Anti-Excessive Filtering Model Based on Sliding Window

Haoang Li[1, a], Weiming Hu[1, b], Jian Yao[1, c] and Wenqiao Zhang[1, d]

[1]School of Remote Sensing and Information Engineering

Wuhan University, Wuhan, Hubei, P.R. China

[a]haoang.li@whu.edu.cn, [b]weiming.hu@whu.edu.cn, [c]jian.yao@whu.edu.cn, [d]wenqiao.zhang@whu.edu.cn

Keywords: LiDAR; Sliding Window; Anti-Excessive Filtering.

Abstract. This paper is purposed on the problem that ground points in cloud of LiDAR (Light Detection and Ranging) has been excessively filtered. And it proposes an improved model basing on the traditional sliding window model, combining with optimized rules on determining standard elevation value, tolerance of elevation difference and dynamic thresholds. The optimized rules on determining standard elevation could increase the accuracy of filtering, by the nearest neighbor and interpolation algorithm whose theoretical base is spatial correlation. Tolerance of elevation difference prevents excessively filtering by skipping some windows with negligible elevation difference, maintaining the accuracy at a relevant higher level. Dynamic thresholds, changing with various terrains and times of iteration, are adopted to further improve the model based on the mentioned factors. To demonstrate our algorithm and prove its effectiveness, the standard data from ISPRS (International Society for Photogrammetry and Remote Sensing) for testing is selected. Results have shown that both robustness and efficiency increase with the optimization.

Introduction

LiDAR is a remote sensing technology that measures distance by illuminating a target with a laser and analyzing the reflected light. LiDAR is popularly used as a technology to make high-resolution maps, with applications in archeology, geography, geology, etc. LiDAR filtering is an important part in data processing. The purpose of filtering is to recognize, as well as to remove, the non-ground points like the returns from buildings and trees, with the left points producing Digital Elevation Model (DEM) [1] which is the core of basic geographic information.

There are generally two ways for filtering: the first one is that beginning with a rather small area, the algorithm filters the whole cloud points in this area and then the area of processing expands; the second one is that beginning with a relatively big area, the algorithm processes this specific area by continuous grid division, contributing to the higher accuracy of the terrain model. Generally speaking, filtering with morphology and filtering based on refinement that contains Triangulated Irregular Network (TIN) and sliding window filtering algorithm, represent the above two ways respectively.

Filtering with morphology is based on shrinking windows. This method first sets some windows for filtering and then uses morphological opening algorithm to remove those non-ground points, leaving the ground points behind to create a simulative earth terrain [2]. This method is valid and convenient to perceive. However, it is hard to simulate the surface with violent changes [3].

The procedure of filtering of the algorithm based on idea of refinement begins with an initial sparse DEM, then selecting points from candidate data, interpolating and refining DEM. TIN filtering algorithm adding cloud points within elevation difference and angle thresholds to the triangles, to achieve the purpose of reconstruction of the ground surface [4]. Sliding window algorithm first chooses a shrinking window, which should depend on a suitable range, and then in the window, constantly searches for the lowest point calculating a rough terrain model. The model is a reference for the next step to eliminate more non-ground points to obtain higher precision of DEM [5]. Hosseini et al. optimized the window shape using strips [6]. Traditionally, window dividing and threshold

determination often rely on experience, which leads to quite different filtering results and various accuracy in different terrains.

Due to the insufficient restrictive conditions of the existing LiDAR filtering algorithm based on sliding window, there is a greater probability of excessive filtering which mistakes the ground points as non-ground points. This paper introduces the Anti-Excessive Filtering Model which optimizes the original sliding window filtering algorithm from the three following views to solve this problem.

Algorithm

Optimum Determination of Standard Elevation Value.

In traditional sliding window filter algorithm, the lowest point among the window is selected as the elevation datum whose elevation E is defined as the unified standard elevation value for the other points in the same window. If the subtraction of the height of the point and E is less than the fixed threshold, this point will be defined as a ground point and vice versa. The further the distance between the points is, the more the errors there would be, especially in sloping fields.

The key to the determination of optimum standard elevation value, in the improved model, for each point lies in two aspects. First according to First Law of Geography [7], the nearer two points are, the more inter-dependent they would be. Thus, the first method finds the nearest standard point to the undetermined point, using the height of nearest point to represent the standard elevation of the undetermined point, rather than using the lowest point within the window. This method is usually adopted when the windows are bigger and points are far from each other. The second method is using Kriging interpolation [8]. Interpolation will not be as effective when the distances among most of the testing points are larger than the Range (a parameter of Semi-variogram). Thus, interpolation is more helpful when the area of each window is small and points are near each other.

To fit Kriging Semi-variogram, all points in the surveyed area are considered [9], calculating and recording the Range of model as C. For every point in the central window, if the distances between this point and the lowest points from its 8-neighbor windows are all greater than C, the first method is adopted. Else, the second method is adopted, the corresponding points are involved into interpolation.

Above two methods of determination of standard elevation value, compared with the traditional algorithm, if used in their respective applicable conditions, they will be more reasonable and accurate, and more effectively curbs excessive filtering due to undue selection of standard elevation value.

Adaptive Adjustment of Tolerance of Elevation Difference.

With the progress of iteration, in some windows there still exits a notable difference in height, indicating that filtering is still needed, while other windows have reached its perfect condition that nearly all the non-ground points has been removed. However, the traditional way, when facing the above situation, continues to filter all the grids with a unified threshold, which would surely contribute to the excessive filtering, decreasing the accuracy of filtering.

To approach this problem, this paper puts forward the concept of tolerance of elevation difference to fit different windows. When the elevation difference in some window has dropped below the tolerance, meaning that the points left are almost ground points, this exact window would be skipped of filtering, and the algorithm automatically deems all the points left as ground points. After several times of iteration, when all the windows are skipped of filtering because of the tolerance, the accuracy would become stable and it is time for the algorithm to end.

The slope line of surveyed area is simplified into a straight line in Fig. 1. This paper divide slope equidistantly along the x direction into n windows W_i ($i = 1, 2, ..., n$), and then select W_k ($1 < k < n$) as an example window. The max elevation of the surveyed area is marked as ΔZ. Z_p which is the max elevation of W_k is defined as follows:

$$Z_p = \Delta Z \times (n - k) / n . \tag{1}$$

The elevation of lowest point in W_k is recorded as H_{min}. The tolerance of elevation difference H_t of W_k is defined as follows:

$$H_t = \frac{n-k}{n} \times \Delta Z - H_{min} .$$

(2)

By using tolerance of elevation difference, excessive filtering is prohibited effectively. In sloping area, the effect of this model is even more significant.

Setting of Dynamic Threshold.

The determination of thresholds in sliding window filtering algorithm is crucial to the results: a high threshold could prevent some low-rise buildings being removed, while a low threshold could dim the terrain, losing all the features that should be kept [10]. Predecessors paid great attention to the rules on threshold determination by the gradient-based changeable thresholds [11] and auto-detection of fracture lines [12]. The determination of threshold is tightly associated with the difference of elevation within the window. The value of elevation difference in a window is referred to as ΔH. Traditionally, a static threshold value T is set as $T = M\% \times \Delta H$ for every window, with M being fixed with a constant I ranging from 50 to 60, leading to the result that, for every iteration, $(100 - M)\%$ of the points of each window are bound to be removed even if they are probably ground points.

To solve this problem, dynamic thresholds are put forward. There are two aspects of the dynamic threshold changes. One is that the threshold is increasing when the iteration goes on, and the other is that the speed of this increase slows down, that is, the increment of thresholds should decrease for each iteration. This paper set M that determines threshold as an increasing variable, with a decreasing increment when the iterations go on. The iteration time is marked as N, and the increment of M of every iteration is considered as R_i ($i = 1, 2, ..., N$). The total increment of M is fixed at $(100 - I)$ via several interpolations:

$$\sum_{i=1}^{N} R_i = 100 - I .$$

(3)

R_i are combined with unit weights $\Delta\rho$ whose number is $(N - i)$, are defined as $R_i = (N - i) \times \Delta\rho$. Calculating $\Delta\rho$ and getting R_i finally as follows:

$$R_i = (N - i) \times \frac{100 - I}{\sum_{i=1}^{N} (N - i)} .$$

(4)

Using the dynamic threshold, with the increase of the number of iterations, if the threshold increases appropriately, the fault that ground point is regarded as non-ground points can be reduced, so as to achieve the goal of excessive filtering resistance.
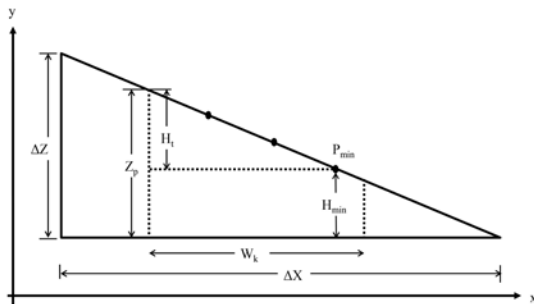

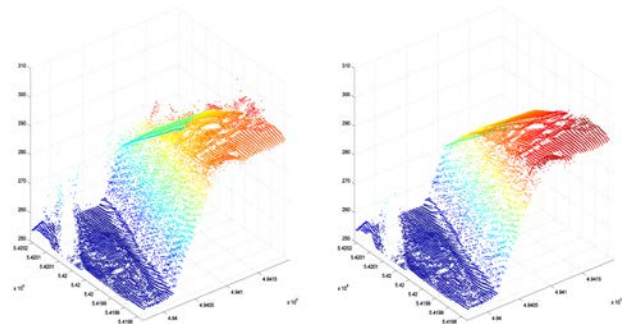Fig. 1.Schematic diagram of the longitudinal section in the sloping fields.

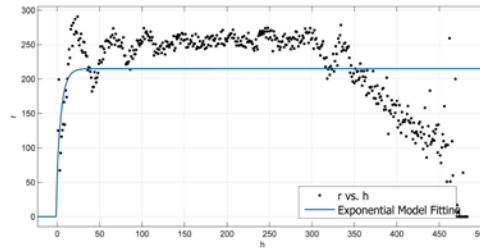Fig. 2. Contrast between Sloping Fields before and after Filtering.

Fig. 3. Half variance function fitting measurement results in sloping fields with vegetation.

Experimental Results

The standard data provided by ISPRS is selected to test the algorithm, indicating two areas, one featuring the sloping fields with vegetation (abbreviated as S) and another one featuring the flat fields with buildings (abbreviated as F). S which containing 42477 points, has an intensity of 0.67 point/m^2 and the gap between points in sloping fields is 2 to 3.5 meters. F which containing 17845 points, has an intensity of 0.18 point/m^2. The following three experiments are aimed to test and evaluate respectively the availability and error of the model, according to the three major parameters: the standard elevation value, the tolerance of elevation difference and the dynamic threshold.

In the figures of experiment results, the sloping field is marked with rectangles and the flat field is marked with circles. In the experiments, the original division of window is 20×20 and after every iteration, the division times of length and width would both double.

In Fig. 2, the first figure on the left is a 3D visualization of original cloud point data, while the second figure on the right is another one of the after-processed results using the Anti-Excessive Filtering Model.

Evaluation on Standard Elevation Value.

In the experiment, take sloping fields for example, The Range of the model is approximately 21 meters in Fig. 3. Traditional model of standard elevation determination (TSE) and optimized model of standard elevation determination (OSE) are used to filter the data. As control, the other two methods are determinations with interpolation only (ISE) and with the nearest neighbor only (NSE). As the Fig. 4 shows, the solid lines represent the results of TSE. The dotted lines represent OSE. The dash-dotted lines represent ISE and the pointed lines represent NSE. Spatial distances and spatial correlation between standard height points are taken into consideration. When the time of iteration reaches 2, the distances among different standard elevation points are getting closer to the Range, which is the requirement of interpolation. Thus as iteration goes on, interpolation is having more obvious advantage and promoting the accuracy. In the other hand, the accuracy of OSE is better than ISE or NSE, indicating that integrated model is superior than using one of them only.

The results indicate that in those areas with big variance in elevation, it is more effective using the distance-based and correlation-based model to determine the standard elevation value, to prohibit excessive filtering and promote accuracy, comparing with the traditional way only using the lowest point as the unified standard point for the whole grid. Furthermore, while the iteration goes on and the divisional windows are getting smaller, interpolation will have a more obvious effect on the accuracy. But it is not necessary to use optimized model in flat fields because the traditional model can get the high accuracy too, and the effect of interpolation is not evident.

Evaluation on Tolerance of Elevation Difference.

Fig. 5 reports the differences in sloping and flat fields, caused by varied tolerance of elevation difference (TED) marked by lines of various patterns. The solid, dotted, dot-dash and dashed lines represent no TED, TED=1, 2 and 3 respectively. In sloping fields, the accuracy culminates when the tolerance of height difference is fixed at 2 meters, an approximate optimum number derived from the formula better than the tolerance is fixed at 3 meters or 1 meter. When the tolerance is not adopted, the data shows that accuracy slides down. In flat fields, the algorithm performs the best with the tolerance fixed at 1 meter which also equals to the theoretical value. When tolerance is not adopted,

accuracy slides down, too. Besides, the best tolerance is higher in the sloping fields than that in the flat fields which is explicable because it is demonstrated by the formula that there is significant positive correlation between optimum tolerance of elevation difference and slope.

Additionally, other parameters under the same conditions, it is better for the algorithm with tolerance of elevation difference to filter a sloping field than to do a flat field. The experiments have proved that the introduction of tolerance of elevation difference has a better performance in the areas with more topographic rises and falls. Overall, adding tolerance of elevation difference contributes to the stability of the algorithm and the improvement of accuracy.

To sum up, adding tolerance of elevation difference con- tributes to the stability of the algorithm and the improvement of accuracy.

Evaluation on Dynamic Threshold.

The improved models mentioned above, with the optimized standard elevation datum and the tolerance of elevation difference, is marked as IM. To go a step further, this paper then adopts dynamic thresholds to achieve further optimization. In Fig. 6, IM have been used to filter data and the corresponding results have been recorded by the dotted lines. Then, this paper introduces dynamic thresholds on this base, and the results are recorded by the solid line. As control, filtering using dynamic thresholds exclusively is marked as dash-dotted line.

From the results of experiments, if dynamic thresholds are added, the accuracy has a growth. In the opposite, without the IM base above, the dynamic thresholds alone could not prevent the accuracy from sliding downwards. Thus, it can be concluded that dynamic thresholds can further promote the accuracy of algorithm in condition that mentioned model is added. They all together consist of the entire and complete model of window-based LiDAR anti- excessive filtering. With all these detailed methods working together, the robustness and accuracy are guaranteed.
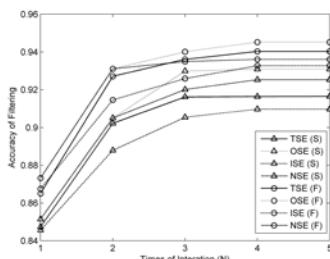
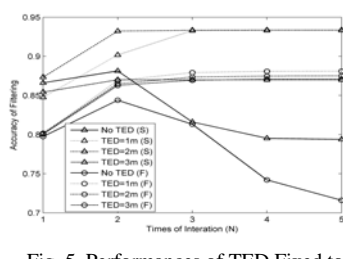Fig. 4. Performances of TSE and OSE in Sloping and Flat Fields.

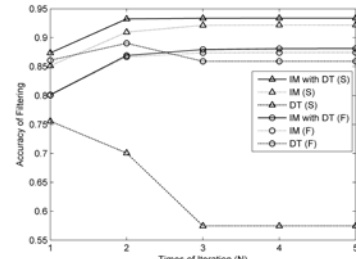Fig. 5. Performances of TED Fixed to 1, 2 and 3, and without TED in Sloping and Flat Fields.

Fig. 6. Performances of DT, IM with DT and IM in the sloping and flat fields.

Summary

This paper introduces an Anti-Excessive Filtering Model Based on Sliding Window. Beginning with the pros and cons between the traditional ways to determine the standard elevation value and the improved one based on the spatial correlation analysis, the paper then talks about how the tolerance of elevation difference determined by slope influences filtering accuracy, as well as iteration times, subsequently illustrating the necessity of dynamic thresholds and figuring out the change law of the increment of the threshold.

The optimized rules on determining standard elevation value can increase the accuracy of filtering in sloping fields significantly, but the spatial distance-based theory is not working evidently when applied to the flat fields. Tolerance of elevation difference prevents excessively filtering by skipping some windows with negligible elevation difference, maintaining the accuracy at a relevant higher level, but some non-ground points like returns from low-rise buildings and artificial statues, would be mistaken as ground points. Dynamic thresholds, changing with various terrains, should be adopted to further improve the model based on the mentioned factors. Generally, the results of experiments have shown that the optimized model can effectively prevent the case of excessive filtering, and achieve the relatively higher accuracy and robustness.

References

[1] X. Meng, N. Currit, and K. Zhao: Ground filtering algorithms for airborne LiDAR data: a review of critical issues, Remote Sensing, vol.2, no. 3 (2010), p. 833–860.

[2] K. Zhang et al: A progressive morphological filter for removing non- ground measurements from airborne LiDAR data, IEEE Transactions on Geoscience and Remote Sensing, vol. 41, no. 4 (2003), p. 872–882.

[3] X. Huang, H. Li, X. Wang and F. Zhang, Filter algorithms of airborne LiDAR data: review and prospects, Acta Geodaetica et Cartographica Sinica, vol. 38, no. 5 (2009), p. 466–469.

[4] P. Axelsson: Dem generation from laser scanner data using adaptive tin models, International Archives of Photogrammetry and Remote Sensing, vol. 33, no. B4/1; PART 4 (2000), p. 111–118.

[5] J. Yu, G. Zhang, P. Li, K. Qin and H. Yang: Research on filter processing of LiDAR data, in Geoinformatics: Remotely Sensed Data and Information (2006).

[6] S. Hosseini, H. Arefi, and Z. Gharib, Filtering of LiDAR point cloud using a strip based algorithm in residential mountainous areas, ISPRS- International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 1 (2014), p. 157–162.

[7] W. Tobler: Smooth pycnophylactic interpolation for geographical regions, Journal of the American Statistical Association, vol. 74, no. 367 (1979), p. 519–530.

[8] W. Beers and J. Kleijnen: Kriging interpolation in simulation: a survey, in Proceedings of the IEEE 2004 Winter Simulation Conference, vol. 1 (2004).

[9] R. Haining: Spatial data analysis, Cambridge University Press Cambridge (2003).

[10] N. Shu: Laser Imaging, Wuhan University Press (2005).

[11] H. Lee and N. Younan, Dtm extraction of LiDAR returns via adaptive processing, IEEE Transactions on Geoscience and Remote Sensing, vol. 41, no. 9 (2003), p. 2063–2069.

[12] W. Schickler and A. Thorpe, Surface estimation based on LiDAR, in Proceedings of the ASPRS Annual Conference (2001), p. 23–27.