

GVF Snake extends the capture regions, and it is able to drag the Snake toward the deep depression areas of the objects. However, both the conventional active silhouette model and the improved GVF Snake model are iterations of time $t$, which are sensitive to initial silhouette positions and noises and demand heavy computation.

2.2 PGVF Snake model

The PGVF[11],[12] Snake active silhouette model replaces the image external force of the GVF Snake model with the solution of the Poisson Equation solved to improve the image’s gradient vector flow, enabling the active silhouette lines to more easily approximate those silhouettes with concave-convex characteristics or shapes with relatively large surface ranges.

The Poisson equation of the PGVF Snake active silhouette model can be expressed as:

$$\nabla^2 \phi(x, y) = f_{edge}(x, y) \quad (4)$$

In the equation, $\nabla^2$ is a Laplacian operator, and $f_{edge}$ is the edge binary image obtained by using the Canny edge detection operator. $\phi(x, y) = (u(x, y), v(x, y))$ can be obtained by solving the Poisson equation (4-1) with numerical difference methods, and then the external energy $E_{ext}$ of the active silhouette mathematical model can be expressed as the Poisson gradient vector flow $\phi(x, y)$. The external force field is $-\nabla \phi(x, y)$. Then, the energy balance equation of the PGVF active silhouette model is:

$$C(s) = \alpha C^n(s) - \beta C^m(s) - \nabla \phi(x, y) \quad (5)$$

The gradient field of the edge image $f(x, y)$ is $\nabla f(x, y)$. Thus, the energy functional of the image active silhouette line is:

$$E = \int \int \mu [\nabla^2 \phi + (|\nabla f|^2)^{-1} (\phi - \nabla f)]^2 dx dy \quad (6)$$

In the above expression, $\mu$ is a control parameter, and $\nabla^2$ is a Laplacian operator. In order to minimize the energy function, the Euler-Lagrange differential equation (4-3) can be obtained by variational calculation:

$$\mu \nabla^2 \phi - (\phi - \nabla f) \nabla f^2 = 0 \quad (7)$$

The iteration of PGVF is:

$$\begin{align*}
\{ u &= u + \mu \nabla^2 u - (u - f_x)(f_x^2 + f_y^2) \\
\{ v &= v + \mu \nabla^2 v - (v - f_y)(f_x^2 + f_y^2) 
\end{align*} \quad (8)$$

The convergent iteration of the silhouette is:

$$\begin{align*}
\{ X &= \text{inv} \times (\gamma' \times X + \kappa \times \mu) \\
\{ Y &= \text{inv} \times (\gamma' \times Y + \kappa \times \mu) 
\end{align*} \quad (9)$$
In the expression, \( X \), \( Y \) are the horizontal and vertical coordinate matrices of the control points, and \( \text{inv} \) is the parameter matrix of \( \alpha, \beta, \gamma, \kappa \) is the computing matrix.

The Poisson equation of the PGVF Snake active silhouette model contains no derivative of time, thus its definite condition contains no initial conditions but only a boundary condition. With the Neumann Boundary Condition set as its definite condition (which means the derivative value of the unknown function is given on the boundary), the PGVF Snake active silhouette model not only improves the vector field but also, at the same time, demands no iteration of time in the process of target silhouette characteristic extraction, which results in a faster speed.

3 APPROACH FOR BUILDING TOP SILHOUETTE EXTRACTION BASED ON THE PGVF SNAKE MODEL

3.1 Procedures for building top silhouette extraction based on the PGVF Snake model

To extract the building top silhouette using the PGVF Snake active silhouette model algorithm, the Canny edge operator was first used for edge detection. Then, by setting appropriate threshold values, pseudo edge points and some small noise edge points were eliminated, and the optimized \( f_{\text{edge}}(x, y) \) was obtained as a prerequisite for the Poisson equation. Meanwhile, the initial silhouette positioning method of the building tops based on block images was adopted, and effective initial silhouette lines of the active silhouette model were set. A flow chart of the entire experiment is shown in Figure 1.

![Figure 1. Flow chart of the building top silhouette extraction.](image)

The main algorithm in the PGVF Snake active silhouette model lies in the iterative diffusion and silhouette convergence of the Poisson gradient vector flow field. After reading the original image \( f(x, y) \), the edge graph \( f_{\text{edge}}(x, y) \) of the original image was calculated. Then, the iterative diffusion equation of the Poisson gradient vector flow was calculated in accordance with the following procedures:

1. Normalize the image’s edge graph \( f_{\text{edge}}(x, y) \);
2. Calculate the gradients of the edge graph \( f_{\text{edge}}(x, y) \) in both directions of \( x \) and \( y \);
3. Calculate the parameters and initialize the parameter array;
4. Use the Neumann Boundary Condition to calculate the Laplacian operator; and
5. Update the computational result of PGVF.

The silhouette convergence used numerical interpolation methods for the convergence of the Snake.

3.2 Initial silhouette positioning of the building tops based on block images

Due to the relatively large quantity of target elements of the surface feature in high-resolution images, as well as the regional concentrative distribution of the buildings, we (on the basis of Canny edge detection) used methods based on block images to perform initial positioning for the building silhouette of every region to extract all of the buildings in the images. The basic algorithm and flow chart are shown in Figure 2 and described below:

1. In the binary images of the Canny edge detection, images are divided into \( n \) blocks by setting roads as the boundary characteristic.
2. Set threshold \( T \), set the regional blocks of the images as units, and eliminate the unnecessary false edges and noise edges in the regional blocks. If the continuous length of the pixel \( l < T \), eliminate this edge; if \( l \geq T \), keep the edge.
3. By setting the regional blocks of the images as units, the starting points, terminal points, and knee points of the edge lines are connected with lines to form initial silhouette lines of the building tops with block images as units.

![Figure 2. Initial silhouette positioning of the building top based on block images.](image)
effective approximation region of the true silhouette, which reduced the iterations greatly and shortened the iteration time.

4 EXPERIMENTAL RESULTS AND ANALYSIS

On the basis of the rough initial silhouette line of the building tops based on block images, the use of the PGVF Snake model effectively extracted the building top silhouette characteristics of every image block automatically by PGVF iteration and silhouette convergence iteration. We conducted the experiments using the QuickBird satellite that took photos of the urban areas of Los Angeles, USA. The original image below was taken on February, 2003. The resolution is 0.6 meters. The experimental procedures and results are shown as figure 3:

![Original image](image1)

(a) Original image.

![Canny edge detection result](image2)

(b) Canny edge detection result.

![Binary edge image after denoising](image3)

(c) Binary edge image after denoising

![Gradient vector flow field of the binary edge image](image4)

(d) Gradient vector flow field of the binary edge image.

![Initial silhouette positioning of the block image building tops](image5)

(e) Initial silhouette positioning of the block image building tops.

![Iterative convergence result of the PGVF Snake model](image6)

(f) Iterative convergence result of the PGVF Snake model.

Figure 3. Results of the building top silhouette extraction based on PGVF Snake optimization model.

Figure 3 shows the computational results of the original image after Canny edge detection, filtering edge noises, gradient vector flow computation, initial image positioning of the block image buildings, and PGVF iteration, respectively. In Figure 3, (f) is the extraction result under the premise of initial silhouette positioning of the building top based on block images, which took 15 iterations and less than 2 s. Parameters of the elastic energy and bending energy in the PGVF active silhouette model were set as $\alpha = 0.05$, $\beta = 0.05$, and $\mu = 0.1$.

From the experimental results in Figure 3, we can see that, in the initial silhouette positioning graph of the block image building tops (Figure 3-(e)), the curvature of the silhouette line in (1) was relatively large, and the silhouette line in (2) was discontinuous. The initial silhouette line is a line that connects two endpoints, whose iterative result was quite ideal in the final result (Figure 3-(f)).

Meanwhile, in order to illustrate the effectiveness of the PGVF active silhouette model, we also selected 20 buildings in specified regions and used CANNY operator for initial silhouette extraction. Then, the conventional, GVF, and PGVF models were each used for building top silhouette extraction. Comparisons were made on the accuracy of the extraction, extraction time, etc. Experimental results are shown in Figure 4.

![Force field of the conventional model](image7)

(a) Force field of the conventional model

![Force field of GVF Snake](image8)

(b) Force field of GVF Snake.

![Force field of PGVF Snake](image9)

(c) Force field of PGVF Snake.

Figure 4. Comparisons on the vector flow force field of the conventional active silhouette, GVF Snake, and PGVF Snake models.

Table 1 contains the experimental data extracted from the same building top silhouette under the premise of the same initial silhouette.

<table>
<thead>
<tr>
<th>Model Comparison</th>
<th>Conventional model</th>
<th>GVF model</th>
<th>PGVF model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completeness</td>
<td>61.5%</td>
<td>84.6%</td>
<td>92.3%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>65.5%</td>
<td>85.5%</td>
<td>89.5%</td>
</tr>
</tbody>
</table>

Table 1. Comparisons of the experimental results from the conventional active silhouette, GVF, and PGVF models.
<table>
<thead>
<tr>
<th>Time (s)</th>
<th>8.52</th>
<th>8.31</th>
<th>2.42</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iterations</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>

From figure 4, we can see that the image force field of the conventional model (Figure 4-a) was obtained by the computation of gray values. Its edge characteristics were quite obscure and disordered. The gradient vector flow force field of GVF Snake (Figure 4-b) was obtained by the computation of edge gradients. Its edge characteristics were obvious. The Poisson gradient vector flow force field of PGVF Snake (Figure 4-c) was obtained by solving the Poisson equation of the binary edge graph. Its force field appeared much the same as that of the GVF Snake, but the overall flow fields of these models differed greatly in corners and gaps. Thus, the gradient vector flow force field of the PGVF model was an improvement over that of the GVF model.

From Table 1, we can see that the PGVF model exceeded the conventional model in both accuracy and speed. Moreover, when compared with the results from the GVF model, accuracy was improved to a certain extent, and the extraction speed was much faster.

5 Conclusion

Building top silhouette extraction based on the PGVF Snake model first used the Canny operator for edge detection and denoising optimization to obtain binary edge images. Then, based on the image block method, the initial silhouette was positioned within the effective ranges of the building tops’ true silhouette. Finally, under the premise of the initial silhouette, the building top was extracted using the PGVF model for iterative convergence. Experimental results showed that, when using the PGVF Snake model for extracting the building top silhouette from high-resolution images, the PGVF Snake model exceeded the conventional and GVF models in accuracy, speed, and degree of automation, and a better result was obtained. However, considering the importance that the building’s initial silhouette positioning played in the extraction results and speed, further research on how to position the building areas and the initial silhouette in a fast and accurate manner is required. Moreover, improving the robustness of the use of the PGVF algorithm for different surface features and images will be another focus of further study.

6 References