BUILDING BOUNDARY EXTRACTION BASED ON LIDAR POINT CLOUDS DATA

SHEN Wei

Ocean College of Shanghai Fisheries University, Shanghai 200090, China - shen_wei@sina.com

Commission III, WG III

KEY WORDS: LIDAR, Laser Scanning, Point Clouds, Algorithms, Building, Extraction, Processing

ABSTRACT:

LIDAR (Light Detection And Ranging) is an active remote sensing system, which could provide fleetly three dimensional information of earth surface with high vertical accuracy. Now the building boundary extraction and normalization is the key approach for 3D modeling of building and city mapping. In this paper, a new algorithm named Alpha Shapes is developed to extract the building boundary. Compared with other algorithms, Alpha Shapes algorithm works effectively in inner and outer boundaries extraction from point clouds data with convex and concave polygon shape. Moreover it can keep fine features of buildings adaptively and filter the footprints of non-building. In addition, an improved boundary simplifying algorithm is suggested to refine the extracted building boundary. And two regularization algorithms are developed to make the refined boundary regular. The experiments proved that the normalized building boundary is generated perfectly by these algorithms.

1 INTRODUCTION

1.1 Building Boundary Extraction using the LIDAR data

LIDAR (Light Detection And Ranging) is an active remote sensing system, which utilizes laser beam for detection and measurement. Such system is also known as Laser Radar or Airborne Laser Scanning (ALS) (Aloysius Wehr, 1999). The advantages of using LIDAR for 3D Urban data capturing are high speed, high density, high vertical accuracy and low cost against traditional photogrammetry. LIDAR system is becoming a revolution, which can be widely applied in the feature extraction and the 3D reconstruction of terrain, building, tree, road, power line etc..

Building is the main element to form a city. Building boundary is the fundament of urban mapping and 3D building modelling. Now, the research of Building Boundary Extraction using LIDAR data is one popular and difficult topic in LIDAR data processing. The LIDAR data can be characterised as sub-randomly distributed 3D point clouds which may contain more information than a 2.5D surface model. But, the random of point clouds data and the complexity & diversity of building lead to the difficulties of the building boundary extracting.

This research focuses on the building boundary extraction and regularization. The data used here is building point clouds which has been extracted from raw LIDAR data by preprocessing. In this paper, “Alpha Shapes algorithm” is developed to extract the building boundary. In addition, an enhanced boundary simplifying algorithm and two other developed Regularization algorithms are used to refine the extracted boundary. Finally, the normalized building boundary is generated perfectly with the series of algorithms.

1.2 Existing Algorithms

Normally, the LIDAR point clouds are rasterized into gray level DSM. The generated DSM is furthered processed to obtain nDSM(normalized DSM). Image processing methods such as image classification or boundary line extracting integrating high resolution images are utilized to extract building outline (Ma Ruijin, 2004; Li Shukai, 2000; Sohn G., 2003).

In Ruijin Ma (2004) doctoral thesis, the first step of boundary extraction was performing image segmentation towards nDSM raster imagery based on the elevation threshold to eliminate the unwanted features. Then, some filters were applied to obtain the boundaries in vector format by using the off-the-shelf software, like ERDAS IMAGINE software.

Li Shukai (2000), described the strategy for building boundary extraction from DSM and satellite imagery. The steps were laser ranging analysis, shadow analysis and reconstruction of building boundary. Sohn G. (2003) performed image fusion as the preprocessing part of the research. Then, edge detection was performed from the imagery. Finally, Burns algorithm (Burns, J, 1986) was applied to extract straight lines which belong to building boundary.

Liang Xinlian (2005) implemented the building boundary extraction by applying the Smallest Deviation Approximation algorithm to LIDAR data. Firstly, the two farthest points within the building area were fond, then the building area was separated into two small parts. Then the process will be continued to separate again until the distance value is smaller than the threshold value.

You Hongjian et al. (2001) developed an extraction algorithm based on predicting the direction of azimuth. His method is to separate the edging points of the nearest building boundary continually. At the same time, by matching the edge's lining and the changes of direction angles to get the grouping of edging points. The grouping results will be then used to calculate the main direction of the building, which its results will be used in edge planning and eventually obtaining the furthest external edge of a building.

Currently, the researches in extracting the building boundary using LIDAR point clouds are not abundant so far. Comparing with the extraction of convex polygon outline, the researches of concave polygons outline extraction were rare and its end results were unsatisfied. Currently, the general extraction methods include Smallest Perimeter Method Based on The Constriction, Line Approximation Method Based on The Grouping of Smallest Variance and Line Approximation Method Based on The Distribution of Variance. All these methods require iterative approximation and tremendous calculation processes. Therefore, it's not an ideal application when the data is huge.

Aiming at the current research actuality and requirements, this study developed a better and efficiently extraction algorithm for the inner and outer boundaries of concave polygons and convex polygons buildings, i.e. “Alpha Shapes Algorithm”.

2 ALPHA SHAPES ALGORITHM

2.1 Algorithm Definition

The extracted point clouds belonging to one building can be represented by a set S of points in 3D or 2D. If we'd like to have question like ``the shape formed by these points.'' , there are probably many possible interpretations, the alpha shape being one of them. α-shape can be used for shape reconstruction from a dense unorganized set of data points. Indeed, an α-shape is demarcated by a frontier, which is a linear approximation of the original shape (Bernardini, F. , 1997.). So Alpha Shapes can be used to perform the boundary reconstruction from the irregular point clouds. The set S of points has an α-shape in polygon (as depicted in Figure 1). This polygon is determined by the S and α. α-shape directly illustrates the S shape, at the same time, the parameter α controls the precision of the boundary.

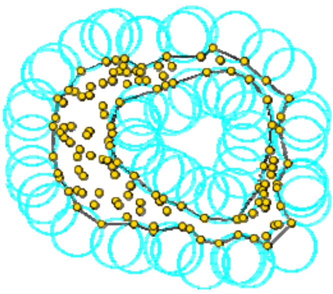


Fig.1 Alpha Shapes algorithm extracting principle

Figure 1 is an example for Alpha Shapes Algorithm. Imagine that a circle with a radius α is rolling around the point-set S. When α value is bigger than a threshold, the circle won’t fall into the area of S The rolling track will form the boundary of this point-set S When the α value is approaching infinity ($\alpha \to \infty$),α-shape must be the convex hull (Nataraj Akkiraju, et al., 1999). On the other hand, when the α value is very small ($\alpha \to 0$), every point might be the boundary. When the point-set S containing evenly distributed points and α approaching an optimal value, the α-shape can extract the inner and outer outlines of the polygon at the same time, as shown in Figure 1. The earliest Alpha Shape concept was been applied to 2D space and now it has been developed into 3D space lately.

2.2 Algorithm Implement

2.2.1 Judgement Condition: In a point-set S which formed by n points, these points can link to form n×(n-1) lines. Our question is which lines are the boundary lines. Alpha Shapes Algorithm is able to give a perfect result.

The judgement condition can be described as followings:
In the point-set S, draw a circle with the radius of α and 2 points (P_1 & P_2) randomly as depicted in Figure 2. If there are no other points within the circle, then both P_1 and P_2 will be defined as boundary points automatically, and the 2 points should be connected to obtain one boundary line.

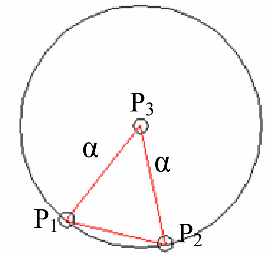


Fig.2 Distance Intersection Algorithm

In this algorithm, the circle center P_3 (X_3, Y_3) need to be calculated based on $P_1(X_1,Y_1)$ & $P_2(X_2,Y_2)$ and the radius α. Actually, this calculation is implemented based on the Distance Intersection Algorithm.

Distance Intersection Algorithm (Zhang Hong, 2006):

$$\begin{cases} x_3 = x_1 + \dfrac{1}{2}(x_2 - x_1) + H(y_2 - y_1) \\ y_3 = y_1 + \dfrac{1}{2}(y_2 - y_1) + H(x_1 - x_2) \end{cases} \quad (1)$$

where

$$H = \sqrt{\frac{a^2}{S^2_{P1P2}} - \frac{1}{4}}$$

$$S^2_{P1P2} = (x_1 - x_2)^2 + (y_1 - y_2)^2, a = \text{ radius}$$

When the circle center P_3 obtained, if the distance between any point of S and P_3 less than α , then there is no other point within the circle.

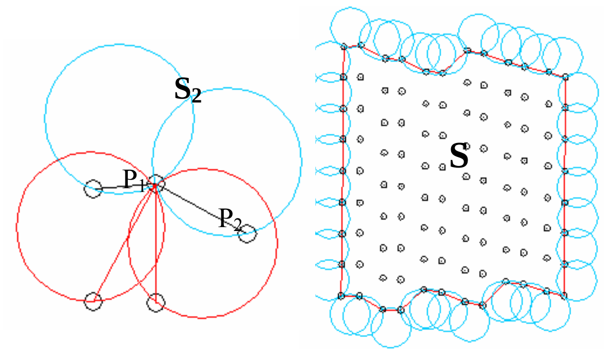
2.2.2 Algorithm Flow: The algorithm can be described as followings:

1. Begin with any point P_1 within point-set S, choose points which distance to P_1 less than 2×α to form a new point-set S_2 (Figure3-a). Within S_2, pick any point P_2 and calculate the circle center P_0 with P_1, P_1 and α.
2. Within S_2, calculate the distance L between every point and P_0:

   ①If $L_{i(i=1-n)} \geq a$, then points P_1, P_2 are boundary points and segment P_1P_2 is boundary line segment.

   ②If $L_{i}\angle a$, then interrupt and turn on step 3.

3. Within S_2, choose next point and repeat step 1 & step 2 until every point Within S_2 finished.
4. Within S, choose next point and repeat step 1 - 3 until every point Within S finished.



(a) (b)

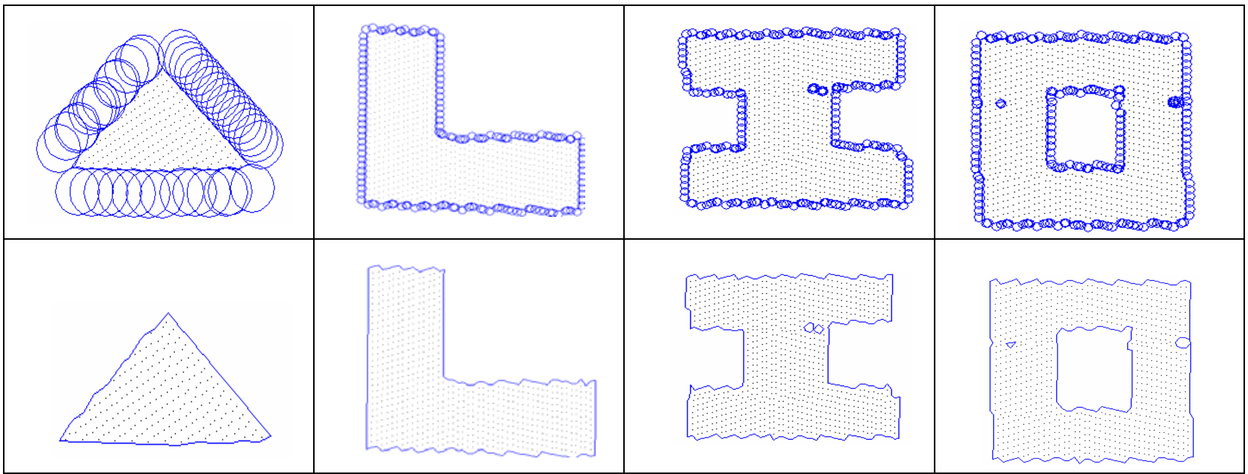
Fig.3 Alpha Shapes Algorithm Judgement and Result

When the flow above finished, all the boundary lines will be extracted, Figure3-b.

2.3 Algorithm Test

This research tested various type of building outline in the city center of Malaysia. The LIDAR data are the first return beam using OPTECH sensor in year 2002. The average density of point clouds is 6points/m^2 .The horizontal accuracy is 15cm..
The extracted result of various type of building boundary using Alpha Shapes algorithm was shown in Figure 4 below.

Alpha Shapes algorithm can be used for building outline extraction towards both convex and concave polygons. The value of α can be used to determine or control the smoothness of the building outline. For convex polygon, the α value can be larger (even $\alpha \to \infty$) for a smoother building outline. (As shown in Figure 4-a, α=10) . However, for concave polygon, the corner and internal of polygon need to take into consideration when deciding the α value. So, the α value should be great than average points distance $\bar{L}$ and less than $\bar{L}\times 2$.(As shown in Figure 4 b-d, $\bar{L}$ =1, α =1.8)



(a) Triangle	(b) L-shape	（c） I-shape	（d） Square doughnut

Fig.4 The extracting result of different building boundary with Alpha Shapes algorithm

3 BOUNDARY SIMPLIFICATION AND REGULARIZATION

The outline obtained from Alpha Shapes algorithm above is very rough which can be defined as raw outline. Raw outline usually is composed of zigzag shape along the derived outline as shown in Figure 3 & 4. In order to get regularized boundaries, the raw outline simplification is necessary. The main work is to retrieve and retaine the inflexion points while eliminate the intermediate points.

3.1 Improved Pipe Algorithm

In this research, the Improved Pipe Algorithm originally proposed by Zhao (2001) is developed to perform the simplification work.The Pipe algorithm retrieves the polygon inflexion points by judging the direction changes of pipe. A pipe with diameter d has been used to interpret the changes of boundary as shown in Figure 5. If different between the points on the boundary line is less than d , it means there is no essential changes on the boundary where the intermediate points can be removed. However, if the different between the points on the boundary is more than d , it indicates that the point is a inflexion point and should to be remained. The advantage of this algorithm is the building boundary points can be processed in flexible way where only related points have been taking into calculation. The algorithm not only considers the azimuth changes but considers the overall changes of the building shape. So, the azimuth changes of line didn't affected the overall boundary, however it help to pick up the inflexion points easily.

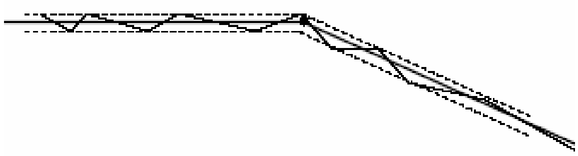


Fig.5 Pipe algorithm simplifying principle

These inflexion and intermediate points of a segment could be used to derive parameters of the segment using a least square regression algorithm.

3.1.1 Algorithm Principle and Realization: As depicted in Figure 6, the 2 parallel lines compose a pipe with diameter d . Now we need to place the triangle (P_1 P_2 P_3) into the pipe as shown in Figure 6. Begining with the segment P_1P_2 , at P_2 a vertical assistant line with length d is drawed. So, the triangle can be placed into the pipe if the highness is less than d otherwise it will be failed to fit in.

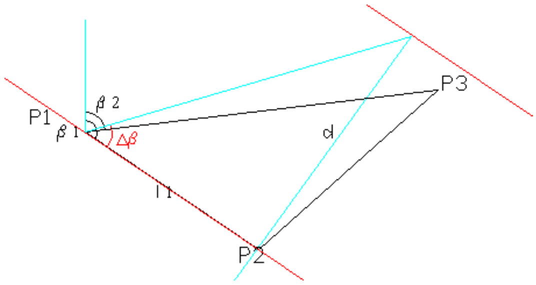


Fig.6 Pipe algorithm geometry principle

Based on the understanding above, starting from the first two points P_1 & P_2, the direction β1 and length l_1 of the line connecting these two points are calculated. Perpendicular to the current line segment, the assistant line length d is draw and $\Delta\beta=\arctan(d/l_1)$ is calculated. Next point P_3 is connected with the starting point P_1 to form a new segment, and its direction β_2 and length l_2 are calculated. Then the condition $\beta_2 \in (\beta_1 \pm \Delta\beta_1)$ is judged.

①If it is ture, consider P_2 is not inflexion point, then the segment P_1P_3 becomes the star segment and judges with next P_4. Now the azimuth range is $[\max(\beta_1-\Delta\beta_1, \beta_2-\Delta\beta_2), \min(\beta_1+\Delta\beta_1, \beta_2+\Delta\beta_2)]$, return last step;

②If it is fault, consider P_2 is inflexion point, then the segment P_2P_3 becomes the star segment and repeat steps above until the last point.

In this method, the diameter d is the only parameter and threshold which can judge the departure distance between point and line segment. Here, $d = a \times 2$. Because the length of any line segment within Alpha Shapes algrithm is less than $a \times 2$, i.e. $P_iP_{i+1} < a \times 2$.

3.2 Boundary Regularization

The remained inflexion points were thus simplified and formed a basic framework of the polygon. But the framework is not regular. This study defined two regularization algorithms which can be applied for four sided and multi-sided polygon.

3.2.1 Circumcircle Regularization Algorithm: The objective of this method is to force the irregular quadrangle to be a rectangle. The geometry principle is that the angle composed by the diameter and the point in the circle is a right angle, as shown in Figure 7.

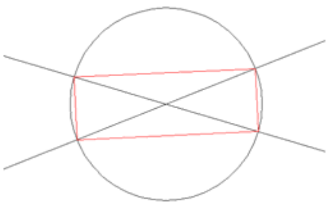


Fig.7 Circumcircle regularization algorithms principle

Based on the understanding above, the algorithm can be described as followings:

1. The intersection of two diagonal lines of the quadrangle is identified firstly. This intersection point works as the geometrical center point of the polygon and the circumcircle.
2. The furthest distance of the intersection point with the four quadrangle nodes is calculated as radius (or the average distance as radius) to draw a circumcircle.
3. Four intersection points are identified between the circle and the two diagonal lines and linking up the four points will result in a new rectangle.

The example of the circumcircle regularization algorithm is shown as in Figure 8.

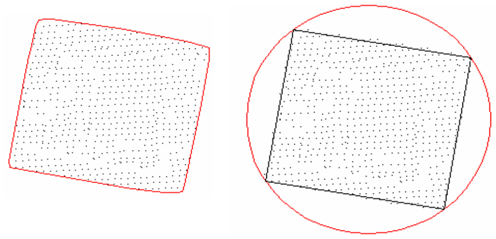


Fig.8 Circumcircle regularization algorithms example

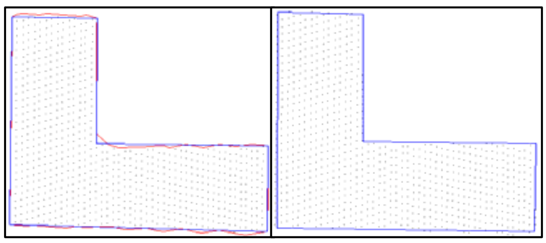
3.2.2 Clustering and Adjustment Algorithm: Aiming at the multi-sided polygon, the Clustering and Adjustment Algorithm is developed in this paper. This method includes clustering and adjustments. The procedures are as follow:

1. Find the azimuth angle of every outline and separate the boundary lines into type A & B based on the azimuth angle (Type A as building main direction and B for adjustment)
2. Calculate the average azimuth angle $\bar{a}$ for type A and B, Weighted average method can be applied to increase the calculation accuracy with length of boundary defined as weight. The formula is:

$$\bar{a} = \sum_{i=1}^{n} (ai \times li)/n$$

3. The average azimuth angle of type A will remain while the average azimuth angle of type B will be adjusted based on type A and obtain a new type B average azimuth angle.
4. Adjust every boundary line by using the middle point as primary axis until they fulfill the new average azimuth angle. The intersection nodes are then identified.
5. These nodes are linked up to obtain an Regularized building outline as shown in Figure

The algorithm will produce an optimized outline of a convex polygon as shown in Figure 9.



(a) Raw outline (b) Optimized Outline

Fig.9 cluster and adjustment algorithms example

Because of the limitation of algorithm, this method just can process polygon which sides number greater than four and must be even.

In addition, the asymmetry & abnormity of LIDAR scattered data and incomplete data all can result in errors. So some time, the manually supervising is needed.

4 CONCLUSIONS AND DISCUSSIONS

4.1 Advantages of Alpha Shapes Algorithm

In this paper, Alpha Shapes Algorithm is firstly applied in building boundary extraction based on LIDAR point clouds data. Alpha Shapes has lots of advantages in LIDAR data processing as described followings:

1. Both convex polygon and concave polygon are applicable. Especially the algorithm can extract the boundary of concave polygon perfectly.
2. High processing speed, avoiding constructing TIN, especially fitting outlines extraction based on scattered data.
3. Both inner and outer boundaries can be extracted by this algorithm, such as dooryard and stadium.
4. Adaptive function can extract the boundary of different shape and size by adjusting α value.
5. Filtering function. Because the reasons of algorithms and LIDAR data, lots of non-building points left in the building data. Alpha Shapes can remove noise and some trees by judging the polygon's size & side number, and points number within the boundary. At the same time, some holes generated from the asymmetric data in the polygon can be removed.
6.

4.2 Accuracy Assessment

Although there are many researches on 3D LIDAR point clouds data processing, however there are very little studies about the accuracy of the extraction and modeling. Some researchers pointed that (Zhang xiaohong, 2006) the accuracy assessment should take the matching of the reconstructed model and raw point clouds data into consideration, which means the fitting accuracy. And not comparing the model with the actual building. This is an accuracy which reflects and focuses on the 3D model reconstruction. Since LIDAR data itself contains some natural limitations, the errors caused by the data itself sometimes might be bigger than the errors caused by the processing methods.

This study utilizes very high resolution satellite imagery (Quickbird) and ground surveying data as main reference to conduct accuracy assessment on the building boundary. The tested data shown that the building boundary extraction and regularization algorithms mentioned in this study can well guarantee the boundary accuracy. The difference between the reconstructed boundary and actual measurement is normally less than 0.5 m. The error caused by the algorithms is proportional to the complexity of the building shape which is normally lesser than the LIDAR data error itself (LIDAR system error and data limitation).

The experiments proved that the algorithms mentioned in this study are effective and very practical in building boundary extraction and regularization. These methods provide a good solution for LIDAR data processing and 3D urban model rebuilding.

REFERENCE

Aloysius Wehr, Uwe Lohr, 1999. Airborne laser scanning—an introduction and overview, ISPRS Journal of Photogrammetry & Remote Sensing Vol.54 , pp. 68-82.

Bernardini F. and Bajaj C. , 1997. Sampling and reconstructing manifolds using alpha-shapes. Technical Report CSD-TR-97-013, Dept. Comput. Sci., Purdue Univ., West Lafayette, IN,.

Burns, J.B., Hanson, A.R. and Riseman, E.M., 1986. Extracting straight lines. IEEE Pattern Analysis and Machine Intelligence, Vol. 8(4),pp.425-455.

Li Shuka, Xue Yongqi , 2000. Efficiently 3D Remote Sensing Integration System. Science Press, Beijing, pp.171-176.

Liang Xinlian, 2005. Airborne LIDAR Data Filtering and Building Model Reconstruction. Dissertation, Chinese Academ y of Surveying and Maping, Beijing, China.

Ma Ruijin, 2004. Building Model Reconstruction from LIDAR Data and Aerial Photographs. Dissertation, The Ohio State University, USA.

Nataraj Akkiraju, et al. Alpha Shapes: Definition and Software http://www.geom.uiuc.edu/software/cglist/ GeomDir/ shapes95def/index.html, 1999.

Sohn G., Dowman I. , 2003. Building Extraction using LiDAR DEMs and IKONOS Images. Proceedings of the ISPRS working group III/3 workshop "3-D reconstruction from airborne laser scanner and InSAR data". Dresden, Germany.

You Hongjian, Li Shuka. ,2001. Extracting Man-made Building from Sparse Laser Sample Point. Journal of the Graduate school of the Chinese Academy of Science, Vol. 18(2), pp. 154-158.

Zhang Hong et al, 2006. GIS Algorithm Elements. Science Press, Beijing, pp.41-43.

Zhang xiaohong et al. 2006. Building Reconstruction from Airborne Laser Altimetry Points Cloud Data Set Based on Invariant Moments Geomatics and Information Science of Wuhan University. Vol. 31, pp. 168–171.

Zhao Zhiyuan, 2002. Line Simplification. http://www-cg-hci.informatik.nioldenburg.de/~da/peters/Kalvin/Doku-CG.htm (accessed 18 Sep. 2006)