IOP Conf. Series: Earth and Environmental Science **18** (2014) 012028 doi:10.1088/1755-1315/18/1/012028

Comparisons of adaptive TIN modelling filtering method and threshold segmentation filtering method of LiDAR point cloud

Lin Chen[1, 2], Xiangtao Fan[1, 3], Xiaoping Du[1]
[1]Key Laboratory of Digital Earth, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, No.9 Dengzhuang South Road, Haidian District, Beijing 100094, China
[2]University of China Academy of Sciences, No.19A Yuquanlu, Beijing 100049, China

E-mail: xtfan@ceode.ac.cn

Abstract. Point cloud filtering is the basic and key step in LiDAR data processing. Adaptive Triangle Irregular Network Modelling (ATINM) algorithm and Threshold Segmentation on Elevation Statistics (TSES) algorithm are among the mature algorithms. However, few researches concentrate on the parameter selections of ATINM and the iteration condition of TSES, which can greatly affect the filtering results. First the paper presents these two key problems under two different terrain environments. For a flat area, small height parameter and angle parameter perform well and for areas with complex feature changes, large height parameter and angle parameter perform well. One-time segmentation is enough for flat areas, and repeated segmentations are essential for complex areas. Then the paper makes comparisons and analyses of the results by these two methods. ATINM has a larger I error in both two data sets as it sometimes removes excessive points. TSES has a larger II error in both two data sets as it ignores topological relations between points. ATINM performs well even with a large region and a dramatic topology while TSES is more suitable for small region with flat topology. Different parameters and iterations can cause relative large filtering differences.

1. Introduction

Automated stereo image matching captures good Digital Terrain Model (DTM) for open terrain, but severe problems occur for regions, like forest areas, wetland, coastal areas and build-up areas [1]. Airborne laser scanning represents a new and independent technology for highly automated generation of DTM and surface models [2]. Research has demonstrated that LiDAR DTM generation is more efficient and accurate as compared to traditional methods [3]. With the progress of precise kinematic positioning of differential GPS and inertial attitude determination of IMU in the 1980s, airborne laser scanning reaches overall vertical system accuracy in the decimeter order. As most technical hardware difficulties and system integration problems have been solved, what much remains is the development of algorithms and method for interpretation and modeling of laser scanner data [4].

In past decade, there have been many filtering algorithms. Some rasterizes the cloud points into height image and processes the image with remote sensing image processing methods. Some build up assumptions and processes the 3D point clouds based on their spatial relationships. Rasterized height image is easier to handle but reduces the precision. The mainly used assumptions are: (1) non-ground points have higher elevations than ground points; (2) slopes in an area don't change dramatically [5].

In the paper, Adaptive Triangle Irregular Network Modelling (ATINM) algorithm and Threshold Segmentation on Elevation Statistics (TSES) algorithm were used. Though these methods had been

[3] To whom any correspondence should be addressed.

widely used, some key details were not clearly explained. In this paper, Gross errors were removed before implementing these algorithms. Parameter selections of ATINM and iterative elevation statistics of TSES were emphasized in details.

2. Methods

2.1. ATINM

ATINM method was first put forward by Axelsson [4, 6, 7], which has been implemented in TerraScan software. It is sensitive to parameter selections.

2.1.1. Removing low error points. There are absolute gross error points made by instrument system. The limited points have a significant effect on the initial TIN construction. For each point, search the points within a defined distance and compare the elevation differences. If current point is obviously lower than the points around, it will be removed as gross error.

2.1.2. Selecting initial points. The whole study area is divided into defined grid, the size of which is based on the largest size of buildings in the area. Then the lowest point in each grid is selected as the seed point. All initial seed points are used to form TIN, e.g. the initial terrain.

2.1.3. Parameters setting. Two parameters, maximum distance to the TIN facet and maximum angel to the nodes are needed to be set. These two parameter thresholds change with the terrain slopes and surface features. Regions with steep slopes and complex artificial constructions need a larger parameter threshold. In original algorithm, these two parameters are derived from data in the form of median values of surface normal angles and elevation differences. However, in most cases, they are estimated by experience and by repeated adjustments.

2.1.4. Densification of the TIN. For each iteration, the nearest TIN facet is searched for each point of the unclassified group. Then the distance from current point to the TIN facet and the angel to the nodes of the triangle are calculated. The points which meet the parameter thresholds are classified as ground points and are added into the TIN for the next iteration. The iteration will continue until no points are added into the TIN.

2.2. Single threshold segmentation

In LiDAR point cloud, different features in local area have different elevation distributions. Cloud points to be extracted are considered as object, other cloud points as background. Absolute elevation difference can be well showed in the elevation histogram [8]. Both object and background points form peaks in the histogram. Spatial statistics is brought in to find the valley value accurately and separate object and background points.

2.2.1. Removing error points. Very high and very low points which have great elevation differences, are mixed with other LiDAR points. These points should be searched and removed before carrying out elevation statistics. Similar procedures are described in ATINM above.

2.2.2. Elevation statistics and segmentation. First, find the lowest and highest elevations of all cloud points, and set the mean value as the initial elevation threshold. Second, calculate the mathematical expectations of both object and background points divided by the threshold, and set the mean value of the expectations of the newly segmented object and background points as new elevation threshold. Third, continue the iterations of last step until the threshold changes within the defined range. The final elevation threshold divides the whole LiDAR data into the final object and background points.

3. Data and experiments

Samp21 and Samp41 provided by ISPRS, where the ATINM performed not very well, were used. Reference data handled semi-automatically and semi-manually are also available.

3.1. Data set 1

The terrain is relatively flat and there is a narrow bridge protruding in the study area which is the key feature to be removed from ground points.

3.1.1. Threshold segmentation. The elevation histogram on the left in figure 1 shows that very high points exist, which can be directly classified as unground points. After removing these points, other unground points show smooth and small distributions at higher elevations mainly between 292 and 295 as shown in the right graph of figure1. One-time segmentation is enough to separate ground points.

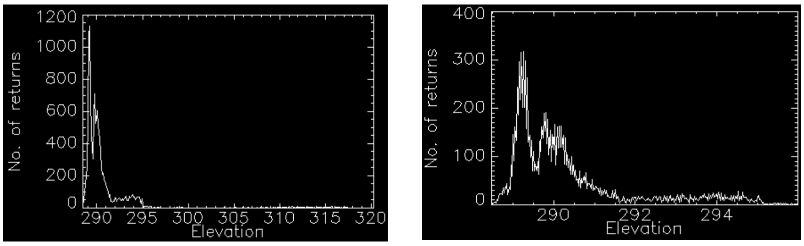

Figure 1. Elevation histograms before and after removing very high points

As the threshold changed within the defined range 0.01 after the fourth iteration, the final threshold was set as 291.67. Four iteration calculations to determine the threshold are shown in table 1. The quantitative comparisons are shown in table 3.

Table 1. **Iterative calculations of elevation threshold.**

Threshold	Object Expectation	Background Expectation
292.23	289.86	293.77
291.81	289.82	293.57
291.69	289.81	293.53
291.67	289.81	293.52

3.1.2. Adaptive TIN. ATINM is a bottom-up searching method, which is unnecessary to remove very high points. To get suitable parameters, set the elevation threshold an empirical constant value, and count the errors changing with the angle threshold. Then set the elevation threshold another empirical constant value, and do the same statistics. For this data, 1.5m, 1.0m, and 0.6m were picked as the elevation threshold. Figure 2 shows extracted numbers of ground and unground points changing with the angles. Figure 3 shows type I and type II errors changing with the angles.

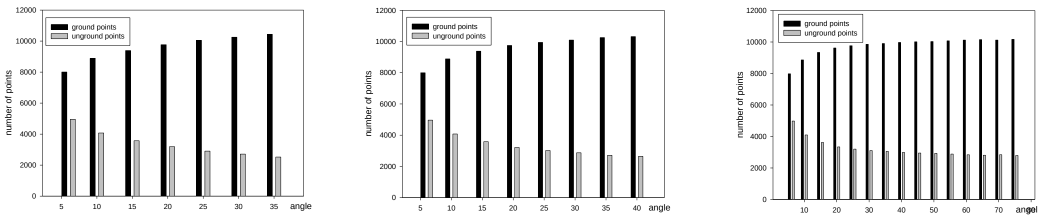

Figure 2. Numbers of ground and unground points under the elevation threshold of 1.5m, 1m, 0.6m

As the angle increased, the number of ground points increased and the number of the unground points decreased. When the elevation threshold was set 0.6m, both the two numbers kept stable even under big angle threshold as shown in figure 2.

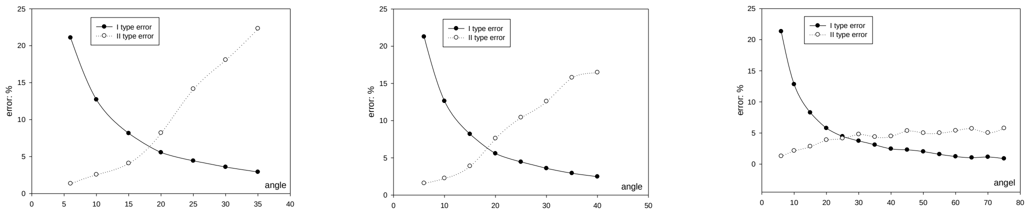

Figure 3. Type I errors and type II errors under the elevation threshold of 1.5m, 1m, 0.6m

The error trends at the height of 1.5m and 1.0m were similar. When the type I error was lowered, the type II error increased quickly. The error trend at 0.6m was smoother, and both the two type errors at larger angle threshold were balanced.

3.1.3. Result comparison. The first graph in figure 4 is the original data. The second is the result data by TSES and the third by ATINM.

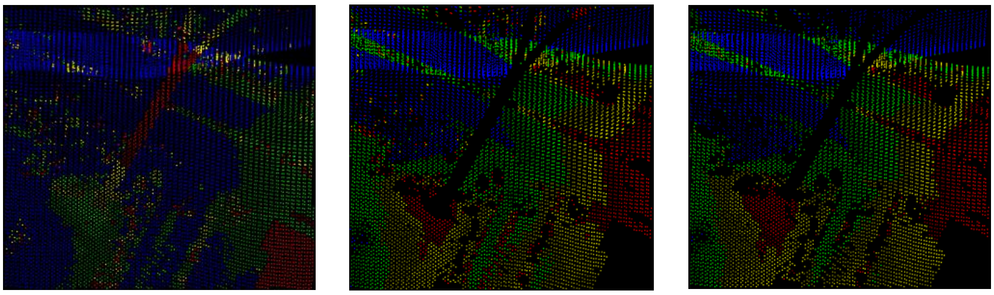

Figure 4. Original data and classified ground points data

The high building in the lower right corner and bridge in the center were well classified by both methods in figure 4. The scattered high points were better removed by ATINM.

Table 2. Quantitative error comparison.

	Threshold Segmentation		Adaptive TIN	
	Ground Points	Unground Points	Ground Points	Unground Points
Reference Ground Points	10050	47	9982	115
Reference Unground Points	559	2316	145	2730
Error Type I	0.46%		1.14%	
Error Type II	19.44%		5.04%	

More ground points were preserved by TSES, but type II error by it was much higher. Type II error by the other method was more acceptable.

3.2. Data set 2
The elevations differ in a large range. Data gaps and clump of low points exist in the study area.

3.2.1. Threshold segmentation.

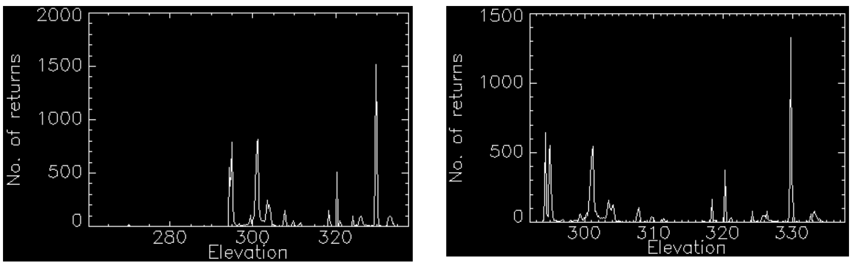

Figure 5. Elevation histograms before and after removing very low points

There are many peaks in the elevation histogram shown below. Ground points could not be well abstracted from one-time segmentation. Repeated segmentation of the segmented outcome is essential.

Four valley thresholds, 313.68, 298.79, 305.35 and 302.38, were successively picked. By the third threshold, type I error was 0, and type II error was 16.99%. In order to reduce type II error, the fourth iterative calculation was implemented, which result was more balanced.

3.2.2. Adaptive TIN. To ensure the initial seed point selections, very low points were removed. Then 1.0m, 1.5m, and 2.0m were picked as the elevation threshold. Figure 6 showed extracted numbers of points and figure 7 showed two type errors changing with the angle.

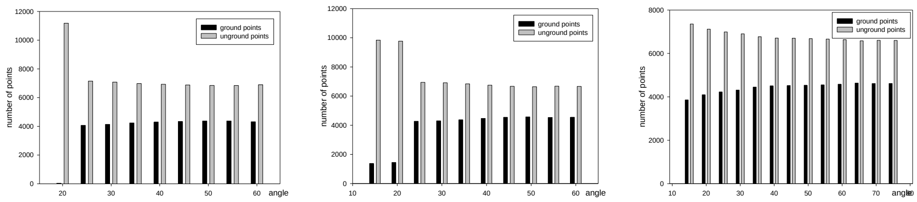

Figure 6. Ground and unground points number under the elevation threshold of 1.0m, 1.5m, 2.0m

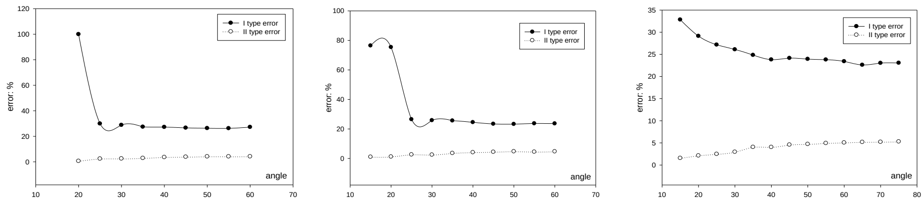

Figure 7. Type I error and type II error under the elevation threshold of 1.0m, 1.5m, 2.0m

When the elevation threshold was set as 1.0m and 1.5m, the results at the angle less than 25° were very bad. Most results with larger angles were relatively stable and balanced.

3.2.3. Result comparison. The first graph in figure 4 is the original data. The second is the result data by TSES and the third by ATINM.

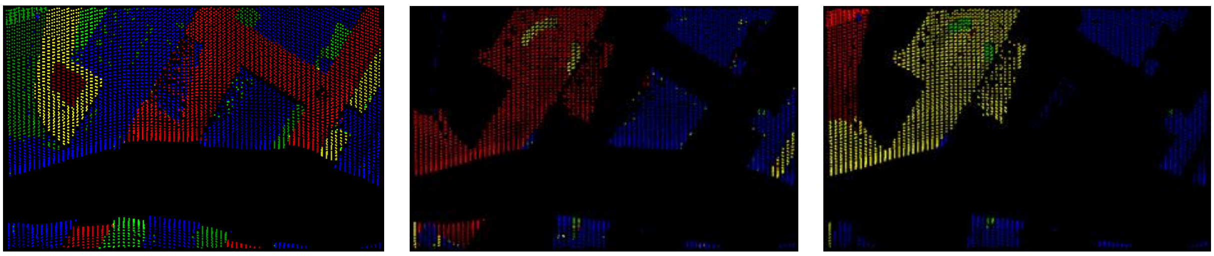

Figure 8. Original data and classified data using TSES and ATINM

The high building in the center were well removed by both methods. Low point area in the center was preserved by the threshold segmentation. High points in local area were well remove by ATINM. The quantitative comparisons are shown in table 3.

Table 3. Quantitative error comparison

	Threshold Segmentation		Adaptive TIN	
	Ground Points	Unground Points	Ground Points	Unground Points
Reference Ground Points	4992	626	4304	1314
Reference Unground Points	462	5133	247	5377
Error Type I	11.14%		23.39%	
Error Type II	8.26%		4.39%	

4. Results
Both results of two data sets by TSES show a lower I type error. If the elevation histogram has clear peaks and valley, or the object and background points have obvious elevation boundary, the method is time-saving of good quality. For area with complex feature distribution and large height changes,

repeated segmentation of previous outcome can reach a satisfactory result. If the study area is very large or of steep topology, different terrain features may make the elevation histogram fluctuated. Ground points on highland and unground points on lowland may have equal elevations. Data partition and sectional elevation statistics are possible to improve the accuracy in some degree.

Both results of ATINM show a lower II type error. For a flat terrain, low elevation threshold can ensure good error control, and large angle threshold can help extract more ground points. For area where the elevations change severely, larger elevation parameter is recommended, and larger angle parameter can reach a balanced error type. In normal data processing, parameter can be repeatedly adjusted by contrasting the visualization result. When the number of extracted points keeps stable, the angle and height parameters are reasonable.

5. Conclusion

The parameter selection and iteration calculation are key procedures of these two methods. The results in this paper are aimed to provide guidance for LiDAR point cloud processing.

ATINM has a universal applicability. Its low II type error can lower the errors in DEM interpolation. TSES is easier to handle. More extracted points preserve regional topology details. As these two methods show complementary characteristics in error type, they may be combined in some degree to reach both lower I type and II type errors.

Elevation data of LiDAR point can also be transformed into height textures and combined with optical images or multiple source data to improve the accuracy.

Acknowledgements

The project is supported by the National Basic Research Program of China through grant number 2009CB723900.

References

[1] Kilian J, Haala N and Englich M 1996 Capture and Evaluation of Airborne Laser Scanner Data, *Int. Arch. of Photogrammetry and Remote Sensing (Vienna, Australia)* 383-8.

[2] Ackermann F 1999 Airborne laser scanning - present status and future expectations *Isprs Journal of Photogrammetry and Remote Sensing* **54** 64-7

[3] Hodgson M E, Jenson J R and Schmidt L 2003 An evaluation of lidar- and IFSAR-derived digital elevation models in leaf-on conditions with USGS Level 1 and Level 2 DEMs *Remote Sensing of Environment* **84** 295-308

[4] Axelsson P 1999 Processing of laser scanner data - algorithms and applications *Isprs Journal of Photogrammetry and Remote Sensing* **54** 138-47.

[5] Xianfeng Huang, Hui Li and Xiao Wang 2009 Filter Algorithms of Airborne LiDAR Data: Reviews and Prospects *Acta Geodaetica et Cartographica Sinica* **38** 466-9

[6] Axelsson P 2000 DEM Generation from Laser Scanner Data Using Adaptive TIN Models *International Archives of Photogrammetry and Remote Sensing* **33** 110-7

[7] Axelsson P 2001 Ground estimation of laser data using adaptive TIN-models *Proceedings of OEEPE workshop on airborne laser scanning and interferometric SAR for detailed digital elevation models (Stockholm, Sweden, 1-3 March 2001) (Royal Institude of Technology Department of Geodesy and Photogrammetry)* **40** 185-208

[8] Chun Liu, Chen Huayun and Wu Hangbin 2009 *Three-dimensional laser remote sensing data processing and feature extraction* (Beijing: Science Press) p 133