BUILDING DETECTION AND RECOGNITION FROM HIGH RESOLUTION REMOTELY SENSED IMAGERY

S. Y. Cui [a, *], Q. Yan [a], Z. J. Liu [a], M. Li [b]

[a] Key Laboratory of Mapping from Space of State Bureau of Surveying and Mapping, Chinese Academy of Surveying and Mapping, Beijing 100039, China - gisyong@126.com;
[b] Geoinformation Science and Engineering College, Shandong University of Science and Technology, Qingdao 266510, China - limin82128@163.com

Commission III, WG III/4

KEY WORDS: Building Extraction, Active Contour Model, Hough Transform, Convex Hull, Matrix Search

ABSTRACT:

This paper supposes a schema to deal with the tough task of building detection and recognition from high resolution remotely sensed imagery. It is a region-based and semi-automatic schema combining with Hough transform and computation of convex hull of the pixels contained in the building areas, which can produce a precise result when the contrast between flat building rooftop and the background is high enough. The first step of this strategy is applying seed region grow algorithm to collect pixels contained in the building region to form the approximation shape of building. In order to retrieve the precise shape of building, we devise two approaches, which are based on Hough transform and convex hull computation, to deal with different scenes. Based on the fact that most buildings in real world can be represented by a convex polygon, the first schema uses this idea to compute the shape of the building. The second schema search the desired shape represented by a related orthogonal corner from the node matrix constructed by the dominate line sets of the building. Extraction result shows this schema supposed is robust and applicable to most high resolution remotely sensed imagery.

1. INTRODUCTION AND BACKGROUND

1.1 Introduction

With the successful launch of some high resolution satellites including IKONOS and Quick Bird in recent years, large mount of high resolution remotely sensed imagery can be utilized to extract man-made objects to update for geographic information system database. And man-made object detection and recognition from remotely sensed imagery is also of significant practical importance for mapping, cartography, photo interpretation, military activities and so on. Traditionally, manual plotting is deployed in man-made object extraction, but it is time consuming and expensive, so automatic or semi-automatic acquisition and update of building data is greatly needed, especially after the availability of high resolution satellite imagery such as IKONOS and QuickBird. In the last three decades, a significant amount of work that has been done in the field of aerial image understanding has concentrated on development of efficient algorithm to automatic or semi-automatic detection(at present, semi-automatic methods are applicable in production) and extraction of typical man-made objects, such as building. Consequently, various strategies and methodologies have been brought forward to deal with the tough task of building extraction. In the following section, we briefly review the previous research in this field.

1.2 Previous works

A collection of state of the art articles can be found in the periodical proceedings edited by Grün et al. (1995), Grün et al. (1997) and Baltsavias et al. (2001b). Mayer (1999) presented a comprehensive survey on the techniques used for image based building extraction. Previous research on the building detection and extraction is briefly reviewed as follows. Morhan and Nevatia (1989) used perceptual organization to detect and describe building in aerial images. They recognize the usefulness of the structural relationships made explicit by perceptual organization in complex image understanding. They first detect linear features, which are then grouped into parallels. Parallel collation with aligned endpoints triggers the formation of a U structure. Two U structures trigger the formation of a rectangle hypothesis. A constraint satisfaction network is used to select the best consistent rectangles by minimizing the cost of the network. This kind of approach is usually comprised of a complicated process of bottom–up grouping. Detecting buildings in aerial images is also the goal of Heurtas and Nevatia (1988). The search for rectangle hypotheses is made by local contour tracing techniques. Shadows are used to confirm hypotheses and to estimate the height of buildings. Contour tracing with some structural guidance as oriented corners and depth from shadows has been used in (A. Huertas, R. Mohan and R. Nevatia, 1986). These kinds of methods are often confronted with the issue of fragmentation of edges. Scott Lee and Jie Shan (2003), etc. use the classification result of IKONOS multi-spectral images to provide approximate location and shape for candidate building hypothesis. Then the fine extraction is carried out in the corresponding panchromatic image through segmentation and building squaring based on the Hough transform. Sohn and Downman (2001) used a local Fourier transformation to analyze the dominate orientation in a building cluster and extract rectilinear building outline from IKONOS imagery based on a binary space partitioning tree. Fua and Hanson (1987) segment the scene into regions, find edges

* Corresponding author. This is useful to know for communication with the appropriate person in cases with more than one author.

lying on-region boundaries, and then see if there is evidence of geometric structure among these edges to classify the region as a man-made object.

Most of these surveyed methodologies on building extraction from monocular images can be classified into two categories: edge-driven and region-driven. An edge-driven approach uses an edge map as its starting point usually followed by a line extraction process in order to reduce the numerous spurious and insignificant edges that are found. However the lines including the building-correspondence lines and some irrelevant lines are fragmented and distributed randomly in the scene. A mass of endeavours (V. Venkateswar and R. Chellapa, 1986; A. Huertas and R. Nevatia 1988; R. Mohan and R. Nevatia, 1989; R. Irvin and D. McKeown 1989; Y. Liow and T. Pavlidis, 1990) have been made to link or group the line segments corresponding to the building to obtain desired building boundaries. This is the primary difficult of the edge-driven approach. And usually it needs a complex bottom-up grouping process. In a region-driven building extraction strategy, the source image is initially segmented entirely into different regions. After segmentation, an attempt is made to determine which regions are corresponding to the building component and to combine these building related sub-regions because the exact building region may be segmented into many sub-regions. At present, some task-specified automatic approaches or semi-automatic approaches to address this issue are applicable which are implemented by the manual selection of the building sub-regions to form the building outline. This approach avoids the complex process of bottom-up shape recognition and formation in the edge-based approach in certain extent. The schema supposed in this paper is fall into this kinds of strategy.

This paper is organized as follows: we begin with an overview of our approach in section 2, while in the following section 3 a briefly demonstration of seed region grow algorithm is given. In section 4, we will elaborate the two schema devised to form the precise shape of the building. In the last section, some extraction result and a discussion will be presented.

2. METHOD OVERVIEW

This schema is a region-based approach. Seeded region grow algorithm (Mat-Isa. N. A, 2005) is first applied to collect pixels inside building regions to form approximate shape of building. In order to extract the regular building boundary, two schemas are supposed which can be applied in different scenes. When the building polygon can be represented as a convex polygon, we can calculate the convex hull of the building from the pixels collected in the growing process. When the building polygon is not convex, the boundary of the building can be obtained by boundary fitting on the condition that the contrast between building and background is large enough. If the contrast is low, Hough transform is applied to the region image derived after the growing process. The dominate line sets of the building, which is perpendicular to another, can be retrieved. The intersections of two lines set construct a node matrix. If there are m × n lines in two dominate orientations, the node matrix is m by n. Based on the classification of orthogonal corners; shape of buildings can be represented by a tag sequence which is a series of symbols of related right angle corners. Based on the assumption that building is comprised of orthogonal corners, a filtering of orthogonal corners can be carried out to eliminate some false corners. The fine extraction process is implemented via matrix search algorithm. The building boundary can be precisely obtained. The whole extraction flow is shown in Figure.1.

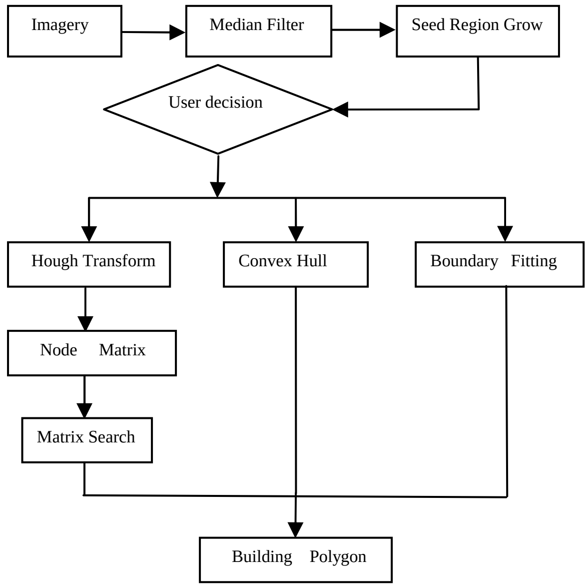


Figure.1 Extraction flow

3. SEED REGION GROW ALGORITHM

In this algorithm, the user needs to determine the region of building by manually selecting the position of the building in the image after filter. And the user also needs to define a threshold.

3.1 Algorithm

1) Manually select the seed points of building region.
2) Chose $N \times N$ neighbourhood of the seed points. Calculate the mean value $\bar{x}$ and the standard deviation σ of the $N \times N$ neighbourhood.
3) Grow the seed points to its neighbour's pixels. Compare the grey level of the seed points with its neighbour pixel. Include the pixel into the region if it satisfy one of the conditions listed below:
    a) If the gradient of the pixels is less than 95% of the equalized histogram and the grey level of the pixels is less or equal to the predefined threshold.
    b) If the gradient of the pixels is more than or equal to 95% of the equalized histogram and the grey level of the pixel is not more than or equal to one standard deviation away from the region mean.
4) Set the neighbour pixel of the seed point, which is added to the region in the previous step.
5) Repeat step 2 to 4 until all the pixels have been considered to be grown or the pixels cannot be grown anymore.

4. FINE EXTRACTION SCHEMA

After the steps above, an approximate region of the building has been derived. In order to extract the precise and regular boundary of the building, the following two schemas are supposed.

4.1 Boundary formation

In high resolution remotely sensed image, the shape of a certain number of buildings is rectangular. For this kind of building, the boundary can be represented as a convex hull. The convex hull of a finite point set S = {P} is the smallest 2D polygon Ω (or polyhedron in 3D) that contains S. That is, there is no other polygon (or polyhedron) Λ with $S \subseteq \Lambda \subset \Omega$. Also, this convex hull has the smallest area and the smallest perimeter of all polygons containing the set S. This idea is explained in the following Figure.2. The black pixels are collected via previous growing approach and the red boundary is the convex polygon of the black pixels. With these pixels, the approximate shape of the building is obvious. The exact boundary of the building can be represented by the convex hull of these pixels.

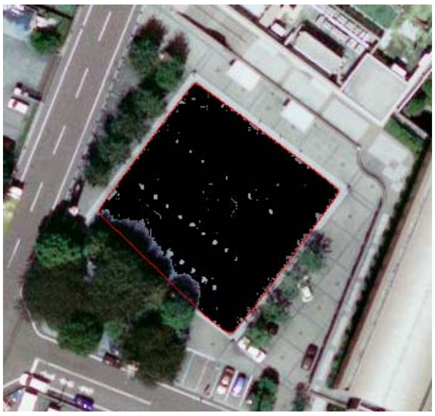


Figure.2 Convex hull of the building pixels

There are various algorithms to compute the convex of point set. Andrew's Monotone Chain Algorithm (A. M. Andrew, 1979) is one of the fast 2D hull algorithms, which is implemented as a stack. We choose it for that it runs in O (n log n) time due to the sort time. First the algorithm sorts the point set **S** = {P_0, P_1... P_{n-1}} by increasing x and then y coordinate values. Let the minimum and maximum x-coordinates are x_{min} and x_{max}. Clearly, $P_0.x = x_{min}$, but there may be other points with this minimum x-coordinate. Let P_{--} be the point in **S** with $P.x = x_{min}$ first and then min y among all such points. Also, let P_{-+} be the point with $P.x = x_{min}$ first and then max y second. Note that $P_{--} = P_{-+}$ when there is a unique x-minimum point. Similarly define P_{+-} and P_{++} as the points with $P.x = x_{max}$ first, and then y min or max second. Again note that $P_{+-} = P_{++}$ when there is a unique x-maximum point. Next, join the lower two points, P_{--} and P_{+-} to define a lower line $\mathbf{L}_{min}$. Also, join the upper two points, P_{-+} and P_{++} to define an upper line $\mathbf{L}_{max}$. These points and lines are shown in the following example diagram.

The algorithm now proceeds to construct a lower convex vertex chain Wmin below $\mathbf{L}_{min}$ and joining the two lower points P-- and P+-; and also an upper convex vertex chain Wmax above $\mathbf{L}_{max}$ and joining the two upper points P++ and P-+ . Then the convex hull W of S is constructed by joining Wmin and Wmax together.

The lower or upper convex chain is constructed using a stack algorithm almost identical to the one used for the Graham scan. For the lower chain, start with P-- on the stack. Then process the points of S in sequence. Only consider points strictly below the lower line $\mathbf{L}_{min}$. Suppose that at any stage, the points on the stack are the convex hull of points below $\mathbf{L}_{min}$ that have already been processed. Now consider the next point P_k that is below $\mathbf{L}_{min}$. If the stack contains only the one point P_{--} then put P_k onto the stack and proceed to the next stage. Otherwise, determine whether P_k is strictly left of the line between the top two points on the stack. If it is, put P_k onto the stack and proceed. If it is not, pop the top point off the stack, and test P_k against the stack again. Continue until P_k gets pushed onto the stack. After this stage, the stack again contains the vertices of the lower hull for the points already considered. The geometric rationale is exactly the same as for the Graham scan. After all points have been processed, push P+- onto the stack to complete the lower convex chain.

The upper convex chain $\mathbf{\Omega}_{max}$ is constructed in an analogous manner, but processes S in decreasing order {P_{n-1}, P_{n-2}... P_0}, starts at P++, and considers only points above $\mathbf{L}_{max}$. Once the two hull chains have been found, it is easy to join them together

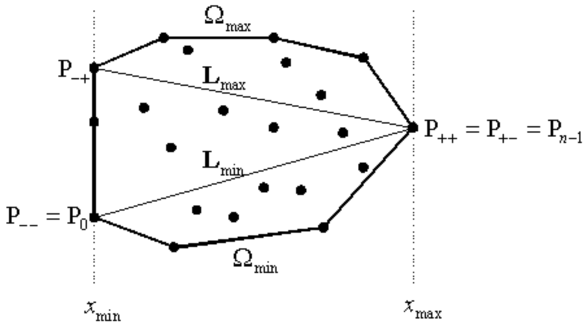


Figure.3 Convex computation algorithm

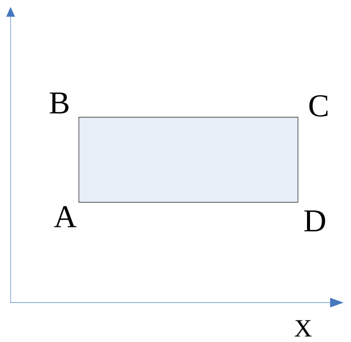
This idea is only applicable when the building shape can be represented by a convex hull and the grow result is well enough. If the building shape is not a convex polygon, the boundary of the building can be calculated by active contour model from the region derived by grown algorithm. Details of ACM algorithm have been demonstrated in the book by Sonka. M (2003). The following figure shows the precise boundary (with blue colour) calculated by ACM algorithm.

Figure.4 Building boundary calculated by ACM

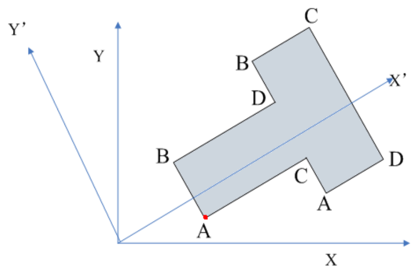
4.2 Building shape representation

In real world, the building boundary is comprised of some straight lines and right angels. The corners whose constructing edges are parallel to the axis in planar coordinate system can be classified into four types which can be labelled ABCD

(Figure.5 (a)). The criterion of classification is the orientation (u v in building model coordinate system, where $v = u + 90^0$) of the corresponding edges constructing the corner. As a result, building shape can be represented by a corner sequence. Base on the classification, the example shape comprised of right angle in Figure.5 (b) can be represented by a tag sequence, which is ABDBCDAC (starting from the red corner in left bottom). In this way, the shape of the buildings in imagery can be expressed fully and readily.



(a)

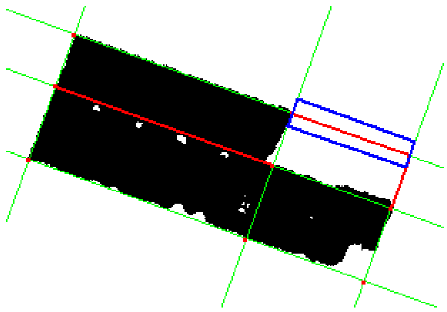


(b)

Figure.5 (a) corner classify (b) shape representation by tag sequence

4.3 Matrix search based on Hough transform

The boundary of building obtained in section 4.1 is not very regular. To obtain the two perpendicular line sets without fragmentation, Hough transform (P. V. C. Hough, 1962) is an ideal alternative as its robustness. Peaks corresponding to these perpendicular line sets in Hough space (ρ , θ) where $0 \le \theta \le \pi$ and $-\sqrt{(m^2+n^2)} \le \rho \le \sqrt{(m^2+n^2)}$ for an image of m by n pixels will have the same value $\theta_1 \theta_2$ (the dominate orientations, $\theta_1 = \theta_2 + 90^0$), causing peaks of two columns aligning vertically in Hough space. All peaks in the two columns surpassing a threshold in Hough space produce two lines set which are perpendicular to another. A node matrix whose elements are intersection nodes of two perpendicular line set can be determined. However, the type of the node can not be determined. We take the following case in Figure.6 as an example to demonstrate our schema to this issue. There are 9 edges which form a 3×3 node matrix. The 3 red lines is not the building edge. In order to remove these false lines, a buffer around each line between two nodes is constructed. We compute the average grey value of the two rectangles in each side of the edge. If the difference of the average grey value exceeds a predefined threshold, we can safely suppose that the edge is not a building edge. By this criterion, some false edges will be removed out. After this process, the corner type can be determined by the criterion in previous section. Obviously, building shape information is contained in the node matrix. To retrieve the building-related node sequence, a search is implemented from all the A type nodes. A search for the next right angle corner in the sequence is carried out by scanning 4-beighbor element in the matrix. In this case, the desired node sequence is ABCACDA (search from the left down corner) which is the aimed tag sequence of the search. However, in the present of significant mount of right angles, it is likely that more than one sequence, which is identical to the required sequence, will be found. The selection of the final tag sequence is based on the similarity between the desired sequences and the searched sequence. The searched sequence with largest similarity may be the same as the desired sequence.



(a)

(b)

Figure.6 (a) Matrix construction (b) search result

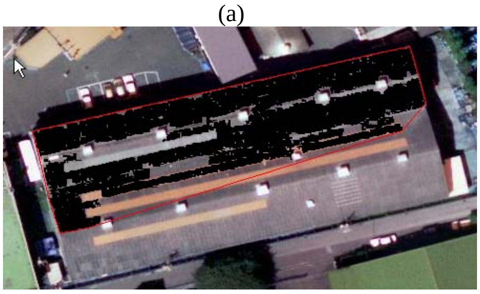
5. RESULTS AND DISCUSSION

In all these experiments, processing was carried out in the Visual C++ 6.0 environment using real world imagery. In addition to these experiments, complexity analysis and conclusion is also discussed in this section.

5.1 Experiment result

To improve the capability of grow algorithm, median filter is first allied to the image. The pixel value of the seed point which is manually selected by the user needs to be the same as the average of all the pixels in the building region. If the pixel value of the seed point is large than the average, the region obtain via the grow algorithm may be not very precise. In addition, the threshold in the grow algorithm is very important. If it is very large or small, the growing result may be not the same as the building region. This is depicted in Figure.7.

(a)



(b)

Figure.7 (a) threshold is 8 (b) threshold is 5

(a)

(b)

Figure.8 (a) (b) Extraction results

5.2 Conclusion

In this paper, we suppose a robust and semi-automatic approach to deal with building extraction. This approach is not restricted by the shape of the building. It can precisely extract the boundary of rectangular building with homogeneous flat rooftop with a very low computation complexity which is important in data production. The second schema (matrix search) supposed introduce a new method to represent shape of rectangle or combination of rectangle. Pivotal reason is that it depends on the robustness of Hough transform and utilizes a new mathematic model to represent shape. However, it has some disadvantages; for example, it heavily depends on the approximate shape derived via grown algorithm. If the grown result is not satisfactory, the extraction may be failed in a large extent. Several aspects of the proposed scheme need further research. Adaptive region grow approach needs to be explored to improve the detection of the building approximate shape. Some other characteristics and processes should be incorporated into this scheme such as image segmentation. In the future research, these issues should be addressed.

ACKNOWLEDGEMENT

This work is carried out under the Project for Young Scientist Fund sponsored by the National Natural Science Foundations of China (40401037) and National Key Basic Research and Development Program of China (2006CB701303).The author would like to thank Dr Xiangguo, L., and Na, J., for their inspiration and encouragement.

REFERENCE

Grün, A., Kuebler, O., and Agouris, P., 1995. *Automatic Extraction of Man-made Objects from Aerial Space Image (I)*, Birkhaeuser Verlag, Berlin, Germany, pp. 321.

Grün, A., Baltsavias, E., and O, Henricsson., 1997. *Automatic Extraction of Man-made Objects from Aerial Space Image (II)*, Birkhaeuser Verlag, Berlin, Germany, pp. 393.

Baltsavias, E., Grün, A., and L, V, Gool., 2001b. *Automatic Extraction of Man-made Objects from Aerial Space Image (III)*, A. A. Balkema Publishers, Lisse, The Netherlands, pp. 415

Mayer, H., 1999. Automatic object extraction from aerial imagery——a survey focusing on buildings, *Computer Vision and Image Understanding*, 74(2), pp. 138 - 149.

Mohan, R., Nevatia. R., 1989. Using perceptual organization to extract 3-D structures, *IEEE Trans, Pattern Analysis and Machine Intelligence*, 11, pp. 1121-1139.

Huertas, A., Nevatia, R., 1988. Detecting buildings in aerial images, *Computer Vision, Graphics, and Image Process,* 41, pp.131-152.

Huertas, A., Mohan, R., Nevatia, R., 1986. Detecting of complex buildings in simple scenes, *Inst. Robotics and Intelligent Systems*, University of South California, Tech. Rep. IRIS 203.

Lee, D. S., Shan, J., and James. S. B., 2003. Class-guide building extraction from IKONOS imagery, *Photogrammetric Engineering& Remote Sensing*, 69(2), pp. 143-150.

Sohn, G., Dowman, I. J., 2001. Extraction of building from high resolution satellite data, *Automatic Extraction of Man-made Objects from Aerial Space Image (III),* A. A. Balkema Publishers, Lisse, The Netherlands. Pp.345-355.

Fua, P., Hanson. A. J., 1987. Using generic geometric models for intelligent shape extraction, *Proc DARPA Image Understanding Workshop*, Los Angles.

Venkateswar. V., Chellappa. R., 1990. A framework for interpretation of aerial images, *Proceeding of the International Conference on Pattern Recognition*, Atlantic City, NJ, pp. 204-206.

Irvin. R., McKeown D., 1989. Methods for exploiting the relationship between buildings and their shadows in aerial imagery, *IEEE-SMC*, 19 (6), pp. 1564 - 1575.

Liow, Y., Pavlidis. T., 1990. Use of shadows for extracting building in aerial images. *Computer Vision, Graphics and Image Processing*, 49, pp. 242-277.

Mat-Isa. N. A., Mashor. M. Y., Othman. N. H., 2005. Seed region growing feature extraction algorithm: its potential use in improving screening for cervical cancer, *International Journal of The Computer, the Internet and Management*, Vol. 13, No.1, pp, 61-70.

Andrew. A. M., 1979. Another Efficient Algorithm for Convex Hulls in Two Dimensions. *Info. Proc. Letters* 9, pp. 216-219.

Sonka. M., Hlavac. V., Boyle. R., Image processing, analysis, and machine vision.

Hough. P. V. C., 1962. Methods and means for recognition complex pattern, U. S. Patent, 3,069,654