Automatic Building Detection in Aerial Images using a Hierarchical Feature Based Image Segmentation

Mohammad Izadi and Parvaneh Saeedi
School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada, V5A 1S6
{mia4, psaeedi}@sfu.ca

Abstract

This paper introduces a novel automatic building detection method for aerial images. The proposed method incorporates a hierarchical multilayer feature based image segmentation technique using color. A number of geometrical/regional attributes are defined to identify potential regions in multiple layers of segmented images. A tree-based mechanism is utilized to inspect segmented regions using their spatial relationships with each other and their regional/geometrical characteristics. This process allows the creation of a set of candidate regions that are validated as rooftops based on the overlap between existing and predicted shadows of each region according to the image acquisition information. Experimental results show an overall shape accuracy and completeness of 96%.

1. Introduction

Automatic 3D map reconstruction with high accuracy requires precise identification and detection of building rooftops. Any inaccuracy in such identification will affect the accuracy and reliability of the generated 3D map. Building identification or extraction in satellite/airborne photogrammetric images has been a subject of interest for several decades. Previous work on this subject varies according to the type and variety of sensors, input's dimension (mono/stereo/ multiscopic), and the amount of manual interaction or supervision. The general theme for most of these methods is to identify buildings using distinctive features and their relationships, active contours, segmentations and prior models. Since in the proposed work our interest is in utilizing monoscopic aerial images, a number of previous works that incorporate monoscopic input images are reviewed to provide the state of art in this research area.

Lin and Nevatia [1] grouped and analyzed line segments to extract rectilinear rooftops in single image. Nosrati and Saeedi [2] extracted polygonal rooftops by examining the relationship between lines and their cross-sections in aerial/satellite images. Jin and Davis [3] provided a differential morphological profile to generate building hypotheses with a verification process that relied on shadows and spectral information. Theng [4] developed an active contour based model utilizing a circular casting algorithm. Peng et al. [5] suggested a snake model to extract buildings from gray-level aerial images using the radiometric and geometric characteristics of buildings. Wei et al. [6] proposed a probability model to extract buildings with simple profiles with homogenous texture from dense urban area in high-resolution images using some discriminative features. Li et al. [7] used a clustering method to first segment the image. The segmented regions are then verified by their dominant lines' length. Peng and Liu [8] presented a method for building detection based on models and context in which building profiles are extracted from segmented regions utilizing shadow cast direction, context and a snake model. Wei et al. [9] proposed a low precision method by clustering the input image into shadow and non shadow regions first. The shadow regions are then used to verify the presence of buildings and a Hough transform is employed to extract the buildings boundaries. Karantzalos and Paragios [10] proposed a recognition-driven variational framework for automatic buildings extraction from monoscopic aerial and satellite images using prior models. While their method shows great potential, it is rather slow and can only deal with a limited number of shapes (8 profiles).

In this paper, we propose a novel method to extract building profiles using a new hierarchical feature based segmentation method in monoscopic aerial images. The building's context and geometrical features are used for local segmentation of image. The main objectives are to extract rooftop boundaries for buildings with complex shapes and profiles including gabled or flat surfaces.

(a) (b)

Figure 1. a) Low value of h_r , b) High value of h_r.

2. Methodology

2.1 General Segmentation Problems

Most segmentation algorithms segment an image into regions that vary in color/intensity regardless of other significant features of the regions [11]. Generally, segmentation algorithms are parameter dependent which adds to the functionality of such segmentations. An example of such case can be seen for Mean Shift Segmentation Algorithm [1,2] where the quality of segmentation is controlled by three parameters: the spatial resolution h_s, the color resolution h_c, and the minimum segment size *M*. These parameters are usually set manually and often a single setting will not be sufficient for a complete segmentation. Figure 1 depicts an example where h_c was set to two different values. It is a challenge (if possible) to find a proper value of h_c to segment all buildings images correctly.

In this paper a feature-based segmentation is presented that incorporates an adaptive range resolution for h_c to segment the image with no parameter variation.

2.1.1 Hierarchical Segmentation

Using the mean shift segmentation algorithm [13], the input image (I) is segmented for Range Resolutions of $h_c = (T_l, T_l + \Delta T, \dots, T_l + (K-2)\Delta T,\ T_h)$. For each value of $i \in K$, $K = (T_h - T_l + 1)/\Delta T)$, a segmented image S_i is generated including a set of regions R_{ij}:

$$S_i = \cup R_{ij} \ and \ j = 1, \dots, N_i \quad (1)$$

Here N_i is the number of regions in S_i. Parameters T_l and T_h have a large difference so they cover all possible regions in the image. In this implementation, parameters ΔT and T_l are set to 1 (the minimum distance of two pixels with different colors) and T_h is set to 15. The spatial resolution h_s and the minimum segment size *M* for mean shift segmentation are 3 and 200. We found that any T_h larger than this value would cause the entire image to be segmented into one piece. An additional segmented image S_{K+1} is also defined by the entire image.

Using the above segmented images and their regions, a tree structure is established. This tree is utilized to identify the best Range Resolution for various regions of the image via a set of rooftop constraints.

2.1.2 Rooftop Constraints

We define several features associated with building rooftops. These features are defined by examining rooftops in image datasets (Pictometry Int. Corp.'s images of areas around Vancouver's International Airport). With the assumption that the boundary of a region R_{ij} is defined as a parametric curve $B_{R_{ij}}(s)\colon [0,1] \rightarrow (x(s), y(s))$, we define:

Total curvature*:* the mean of spatial change of a point on the boundary. It is measured by the integral of the absolute value of the second order derivative which is normalized by the length of the boundary (the integral of the absolute of the first order derivative) [12].

$$Tc(R_{ij}) = \left(\int_s \left|\frac{d^2 B_{R_{ij}}}{ds^2}\right| ds\right) / \left(\int_s \left|\frac{dB_{R_{ij}}}{ds}\right| ds\right) \quad (2)$$

In man-made structures, the total curvature tends to be small since such structures include smooth boundaries and/or straight lines. Therefore this constraint prevents the proposed system from identifying natural objects such as trees as buildings.

Compactness: the ratio of the area of a rooftop region over the square of its boundary length. One of the observations made when modeling rooftops is that rooftops mostly have symmetrical shapes with comparable width and length values. The compactness measure encapsulates such characteristics.

$$Co(R_{ij}) = (Area(R_{ij}))/\left(\int_s \left|\frac{dB_{R_{ij}}}{ds}\right| ds\right)^2 \quad (3)$$

where $Area(R_{ij}) = \iint_{x,y \in R_{ij}} dxdy$. The compactness measure prevents narrow regions such as roads and grass lines (low *Compactness*) to be identified as rooftops. Moreover, this constraint prohibits the algorithm from identifying small components of gabled rooftops as individual regions.

Bounding Prevention factor: this indicates the tendency of a region to split into internal sub-regions. This factor also measures the tendency of the regions internal sub-regions in merging with neighboring regions around it. Generally, rooftops include strong edges that could cleanly separate a rooftop from its neighboring regions such as grass, trees, roads or other rooftops with different color/intensities. The n-depth *Bounding Prevention* factor (in our implementation n is set 1) of a region R_{ij} is defined by:

$$BP(R_{ij}) = \frac{1}{n} \sum_{k=1,\dots,n} \left(\frac{\sum_{\forall t,\, R_{ij} \cap R_{i-k,t} \neq \emptyset} \frac{\left(Area(R_{ij} \cap R_{i-k,t})\right)^2}{Area(R_{i-k,t})}}{Area(R_{ij})} \right) \quad (4)$$

This measure estimates the tendency of a region to merge with its outside neighboring regions. Pixels of

each region R_{ij} are weighted by the normalized ratio of the overlap between R_{ij} and $R_{i-k,t}$. When all common sub-regions $R_{i-k,t}$ fall completely inside R_{ij}, the external bounding prevention factor will be its maximum (1).

Dividability factor: this is the tendency of a region to be divided into two or more sub regions. The *Dividability* of a region is high if the average tendency to split for n consecutive image levels is low:

$$Di(R_{ij},n)=\frac{1}{n}\sum_{k=1,\dots,n}\left(E(R_{ij},i-k)+E(R_{ij},i+k)\right) \quad (5)$$

E is the entropy of region R_{ij} at level k:

$$E(R_{ij},k)=-\sum_{\forall t,\,R_{ij}\cap R_{kt}\neq\emptyset}P(R_{ij},R_{kt})log(P(R_{ij},R_{kt})) \quad (6)$$

$$P(R_{ij},R_{kt})=\frac{Area(R_{kt})}{Area(R)} \quad (7)$$

$$R=\{\cup R_{kt}|\forall t,\ R_{ij}\cap R_{kt}\neq\emptyset\} \quad (8)$$

This constraint aims to give priorities to regions that when split maintain a larger part of their initial region.

Considering all above features, the following set of rules measures the relative potential of *rooftopness* for any two regions. The region R_{ij} is more probable to be a rooftop than the region R_{kt} if $RS(R_{ij},R_{kt})$ is $True$. For $RS(R_{ij},R_{kt})$ to be *True*, at least three of the following conditions must be satisfied:

$$Tc(R_{ij})<Tc(R_{kt}), \quad (9)$$

$$Co(R_{ij})>Co(R_{kt}), \quad (10)$$

$$Di(R_{ij},n)<Di(R_{kt},n), \quad (11)$$

$$BP(R_{ij},n)>BP(R_{kt},n), \quad (12)$$

$$Area(R_{ij})>Area(R_{kt}). \quad (13)$$

Note that in the case of equal conditions, the area condition is the deciding one.

2.1.3 Search Tree Creation

$Tree<V,E,Root>$ is defined in the following using regions in segmented images (Figure 2).

$$V=\{R_{ij}|\forall R_{ij}\in S_i\},\quad for\ i=1,\dots,K+1\ and\ j=1,\dots,N_i$$

$$E=\left\{(R_{ij},R_{i+1,x})\middle|\forall R_{ij}\in S_i,\exists R_{i+1,x}\in S_{i+1},x=\max_t(R_{ij}\cap R_{i+1,t})\right\}$$
$$for\ i=1,\dots,K+1\ and\ j=1,\dots,N_i$$
$$Root=R_{K+1,1} \quad (14)$$

In this tree, a node (region) R_{ij} at level *i* is the child of node $R_{i+1,x}$ at level *i+1*, only if $R_{i+1,x}$ has the highest overlap (among all neighboring nodes) with R_{ij}.

2.1.4 Searching for Potential Rooftop Regions

To extract all potential rooftops, a recursive algorithm (Table 1) is utilized. The output of this algorithm is a set of candidate rooftop regions. The search tree algorithm relies on five thresholds ($T_{areamin}, T_{Tc}, T_{Co}, T_{Di}, T_{BP}$) that are set to *(1000, 0.5, 0.01, 0.5, 0.5)*. These values are determined empirically through checking various rooftops in our dataset images.

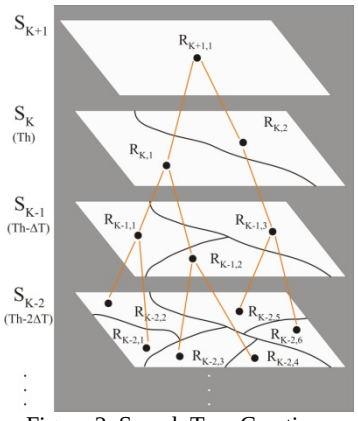

Figure 2. Search Tree Creation.

2.2 Rooftop Verification

A verification step is incorporated in which the existence of shadows is used for verifying candidate regions as rooftop. Shadow regions are extracted on the image by the method in [14]. Using each region's definition, the acquisition geometry, and a fixed height value (2.7 meters), the expected shadows of the region is predicted. If this prediction highly overlaps with the existing shadows, the region is identified as a rooftop.

Table 1. Pseudo code for searching the tree.

```
Rooftops = SearchTree(Tree < V, E, Root >){
 if StoppingCriterion(Root) == True then
    Rooftops = ∅
 else
    R_c = ∅
    for each of the Root'schildren (v_i) do
       R_c = R_c ∪ SearchTree(Tree < V, E, v_i >)
    end
    if MinimumConditions(Root) == True then
       if ∃ v_j ∈ R_c so RS(v_j, Root) == True then
          Rooftops = R_c
       else
          Rooftops = Root
       end
    else
       Rooftops = R_c
    end
 end
}
StoppingCriterion(Root)
 = { True   Root ∈ S_1 or Area(Root) < T_areamin
   { False  Otherwise

MinimumConditions(Root)
 = { True   Root ∉ S_K+1 & Tc(Root) < T_Tc & Co(Root) > T_Co
   {                & Di(Root, n) < T_Di & BP(Root, n) > T_BP
   { False  Otherwise
```

3. Experimental Results

The proposed method is tested on our aerial image dataset images (Pictometry Int. Corp.'s,resolution of 0.15 meter/pixel). Seven random images (due to the limited space) show the typical results from the dataset with a total number of 140 rooftops. In this work, only

one set of parameter values (as detailed in the paper) are used. Figure 3 shows results for four images.

Figure 3. Extracted rooftops in 4 images.

The method is asses quantitatively using 4 metrics:

$$Shape\ Accuracy\ (ShAc) = 1 - |A_G - A_E|/A_G \quad (15)$$

$$Completeness\ (Comp) = TP/(TP + FN) \quad (16)$$

$$Correctness\ (Corr) = TP/(TP + FP) \quad (17)$$

$$Quality\ (Qua) = TP/(TP + FP + FN) \quad (18)$$

A_G is the (manually found) ground truth rooftop's area and A_E the area of the same rooftop detected by this work. TP, FP and FN represent True Positives (no. of correctly extracted pixels), False Positives (no. of incorrectly extracted pixels) and False Negatives (no. of correctly not extracted pixels) of a rooftop.

Table 2 shows the mean quantitative values for each image. These results are better than values reported in recent works by [10] (correctness: 93%, completeness: 88%, quality: 82%) and [5] (shape accuracy: 83.6%).

Table 2. Quantitative results for test samples.

Img. No.	Shape Accuracy	Complete -ness	Correct -ness	Quality	No. of Buildings
1	90.16	99.26	90.37	89.76	14
2	96.08	96.68	93.03	90.16	15
3	99.71	94.96	94.69	90.16	51
4	98.59	95.99	94.67	91.07	9
5	93.37	98.43	92.31	90.96	23
6	92.80	99.53	92.85	92.45	14
7	96.51	95.43	92.21	88.31	14
Overall	**96.24**	**96.65**	**93.25**	**90.34**	140

4. Conclusion

An automatic segmentation based approach was introduced to detect rooftops in aerial images. The method segments input image into potential rooftops using a number of geometrical/regional attributes in multiple layers of segmented images. A search tree is utilized using different segmented image levels. The tree is searched for regions that maximize a set of rooftop definition measures. Candidate regions are verified through shadow evidence. The performance of the method was evaluated using aerial images. Results show the shape accuracy and completeness of 96%.

References

[1] C. Lin and R. Nevatia, "Building detection and description from a single intensity image," Comp. Vis. Image Understanding, vol. 72, no. 2, pp. 101–121, 1998.

[2] M. S. Nosrati and P. Saeedi, "A Novel Approach for Polygonal Rooftop Detection in Satellite/Aerial Imageries," IEEE ICIP 2009, pp. 1709–1712.

[3] X. Jin and C. Davis, "Automated Building Extraction from High-Resolution Satellite Imagery in Urban Areas Using Structural, Contextual, and Spectral Information," EURASIP 2005:14, 2196–2206.

[4] L. B. Theng, "Automatic Building Extraction from Satellite Imagery," International Journal of Computer Science, IAENG, vol 2006:4.

[5] J. Peng, D. Zhang, and Y. Liu, "An improved snake model for building detection from urban aerial images," PRL (Elsevier), 2005, vol. 26, no. 5, pp. 587–595.

[6] L. Wei and V. Prinet, "Building detection from high-resolution satellite image using probability model," IGARSS 2005, vol. 6, pp. 3888- 3891.

[7] H. Y. Li, H. Q. Wang and C. B. Ding, "A New Solution of Automatic Building Extraction in Remote Sensing Images," IGARSS 2006, pp. 3790-3793.

[8] J. Peng and Y. C. Liu, "Model and context-driven building extraction in dense urban aerial images," IJRS vol. 26, no. 7, 2005, pp. 1289-1307(19).

[9] Y. Wei, Z. Zhao, and J. Song, "Urban building extraction from high-resolution satellite panchromatic image using clustering and edge detection," IEEE IGARSS, vol. 3, pp. 2008-2010.

[10] K. Karantzalos, and N. Paragios, "Recognition-Driven Two-Dimensional Competing Priors Toward Automatic and Accurate Building Detection, IEEE IGARSS, vol. 47, issue 1, pp. 133-144.

[11] Y.J. Zhang, "Advances in Image and Video Segmentation," IRM Press 2006 ISBN: 1-59140-753-2.

[12] R. J. Gardner, A. Hobolth, E. B. V. Jensen, and F. B. Sørensen, "Shape discrimination by total curvature, with a view to cancer diagnostics," Journal of Microscopy 217, 1, 49-59.

[13] D. Comaniciu, and P. Meer, "Mean-Shift: A robust approach toward feature space analysis," IEEE Trans. on PAMI, vol.34, pp.1-18, 2002.

[14] M. Izadi and P. Saeedi, "Height Estimation for Buildings in Monocular Satellite/Airborne Images based on Fuzzy Reasoning and Genetic Algorithm," WIAMIS 2009, pp. 1-4.