A CLUSTERING APPROACH FOR DETECTION OF GROUND IN MICROPULSE PHOTON-COUNTING LIDAR ALTIMETER DATA

Jiashu Zhang, John Kerekes

Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology
Rochester, NY, 14623

Beata Csatho, Toni Schenk, Robert Wheelwright

Department of Geological Science
University at Buffalo
Buffalo, NY, 14260

ABSTRACT

Observations from satellite lidar instruments have provided evidence in the remarkable changes in polar ice sheets on a global scale. The Ice, Cloud and land Elevation Satellite-2 (ICESat-2) is scheduled for launch by NASA in 2018 and will monitor the elevation changes of polar ice sheets and vegetation canopy. To validate ICESat-2's approach of photon-counting laser altimetry, measurements obtained from the Multiple Altimeter Beam Experimental Lidar (MABEL) instrument are critical. In support of the ICESat-2 mission, this paper derives an algorithm for the detection of ground and vegetation canopy in photon-counting laser altimeter data. This approach uses a density-based clustering model and modifies the shape of search area. Based on results from MABEL observations, the proposed approach is seen to be robust in detecting ground and vegetation canopy as well as background noise reduction. In addition, this approach can be quickly implemented and adaptive to photon-counting lidar data sets with different point cloud densities.

***Index Terms*—** LiDAR, photon-counting, clustering, MABEL

1. INTRODUCTION

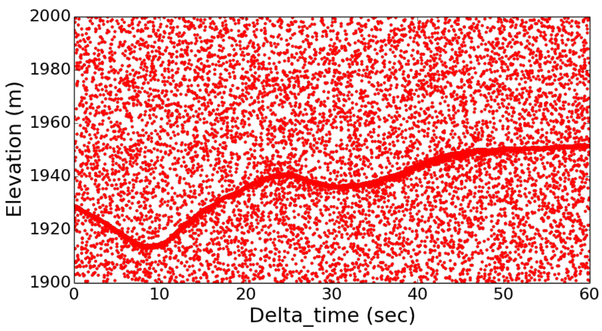
In recent years quantifying changes in polar ice sheets remains an earth science priority. These changes could contribute a large part in terms of sea level rise and global climate change. To monitor the elevation changes of Greenland and Antarctic ice sheets, the Ice, Cloud and land Elevation Satellite-2 (ICESat-2) is currently scheduled for launch in 2018. It is also intended to measure land topography and vegetation characteristics [1]. To simulate ICESat-2-like data, NASA is currently conducting flights over areas of interest using the Multiple Altimeter Beam Experiment Lidar (MABEL) laser altimeter. Measurements from MABEL provide a capability for airborne photon-counting altimetry and therefore serves as a prototype and simulator for the upcoming ICESat-2 mission [2].

The MABEL instrument uses a high-repetition-rate pulsed laser variable from 5 to 25 kHz, with a pulse length of 2 ns. The laser generates both 1064- and 532- nm outputs. MABEL records the time-position of each individual photon via detectors with single-photon sensitivity. The increased sensitivity often results in a more noisy data set, since background photons and system noise can also trigger the detector. While different methodologies have been developed to process lidar elevation data [3], an effective noise reduction and ground detection approach is required for micropulse photon-counting lidar altimeter data.

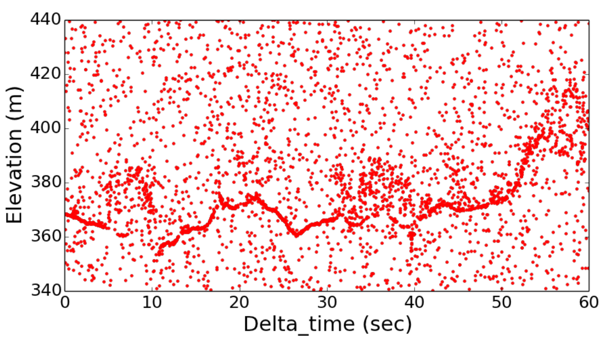
Previous work has shown the main factors affecting performance of photon-counting lidar on ice sheets [4], as well as noise filtering techniques for simulated ICESat-2 [5] and MABEL data [6]. In this paper, a clustering method is modified and used for the detection of the ground surface in MABEL data. This approach is based on the concept of Density Based Spatial Clustering of Applications with Noise (DBSCAN) [7]. Due to the higher density in the horizontal direction in photon-counting lidar point clouds, the shape of searching area is modified from a circle to an ellipse. This will have high accuracy in surface finding and is computationally efficient.

2. DATA SETS

Two example data sets from MABEL will be used in this study. The first one was collected near the Jakobshavn Glacier on April 19, 2012 under clear sky condition in daytime. The other one was collected in Wisconsin, USA on Spetember 26, 2012 under clear sky condition in nighttime. The data set used in this study (L2A, Release 8) consists of range and positional information (corrected for aircraft pitch, roll and yaw) of all received photon detection events, as calculated by the sensor based on time of departure/arrival. Surface elevation can then be inferred from the detected range and altitude of the aircraft. In Figure 1, a 2D elevation profile of a section of (a) Jakobshavn Glacier, and (b) Wisconsin are shown using the complete set of photon detections (red dots) from MABEL data. The total flight time is 1 min.



(a) Jakobshavn Glacier



(b) Wisconsin

Fig. 1. 2D elevation profile of a section of (a) Jakobshavn Glacier and (b)Winconson, using the complete set of photon detections (red dots).

The two example data sets here represent different scenes with different solar conditions. The one from Jakobshavn glacier is for snow/ice covered ground with high noise rate, while the one from Wisconsin demonstrates hilly terrain covered by canopies with low noise rate. A fast algorithm is required for the detection of photons reflected from the ground as well as the vegetation canopy. Here we will use a clustering algorithm relying on a density-based notion of clusters to identify clusters consisting of ground and canopy returns. The test algorithm is based on Density Based Spatial Clustering of Application with Noise (DBSCAN).

3. APPROACH FOR DETECTION OF GROUND

3.1. Introduction to DBSCAN

The key idea of DBSCAN is that for each point of a cluster the neighborhood of a given radius has to contain at least a minimum number of points, i.e. the density in the neighborhood has to exceed some threshold [7]. The shape of a neighborhood is determined by the choice of a distance function for two points p and q, denoted by $dist(p,q)$. Two parameters mentioned here are a Eps-neighborhood of a point, defined by $dist(p,q) \leq Eps$, and the minimum number of points ($MinPts$) in that Eps-neighborhood.

3.2. A modified DBSCAN for surface detection

For our datasets in two dimensions, the distance between two points $p(t_p, h_p)$ and $q(t_q, h_q)$ is defined as:

$$dist(p,q) = [\frac{(t_p - t_q)^2}{t_{scale}^2} + \frac{(h_p - h_q)^2}{h_{scale}^2}]^{\frac{1}{2}} \quad (1)$$

where: t represents $delta\_time$ in Figure 1, which can be considered as along-track distance, and h represents elevation. t_{scale} and h_{scale} are used for normalization so that the points in test data set have comparable order over t and h axis. Hence $dist(p,q)$ is now unit less.

In our algorithm, since most of the clusters (surface returns) have higher density in horizontal than in vertical direction, it is reasonable to modify the shape of search area accordingly. Therefore, the distance between point $p(t_p, h_p)$ and $q(t_q, h_q)$ is now modified as:

$$dist(p,q) = [\frac{(t_p - t_q)^2}{t_{scale}^2 a^2} + \frac{(h_p - h_q)^2}{h_{scale}^2 b^2}]^{\frac{1}{2}} \quad (2)$$

As can be seen in Figure 2, the search area is modified as an ellipse with centroid p, major axis with length $2a$ and minor axis with length $2b$, while $a > b$. Due to the change

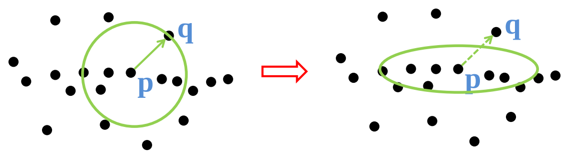


Fig. 2. Modification of searching area using DBSCAN. In left, by using a circular searching area, point q is density-connected to point p, also classified as part of the cluster. While in right, since the searching area is modified as ellipse, point q is no longer density-connected to point p, therefore q is now classified as noise.

in search area, points in the horizontal direction have more weight with respect to the search area center than points in the vertical direction. Therefore, continuous points in a roughly horizontal direction are more likely to be classified as belonging to the cluster. That is also the same as in the detection of ground for MABEL lidar point clouds.

3.3. Estimation of clustering parameters

As the ellipse shape is determined by a and b in Eq.(2), two parameters are needed for modified DBSCAN implementation: $MinPts$ and Eps. Here we develop a simple but effective heuristic way to determine the two parameters. For simplicity, Eps=2 is used all the time so that only $MinPts$ will be modified. It can be done by estimating the average point density within the search ellipse.

(1) A partition of points from test data set is first extracted. This example covers a flight time of δt and an elevation range of δh. The Area S of this sample data set is:

$$S = \delta t \cdot \delta h; \qquad (3)$$

(2) For an ellipse with $dist(p, q)$=Eps, its area $s1$ is:

$$s1 = \pi \cdot Eps^2 \cdot t_{scale} h_{scale} \cdot ab \qquad (4)$$

where: a=0.5, b=0.2. Hence, the number of ellipses within the example data set is roughly estimated as $S/s1$;

(3) The number of points in the example data set is found to be N. Therefore, the average point density (ρ) within the search ellipse can be calculated:

$$\rho = N/S \cdot s1; \qquad (5)$$

(4) To better estimate ρ, more than one example data sets are extracted from test data set and processed through steps (1) to (3) and then averaged. In the proposed clustering method, point density for clusters should be higher than the average density of the whole data set. $MinPts$ can be empirically estimated as:

$$MinPts \geq 4 \cdot \rho \qquad (6)$$

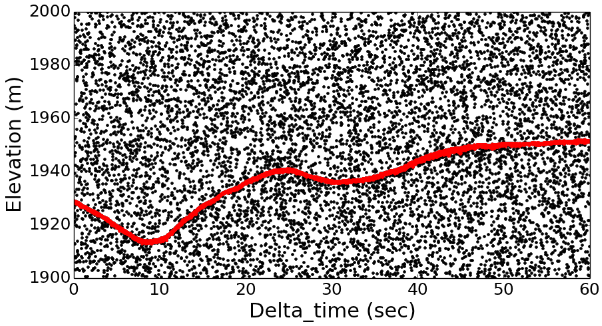
Practically we can always start with the minimum integer larger than 4ρ and increase by 1 gradually. For the MABEL photon-counting lidar data sets as in Figure 1(a), $\rho \approx 0.36$ and $MinPts = 4$ is finally applied. For the other data set as in Figure 1(b), $\rho \approx 3.85$ and $MinPts = 16$ is used. This proposed clustering algorithm can be quickly implemented and adaptive to photon-counting lidar data sets with different point densities.

4. RESULTS AND DISCUSSION

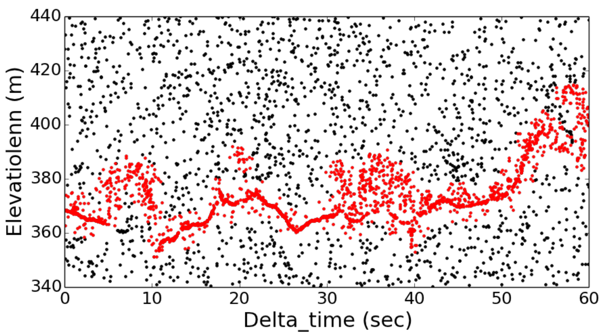
The results for detection of ground for MABEL data are shown in Figure 3. Parameters used in modified DBSCAN are: a=0.5, b=0.2, Eps=2. In addition, $MinPts$=4 is selected for Jakobshavn Glacier and $MinPts$=16 is used for Wisconsin. Here red dots represent classified surface returns while black dots represent classified noise. It is shown that the profile of ground is reliably extracted from point cloud, as can be seen in Figure 3(a). Meanwhile, both the ground surface and canopy can be detected from background noise, as can be seen in Figure 3(b). The proposed algorithm is seen to be robust in detecting ground and vegetation canopy and adaptive for data sets with different point cloud densities. However, since the vegetation canopy would partially block the ongoing and returning photons from ground, the point density of ground in that region is lower than ground without canopy coverage. Therefore, that part of the ground is hard to detect using the proposed method.

5. SUMMARY AND FUTURE WORK

In this paper, an algorithm is proposed for the detection of ground and vegetation canopy for photon-counting laser altimetry data. Two data sets from MABEL in different solar conditions were reviewed. A clustering method based on the



(a) Jakobshavn Glacier



(b) Wisconsin

Fig. 3. Result for detection of ground and vegetation canopy for MABEL dataset collected over (a) Jakobshavn Glacier; and (b) Wisconsin. Here red dots represent classified surface returns while black dots represent classified noise. Parameters used in clustering are: a=0.5, b=0.2, Eps=2, (a)$MinPts$=4, (b)$MinPts$=16.

concept of DBSCAN was introduced. The area shape of a data point search for its nearest neighbors was modified to be an eclipse to match general characteristics of terrain or vegetation canopy. Results showed that the proposed algorithm works well for surface detection in point cloud with variable noise rates. The surface and canopy can be expected to be observable during the ICESat-2 mission. In the future, performance assessment will be studied to quantitatively evaluate the proposed algorithm. The algorithm derived here can be used as a basis for the analysis of data from the ICESat-2 mission, MABEL, and other photon-counting lidar altimeter data in general.

Acknowledgment

This work is supported by NASA under award number NNX11AK77G.

6. REFERENCES

[1] W. Abdalati, H. Zwally, R. Bindschadler, B. Csatho, S. Farrell, H. Fricker, D. Harding, R. Kwok, M. Lefsky, T. Markus, A. Marshak, T. Neumann, S. Palm, B. Schutz, B. Smith, J. Spinhirne, and C. Webb, "The icesat-2 laser altimetry mission," *Proc. of IEEE*, vol. 98, no. 5, pp. 735–751, May 2010.

[2] M. McGill, T. Markus, V. S. Scott, and T. Neumann, "The multiple altimeter beam experimental lidar (mabel): An airborne simulator for the icesat-2 mission," *Journal of Atmospheric and Oceanic Technology*, vol. 30, no. 2, pp. 345–352, 2013.

[3] T. Schenk and B. Csathó, "A new methodology for detecting ice sheet surface elevation changes from laser altimetry data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 50, no. 9, pp. 3302–3316, 2012.

[4] J. Kerekes, A. Goodenough, S. Brown, J. Zhang, B. Csathó, A. Schenk, S. Nagarajan, and R. Wheelwright, "First principles modeling for lidar sensing of complex ice surfaces," in *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*. IEEE, 2012, pp. 3241–3244.

[5] U. C. Herzfeld, B. W. McDonald, B. F. Wallins, T. A. Neumann, T. Markus, A. Brenner, and C. Field, "Algorithm for detection of ground and canopy cover in micropulse photon-counting lidar altimeter data in preparation for the icesat-2 mission," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 4, pp. 2109–2125, 2014.

[6] K. Horan and J. Kerekes, "An automated statistical analysis approach to noise reduction for photon-counting lidar system," in *Proc. 2013 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Melbourne, Australia, 2013.

[7] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. Knowledge Discovery and Data Mining (KDD'96)*, 1996, pp. 226–231.