Infrared Physics & Technology 75 (2016) 87–92

Contents lists available at ScienceDirect

Infrared Physics & Technology

journal homepage: www.elsevier.com/locate/infrared

An entropy-based filtering approach for airborne laser scanning data

Zhe Zeng [a], Jiaxin Wan [b], Hui Liu [c,*]

[a] School of Geosciences, China University of Petroleum, Qingdao 266555, China
[b] Faculty of Geomatics, East China Institute of Technology, Nanchang 330013, China
[c] Xiamen University, Xiamen 361005, China

ARTICLE INFO

Article history:
Received 12 August 2015
Available online 23 January 2016

Keywords:
DEM
LiDAR point clouds
Filtering
Entropy optimization

ABSTRACT

Parameter-tuning is a challenging task when generating digital terrain models from airborne laser scanning (light detection and ranging, LiDAR) data. To address this issue, this paper presents a filtering method for near-infrared laser scanning data that exploits the principle of entropy maximization as the optimization objective. The proposed approach generates ground elevation of point cloud by constructing a triangulated irregular network, calculates the entropy of the elevation from different parts, and automatically separates ground and non-ground points by the principle of entropy maximization. Experimental results from different ground surfaces show that the proposed entropy-based filtering method can effectively extract bare-earth points from the point cloud without adjusting thresholds.

1. Introduction

In recent years, airborne laser scanning (light detection and ranging, LiDAR) technology, which commonly employs near infrared (NIR) band lasers (e.g. 960 nm or 1064 nm) in a pulsed mode, has become an important means of constructing high-resolution digital elevation models (DEMs) [1,2]. LiDAR has been applied to various fields, such as oil exploration, forestry, geology, and archaeology, because high-resolution DEMs are indispensable data for these practical applications. However, constructing high-resolution DEMs from LiDAR point clouds remains difficult, primarily because accurately filtering non-ground points from point cloud requires some thresholds to be manually adjusted.

Current filtering algorithms include mathematical morphological filtering, gradient, interpolation, triangulated irregular network (TIN), segmentation, and skewness balancing methods, amongst many others. Lindenberger proposed mathematical morphological filtering, which is mainly used opening and closing operator [3]. Determining the thresholds and moving window size were key issues of the algorithm. Kilian presented an approach that exploited multiple filtering by assigning different weights to different window sizes [4], Zhang proposed a strategy that adaptively adjusted thresholds by terrain steepness and window size [5]. Vosselman proposed gradient based filtering [6], where the filtering process was determined by height difference and distance. However, thresholds were still adjusted manually according to the prior knowledge about the area of LiDAR point cloud data. Axelsson proposed TIN-based filtering which initially used a sparse TIN, and then iteratively estimated ground surface [7]. Kraus proposed an interpolation-based method, which exploited least square interpolation through adaptive weight functions [8]. The segmentation-based filtering algorithm separated the non-ground point cloud from the ground point cloud using a variety of segmentation rules which were imposed on the depth image generated from LiDAR point cloud data [9,10].

The algorithms discussed above can successfully filter non-ground points from LiDAR point cloud. However, the procedures require manually adjusted thresholds for different circumstances, which is a strenuous task for generating DEM data. Therefore, threshold-free filtering algorithms began to be considered by many scholars. Bartels proposed an unsupervised method based on the skewness of height information to achieve threshold-free filtering [11,12]. However, the ground points after filtering still contained non-ground points such as low vegetation and building walls.

This paper proposes a filtering method based on the entropy maximization principle, which can separate non-ground points from ground points cloud without manually adjusting thresholds. The proposed method initially builds TIN from seed points by moving window, and then generates the height difference of the point cloud data. The entropies of height difference can then be evaluated and the ground and non-ground parts of the point cloud can be divided by the principle of maximum entropy.

* Corresponding author.

http://dx.doi.org/10.1016/j.infrared.2015.12.021
1350-4495/

Experimental results show that the entropy-based filtering method separates non-ground from the point cloud without manually adjusting thresholds, compared with traditional methods.

2. Shannon entropy-based filtering method for lidar point cloud

2.1. Shannon entropy maximization principle

Entropy is a concept from thermodynamics, which describes the disorder of an irreversible process. Shannon redefined the entropy concept of Boltzmann or Gibbs, which was considered as a measure of uncertainty, to apply to the information of a system [13]. If a system has n discrete statuses, then the entropy, S, of the system can be defined the probability distribution of these statuses,

$$S = -\sum_{i=1}^{n} p_i \ln(p_i) \tag{1}$$

where $p = \{p_i\}$ is the probability of the i-th status, and $\sum_{i=1}^{n} p_i = 1, 0 \leqslant p_i \leqslant 1$. Jaynes proposed the entropy maximization principle [14], and entropy optimization is based on entropy maximization as the object function. Entropy optimization has been extensively studied and applied in many research fields, especially to achieve better results in the field of image segmentation. However, the entropy optimization method has not been employed for LiDAR point clouds filtering previously.

2.2. Shannon entropy-based filter principle for LiDAR point cloud filter

Filtering methods for LiDAR point cloud data identify the different characteristics of the ground and the non-ground points based on the point status information. Current filtering methods generally separate the ground and non-ground points by tuning a threshold relevant to the characteristics considered. In contrast, we propose a filtering method which employs entropy maximization and avoids the need for a threshold.

Let x be an information type, such as height and intensity Fig. 1 shows an example probability distribution of the values. In general, x is $X_{\min} - X_{\max}$, and there are n LiDAR points in the data cloud. The data range can be equally divided into L levels, with

$$n = \sum_{i=0}^{L} c_i \tag{2}$$

where c_i is the number of the i-th level. The probability of the ith level is

$$p_i = \frac{c_i}{n}, \quad (0 \leqslant i \leqslant L) \tag{3}$$

and the entropy of status x can be evaluated by Eq. (1).

Assume that A represents the ground and B the non-ground point cloud. The principle of the Shannon entropy-based filtering is that some level, t, can divide the point cloud into A and B, such that the sum of the entropies of A and B is maximized (see Fig. 1).

2.3. Implementation of entropy-based filtering

The height value is selected as the variable for filtering, and entropy-based filtering was implemented (see Fig. 2) as follows:

(1) Preprocessing the point cloud.
Divide the point cloud (see Fig. 3a), extract seed points (Fig. 3b–d) and construct the initial TIN from the seed points.
(2) Calculating the height difference.
Interpolate height values and calculate height difference. The interpolation is based on each triangle of the TIN and may be interpolated by the heights of the triangle vertices. Height difference can be calculated by subtracting the interpolated height from the original height of a LiDAR point.
(3) Filter using entropy maximization.
The steps are shown in Fig. 4.
  (i) The maximum height difference is set to be L. Calculate $p_i (i \in [0, L])$ from Eq. (3) and create a histogram of height difference.
  (ii) Assign the initial level, t_{init}, to divide ground and non-ground points.
  (iii) A and B are divided by level t. Create separate histograms and calculate the Shannon entropies, $S_A(t)$, $S_B(t)$, respectively. The entropy sum is just of the two parts can be evaluated by $S_{A+B}(t) = S_A(t) + S_B(t)$. Put the entropy sum into an entropy queue Q.
  (iv) Increase the splitting level t by 1. If t is less than L, then re-perform step (iii). If t is equal to L, then perform the subsequent step.
  (v) Find the maximum entropy sum in the queue Q and corresponding splitting level t, denoted by t_{OPT}.

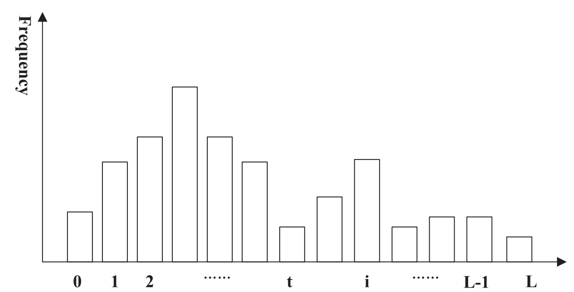

Fig. 1. Histogram of levels for a given information type.

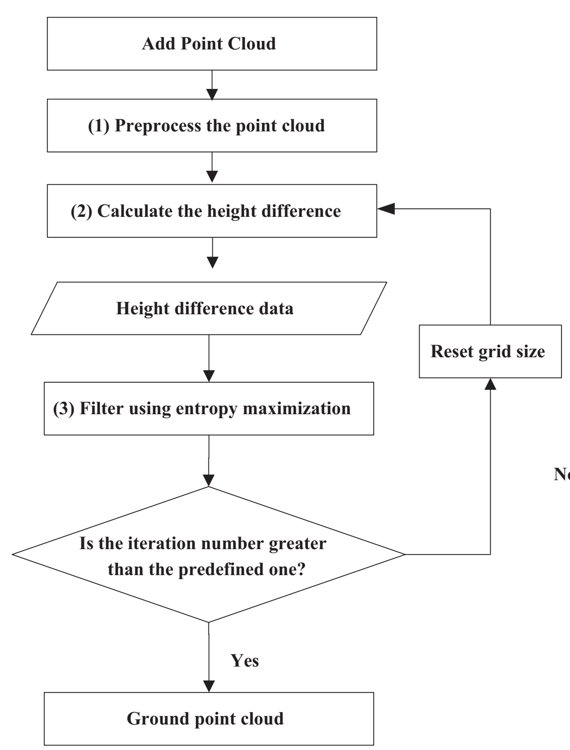

Fig. 2. Flow chart of entropy-based filtering method.

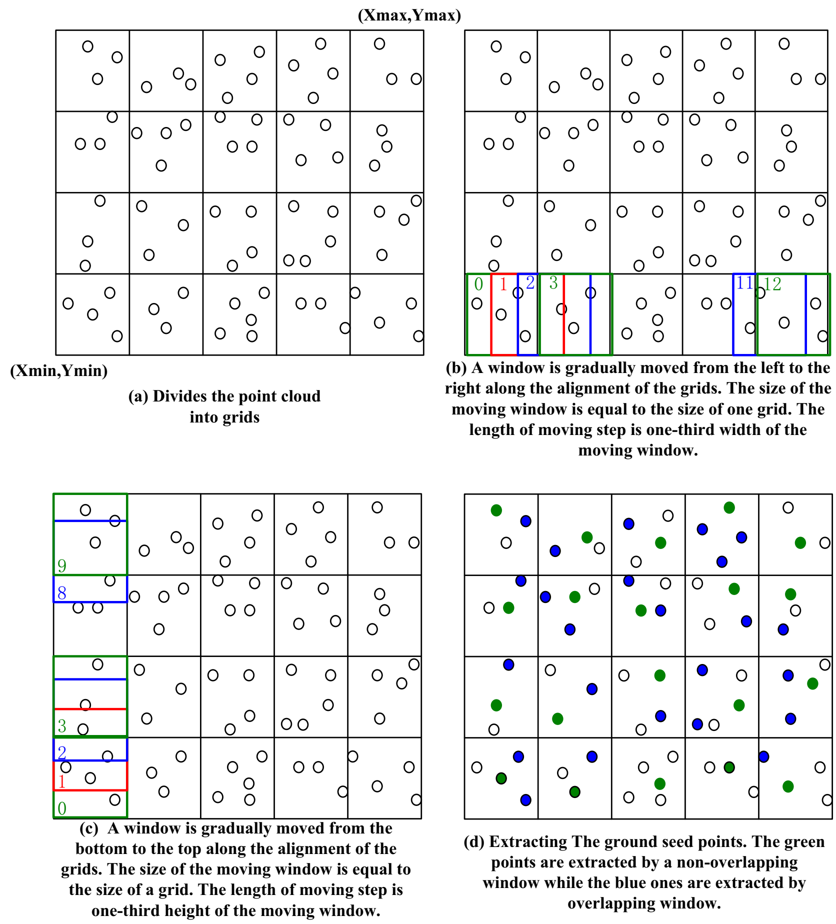

Fig. 3. Preprocessing the point cloud by moving window in the grids.

(vi) Filter the non-ground points by deleting the points of which height difference is less than level t_{OPT}.

(vii) The process is conducted iteratively (see Fig. 2). If the iteration count is less than a predefined cut-off, then reset the grid size and repeat steps (2) and step (3).

3. Experimental results

To evaluate our proposed method, we performed experiments using six NIR band airborne laser scanning datasets which included plain and slope with buildings, vegetation and small objects, as indicated in Fig. 5. All points of these datasets had been labeled Bare Earth or Object [15]. Filtering was implemented as discussed above, using the height difference. After filtering, the remained points were bare earth and thus the filtered points were objects. The output from the proposed filtering method can be compared against the reference for Type I and II errors.

Fig. 6 shows the resulting dataset after applying our proposed filter system on the six sample datasets, Grey points represent (predicted) bare earth, black ar non-ground, blue are Type I errors, and red are Type II errors. The proposed entropy-based filtering approach can adequately identify various objects, particularly large buildings (e.g. buildings in Samples 12 or 41), low objects (e.g. low cars in Sample 12, vegetation on slopes in Sample 51) and attached objects (e.g. bridges in Sample 21). The substantial shape of the terrain is retained, although a few non-ground points are not identified (e.g. a1, b1, c1 in Sample 12; a2 in Sample 21; a3, b3 in Sample 41; a4 in Sample 51; and a6, b6 in Sample 61). On the other

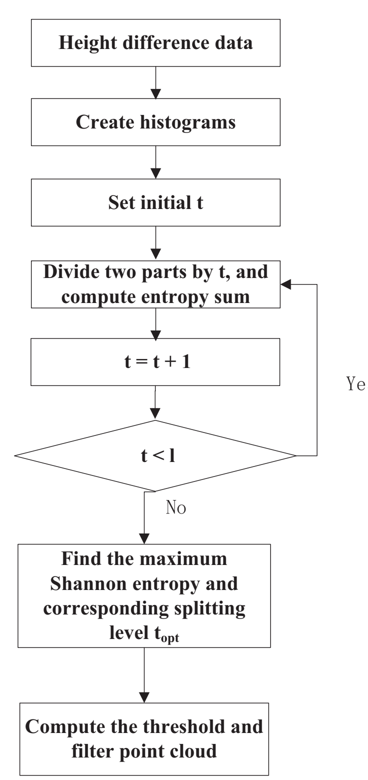

Fig. 4. Flow chart for the proposed filtering method.

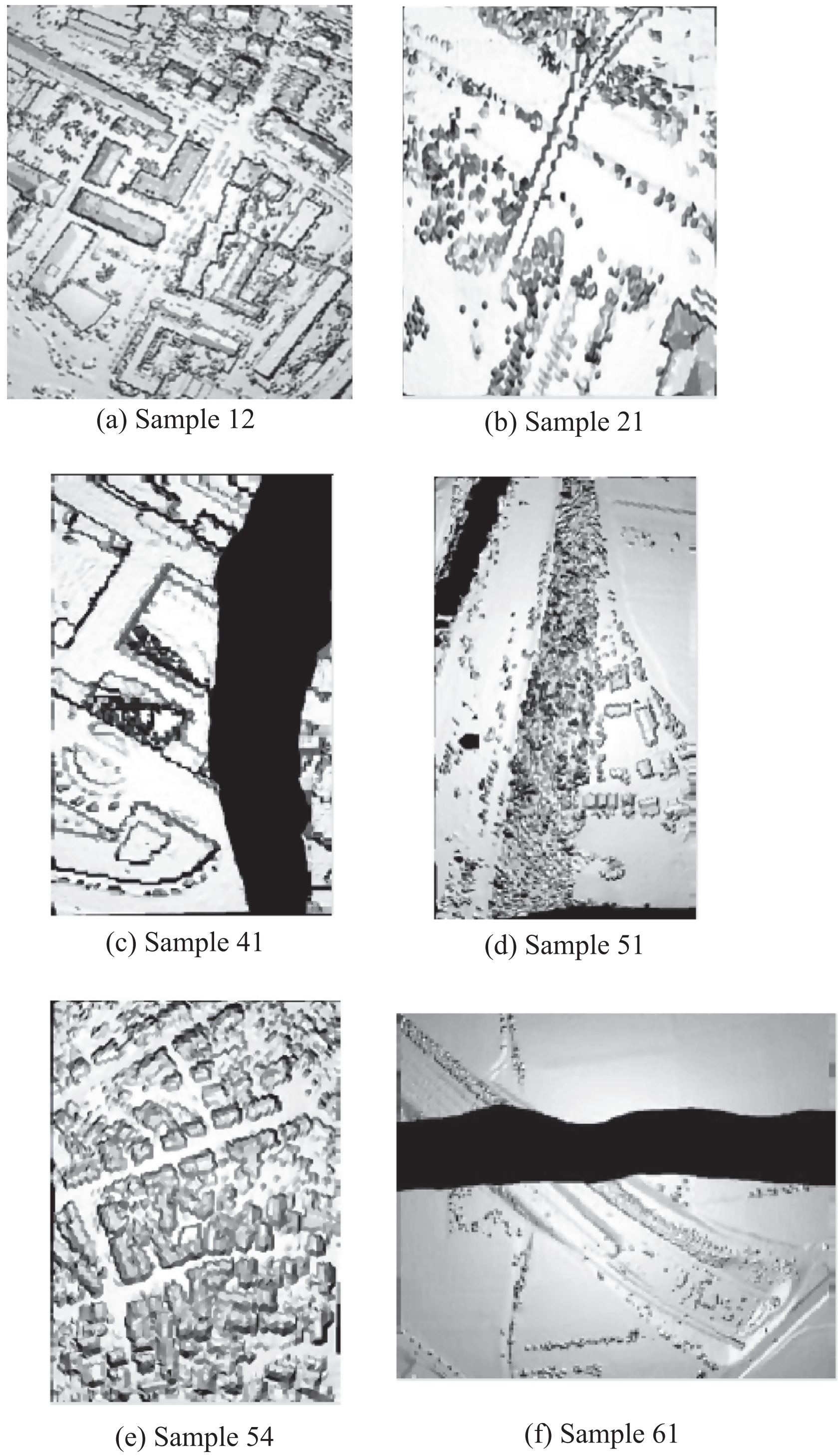

Fig. 5. Six airborne LiDAR data sets chosen for experimental verification of our proposed filtering methods. The data is part of the ISPRS dataset, which is acquired by NIR band (960 nm or 1064 nm) laser [15].

hand, a small number of ground points are also erroneously removed (e.g. d1 in Sample 12; b2 in Sample 21; c3, d3 in Sample 41; b4 in Sample 51; and a5 in Sample 54).

Fig. 7 also shows the errors arising from filtering. Type I errors are approximately 1.4–6.8%, with an average of approximately 2.8%. Thus, the proposed entropy-based filtering method successfully identifies ground points. The filtering method also successfully distinguished most objects, such as buildings and vegetation, from ground points. Type I errors arise primarily due to terrain discontinuity. Type II errors range from 1.3% to 4.8%, with an average of approximately 3.3%. Type II errors for three sample datasets (Samples 12, 21 and 61) are less than 3% while other dataset (Samples 41, 51 and 54) are approximately 4.5%. This shows that the entropy-base filtering method can filter

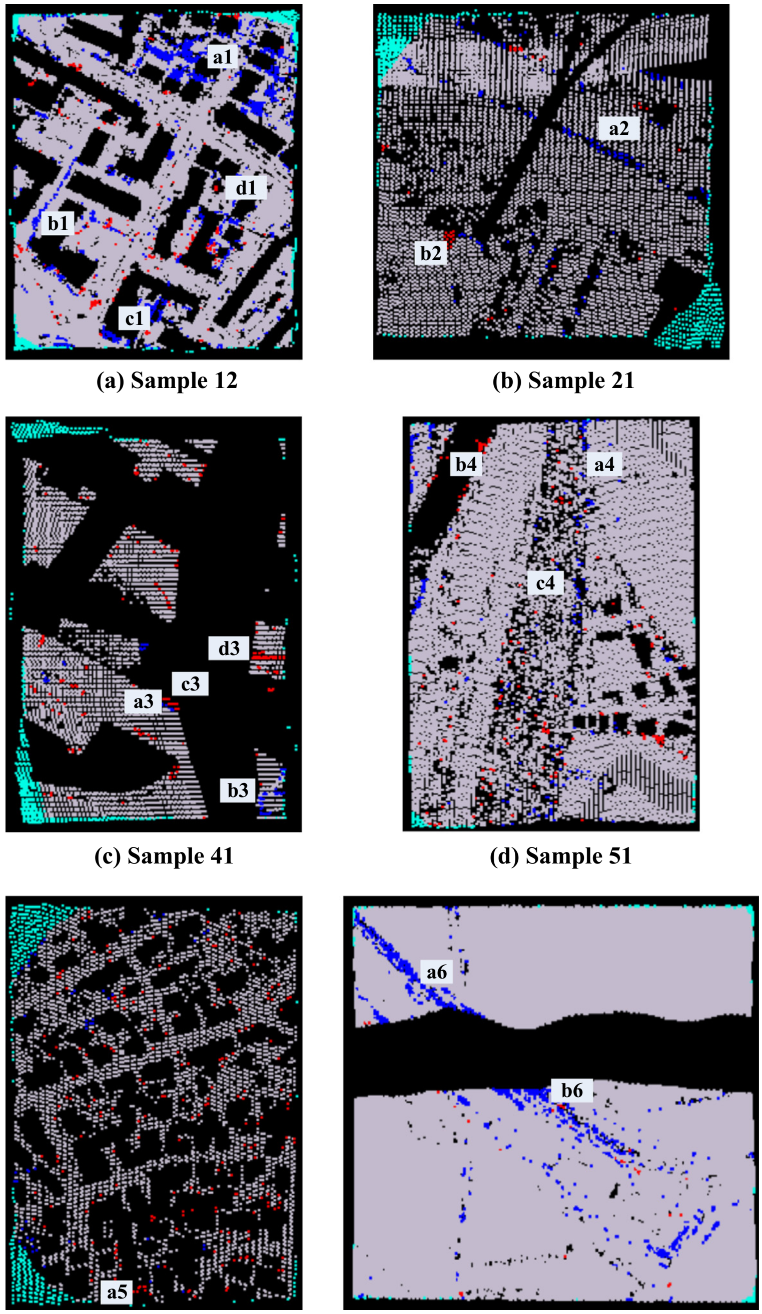

Fig. 6. The datasets of Fig. 5 filtered by our proposed method showing the resultant ground dataset.

the objects such as large buildings, low vegetation and bridges. Total error (a ratio of the count of incorrectly identified points to the total number of points) ranges from 2.1% to 4.1%, with an average of 2.9%.

Table 1 shows the total error comparison. The total errors of the proposed filtering method are less than the minimal total error values of the eight classical filtering methods for most of the datasets (Samples 21, 41, 51 and 54). The total errors of the proposed method are also small values, for the two other datasets (Samples 12 and 61), which are still slightly large than the minimal total error values of the eight classical methods. Thus, in general, the NIR-band LiDAR point cloud can be accurately filter proposed entropy-based filtering method.

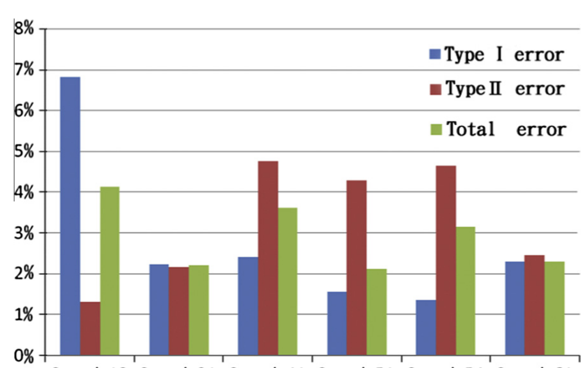

Fig. 7. Errors from the proposed filtering method.

Table 1
Total error comparation between the proposed filtering method and eight classical filtering methods [15]. The minimal values of the eight classical filtering methods are bold font for a dataset.

Dataset	Method							
	Elmqvist	Sohn	Axelsson	Pfeifer	Brovelli	Roggero	Wack	Sithole	Entropy-based
Sample 12	8.18	8.39	**3.25**	4.50	16.28	6.61	6.61	10.21	4.13
Sample 21	8.53	8.80	4.25	**2.57**	9.30	9.84	4.55	7.76	2.22
Sample 41	**8.76**	11.27	13.91	10.75	17.03	12.21	9.01	23.67	3.61
Sample 51	21.31	9.31	**2.72**	3.71	22.81	3.01	11.45	7.02	2.14
Sample 54	21.26	5.68	**3.23**	5.47	23.89	4.96	7.63	6.33	3.15
Sample 61	35.87	2.99	**2.08**	6.91	21.68	18.99	13.47	21.63	2.30

4. Conclusion

This paper proposes a filtering method for airborne LiDAR points cloud based on Shannon entropy optimization. The proposed entropy-based approach can filter the datasets without requiring thresholds. Experiments on NIR-band airborne imaging LiDAR show that the proposed method is accurate and feasible for filtering non-ground from ground point clouds.

Conflict of interest

No conflict of interest.

Acknowledgments

This work was supported by the Fundamental Research Funds for the Central Universities (Grant no. 13CX02034A). The authors would also like to thank the anonymous reviewers for their helpful suggestions that improved the paper.

References

[1] J.P. Wilson, Digital terrain modeling, Geomorphology 137 (2012).
[2] M. Rutzinger, B. Hofle, K. Kringer, Accuracy of automatically extracted geomorphological breaklines from airborne LiDAR curvature images, Geografiska Ann. Ser. A – Phys. Geogr. 94A (2012) 33–42.
[3] J. Lindenberger, Laser-profilmenssungen zur Topographischen Gelandeaufnahme, Universitat Stuttgart, 1993.
[4] N.H. Johannes Kilian, Markus Englich, Capture and evaluation of airborne laser scanner data, Int. Arch. Photogr. Remote Sens. XXXI (1996) 383–388.
[5] K. Zhang, S.-C. Chen, M.-L. Shyu, J. Yan, C. Zhang, A progressive morphological filter for removing nonground measurements from airborne LiDAR data, IEEE Trans. Geosci. Remote Sens. 41 (2003) 872–882.
[6] G. Vosselman, Slope based filtering of laser altimetry data, Int. Arch. Photogr., Remote Sens. Spat. Inf. Sci. XXXIII (2000) 935–942.
[7] P. Axelsson, DEM generation from laser scanner data using adaptive TIN models, Int. Arch. Photogr., Remote Sens. Spat. Inf. Sci. XXXIII (2000) 110–117.
[8] K. Kraus, N. Pfeifer, Determination of terrain models in wooded areas with airborne laser scanner data, ISPRS J. Photogr. Remote Sens. 53 (1998) 193–203.
[9] G. Sithole, G. Vosselman, Filtering of airborne laser scanner data based on segmented point clouds, in: ISPRS WG III/3, III/4, V/3 Workshop "Laser Scanning 2005", Enschede, the Netherlands, 2005.
[10] J. Zhang, X. Lin, Filtering airborne LiDAR data by embedding smoothness-constrained segmentation in progressive TIN densification, ISPRS J. Photogr. Remote Sens. 81 (2013) 44–59.
[11] M. Bartels, H. Wei, Segmentation of LiDAR data using measures of distribution, Int. Arch. Photogr. Remote Sens. Spat. Inf. Sci. XXXVI (2006) 426–431.
[12] M. Bartels, H. Wei, Threshold-free object and ground point separation in LiDAR data, Pattern Recogn. Lett. 31 (2010) 1089–1099.
[13] Claude E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (3) (2001) 379–423.
[14] J.N. Kapur, H.K. Kesavan, Entropy Optimization Principles with Applications, Academic Press Inc, 1992.
[15] G. Sithole, G. Vosselman, ISPRS test on extracting DEMs from point clouds: a comparison of existing automatic filters, 2003. Available: <http://www.itc.nl/isprswgiii-3/filtertest/>.