

A FILTERING METHOD OF AIRBORNE LASER SCANNER DATA FOR COMPLEX TERRAIN

Hiroshi Masaharu^{a,*} and Kazuyuki Ohtsubo^a

^a Geographical Survey Institute, Kitasato-1, Tsukuba-shi, 305-0811 Japan
masaharu@gsi.go.jp

Commission III, Working Group III/3

KEYWORDS: Laser scanning, LIDAR, DEM/DTM, Surface, Urban, City, Three-dimensional, Model

ABSTRACT:

A new filtering method of airborne laser scanner data has been developed that is suitable for complex terrain, that is flat but with artificial depression such as openings of underground tunnels, typical in highly developed urban areas. The method firstly selects lowest points in small patches from laser scanner data. Then inappropriate points for ground estimation is removed based on statistics within a buffer around the point. The method was tested and gave fairly good result with the laser scanner data of Shinjuku, Tokyo where many skyscrapers and openings to underground tunnel exist.

1. INTRODUCTION

Airborne laser scanner (ALS) data is becoming widely used for various applications. The original data obtained by ALS express the surface of ground objects, not only the ground surface but also trees and roofs of buildings. These data are called digital surface model (DSM). For many applications, it is necessary to distinguish these ground objects and make digital terrain model (DTM) that expresses the ground elevation by removing trees and buildings from DSM. This process is called "filtering." Therefore automatic filtering processing is a very important research issue.

Many researches have been done. Kraus and Pfeifer (1998) developed a filtering method suitable for wooded and hilly areas. Petzold et al. (1999) noted that filtering can be done by applying minimum filter iteratively by changing filter size. Axelsson (2000) uses TIN. Lohman et al. (2000) use dual-rank-filtering. Vosselman (2000) uses slope as criteria. Briese and Pfeifer (2001) use hierarchical approach.

We started to develop filtering methods aiming at 3D city modeling application. At first we developed two-stage method to estimate DTM (Masaharu et al., 2001). In the first step, high buildings are deleted comparing with the existing DTM. We expected to delete large building roof data from the DSM by this step. In the second step the lowest points in 15 m square patches are selected and interpolated.

But we experienced that this filtering algorithm resulted in erroneous estimation around an opening of an underground tunnel. This showed that this kind of algorithm using lowest points is vulnerable to the anomalies appearing in the direction of lower height. But artificial depressions of this kind are rather common in highly developed urban area. Therefore we aimed at developing filtering methods applicable to highly developed urban areas with complex city structure and multilevel land use.

In this paper, a filtering method is presented together with the result applied to the ALS data of Shinjuku subcenter of Tokyo Metropolitan area.

2. METHOD OF THE FILTERING

2.1 Outline of the method

The basic idea is that we employ the lowest points in regularly divided patches of the area (primary selection) but further selection is applied to these points according to whether they are within one sigma from the mean of the neighbouring lowest points of patches. If the point is judged out of the range by this statistical test, the point is removed (secondary selection). This secondary selection is applied three to four times. Through this process, points on the roofs of buildings and points at the bottom of underground tunnels that may remained in the primary selection are both removed.

2.2 Process of the filtering method

The process of the method is as follows.

- (1) To divide the target area into regular interval patches (for example, 50 m or 20 m squares) and select the lowest height data points in each patch. The x , y , z coordinates of the points are recorded. This primary selection can be applied not only regular grid type DSM but also randomly distributed DSM with x , y , z coordinates (point cloud). This means the filtering method is applicable to the randomly distributed point data of ALS.

As the patch size is not large enough to avoid selecting points on the building roofs, roof points as well as points taken from opening of underground tunnel are included in the primary selection result. These points are removed in the following processes.

* Corresponding author

- (2) For each selected point, a point buffer with a predefined radius (for example, 100 m or 50 m) is generated and the mean and the standard deviation of height of points (Here we refer to only the selected points in the previous process.) inside the buffer are calculated. If the center point of the buffer is out of the range of one sigma in both lower and higher directions, then this point is removed. But the removal is done after this processing is done for all the points. This means that the points that are judged to be removed are also used for the calculation of the statistics of all other points. We call this the secondary selection.
- (3) The secondary selection is repeated until change of statistic values become small. This is done three or four times in most cases.
- (4) An approximate DTM (ground surface) is made by interpolating elevation of every regular grid point from the remaining points of the above processing. Then this approximate DTM is compared with the original DSM and if the difference is below a predefined threshold (for example, 1 m) the original DSM points are selected and these points make the DTM data points. Otherwise, the location is marked as no data. To make the final DTM, the no data points are interpolated using the selected DSM data.

3. TEST RESULTS

3.1 Data used

The data was taken by a RAMS laser scanner on a fixed wing airplane. The test area is Shinjuku subcenter of Tokyo Metropolitan area with 1.19 km² area. Random point data were used for the test. Average point density of the measurement is 0.25 point/m². The data was provided by the PASCO Corporation.

3.2 Filtering of the test area

3.2.1 Result of 50 m patches and 100 m buffers: The results when the patch size of primary selection is 50 m and the buffer size of secondary selection is 100 m, are shown in Table 1 and Figure 1. The total number of the observed points in the test area was 300,000. From the total, 455 points (1.5%) were selected in the primary selection. After three secondary selection processes, 260 points remained. The highest elevation after primary selection is 56.275 m. This means that no roof points of skyscrapers of the test area was selected if we use 50 m patch size. The lowest elevation of -8.682 is eliminated by the first secondary selection process. We consider that this means the secondary selection worked effectively to remove anomaly points. There is no large difference between the second and third secondary selection results. Therefore it almost converged in the second iteration.

No. of secondary selection	No. of points	elevation		
		average	maximum	minimum
0	455	33.191	56.275	-8.682
1	364	33.683	42.412	20.728
2	298	33.918	40.029	24.459
3	260	33.999	40.022	25.915

Table 1. Statistics of filtering process of 50 m patches and 100 m buffers

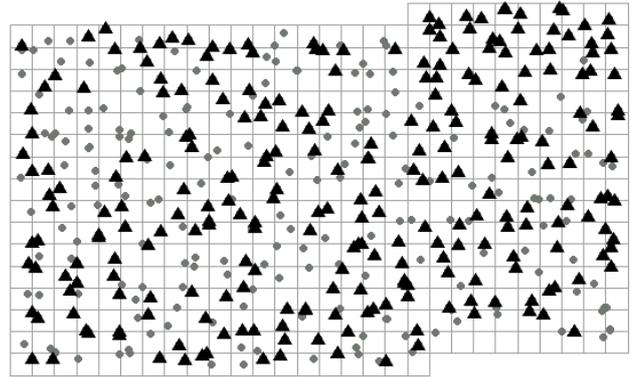


Figure 1. Distribution of selected points

Grid lines represent 50 m patches.

Result of primary selection is shown as small gray disks and black triangles. Remaining points after three secondary selections are shown as black triangles.

3.2.2 Result of 20 m patches and 100 m buffers: To utilize as many laser points as possible for the estimation of DTM, we tested 20 m patch size (Table 2 and Figure 2). In total 2704 points (9.0%) are selected by the primary selection. Maximum elevation of 254 m is included among these points. This means that some patch(es) are completely inside the roof area of high-rise buildings. The highest point is removed by the first secondary selection. We carried out the iteration four times but we consider it has nearly converged in the third iteration. Figure 2 shows the result after three iterations.

No. of secondary selection	No. of points	elevation		
		average	maximum	minimum
0	2704	40.838	254.448	-8.682
1	2425	36.470	71.216	9.950
2	2019	36.414	47.267	27.266
3	1600	36.576	42.412	30.325
4	1308	36.767	41.093	31.717

Table 2. Statistics of filtering process of 20 m patches and 100 m buffers

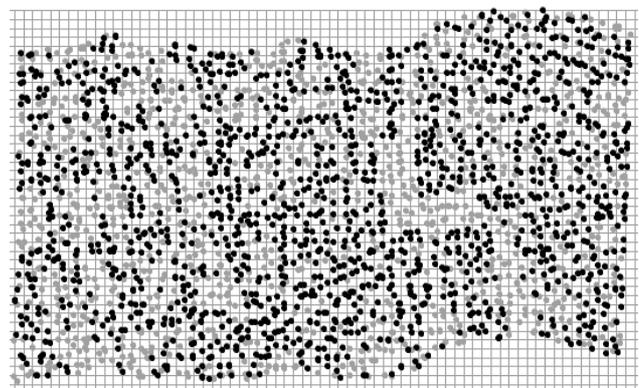


Figure 2. Distribution of selected points

Grid lines represent 20 m patches.

Result of primary selection is shown as small gray disks and black disks. Remaining points after three secondary selections are shown as small black disks.

3.2.3 Approximate DTM: The points selected by the previous process represent estimation of the laser points lying on the ground. The approximate DTM is made by smoothly interpolating these points. We made regular grid type DTM with 2 m interval by weighted mean of six nearest points of black points in Figure 2, where weights are inversely proportional to distances.

3.2.4 Comparison of DSM with approximate DTM and making final DTM: Then original DSM is compared with this approximate DTM. The nearest DSM point to the 2 m interval grid is used for the comparison. Figure 3 shows the difference between DSM and approximate DTM. Figure 4 shows the points whose difference are within ± 1 m. The mean difference of these points was 0.327 m and the standard deviation was 0.359 m. Laser-observed points extracted here are the estimation of the points reflected on the ground. The number of the points was 64092. These points are considered the final estimation of the DTM and when necessary “no-data” area can be interpolated from surrounding points.

The original DSM and the result of the filtering are shown as bird’s-eye views in Figure 5 and Figure 6.



Figure 3. Difference of DSM and approximate DTM
Smaller difference is shown in lighter color. Note that “no-data” areas, especially in the shadow of buildings, are also expressed in white.



Figure 4. Laser-observed points whose height difference from the approximate DTM is within ± 1 m

4. DISCUSSION

4.1 Patch size and buffer size

From the comparison of results of sections 3.2.1 and 3.2.2, the mean height is lower for 50 m patch than for 20 m patch. It was also found from histogram of the difference between DSM and

approximate DTM that approximate DTM distributed lower than the DSM. This can be expected because the approximate DTM is made from the lowest points within patches. From these observations, we consider 20 m patch is better to represent DTM than 50 m patch. One problem of 20 m patch is that there are more chances to take points on the building roofs than 50 m patch. But this problem is solved by applying secondary selection and is not a real problem.

We compared buffer size of 100 m and 50 m. Patch size was fixed as 20 m in this comparison. It was found that some points on the roof could not be removed when we use 50 m buffer size. We judged this from the fact that the maximum elevation did not change after second and third iterations. Therefore we consider buffer size must be larger than 50 m for this area and 100 m is better choice.

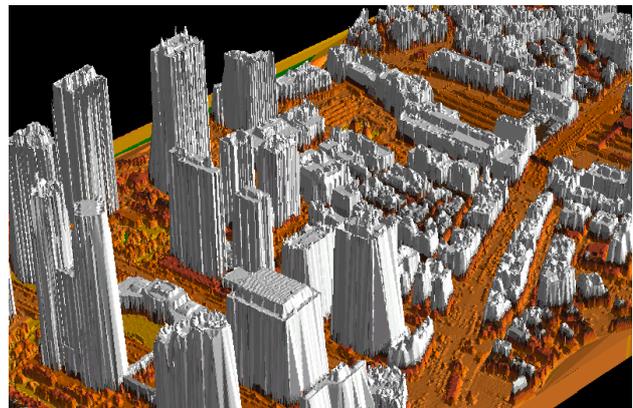


Figure 5. Bird’s-eye view of laser scanner DSM at Shinjuku, Tokyo (before filtering)

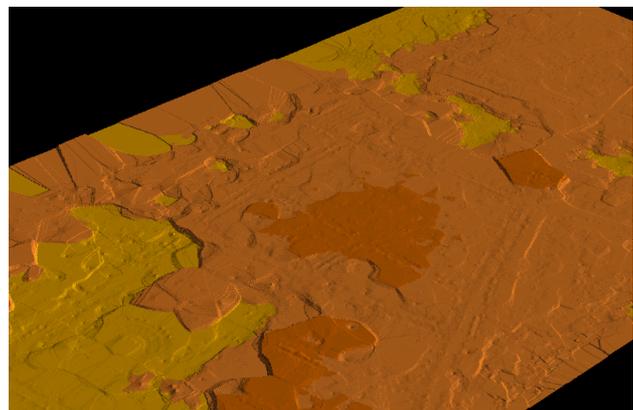
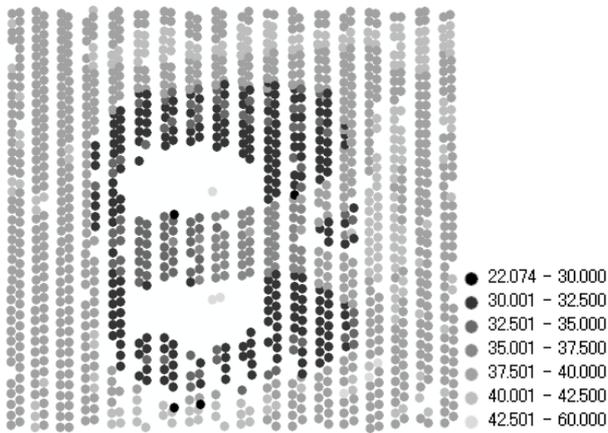


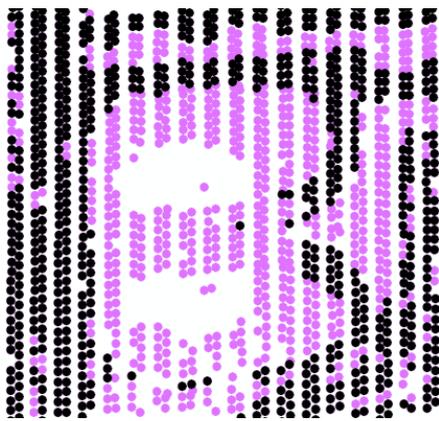
Figure 6. Bird’s-eye view of DTM of the same area from the same viewpoint (after filtering)

4.2 Effectiveness of the filter to remove tunnel openings

We confirmed the filter can remove tunnel openings from estimation of ground surface. Figure 7 shows a portion around a tunnel opening near the railway station as (a) DSM points in grey scale and (b) filtered result. Figure 8 shows the location of this tunnel opening and quantitative analysis test area (Section 4.3) on the 1:10,000 topographic map “Shinjuku” published by the Geographical Survey Institute.



(a)



(b)

Figure 7. Filtering effect at around a tunnel opening
 (a) DSM points measured by ALS, height shown in grey scale
 (b) Filtering result; black: ground surface, light purple: others

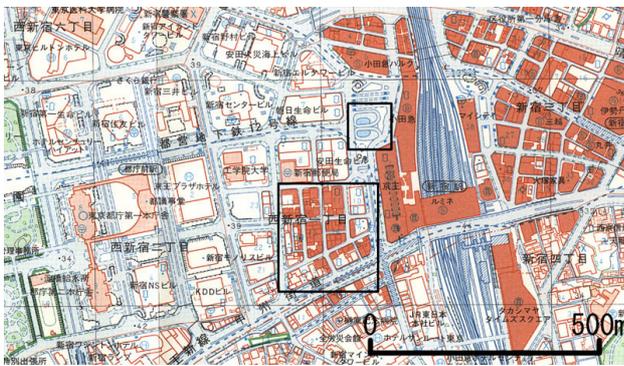


Figure 8. A portion of 1:10,000 Topographic Map “Shinjuku” showing the whole test area for filtering

The rectangles show the test areas for sections 4.2 and 4.3

4.3 Quantitative Analysis

In order to make quantitative analysis of the filtering result, we compared the filtering result with elevation of surface that is considered to represent true ground surface. The latter surface was made manually selecting as many points as possible that are interpreted to be on the ground from the ALS observed DSM points. The test was carried out for an area shown as the lower rectangle in Figure 8.

The test area consists of 8534 ALS observed points. 337 points that represent ground surface were selected manually. The mean interval was about 10 m within ground area. On the other

hand, the filtering method gave 103 points as representing ground surface in the test area. These are the result of secondary selection. Figure 9 shows distribution of manually selected points and points given by the filtering. The mean height of manual selection is 39.95 m and that selected by the filtering is 39.71 m. The difference between them is 0.24 m.

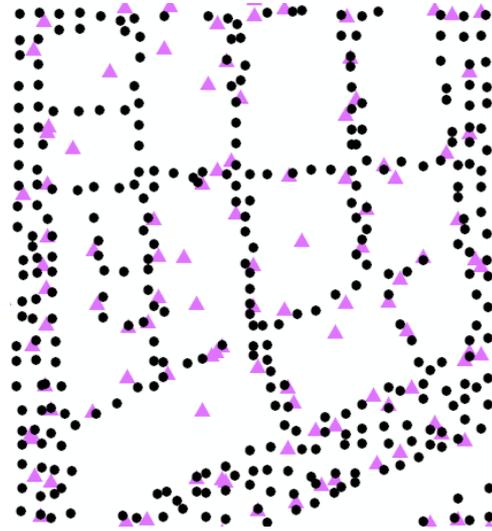


Figure 9. Ground points of manual selection and filtering selection

- (a) Manual selection: black circle point
- (b) Filtering selection: light purple triangle point

Interpolated surface were made from these selected points and all DSM points were classified according to the height difference from the surface. Table 3 shows this classification for both manually selected (true) ground and for the filtering (interpolated surface of points after secondary selection).

Height difference	Manually selected ground surface (a)	Approximate DEM by the filtering (b)	Difference (a) - (b)
Below -1 m	21	20	1
-1 to +1 m	2753	2646	107
+1 to +3 m	508	573	-65
Over +3 m	5252	5295	-43
Total	8534	8534	0

Table 3. Number of ALS observed points according to height difference from estimated terrain by manual selection and the filtering method

The mean height difference explains why some DSM points are classified to different categories. The thresholds of course affect the classification. We used ± 1 m as threshold considering the measurement accuracy of the ALS and heights of ground objects.

To summarize, 108 points out of 2753 points were misclassified not to be on the ground. Error ratio was 3.9%. But there was no error of interpreting building roof points as ground points. On the other hand, one “underground” point was misclassified as ground point. Finally, the mean height difference between extracted points by both methods (points height difference of which are within ± 1 m) was 2.5 cm. This means that misclassified 107 points did not cause very large height differences in the result.

4.4 Characteristics of the proposed method

Basic filtering principle is to smoothly interpolate locally lower points because laser scanner DSM includes buildings and trees that appear higher than the ground height. But this is vulnerable to the anomaly points that appear in lower direction. In the proposed method, we first use lowest points in the patches but after that the processing is symmetrical in both higher and lower directions. This was effective to avoid distortion caused by lower anomalies.

The interpolated approximate DTM results in a little lower height on average. But this is used to judge whether the laser DSM points are on the ground or not. Therefore the final DTM uses the DSM points that is judged to be on the ground. This process corrects the deviation appeared in approximate DTM.

The secondary selection uses mean and standard deviation in a buffer. This means that the terrain is assumed to be flat. In order to apply this method to non-flat area, it would be necessary to calculate trend surface of the terrain and subtract it from the DSM before applying this filtering method. This is an issue of future study. In conclusion, the proposed filtering method is suitable for the urban area with complex structure and the terrain is basically flat.

5. CONCLUSION

A filtering method has been developed that is suitable for urban area with complex structures on flat terrain. The method was tested at the Shinjuku subcenter of Tokyo Metropolitan area and gave fairly good result. We applied this method to another city and used it as a basis for deriving 3D city model.

This method is only applicable to flat terrain at the moment. It is a future study issue to improve the method for hilly area.

REFERENCES

- Axelsson, P., 2000. DEM generation from laser scanner data using adaptive tin models. In: *International Archives of Photogrammetry and Remote Sensing*. Vol. XXXIII, Part B4, Amsterdam, The Netherlands, pp. 110-117.
- Briese, C., Pfeifer, N., 2001. Airborne laser scanning and derivation of digital terrain models. *Optical 3-D Measurement Techniques V*, Vienna, Austria, pp. 80-87.
- Kraus, K. and Pfeifer, N., 1998. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol.53, pp.193-203.
- Lohmann, P., Koch, A., Schaeffer, M., 2000. Approaches to the filtering of laser scanner data. In: *International Archives of Photogrammetry and Remote Sensing*. Vol. XXXIII, Part B3, Amsterdam, The Netherlands, pp. 540-547.
- Masaharu, H., Koarai, M. and Hasegawa, H., 2001. Utilization of airborne laser scanning in Japan. In: *Proceedings of SPIE*, Vol. 4309, San Jose, USA, pp.81-92.
- Petzold, B., Reiss, P. and Stössel, W., 1999. Laser scanning—surveying and mapping agencies are using a new technique for

the derivation of digital terrain models. *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol.54, pp.95-104.

Vosselman, G., 2000. Slope based filtering of laser altimetry data. In: *International Archives of Photogrammetry and Remote Sensing*. Vol. XXXIII, Part B3, Amsterdam, The Netherlands, pp. 935- 942.