A STAR-CORNER ALGORITHM FOR BUILDING EXTRACTION IN SATELLITE/AERIAL IMAGES

Parvaneh Saeedi and Melissa Cote

School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada
{psaeedi, melissa_cote}@sfu.ca

ABSTRACT
This paper presents a novel approach for extraction of rooftops in satellite/aerial imageries. This method is semi-automatic in which a point inside each rooftop must be identified by the user. The method utilizes corners and star algorithm to generate a set of rooftop outlines that are assessed and refined through an energy minimization process. The energy of each rooftop candidate is computed using color invariant models of rooftops’ inside and outside regions. A refinement process is incorporated that employs color values to fit rooftop outlines to their best potential locations. The proposed algorithm is a clean and efficient method that can cope with angular surfaces (gabled rooftops) or arbitrary illumination conditions. Experimental results for images of Richmond, BC, verify an average shape accuracy of 95%.

Index Terms—Building detection, 3D building model reconstruction, curve evolution, energy minimization, Gaussian color invariant

1. INTRODUCTION
Automatic 3D building model reconstruction from satellite/aerial imageries has been an active research that has attracted more attention over the last two decades due to abundance of commercially provided high-resolution fast-updating images. Building extraction as the basis of 3D building reconstruction has been addressed through many approaches. These approaches vary according to sensor modalities and their properties, supplementary information, prior knowledge and model/shape assumptions [1-5].


Perhaps one of the reasons for the popularity of curve evolution based methods is their flexibilities in fitting into complicated profiles comparing to the methods based on constrained templates or simple shapes. However, conditions such as shadows, arbitrary illumination and/or variant reflection and sloped surfaces can affect the performance of curve evolution based methods. These conditions usually exist in the input images and to develop methods that are capable of extracting buildings under such conditions proves vital to the usability of these methods.

In this paper we propose a new semi-automatic approach for rooftop extraction in satellite/aerial imageries. The proposed method combines the strength of energy based approaches with corners to extract building with complex profiles. Details of the proposed method are presented next.

2. METHODOLOGY
Figure 1 shows the flow diagram of the proposed method. Details of each part are explained in the following sections.

2.1. Color Enhancement
Arbitrary illumination and/or varying reflection could raise discontinuities on intensity values of rooftop components (e.g. gabled rooftops). Geusebroek et al. [12] proposed a Gaussian color invariance model under different assumptions (object and illumination properties) for
creating illumination and geometrical invariant image models. In particular, one of their models is for matte/dull objects that are illuminated with arbitrary illumination sources. This model is adopted in here since:

1. Most rooftops are made of matte/dull materials,
2. Varying reflection by slopped surfaces has similar visual effects as even surfaces illuminated arbitrarily, and
3. Rooftop surfaces are generally colored.

Based on the above model $C_{\lambda}$ (object reflectance properties) and $C_{\lambda,\lambda}$ are computed:

$$C_{\lambda} = \frac{E_{\lambda}}{E}, \quad C_{\lambda,\lambda} = \frac{E_{\lambda,\lambda}}{E}$$

Here $E$ is the reflected spectrum, and $E_{\lambda}$ and $E_{\lambda,\lambda}$ are the first and the second order differentiations with respect to the wavelength $\lambda$. The Gaussian color model is approximated [12] from the RGB model by:

$$\begin{bmatrix}
E \\
E_{\lambda} \\
E_{\lambda,\lambda}
\end{bmatrix} =
\begin{bmatrix}
0.3 & 0.58 & 0.11 \\
0.25 & 0.25 & -0.5 \\
0.5 & -0.5 & 0
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}$$

Figure 1 (a) and (b) show an aerial RGB image ($I$) and its Gaussian color model ($\Gamma$). To visualize the Gaussian color model, $C_{\lambda}$ and $C_{\lambda,\lambda}$ are placed into G and R channels while channel B is set to zero.

### 2.2. Corner Detection

Harris corner detector (on gray scale version of the input images) is utilized to detect points with low self-similarities. In order to reassure that all distinctive points are detected the sensitivity of the Harris corner detector is increased.

### 2.3. Star-Corner algorithm

In this step a starting point inside the target rooftop is manually selected. A circular area, called reference region ($R$), with radius $r$ (8 pixels) centered at the starting point is used to compute the means ($\mu_R=[H_{Rc,0},H_{Rc,1}]$) and standard deviations ($\sigma_{Rc}=\sigma_{Rc,0},\sigma_{Rc,1}$) for $C_\lambda$ and $C_{\lambda,\lambda}$. From the starting point, several rays with angular separation of $\theta$ degrees ($10^\circ$) are radiated. Here we define a search sector ($S$) that is a region centered on each ray sweeping from $\theta/2^\circ$ to $-\theta/2^\circ$. This sector moves in the outward direction (with respect to the starting point) and each time the algorithm searches for corners that fall inside $S$. At each position a measure of $D_S$ is computed that measures the similarity between the search sector and reference region.

$$D_S = \frac{1}{\text{Area}(S)} \sum_{x,y} \left\| I(x,y) - \mu_I \right\|$$ (4)

Here a large $D_S$ indicates that perhaps the borders of a rooftop have been reached and therefore the searching task for the current ray should not be further extended. This is shown in Figure 2.

### 2.4. Corner Grouping

Extracting all corners in input images is a critical issue for reliability of the proposed algorithm, yet over sensitized corner detection could impact time complexity by increasing the number of potential rooftop candidates. For a faster performance a “Corner grouping” step is incorporated in which the image $\Gamma$ is tessellated by blocks of non-overlapping $w_c \times w_c$ pixels ($w_c$ is 16). For each block with more than one corner, one corner representative is chosen based on the Harris corner response. All the other corners are discarded, Figure 3.

### 2.5. Energy Minimization

In this step all possible combinations (minimum $n$ (3) and maximum $m$ (9)) of the representative corners ($N$) are computed and placed in set $H$.

$$H=\{N_C_m, \ldots, N_C_n\}$$ (5)

Those hypotheses that do not contain the starting point are filtered out. An energy function ($EF$) computes the difference between the two internal and external distances for each hypothesis:

$$EF = \frac{\gamma_1}{\text{internal region of } h \in H} - \frac{\gamma_2}{\text{external region of } h \in H}$$ (6)

Here $\gamma_1, \gamma_2 > 0$ are constants that control priorities of the internal similarity and the external dissimilarity and both are set to 1 for this work. In Equation (6), the first term computes the normalized $\Gamma$ similarity between hypothesis’s internal and reference regions and the second term estimates the normalized $\Gamma$ dissimilarities between hypothesis’s external and reference regions. The hypothesis with the
lowest energy value is selected as the best rooftop candidate.

In order to improve the speed of the algorithm, two lookup tables are employed that contain the internal or external energy terms in Equation (6) between every two corners in the set \( N \). This means that regardless of the number of hypotheses, the energy terms (for every two corners) are computed only once. Figure 6 shows a sample results overlaid on the original aerial image.

![Figure 6](image)

**Fig.6.** Typical results after energy minimization.

### 2.6. Refinement

In this step the boundary of each hypothesis is refined to fit to the exact profile of the rooftop. Since each hypothesis is built using representative corners and that may not necessarily be the best choice among all corners that it represents, some inaccuracies in the contour of extracted hypothesis are anticipated. Also, the Gaussian invariant model does a good job in removing color variations on rooftops. However it could also remove color variations of neighboring or background regions if their colors are fairly similar to the rooftop’s, Figure 7.

![Figure 7](image)

**Fig.7.** While boundary (a) of the rooftop is clear on the RGB image (b) it is not easily separable from the background in (c).

This partially originates by the projection from the three-channel RGB onto the two-channel Gaussian color invariant model. Therefore this refinement process utilizes the original RGB image values (transformed into the CIE \( L^*a^*b^* \) space and is referred by \( \Psi \)). A binary mask \( (M) \) for each energy minimized hypothesis \( (H) \) is generated. This mask is then dilated by a structuring element of \( 11 \times 11 \) pixels. The extended mask is applied on the image to isolate the hypothesis and its neighboring regions from the rest of the image. We refer to this image patch as extended hypothesis \( (EH) \). The mean square error, \( MSE \), (on a window of \( 3 \times 3 \) pixels) is computed for all pixels on the \( EH \) to generate an error patch of \( EH_{MSE} \).

\[
MSE = \frac{1}{9} \sum (\Psi(x,y) - \mu_{\Psi R})^2
\]

Here \( \mu_{\Psi R} \) is the average color of the reference area in \( \Psi \). \( (x,y) \) are the coordinates of points in the \( EH \). At this stage Otsu thresholding is applied on \( EH_{MSE} \). For every boundary point of the hypothesis \( (H) \), a window of \( 11 \times 11 \) pixels is used for which the \( EH_{MSE} \) is segmented into two regions. Any value above the Otsu threshold is set to zero and otherwise it is set to 1. These segments are then overwritten onto the initial hypothesis’ mask \( (M) \), Figure 8(d). A connectivity check is then applied to remove any unattached pieces (caused by neighboring rooftops with similar colors), followed by a hole filling process that removes any hole inside the updated mask. The final hypothesis is defined by the boundary points of the updated mask, Figure 8(f).

![Figure 8](image)

**Fig.8.** (a) Primary contour, (b) mask of the primary contour, (c) \( MSE \) image, (d) updated mask, (e) final mask after removing disconnected pieces and hole filling. (f) Final hypothesis.

### 3. EXPERIMENTAL RESULTS

The proposed method is tested on 14 aerial (Pictometry Int. Corp.'s, resolution of 0.15 pixel/meter) and satellite (QuickBird, 0.6 pixel/meter) images. All experiments are conducted on a PC with CPU Intel Core2 2.4GHz with 2GB RAM and all programs are implemented in MATLAB 7.6 images (due to the space limitation) are used to show the typical results.

Note that only one set of parameters are used for all satellite and one set for all aerial images. The difference between these parameters is only in the sensitivity of the Harris corner detector (was set higher for satellite images).

The method is assessed quantitatively using 4 metrics:

\[
\text{Shape Accuracy} (\text{ShAc}) = 1 - \frac{|A_{G} - A_{F}|}{A_{G}}
\]

(8)

\[
\text{Completeness} (\text{Comp}) = \frac{TP}{(TP + FN)}
\]

(9)

\[
\text{Correctness} (\text{Corr}) = \frac{TP}{(TP + FP)}
\]

(10)

\[
\text{Quality} (\text{Qua}) = \frac{TP}{(TP + FP + FN)}
\]

(11)

\( A_{G} \) is the (manually found) ground truth rooftop’s area and \( A_{F} \) the area of the same rooftop detected by this work. TP, FP and FN represent True Positives (correctly extracted),
False Positives (incorrectly extracted) and False Negatives (correctly not extracted) pixels of a rooftop.

Table 1 shows the mean quantitative values for each image. Based on these results, the shape accuracy for both satellite and aerial images is about 95%.

This value is better than those reported by [6] (80%) and [7] (84%). The quality of the extraction drops from 91% (aerial) to 87% (satellite) which can be explained by the lower quality (resolution) of satellite images.

Table 1: Quantitative results for aerial and satellite images.

<table>
<thead>
<tr>
<th>Img. No.</th>
<th>No. of Bldgs</th>
<th>Ave. ShAc [%]</th>
<th>Ave. Comp [%]</th>
<th>Ave. Corr [%]</th>
<th>Ave. Qua [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aerial images</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>6</td>
<td>95.29</td>
<td>97.84</td>
<td>93.44</td>
<td>91.56</td>
</tr>
<tr>
<td>A2</td>
<td>13</td>
<td>91.69</td>
<td>99.48</td>
<td>93.86</td>
<td>91.42</td>
</tr>
<tr>
<td>A3</td>
<td>10</td>
<td>97.59</td>
<td>93.75</td>
<td>91.58</td>
<td>86.32</td>
</tr>
<tr>
<td>A4</td>
<td>6</td>
<td>95.12</td>
<td>97.97</td>
<td>94.32</td>
<td>91.65</td>
</tr>
<tr>
<td>A5</td>
<td>6</td>
<td>91.66</td>
<td>98.13</td>
<td>90.59</td>
<td>89.06</td>
</tr>
<tr>
<td>A6</td>
<td>6</td>
<td>89.92</td>
<td>99.01</td>
<td>89.95</td>
<td>89.15</td>
</tr>
<tr>
<td>A7</td>
<td>10</td>
<td>96.42</td>
<td>97.56</td>
<td>94.19</td>
<td>92.03</td>
</tr>
<tr>
<td>A8</td>
<td>11</td>
<td>98.26</td>
<td>97.67</td>
<td>96.01</td>
<td>93.87</td>
</tr>
<tr>
<td>A9</td>
<td>13</td>
<td>96.99</td>
<td>97.83</td>
<td>94.98</td>
<td>93.03</td>
</tr>
<tr>
<td>A10</td>
<td>18</td>
<td>98.85</td>
<td>96.90</td>
<td>95.82</td>
<td>92.98</td>
</tr>
<tr>
<td>Total</td>
<td>99</td>
<td>95.17</td>
<td>97.61</td>
<td>93.18</td>
<td>91.10</td>
</tr>
<tr>
<td>Satellite images</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>8</td>
<td>95.00</td>
<td>97.87</td>
<td>93.22</td>
<td>91.37</td>
</tr>
<tr>
<td>S2</td>
<td>9</td>
<td>91.90</td>
<td>94.47</td>
<td>90.56</td>
<td>88.42</td>
</tr>
<tr>
<td>S3</td>
<td>9</td>
<td>94.04</td>
<td>91.79</td>
<td>91.10</td>
<td>84.24</td>
</tr>
<tr>
<td>S4</td>
<td>7</td>
<td>97.91</td>
<td>92.68</td>
<td>90.79</td>
<td>84.72</td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>94.71</td>
<td>94.2</td>
<td>91.41</td>
<td>87.18</td>
</tr>
</tbody>
</table>