FILTERING PHOTOGRAMMETRIC POINT CLOUDS USING STANDARD LIDAR FILTERS TOWARDS DTM GENERATION

Z. Zhang [1,*], M. Gerke [2], G. Vosselman [1], M. Y. Yang [1]

[1] Dept. of Earth Observation Science, Faculty ITC, University of Twente, Enschede, The Netherlands - (z.zhang-1, george.vosselman, michael.yang)@utwente.nl
[2] Institute of Geodesy and Photogrammetry, Technical University of Brunswick, Germany - m.gerke@tu-bs.de

Commission II, WG II/3

KEY WORDS: Point Cloud Filtering, Digital Terrain Models (DTMs), Dense Image Matching, Accuracy Evaluation

ABSTRACT:

Digital Terrain Models (DTMs) can be generated from point clouds acquired by laser scanning or photogrammetric dense matching. During the last two decades, much effort has been paid to developing robust filtering algorithms for the airborne laser scanning (ALS) data. With the point cloud quality from dense image matching (DIM) getting better and better, the research question that arises is whether those standard Lidar filters can be used to filter photogrammetric point clouds as well. Experiments are implemented to filter two dense matching point clouds with different noise levels. Results show that the standard Lidar filter is robust to random noise. However, artefacts and blunders in the DIM points often appear due to low contrast or poor texture in the images. Filtering will be erroneous in these locations. Filtering the DIM points pre-processed by a ranking filter will bring higher Type II error (i.e. non-ground points actually labelled as ground points) but much lower Type I error (i.e. bare ground points labelled as non-ground points). Finally, the potential DTM accuracy that can be achieved by DIM points is evaluated. Two DIM point clouds derived by Pix4Dmapper and SURE are compared. On grassland dense matching generates points higher than the true terrain surface, which will result in incorrectly elevated DTMs. The application of the ranking filter leads to a reduced bias in the DTM height, but a slightly increased noise level.

1. INTRODUCTION

As basic topographical data, Digital Terrain Models (DTMs) are widely used in ortho image rectification, scene classification, 3D reconstruction, etc. Currently, DTMs can be obtained by airborne laser scanning (ALS), digital photogrammetry and interferometric synthetic aperture radar (InSAR) (Chen et al., 2016). During the last two decades, much effort has been paid to filtering the ALS points and obtaining DTMs. DTMs are derived by point cloud filtering followed by interpolation. The second method for DTM generation is aerial photogrammetry. The 3D object coordinates are obtained by matching two or more overlapping images, for instance by dense image matching (DIM). The resulting point clouds can also be used as the basis for DTM production.

While the technique of DTM generation from ALS data is relatively mature after 20 years of development, it is still valuable that we look into the technique of DTM generation from aerial imagery. Taking the Netherlands as example, normally, a period of five years is required to update the whole national DTM using ALS data. In contrast, aerial images over the country are obtained yearly. Therefore, generating DTM from aerial imagery can significantly shorten the interval for data updating.

Advances in aerial image quality and dense matching techniques provide the feasibility of extracting high quality DTMs from aerial images. Firstly, aerial images are obtained with higher radiometric quality. On-board GPS and Inertial Measurement Unit (IMU) allow to obtain more and more accurate orientation elements for bundle adjustment. Development in dense matching algorithms, e.g. Patch-based Multi-View Stereo (PMVS) (Furukawa and Ponce, 2010) and Semi-global Matching (SGM) (Hirschmüller, 2008) makes it possible to obtain accurate point cloud. nFrames SURE states that the vertical accuracy of their products can be better than 1 pixel. Pix4Dmapper ("*Pix4D*" are used below) also reports 1-3 GSD vertical accuracy. The evaluation based on roof segments in (Zhang et al., 2017) also confirms that the vertical accuracy achieved by Pix4D is better than 2 GSD. These numbers give rise to the assumption that it is possible to generate accurate DTMs from dense matching points.

The aim of this paper is to study whether the standard Lidar filters can be used to filter DIM points towards DTM generation. Some previous studies have compared the characteristics of point clouds from laser scanning and dense matching. Accuracy and noise level are the two critical factors that influence the final DTM quality. In the airborne cases, the vertical accuracy of dense matching is usually worse than the accuracy from laser scanning. Compared to the ALS point cloud, the noise level of the DIM data depends on the dense matching algorithm and denoising method (Ressl et al., 2016; Zhang et al., 2017). In ALS points data gaps may appear on wet terrain surface while in DIM points data gaps appear due to failing image matching. These data gaps will cause problems in DTM interpolation.

The paper is structured as follows: In Section 2, we review some work of DTM generation from ALS data and DIM data. Section 3.1 introduces the data and experimental setup. Section 3.2 studies the robustness of standard Lidar filter to DIM noise and artefacts. Section 3.3 evaluates the filtering result on the DIM points in urban scenes. Based on the filtering result in Section 3, Section 4 evaluates the potential DTM accuracy derived from DIM point clouds. Section 5 concludes the paper. The paper not only shows the deficiencies within the DIM points compared to ALS points, but also discusses the research problems related to generating accurate DTMs from DIM points.

2. RELATED WORK

Since the end of 1990s, optical sensors, radar systems and laser scanning systems have been widely used to capture topographic data (Li, 2004). 3D object coordinates are commonly obtained by photogrammetry and laser scanning. DTMs are generated through filtering point clouds and then interpolating on the ground points. It has been a hot research topic to develop robust algorithms for filtering ALS points (Meng et al., 2010; Chen et al., 2017).

Point cloud filtering is the process of discriminating between ground and non-ground points. Generally, the filtering algorithms can be divided into five categories: morphological filtering (Kim and Shan, 2011), surface-based filtering (Kraus and Pfeifer, 1998), progressive TIN (Triangulated Irregular Network) densification (Axelsson, 2000), segment-based filtering (Lin and Zhang, 2014), classification-based filtering (Hu et al., 2016). A quantitative comparison of eight filtering methods can be found in (Sithole and Vosselman, 2004). They found that filtering based on the local surface estimation was generally better than global filtering. Also no filter worked perfectly on various scene complexity. Nowadays, these standard Lidar filters are relatively mature and have already been implemented in many commercial software for laser scanning data processing, e.g. LAStools, SCOP++, Terrasolid.

Recently some studies concerning DTM generation from dense image matching data were published. Among these studies, it is quite common that the filtering operation is run on the DSM instead of on the raw point clouds. The reason is that DSM interpolated from the DIM points is less noisy than the raw points while it still retains a similar accuracy (i.e. the bias level to the ground truth). Perko et al. (2015) and Mousa et al. (2017) filtered DSMs using a Multi-directional and Slope Dependent filtering algorithm. Their DSMs were generated from satellite images and airborne images, respectively. Zhang et al. (2016) filtered a medium resolution DSM from satellite images by using a two-step semi-global filtering method. Beumier and Idrissa (2016) tried to recognize the ground locations from the DSMs using a mean shift segmentation followed by a local regional filtering. In the DTM generation module of Pix4D, the software takes DSMs as input. The ground objects (e.g. buildings and trees) are identified and removed based on the local height gradient. Then the DSM is smoothed and interpolated into the final DTM.

In addition, there are also a few studies filtering the raw DIM points. In general, the standard Lidar filter requires a precise point cloud with little noise as input. Yilmaz and Gungor (2016) compared the effects of five standard filters on the raw DIM points derived from UAV images. Debella-Gilo (2016) filtered the DIM points based on slope-based filtering aided by an existing lower-resolution DTM. However, they didn't report on the noise level of the DIM point cloud or any denoising operation.

Among the studies of generating DTM from photogrammetric point clouds, it is common to use the standard Lidar filtering algorithms or ideas to filter DIM points or DSM. Obviously, the noise level in the point clouds or DSMs has a major impact on the filtering result. However, no study has studied the impact of point cloud noise on the filtering result and thus the final DTM accuracy. In this paper, a comprehensive evaluation of the impact of noise level on the filtering result is implemented. We also evaluate the potential DTM accuracy that can be achieved in case that DIM points are filtered and then interpolated.

3. FILTERING DIM POINTS USING STANDARD LIDAR FILTER

In this section, we present some observations on filtering DIM points using the standard Lidar filter - LASground. The filtering algorithm used in LASground is a modification of the TIN-based approach by (Axelsson, 2000) [1,2]. The lowest points at the initial grid cells in the point cloud are selected as seed points; and then TIN facets are built using these seed points. The coarse TIN surface is densified with the remaining points by judging distance and angle - related criteria. LASground is widely used to filter ALS point cloud. It has been used to create DTM from photogrammetric DSM [3]. In contrast, in this paper it is used to filter the raw photogrammetric point cloud. The research question is whether LASground can be used to filter point cloud from dense matching in which there are usually more random noise than in the Lidar data.

3.1 Study Area and Experimental Setup

The study area lies in the city center of Enschede, The Netherlands as shown in Fig. 1. 510 aerial images including 102 nadir images and 408 oblique images were obtained by Slagboom en Peeters in 2011. The Ground Sampling Distance (GSD) of nadir images is 10 cm. Bundle adjustment was run in Pix4D Pix4Dmapper (version 3.2) using the initial exterior orientations (EOs) and 15 evenly distributed GCPs. After bundle adjustment, the same EOs are used for dense matching in nFrames SURE (version 2.1.0.33) and Pix4D, respectively. Some dense matching parameters are set as below: in both software, the image scale is set to 1/2 resolution; the Minimum Model Count (MMC) in SURE is set to 2; the Minimum Number of Matches (MNM) in Pix4D is set to 3. Note that MMC and MNM in the two software are not comparable because the dense matching algorithms in them are different: SURE employs the tube-shape Semi-global matching (tSGM) (Rothermel et al., 2012) while Pix4D employs patch-based multi-view stereo. Our criterion for adjusting MMC and MNM is to balance the noise level and data gap level in the point cloud by visual inspection.

The ALS data of the same area were acquired by FLI-MAP 400 system mounted in a helicopter in 2007. The point cloud density is 10 points/m^2 and the maximum systematic error in height is

Figure 1. Orthoimage of the study area. The two regions within the yellow rectangles are used in Section 3.1. The region within the red rectangle is used in Section 3.2. The potential DTM accuracy of the whole area is evaluated in Section 4. The area for the two yellow regions, red regions and the whole study region is 880 m^2, 6624 m^2, 0.04 km^2, 1.6 km^2, respectively.

5 cm (van der Sande et al., 2010). The ALS data will be used as reference when evaluating the filtering result in Section 3.3 and when evaluating the potential DTM accuracy in Section 4.

3.2 Robustness of Lidar Filter to Point Cloud Noise

Similar to DTM extraction from ALS point cloud, we assume that DTM sample points can be obtained from two land cover types: paved (or bare) ground and grassland. In this section, we only select pieces of smooth terrain and homogeneous grassland for evaluating the impact of random noise on the filtering. The filtering effect on the bumpy terrain or other small objects is not studied here. Two homogeneous and smooth regions marked by the yellow rectangles in Fig. 1 are used for tests: the left one is smooth ground paved by concrete; the right one is grassland.

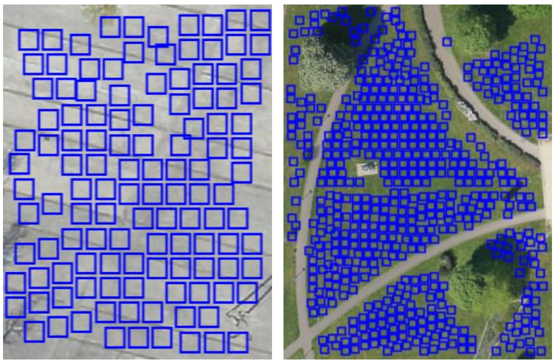


Figure 2. Selected patches on the paved ground (left, 112 patches) and grassland (right, 527 patches) for evaluating the filtering performance. The patch sizes are 2 m × 2 m.

Several parameters in LASground affect the filtering performance. Since our study area is in urban area, the scene is set to "city or warehouses" (i.e. a step size of 25 m) and the parameter for controlling the initial ground points density is set to "default". In addition, we also experimented with the parameters "spike size" and "bulge size". Since the surface of the paved ground and grassland is smooth with little spike (often outliers), these two parameters do not make a difference on the filtering. We also try to adjust the parameter "stddev" which controls the maximal standard deviation for planar patches to be retained. Interestingly, tuning "stddev" did not bring a remarkable change to the filtering result. Therefore, we adopt the "10 cm" suggested by the software.

In order to study the impact of the noise level on the filtering performance, a local evaluation method is used. Square patches of 2 m × 2 m are selected from the ALS data of the area. The patches are selected randomly as evaluating units. The Residuals of Plane Fitting (*RPF*) is calculated using all the points inside the patch.

$$RPF = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \Delta H_i^{\,2}} \tag{1}$$

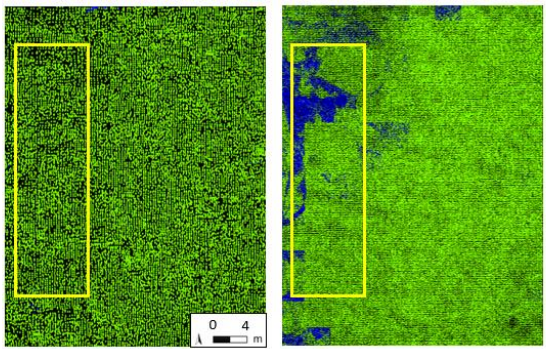
N is the number of points in this patch. ΔH_i is the distance from the i th point to the plane which is fitted to all the points within this patch. The patch will be valid only if *RPF* is smaller than 2 cm. When RPF of the ALS data in a certain patch is smaller than 2 cm, we can say that the terrain in this patch is quite smooth and planar. The patches selected on the paved ground and grassland are shown in Fig. 2. 112 and 527 patches are selected on the paved ground and grassland, respectively. Note that on the grassland in Fig. 2 the patches are all selected on smooth grassland. No patch lies on the bushes or trees. After patch selection, the filtering result and noise level are quantized locally within each patch:

1) Filtering effect: Ideally, all points in every patch in Fig. 2 should be classified as ground points by LASground. In consideration of scarce outliers or misclassifications, if more than 95% of the points within a patch are classified as ground points, we still take it as correct filtering; if the ratio is less than 95%, the filtering in this patch is incorrect.

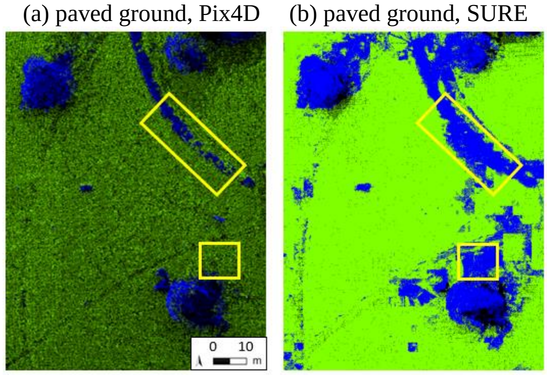
2) Noise level: Height Ranking Range (*HRR*) is used to represent the noise level. It is calculated by sorting the heights of all points within a patch. The *HRR* is obtained by subtracting the *m* percentile from the *n* percentile (*m*<*n*). HRR represents the height range in the vertical direction. Generally, it is robust to blunders in the point cloud. In this paper, m and n are set to 5% and 95%, respectively.

The filtering results from LASground are shown in Fig. 3. In Fig. 3(a-d), the percentage of correctly classified patches is 100%, 80%, 100% and 89%, respectively. LASground performs very well on Pix4D point cloud because the point cloud is precise with little noise. Compared with filtering Pix4D point cloud, Fig. 3(d) shows that filtering SURE point cloud meets more difficulty along the bush and in the shadow. The SURE point cloud is much noisier than Pix4D and this brings problems during filtering.

In order to evaluate the robustness of LASground to point cloud noise, the distribution of the HRR values for all the patches correctly filtered are shown in Fig. 4. The HRR values in the four histograms range from approximately 0.05 m to 0.40 m which indicates that LASground performs well in filtering a point cloud



(a) paved ground, Pix4D (b) paved ground, SURE



(c) grassland, Pix4D (d) grassland, SURE

Figure 3. Filtering results on the paved ground and grassland. The green indicates the identified ground points; blue indicates non-ground points; black indicates data gaps. In (c) and (d), blue indicates identified non-ground points not only on the grassland, but also on the trees and bushes (cf. Fig. 2). Generally, the Pix4D point clouds in Fig. 3(a) and (c) are darker than SURE point clouds in Fig. 3(b) and (d) due to a lower point density.

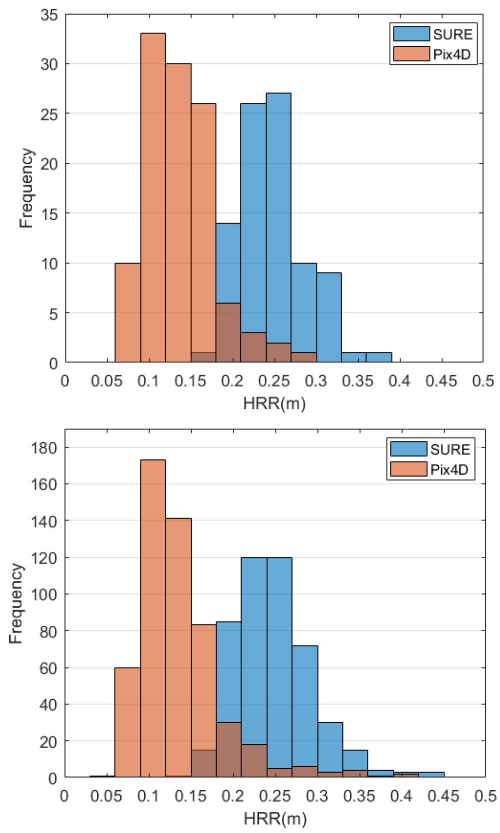


Figure 4. HRR Distribution for all the correctly filtered patches. Bin width is 3 cm. Top: paved ground; Bottom: grassland. The dark brown between the blue and light brown histograms is the overlap of the two histograms.

with a HRR smaller than 0.40 m. In addition, the mean of HRRs for paved ground-Pix4D, paved-ground-SURE, grassland-Pix4D, grassland-SURE are 0.14 m, 0.24 m, 0.14 m, 0.24 m, respectively. This indicates that the noise level of the dense matching point clouds on paved ground and grassland are the same, for either Pix4D or SURE. To the best of our knowledge, the noise level of the point cloud from SURE is dependent on the image quality, image overlapping, orientation accuracy and dense matching algorithm. SURE does not implement any post-processing on the dense matching point cloud.

Now we study the patches which are wrongly filtered, i.e. less than 95% points within the patch are classified as ground points. Fig. 5 visualizes the HRR values of these wrongly filtered patches. The color coding from blue to red indicates that the HRR increases. HRR in these wrongly filtered patches ranges from 0.2 m to 0.59 m. The right figure of Fig. 5 shows that DIM point cloud from SURE is relatively noisy and contains more artefacts in the shadow than other areas. So these areas in the shadow are challenging for LASground.

Fig. 6 shows the two profiles on paved ground and grassland drawn in Fig. 5 (along the yellow lines). Checking the orthoimages and laser points shows that the profile in the left paved ground of Fig. 5 is smooth ground with no bumps or spikes. The profile in the right grassland of Fig. 5 is the grassland in shadow. The length of the point cloud profile is approximately 2 m and the vertical depth is 20 cm. Fig. 6 shows that some artefacts exist in the SURE point cloud. Note that the blue points and green points together form the SURE points. In the top figure of Fig. 6, the ALS point cloud distributes between the “ground points” and “non-ground points” identified by LASground. The

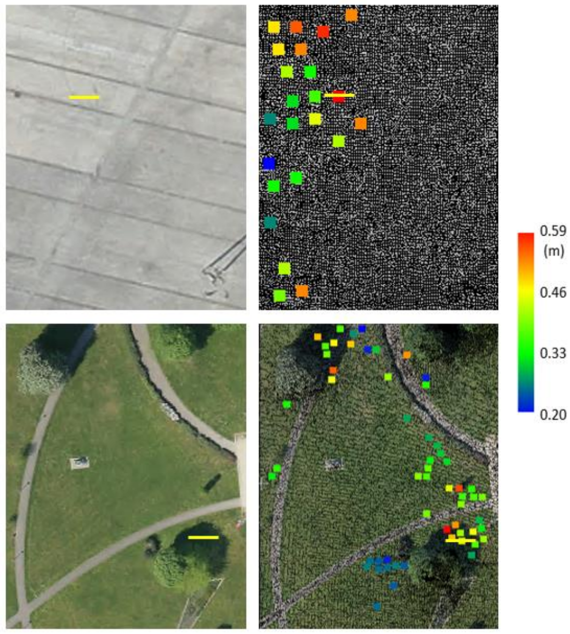


Figure 5. Visualization of the HRR values for the wrongly filtered patches in the SURE points. Top: 23 patches on the paved ground; Bottom: 58 patches on the grassland.

HRR is about 0.5 m. As the higher DIM points are classified as non-ground, the average height of the ground points shows a bias w.r.t. the average height of the ALS points.

In the bottom figure of Fig. 6, hollow space can be found inside the SURE points and the points show two layers. LASground simply takes the points in the top layer as the non-ground points. The HRR is about 0.8 m. Along this grassland profile, the ALS points are located at the bottom of the DIM points.

3.3 Filtering Photogrammetric Points in Urban Scene

In this section, a 0.04 km^2 study area (red rectangle in Fig. 1) is filtered using LASground. This area is mainly covered with buildings, streets, paved ground and individual trees. In some locations, the streets are narrow and covered with shadow. Concerning the filtering parameters in LASground, “step size” shows a large impact on the filtering result: if it is set very large, some roof points will also be taken as ground points. After some trials, we set the parameter according to the scene - “city or warehouses”. That is, the step size is fixed to 25 m in this section.

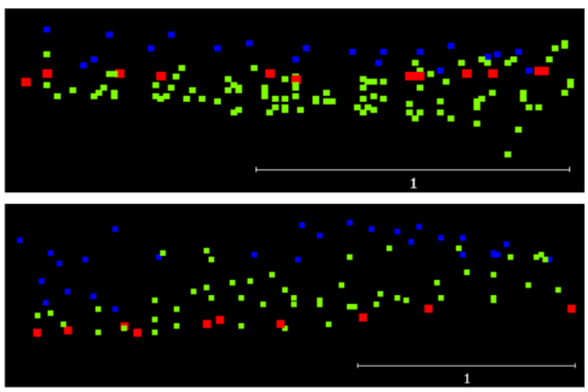
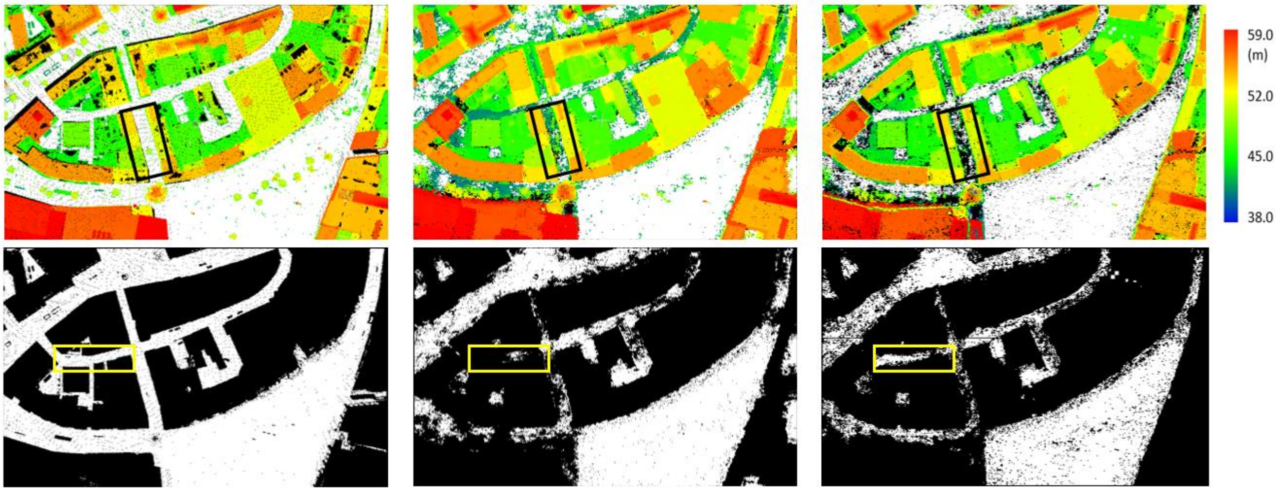


Figure 6. Profiles of three point clouds: ALS points (red), SURE ground points identified by LASground (green) and non-ground points (blue). Top: Profile of the line on paved ground; Bottom: Profile of the line on grassland.



(a) ALS data (b) DIM-raw (c) DIM-RF

Figure 7. Filtering results of a city block. The top row shows both the ground and non-ground points. White indicates the ground points identified by LASground; black indicates data gaps. Non-ground points are colored based on the height value. The bottom row shows only the ground points. The two figures in the first column entitled (a) is the filtering effect of ALS data; (b) shows the filtering effect of the raw point clouds generated by Pix4D; column (c) shows the filtering effect of the Pix4D point cloud processed by a ranking filter. For the meaning of black and yellow boxes, please refer to the text.

Considering the possible artefacts and random noise in the DIM point cloud, a ranking filter is used to refine the raw point clouds. The rationale of ranking filter is to rank the heights of all points within a vertical raster cell. In our case, the median of the heights (i.e. 50% percentile) is taken as the final value assigned to this cell. The cell size is set to 0.5 m × 0.5 m based on heuristics. The cell size should be set small enough to contain sufficient terrain details and should be set large enough to contain points in most cells. If less than 3 points exist in a certain cell, this cell will not be assigned any value but just left empty.

Three point clouds are filtered as shown in Fig. 7: ALS data, raw Pix4D point cloud (DIM-raw), Pix4D point cloud processed by a ranking filter (DIM-RF). We do not present the filtering results of SURE points because the filtering delivers more mistakes when the points are too noisy, especially on the narrow streets. Fig. 7(a) shows the filtering result of ALS data. Building and individual trees are filtered out successfully. The black rectangle shows the filtering result on the narrow street. Here LASground works well.

Fig. 7(b) shows the filtering result of the raw Pix4D point cloud. Dense matching is challenging in shadow area due to poor texture and low contrast in images. Ideally, all the ground points should be labelled as “ground”, including ground points in the shadow. The black rectangle shows the filtering in the shadow. Some points are identified as ground and some are identified as non-ground. In the yellow rectangle, most of the locations are identified as non-ground. Fig. 7(c) shows the filtering result of a Pix4D point cloud processed by ranking filtering.

Fig. 7(b) and (c) show that LASground performs well at filtering individual trees on both the DIM-raw and DIM-RF data, especially on the southeast open square. In the black rectangles, there are more ground locations identified in DIM-raw than in the DIM-RF. This narrow street is located in shadow. Checking the data profile shows that the heights of the DIM points are higher than the real ground surface by approximately 30 cm, and the DIM points are randomly distributed because of remaining matching errors. The DIM-RF identifies fewer ground points than DIM-raw but the identified ground points are more likely to be reliable ground locations.

The yellow rectangles show the filtering effect of a road, which is not in the shadow. LASground filters classified most of the points in Pix4D-raw data as non-ground. In contrast, many locations are taken as ground points in the DIM-RF data.

In both the black and yellow rectangles, LASground tends to deliver better filtering results on the DIM-RF data than the DIM-raw data. It can be explained by the fact that median ranking filter can reduce the noise in the DIM points. The DIM point cloud after pre-processed by a ranking filter is getting more similar to the ALS data in terms of ground representation. Moreover, the noise is removed very considerably and height jumps from ground to above-ground objects are more or less better retained because of the relatively large raster. In this case, LASground can better discriminate ground and non-ground cells because outliers and noise are not affecting the TIN densification step.

Apart from the qualitative comparison above, the filtering results are also evaluated quantitatively using the measures from (Sithole and Vosselman, 2004). The filtering result of ALS data after manual check is taken as the reference. The ALS data and Pix4D-raw data are both 3D while the Pix4D-RF is 2.5D. The filtering result on Pix4D-raw is evaluated as below: Take the surface through the ALS ground points and label the DIM ground points as correct if they are within some margin of the ALS ground surface. To evaluate the 2.5D filtering result, the ALS data are also converted to 2.5D and only the label of the highest point in each bin is taken as the true label. Three quantitative measures are calculated: Type I error is the percentage of bare ground points actually labelled as non-ground points by LASground; Type II error is the percentage of non-ground points labelled as ground points; Total error is the overall statistics of points being wrongly classified. The filtering results are shown in Table 1.

Dataset	Type I	Type II	Total error
DIM-raw	22.3%	5.2%	8.7%
DIM-RF	12.0%	7.0%	8.4%

Table 1. Quantitative evaluation of the filtering results

Table 1 shows that the total error by filtering DIM-raw (8.7%) and DIM-RF (8.4%) are similar. Type I error of DIM-raw is much larger than if DIM-RF is used. The reason is that many ground points on the narrow streets in shadow are misclassified as non-ground points. These DIM points are usually a mixture of real ground points and blunders. LASground will filter out the above points and only the lowest points will be taken as ground points. In addition, the level of Type II errors is smaller than Type I errors. Type II error of DIM-RF is slightly larger than DIM-raw. If we check the filtering effect of individual trees and objects (e.g. chairs and dustbins) on the southeast square in Fig. 7(c), the reason for a relatively high Type II error is that some small objects are smoothed by using a median ranking filter. LASground will classify these locations into ground while the ground truth is non-ground. In contrast, the details of small objects can be better retained in the DIM-raw data. When filtering DIM-raw data, the ground and non-ground points can be better separated.

In summary, the advantage of using a ranking filter on the point cloud is that the filtered point cloud contains less noise. When filtering the points after ranking filtering, LASground performs better in avoiding non-ground points. That is, compared to filtering the raw DIM points, filtering DIM-RF will deliver less ground locations with higher reliability. On the other hand, the disadvantage of using ranking filter is that some low objects may be smoothed. These non-ground locations are thus likely to be misclassified as ground by LASground. In contrast, the details of small objects can be better retained in the DIM-raw data. When filtering the DIM-raw data, the ground and non-ground points can be better separated by LASground.

4. EVALUATING THE POTENTIAL ACCURACY OF DTMS

4.1 Comparison of DTM Accuracy Derived from Pix4D and SURE Point Clouds

The observations in Section 3 indicated that LASground is quite tolerant to the random noise when filtering the DIM points. In particular, all the DIM points on the paved ground, bare earth and grassland are likely to be taken as terrain points by LASground. In this section, we explore the potential accuracy that can be obtained by DTM derived from dense matching. We do not interpolate on the point cloud but we directly calculate the deviation of the DIM point cloud from the reference. The ALS data are taken as reference data and only the vertical accuracy is studied. In the evaluation stage, the square patches of 2 m × 2 m are taken as the evaluation unit. Compared to the point-to-point comparison, the accuracy measures calculated based on each patch are more robust to local blunders and random noise. The study area is the whole region shown in Fig. 1 (1.6 km^2).

First, the ALS data are filtered using LASground. Then, square patches are detected from the ground points. A patch is valid if it meets two conditions: (1) The number of points in this patch is larger than a certain threshold; (2) The RPF (Eq. 1) is better than 2 cm. The patches in shadow are eliminated. The shadow mask is calculated from an orthoimage based on a grayscale histogram (Sirmacek and Unsalan, 2009). Only if all the four corners and the center location of a certain patch lie in the non-shaded locations, the patch will be taken as valid. The selected patches are divided into two categories based on the green index on the ortho image: ground and grassland. Finally 24,634 ground patches and 7381 grassland patches are selected for accuracy evaluation.

After the patches are detected from the ALS point cloud, the DIM points within the square patch boundary in 2.5D space are cropped for evaluation. Concerning a certain patch, a plane is fitted to the ALS points, the mean deviation from the DIM points to the plane is calculated as the accuracy measure as shown in Eq. (2). μ_i denotes the mean deviation between the DIM points and the ALS points for the j*th* patch. i denotes the i*th* patch in the whole study area, j denotes the j*th* point in this patch. There are n_i points in this patch. Δh_{ij} is the distance from the j*th* point to the fitted ALS plane. μ_i is the mean deviation between the DIM points and the ALS points for the j*th* patch.

$$\mu_i = \frac{1}{n_i}\sum_{j=1}^{n_i} \Delta h_{ij} \tag{2}$$

The distribution of mean deviations is shown in Fig. 8. Interestingly, the distribution of the deviations for Pix4D and SURE are quite different even though the same EOs were used for dense matching. Fig. 8 also shows that there is only one peak in the SURE histograms but there are two peaks in the Pix4D histograms. The mean deviation on the ground ranges in [-0.18 m, 0.18 m] for Pix4D data, and ranges in [-0.15 m, 0.15 m] for SURE data. The mean deviation on the grassland ranges in [-0.2 m, 0.2 m] for Pix4D data, and ranges in [-0.15 m, 0.15 m] for SURE data.

In order to make quantitative evaluation of the DIM accuracy in the whole study area, the following two accuracy measures are calculated considering all the patches:

- Mean of mean deviations:

$$\bar{\mu} = \frac{1}{m}\sum_{i=1}^{m} \mu_i \tag{3}$$

- Standard deviation of mean deviations:

$$\sigma_{\mu_i} = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m} (\mu_i - \bar{\mu})^2} \tag{4}$$

$\bar{\mu}$ is calculated by averaging the mean deviations in the whole area. m is the number of patches in the whole study area. The σ_{μ_i} is calculated to represent the standard deviation of the mean deviations from the $\bar{\mu}$. The accuracy measures at the whole block are shown in Table 2.

Dataset	$\bar{\mu}$	σ_{μ_i}
ground-pix4d	0.057	0.056
ground-sure	0.016	0.048
grassland-pix4d	0.078	0.077
grassland-sure	0.030	0.056

Table 2. Accuracy measures of DIM point cloud in the whole block. (Unit: m)

Table 2 shows that $\bar{\mu}$ of SURE point cloud is better than for the Pix4D point cloud on both ground and grassland as could already be seen in the histograms of Fig. 8. In addition, the σ_{μ_i} of SURE point cloud is better than Pix4D point cloud on both ground and grassland.

https://doi.org/10.5194/isprs-annals-IV-2-319-2018 | © Authors 2018. CC BY 4.0 License.

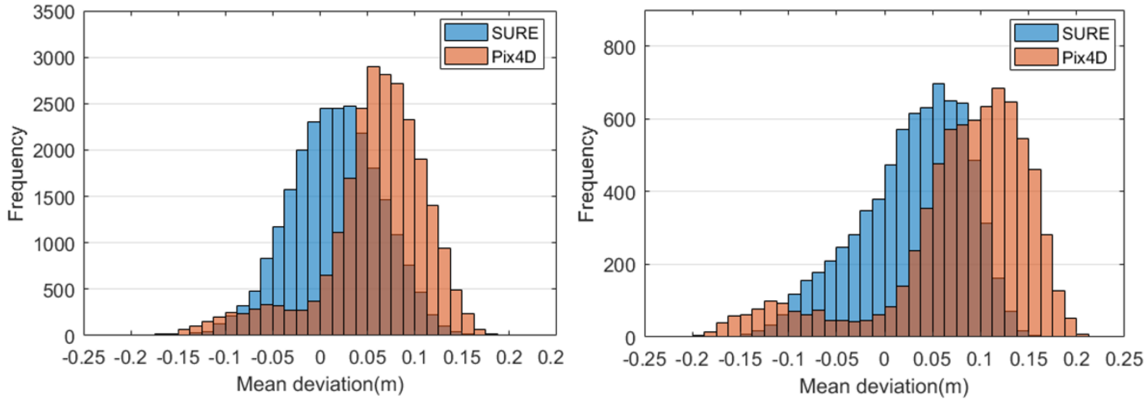


Figure 8. Distribution of mean deviations for the DIM points generated by Pix4D and SURE. Left: 24,634 ground patches; (b) 7381 grassland patches. Note that the dark brown between the blue and light brown histograms is actually the overlapping of the two histograms.

Table 2 also shows that the bias between the DIM data and the ALS data on the grassland is larger than the bias on the ground. That is, the accuracy on the grassland is worse than the ground. This can be explained by that dense matching usually delivers the points on the top surface of the grassland but laser scanning can penetrate the shallow grass and record the points on the real terrain. So the bias on the grassland includes not only the dense matching errors but also the grass height (Ressl et al., 2016).

When filtering the DIM point clouds in the urban scene using LASground, all the points on the ground and grassland will probably be classified as ground points without the negative impact of artefacts. However, the problem is that dense matching will deliver some points higher than the true terrain on the grassland, which will result in incorrect elevated DTMs.

4.2 The Impact of Ranking Filter on The Potential DTM Accuracy

In Section 3, we found that a ranking filter leads to improvements in the ground point filtering. In this section, we check whether the ranking filter would have an impact on the potential DTM accuracy achieved by the Pix4d point cloud. Similar to Section 4.1, the mean deviations for 24,634 ground patches and 7381 grassland patches are calculated and incorporated into the mean of mean deviations $\bar{\mu}$ and standard deviation of mean deviations σ_{μ_i} as shown in Table 3. RF indicates that this point cloud is preprocessed by a ranking filter.

Dataset	$\bar{\mu}$	σ_{μ_i}
ground-pix4d-RF	0.048	0.063
grassland-pix4d-RF	0.067	0.085

Table 3. Accuracy measures of DIM point cloud after pre-processed by a ranking filter. (Unit: m)

Table 3 shows that for both the ground and grassland, when RF is used in a preprocessing step, $\bar{\mu}$ gets improved by around 1 cm. However, σ_{μ_i} increases slightly. That is, when the point cloud is pre-processed by a ranking filter, generally the potential DTM accuracy will improve but the ranking filter will also bring more variation to the DTM errors at the whole photogrammetric level. In addition, we can study the impact of a ranking filter onto the point cloud accuracy by calculating the deviation between DIM-RF and DIM-raw for every patch. Fig. 9 shows the distribution of deviation values for ground patches and grassland patches, respectively. According to statistics, on 13.3% grassland patches and 8.6% patches the deviations between DIM-RF and DIM-raw are larger than 10 cm. The deviation values are relatively small compared to the large patch size (2 m × 2 m). In addition, the deviations between DIM-RF and DIM-raw on the paved ground is generally smaller than on the grassland, which can be explained by the fact that there are usually more artefacts and surface fluctuation on grassland.

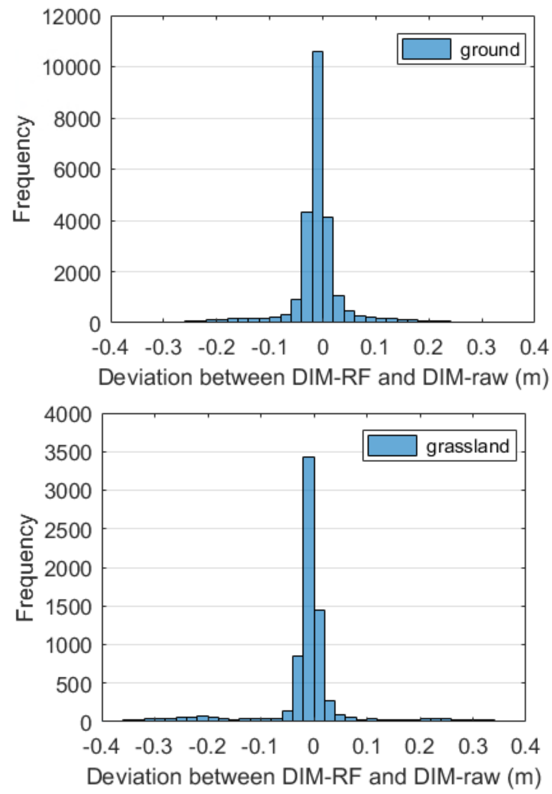


Figure 9. Distribution of deviation between DIM-RF and DIM-raw. Top: paved ground; Bottom: grassland.

5. CONCLUSIONS

This paper studies the question whether the standard Lidar filters can be used to filter dense matching points in order to derive accurate DTMs. Filtering results on the homogeneous ground and grassland show that the filtering performance depends on the noise level and scene complexity. LASground is verified to be relatively robust to random noise. However, filtering algorithms may only select the lower points as ground points in case of a large amount of noise. In addition, artefacts and blunders may appear in the dense matching points due to low image contrast or poor texture (e.g. in the shadow, along the narrow street, etc.). In these cases, LASground will probably classify some noisy ground points as non-ground points. Filtering results on a city block show that LASground performs well on the grassland, along bushes and around individual trees if the point cloud is sufficiently precise. In addition, a ranking filter can be used to filter the DIM point cloud before LASground filtering. LASground will identify fewer but more reliable ground locations. However, a ranking filter will also smooth some ground details so some small objects on the terrain will be filtered out. Since we aim at obtaining accurate DTMs, the ranking filtering shows its value in identifying only reliable ground points.

The accuracy of the point cloud determines the final DTM accuracy. The accuracy of the DIM point clouds is evaluated using a patch-based method. The bias from the reference is studied in the whole study area. Although the same EOs are used for dense matching, the vertical accuracy of SURE point cloud on the ground is better than the Pix4D point cloud. In addition, we also verify that the error on the grassland is larger than the error on the paved ground. We also found that the ranking filter brought only very small deviation to the point cloud. Therefore, the ranking filter might be taken a useful pre-processing tool before filtering noisy photogrammetric point clouds. Future work may focus on modifying the previous Lidar filtering algorithms so that they can be used on relatively noisy DIM point clouds.

REFERENCES

Axelsson, P., 2000. DEM generation from Laser scanner data using adaptive TIN models. Int. Arch. Photogram. Remote Sens. Spatial Inf. Sci., 33, pp. 110-117.

Beumier, C. and Idrissa, M., 2016. Digital terrain models derived from digital surface model uniform regions in urban areas. Int. J. of Remote Sens., 37(15), pp. 3477-3493.

Chen, Q., Wang, H., Zhang, H., Sun, M. and Liu, X., 2016. A point cloud filtering approach to generating DTMs for steep mountainous areas and adjacent residential areas. Remote Sens., 8(1), pp. 71.

Chen, Z., Gao, B. and Devereux, B., 2017. State-of-the-Art: DTM Generation Using Airborne LIDAR Data. Sensors, 17(1), pp.150.

Debella-Gilo, M., 2016. Bare-earth extraction and DTM generation from photogrammetric point clouds including the use of an existing lower-resolution DTM. Int. J. of Remote Sens., 37(13), pp. 3104-3124.

Furukawa, Y. and Ponce, J., 2010. Accurate, dense, and robust multiview stereopsis. IEEE Trans. Pattern Anal. Mach. Intell., 32(8), pp.1362-1376.

Hirschmuller, H., 2008. Stereo processing by Semi-global matching and mutual information. IEEE Trans. Pattern Anal. Mach. Intell., 30(2), pp. 328-341.

Hu, X. and Yuan, Y., 2016. Deep-Learning-Based Classification for DTM Extraction from ALS Point Cloud. Remote Sens., 8(9), pp.730.

Meng, X., Currit N., and Zhao K., 2010. Ground Filtering Algorithms for Airborne LiDAR Data: A Review of Critical Issues. Remote Sens., 2 (3), pp. 833–860.

Mousa, A.K., Helmholz, P. and Belton, D., 2017. New DTM extraction approach from airborne images derived DSM. Int. Arch. Photogram. Remote Sens. Spatial Inf. Sci., pp. 42.

Kim, K. and Shan, J., 2011. Adaptive morphological filtering for DEM generation. In IEEE Geoscience and Remote Sensing Symposium (IGARSS), pp. 2539-2542.

Kraus, K. and Pfeifer, N., 1998. Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS J. Photogram. Remote Sens., 53(4), pp.193-203.

Li, Z., Zhu, C. and Gold, C., 2004. Digital terrain modeling: principles and methodology. CRC press.

Lin, X. and Zhang, J., 2014. Segmentation-based filtering of airborne LiDAR point clouds by progressive densification of terrain segments. Remote Sens., 6(2), pp.1294-1326.

Perko, R., Raggam, H., Gutjahr, K.H. and Schardt, M., 2015. Advanced DTM generation from very high resolution satellite stereo images. ISPRS Ann. Photogram. Remote Sens. Spatial Inf. Sci., 2(3), pp.165.

Ressl, C., Brockmann, H., Mandlburger, G. and Pfeifer, N., 2016. Dense Image Matching vs. Airborne Laser Scanning-Comparison of two methods for deriving terrain models. PFG Photogrammetrie, Fernerkundung, Geoinformation, 2, pp. 57-73.

Rothermel, M., Wenzel, K., Fritsch, D. and Haala, N., 2012, December. Sure: Photogrammetric surface reconstruction from imagery. In Proceedings LC3D Workshop, Berlin (Vol. 8), pp. 1-8.

Sirmacek, B. and Unsalan, C., 2009. Damaged building detection in aerial images using shadow information. Recent Advances in Space Technologies, pp. 249-252.

Sithole, G. and Vosselman, G., 2004. Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. ISPRS J. Photogram. Remote Sens., 59(1), pp. 85-101.

van der Sande, C., Soudarissanane, S. and Khoshelham, K., 2010. Assessment of relative accuracy of AHN-2 laser scanning data using planar features. Sensors, 10(9), pp. 8198-8214.

Yilmaz, C. S. and Gungor, O., 2016. Comparison of the performances of ground filtering algorithms and DTM generation from a UAV-based point cloud. Geocarto Int., pp. 1-16.

Zhang, Y., Zhang, Y., Zhang, Y. and Li, X., 2016. Automatic extraction of DTM from low resolution DSM by two-steps semi-global filtering. ISPRS Ann. Photogram. Remote Sens. Spatial Inf. Sci., 3(3), pp. 249-255.

Zhang, Z., Gerke, M., Peter, M., Yang, M.Y. and Vosselman, G., 2017. Dense matching quality evaluation - an empirical study. IEEE Joint Urban Remote Sensing Event (JURSE), pp. 1-4.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-IV-2-319-2018