

Two-step adaptive extraction method for ground points and breaklines from lidar point clouds

Bisheng Yang ^{a,b,*}, Ronggang Huang ^{a,*}, Zhen Dong ^a, Yufu Zang ^a, Jianping Li ^a

^a State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

^b Beijing Advanced Innovation Center for Imaging Technology, Beijing 100048, China



ARTICLE INFO

Article history:

Received 5 May 2015

Received in revised form 2 June 2016

Accepted 4 July 2016

Available online 21 July 2016

Keywords:

Airborne lidar point clouds

Breakline extraction

Multi-scale morphological filtering

Point cloud segmentation

Segment-based filtering

ABSTRACT

The extraction of ground points and breaklines is a crucial step during generation of high quality digital elevation models (DEMs) from airborne LiDAR point clouds. In this study, we propose a novel automated method for this task. To overcome the disadvantages of applying a single filtering method in areas with various types of terrain, the proposed method first classifies the points into a set of segments and one set of individual points, which are filtered by segment-based filtering and multi-scale morphological filtering, respectively. In the process of multi-scale morphological filtering, the proposed method removes amorphous objects from the set of individual points to decrease the effect of the maximum scale on the filtering result. The proposed method then extracts the breaklines from the ground points, which provide a good foundation for generation of a high quality DEM. Finally, the experimental results demonstrate that the proposed method extracts ground points in a robust manner while preserving the breaklines.

© 2016 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

1. Introduction

Airborne LiDAR is integrated by the global positioning system, the inertial navigation system and the laser scanning sensor, and it could capture the large-scale, dense, three-dimensional point clouds of the earth's surface for many applications, such as generation of digital elevation models (DEMs) and forest investigations. In general, a filtering operation for separating ground points and non-ground points is the first step in the processing pipelines of airborne LiDAR point clouds, particularly the generation of DEMs. In addition, the breaklines should be well extracted for generating high quality DEMs.

Many studies have investigated the filtering of airborne LiDAR point clouds in the last decade (Sithole and Vosselman, 2004; Liu, 2008; Meng et al., 2010; Shan and Toth, 2008; Vosselman and Maas, 2010), but most of filtering methods still require parameters tuning to adapt for various types of terrain (e.g., urban, mountains), and incurring heavy manual editing costs. Therefore, automated filtering still faces great challenges, and a further study

on the filtering is necessary. According to the definitions of entities in Xu et al. (2014), the previously reported filtering methods can be mainly classified into two types, i.e., filtering methods based on point entities and filtering methods based on segment entities.

Filtering methods based on point entities calculate the geometric properties of each point and its neighboring points to determine whether one point belongs to be ground or non-ground, e.g., slope-based filtering (Vosselman, 2000; Sithole, 2001), surface-based filtering (Kraus and Pfeifer, 1998; Axelsson, 2000; Mongus and Žalík, 2012; Mongus et al., 2014), and morphological filtering (Kilian et al., 1996; Zhang et al., 2003; Chen et al., 2007; Cui et al., 2013; Li et al., 2013; Pingel et al., 2013; Mongus et al., 2014). The main differences between these methods are the selection of the geometric properties and the filtering rules employed. For example, slope-based filtering assumes that the gradient between ground points is smaller than that between ground and non-ground points, which yields good quality filtering results in relatively flat areas (Sithole and Vosselman, 2004; Liu, 2008). Surface-based filtering iteratively removes non-ground points or detects ground points based on distances, angles or other measures between points and a reference surface (Kraus and Pfeifer, 1998; Axelsson, 2000; Sohn and Dowman, 2002). Generally, these methods may fail to remove non-ground points with low elevations or to detect ground

* Corresponding authors.

E-mail addresses: bshyang@whu.edu.cn (B. Yang), hrgwhu@gmail.com (R. Huang).

points around breaklines (Meng et al., 2010). To improve the adaptability of the filtering, multi-scale surface-based filtering was proposed (Evans and Hudak, 2007; Mongus and Žalik, 2012; Chen et al., 2013; Hu et al., 2014). However, these methods incur heavy calculation costs. Morphological filtering is robust in detecting some cliffs and exhibits a high time efficiency, but the filtering quality is sensitive to the window size. A large window size may degrade rugged terrain features (e.g., mountain peaks), which is known as the cut-off problem, whereas a small window size may fail to remove large objects. Many studies have tried to reduce the effect of the window size on the final filtering quality (Zhang et al., 2003; Chen et al., 2007; Li et al., 2013; Pingel et al., 2013; Mongus et al., 2014), but morphological filtering is still affected by feature preservation issues in rugged terrain and failures in removing large objects.

Filtering methods based on segment entities firstly partition airborne LiDAR point clouds into segments based on the smoothness constraint (Tóvári and Pfeifer, 2005), the slope (Filin and Pfeifer, 2006), RANSAC (Yang et al., 2013) and so on. Then, the properties of each segment (e.g., shape, size and completeness) and the topological relationships between segments are calculated to remove non-ground segments (Sithole and Vosselman, 2005; Shen et al., 2012; Yan et al., 2012). This type of approach works well in urban areas, especially in terms of the preservation of terrain features. In addition, Zhang and Lin (2013) embedded point cloud segmentation in surface-based filtering to improve the filtering performance of forest areas. However, it is still a challenge of removing non-ground points from the segments mixed ground and non-ground points.

In general, a single filtering method has difficulties to filter airborne LiDAR point clouds with various complex scenes and various terrain types (Sithole and Vosselman, 2004; Podobnikar and Vrečko, 2012). Podobnikar and Vrečko (2012) regionalized the entire test area into four different subareas, where all of the parts were filtered using different filters. Deng and Shi (2013) integrated progressive triangulated irregular network (TIN) densification and a hierarchical robust interpolation method. The fusion of different filtering methods could improve the filtering performance, but this fusion is only based on point entities, and thus this method may still perform poorly when differentiating a cliff from a building in a local area (Sithole and Vosselman, 2005). Fortunately, the fusion of filtering methods based on different entities (e.g. points and segments) can perform better in distinguishing various differences between ground points and non-ground points, thereby improving the filtering performance and preserving the terrain features better. However, the existing filtering methods do not consider the fusion of filtering methods based on different entities.

To generate high quality DEMs, breaklines should be considered as constraints in the interpolating the grid DEMs or fixed edges in the TINs. Breaklines could be classified into jump breaklines, crease breaklines, and curvature breaklines (Brügelmann, 2000). Brügelmann (2000) extracted breaklines based on a range imagery interpolated from ground points. However, the accuracy of breakline extraction is affected by the interpolation of the range imagery. Then, Several researchers extract breaklines from LiDAR point clouds by iteratively intersecting patch pairs within a buffer zone around the approximate breaklines (Kraus and Pfeifer, 2001; Briese, 2004). Briese and Pfeifer (2008) improved the method for extracting different types of breaklines by using different solutions.

In this study, we propose a method to extract ground points from airborne LiDAR point clouds in a robust manner by fusing two filters based on different entities, and to detect multiple types of breaklines from the extracted ground points, thereby laying a good foundation for generating high quality DEMs.

The main contributions of the proposed method are as follows.

- Improve the adaptabilities of the filtering in areas with various types of terrain by classifying points into a set of segments and one set of individual points before removing the non-ground points using segment-based filtering and multi-scale morphological filtering;
- Eliminate the effect of the maximum scale in multi-scale morphological filtering, resulting in improved performance in filtering the areas with rugged terrain; and
- Generate high quality DEMs with the preservation of breaklines (i.e., jump breaklines, crease breaklines, and curvature breaklines).

The remainder of this paper is organized as follows. The proposed method is elaborated in Section 2. In Section 3, the experimental studies were undertaken to evaluate the proposed method. Finally, conclusions are drawn at end of this paper.

2. Two-step extraction method

Fig. 1 illustrates the workflow of the proposed method. The proposed method firstly removes low outliers by the method of Shao and Chen (2008). And, the pseudo-grids with a certain size W_{grid} are generated from airborne LiDAR point clouds based on the literature of Cho et al. (2004). The point of the lowest elevation in each pseudo-grid is labeled as a grid point and the other points are defined as unlabeled points, as shown in **Fig. 2**. When the number of points in a pseudo-grid is zero, a grid point will be interpolated by the points in its neighboring pseudo-grids. Secondly, the proposed method classifies the grid points into a set of segments and one set of individual points, which are processed by a segment-based filtering and an improved multi-scale morphological filtering method respectively, to generate a provisional DEM. Thirdly, the unlabeled points are processed by a back selection procedure based on the provisional DEM, and the final ground points are obtained. Finally, in order to generate high quality DEMs, the breaklines are extracted from the final ground points through the point cloud segmentation.

2.1. Extracting ground points from grid points

2.1.1. Classification of grid points

To better describe the relationship between each grid point and its neighboring grid points, the proposed method employs point cloud segmentation with smoothness (e.g., normal vector and residual) constraints (Tóvári and Pfeifer, 2005) to classify the grid points into a set of segments and one set of individual points. Each segment is labeled as ground or non-ground, each point in the set of individual points will be further processed to be labeled as ground or non-ground one by one. Based on the method of Tóvári and Pfeifer (2005), the grid points are firstly segmented into disjoint segments $S = \{S_1, S_2, \dots, S_{n-1}, S_n\}$, where n is the number of segments. **Fig. 3a** and b illustrate the grid points and the corresponding results of point cloud segmentation. It shows that some areas (e.g. vegetation areas) are covered by small segments. However, grid points of a small segment are more likely to be heterogeneous, and the segment should be removed from S to decrease the omission and commission errors in the filtering. Hence, the number of points in each segment is calculated. Then, the proposed method applies a threshold N_t to classify all grid points into two classes. If the number of grid points in one segment is more than N_t , the segment is retained; otherwise, this segment is removed from S , and grid points of the segment are expressed by point entities. In general, N_t equals to the minimum area of the remaining segments divided by the square of the grid size (W_{grid}). After this classification, two classes are generated, a set of segments consist

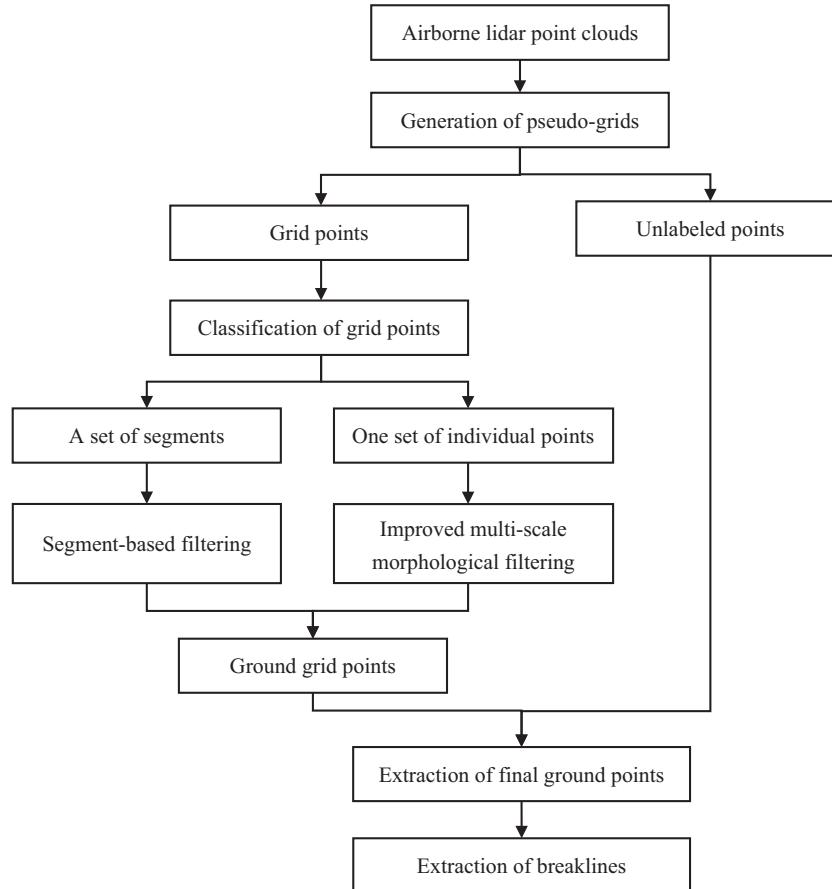


Fig. 1. Workflow of the proposed method.

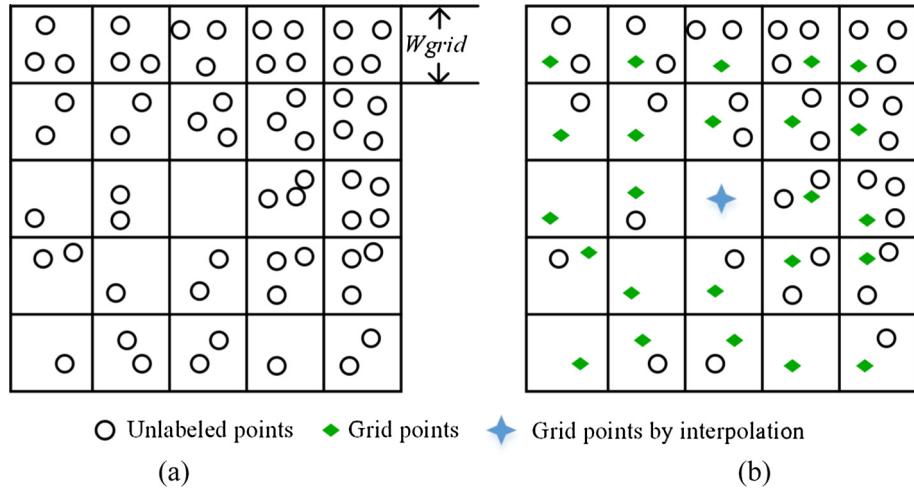


Fig. 2. Generation of pseudo-grids. (a) Unlabeled points in pseudo-grids. (b) The determination of grid points.

of $S_{homo} = \{S_1, S_2, \dots, S_{k-1}, S_k\}$, where k is the number of the segments, and one set of individual points, as shown in Fig. 3c.

2.1.2. Identification of ground grid points based on two filtering methods

After point clouds are classified into two sets, the proposed method uses segment-based filtering based on the segment entity to extract ground segments from the segments, and an improved multi-scale morphological filtering based on the point entity is

employed to detect ground grid points from the individual points. It is because that segment-based filtering could well preserve the discontinuous terrain and terrain details in the smooth areas (i.e., the set of segments), but it is not suitable in the areas where there are many small segments, i.e., the areas of the individual points. Therefore, an improved multi-scale morphological filtering is used to process the individual points. In this way, the proposed method could make use of the advantages of two filtering methods.

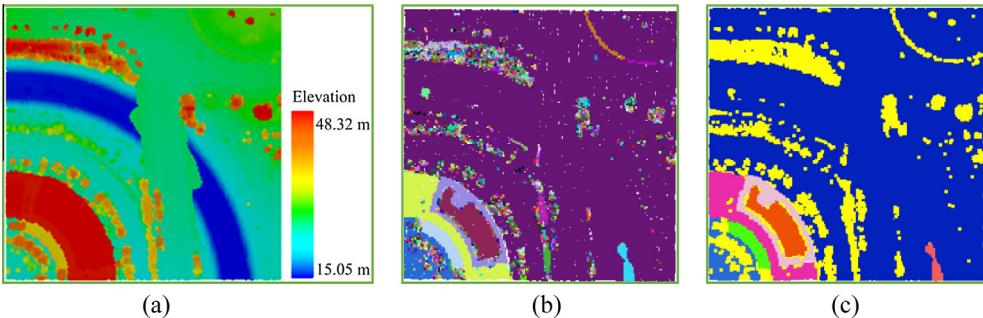


Fig. 3. Classification of grid points. (a) Grid points. (b) Segmentation of grid points, each segment is dotted in one color. (c) The result of classification for the grid points, where the segments are dotted in different colors, and the individual points are dotted in yellow. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

- Segment-based filtering for the segments.

The segments S_{homo} comprise ground and non-ground segments. To remove non-ground segments, segment-based filtering is applied. In this filtering, a provisional DEM is generated as the reference surface, and each segment will be labeled as ground or non-ground according to the elevation differences between boundary points of the segment and the provisional DEM.

First, the ground seed points are detected from all grid points to generate a provisional DEM as the reference surface. The proposed method utilizes a square window (the size is denoted as W_{seed}) to move in the horizontal or vertical direction over the grid points with a step of half of W_{seed} . The grid point with the lowest elevation in each sliding window is regarded as a ground seed point. Theoretically, W_{seed} should be larger than the object with the maximum size. However, W_{seed} is difficult to be estimated and may be slightly smaller than the size of the maximum object, and thus a few points on the large objects may be erroneously classified as ground seed points, called false ground seed points. Therefore, the proposed method employs the mean and standard deviation of elevations and slopes in the neighboring points to iteratively remove false ground seed points. Then, if one ground seed point belongs to one segment, all the grid points in the segment are classified as ground seed points. Based on the extracted ground seed points, a provisional DEM (Z_{rtm}) with a resolution of W_{grid} could be generated by the inverse distance weighting (IDW).

Second, the boundary points of each segment are extracted, and the elevation difference between each boundary point and the provisional DEM is calculated. After generating the provisional DEM, the α -shape method (Edelsbrunner and Mücke, 1994) is used to extract the boundary of each segment. The boundaries of segments comprise the set $ND = \{ND_1, ND_2, \dots, ND_i, \dots, ND_k\}$. $ND_i = \{p_1, p_2, \dots, p_j, \dots, p_{ik}\}$ represents points of one boundary, and p_j is a boundary point of the segment, ik is the number of boundary points in the segment ND_i . And then, the proposed method traverses points of each boundary one-by-one to calculate the elevation difference between the point and the provisional DEM Z_{rtm} . The elevation differences of the points in one boundary comprise the set $ED_i = \{h_1, h_2, \dots, h_j, \dots, h_{ik}\}$. Hence, the elevation differences in all of boundaries comprise the set $ED = \{ED_1, ED_2, \dots, ED_i, \dots, ED_k\}$.

And finally, according to the fact that points on the boundary of a ground segment will generally have small elevation differences (Chen, 2009), we propose the following rules to distinguish ground segments from non-ground segments:

$$\left\{ \begin{array}{l} GS = \{S_i \in S_{homo} \mid \max(ED_i) < h_1 \text{ and } \text{Num}(ED_i < h_2) > 50\% * \text{Num}(ND_i)\} \\ OS = \{S_i \in S_{homo} \& S_i \notin GS\} \end{array} \right. \quad (1)$$

where GS denotes ground segments, OS denotes non-ground segments, h_1 and h_2 are thresholds, $\max(ED_i)$ is the maximum elevation

difference of ED_i , $\text{Num}(ED_i < h_2)$ is the number of boundary points less than h_2 , and $\text{Num}(ND_i)$ is the total number of boundary points in the segment ND_i . Empirically, h_1 is about 0.5–0.8 m, and h_2 is about 0.2–0.4 m. h_2 describes only some boundary points of the segment attached to ground, which is less than h_1 . Fig. 4 shows an example of the segment-based filtering. Fig. 4a shows 7 unclassified segments. Fig. 4b illustrates the result of ground seed points selection, 4 segments are classified as ground segments, and the elevation differences of points in the boundary of each unclassified segment are calculated. Fig. 4c shows the filtering result according to Eq. (1). All elevation differences of boundary points in segment A are larger than h_1 , it thus is classified as an object. All the elevation differences of boundary points in segment B are less than h_1 , then it is classified as a ground segment. The elevation differences of a lot of boundary points in segment C are less than h_2 , it thus is classified as a ground segment.

- Improved multi-scale morphological filtering for the individual points

Multi-scale morphological filtering is a filtering method based on the point entity, and it could be used to rapidly remove non-ground points from the individual points by analyzing the difference between a point and its neighboring points (Zhang et al., 2003). However, the maximum scale is still a non-trivial issue, although many studies have tried to address this problem (Zhang et al., 2003; Chen, 2009; Pingel et al., 2013; Mongus et al., 2014). If the maximum scale needs to be larger than the size of each amorphous object (e.g., unpenetrated large trees) in the individual points, then the cut-off problem will affect the quality of filtering. In addition, the maximum size of amorphous objects is also difficult to be estimated because of the indistinct boundaries of amorphous objects. To overcome the cut-off problem, the proposed method selects a small value for the maximum scale W_{scale} , and to firstly remove the amorphous objects larger than the maximum scale W_{scale} according to the following steps.

Step 1: Apply an open operator with a disc window (the diameter is W_{scale}) to smooth the individual points. In general, W_{scale} ranges from 5 m to 8 m. For example, the tree points become smoother after applying the open operator, as shown in Fig. 5a.

Step 2: Segment the smoothed individual points using the smoothness constraints (Tóvári and Pfeifer, 2005) to obtain a set of segments.

Step 3: Extract the boundary of each segment and remove non-ground segments according to Eq. (1), as illustrated in Fig. 5b.

By following these steps, the points of each amorphous object larger than W_{scale} are filtered. Next, except for the points of removed amorphous objects, the elevations of the other individual points are recovered to the corresponding elevations before apply-

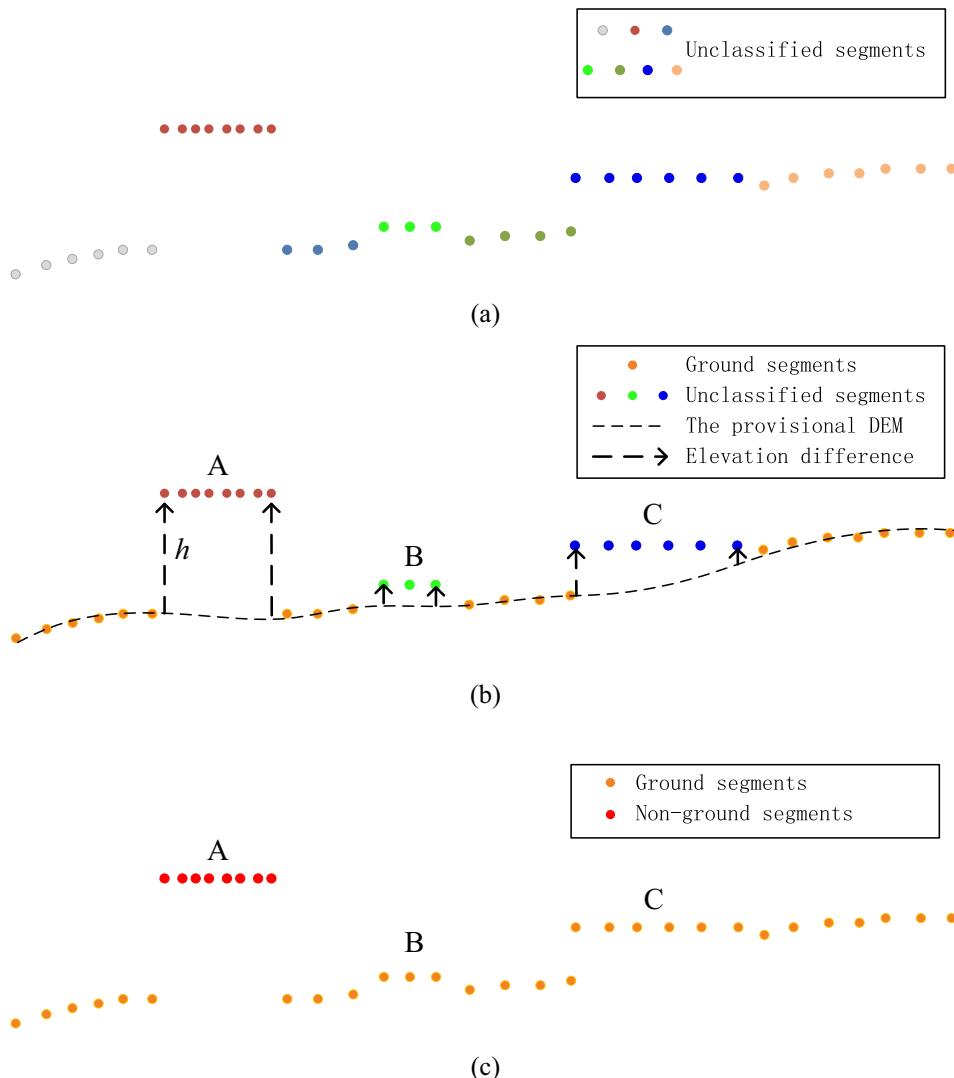


Fig. 4. An example of segment-based filtering. (a) Seven unclassified segments. (b) Generation of the provisional DEM and computation of the elevation difference. (c) The result of segment-based filtering.

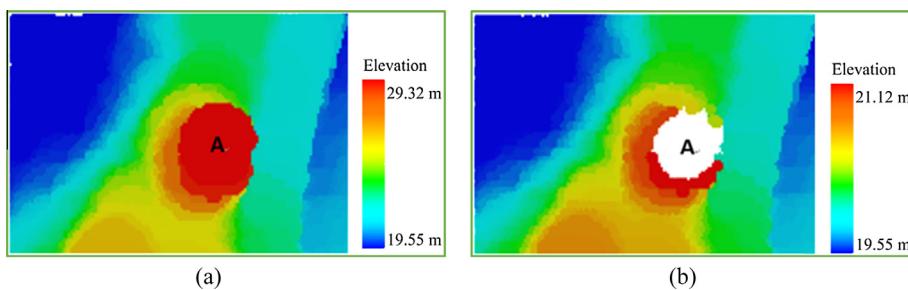


Fig. 5. Removing an amorphous object from the individual points. (a) The open operator is applied to the region of individual points. (b) The obtained result of removing the amorphous object.

ing the open operator. Finally, the multi-scale morphological filtering described by Pingel et al. (2013) is used to filter the non-ground points with the maximum scale of W_{scale} .

2.2. Extracting the final ground points

After the grid points are filtered using segment-based filtering and multi-scale morphological filtering respectively, some ground

points have been obtained. However, the unlabeled points (including ground and non-ground points) in the generation of pseudo-grids should be processed further. First, the provisional DEM is generated using the extracted ground points with the IDW method. Then, the unlabeled points are assessed one by one according to Eq. (2), which is inspired by the methods of Chen et al. (2007) and Pingel et al. (2013):

$$h = h_v * \cos(s) \quad (2)$$

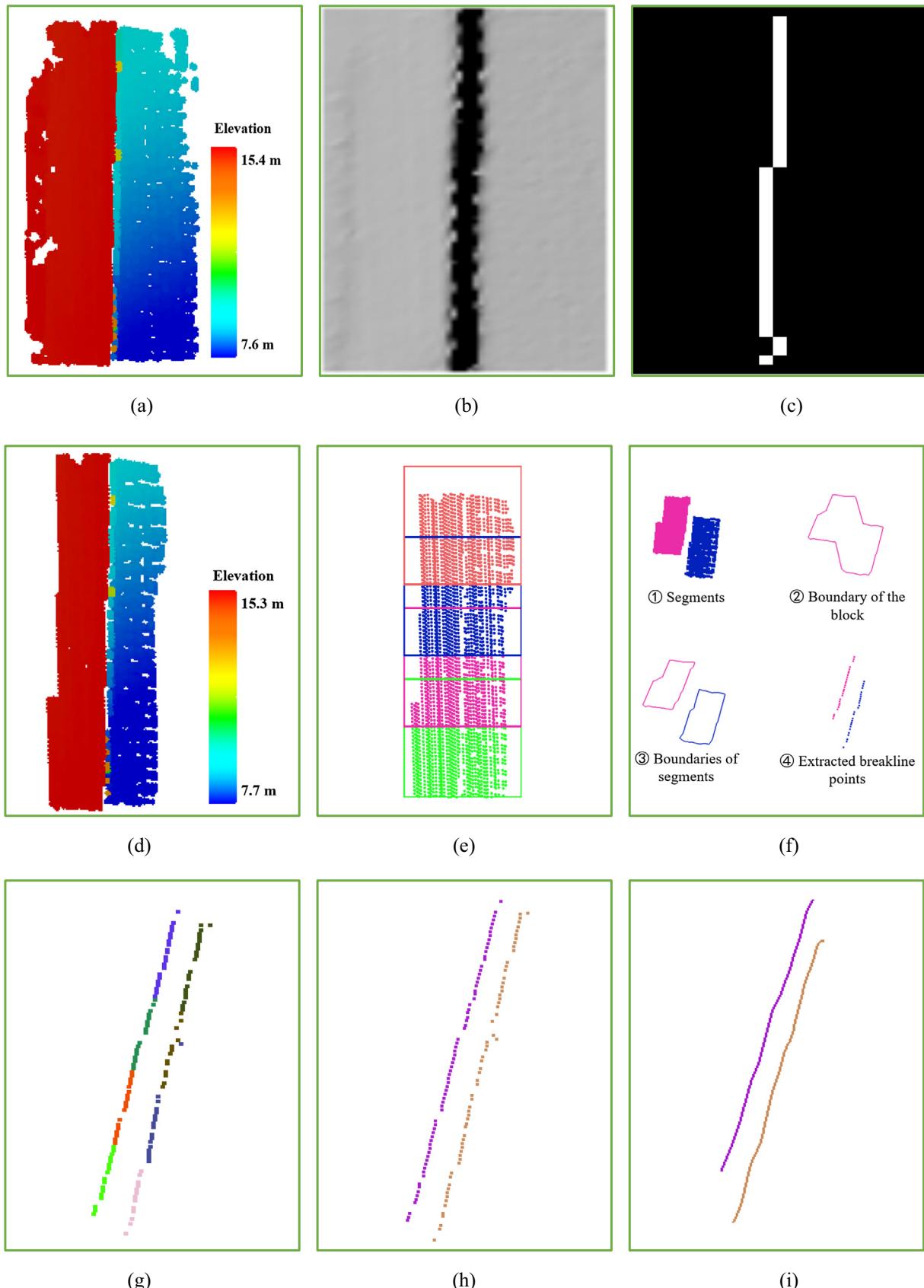


Fig. 6. An example of extracting breaklines from ground points. (a) The extracted ground points. (b) The generated range image. (c) The approximate 2D breakline. (d) The candidate points around the approximate breakline. (e) Generation of the overlapping blocks. (f) The process of the breaklines extraction from one block. (g) The extracted breaklines of all blocks. (h) The linked breaklines across blocks. (i) The fitted breaklines.

Table 1

Parameters settings for the 15 sample datasets.

Parameters	W_{grid} (m)	N_t (pts)	W_{seed} (m)	h_1 (m)	h_2 (m)	W_{scale} (m)	h_3 (m)
Values	1.0	10	25	0.5	0.3	6.0	0.3

where h_V is the vertical distance from an unlabeled point to the provisional DEM, and s is the terrain slope derived from the provisional DEM.

If the elevation difference h of an unlabeled point is larger than a threshold h_3 , the point is labeled as a non-ground point. The threshold h_3 is related to the coordinate errors and the point span, and it is specified as about 0.2–0.4 m empirically.

2.3. Extracting breaklines from the ground points

The extracted ground points are firstly interpolated as a range image to extract the approximate breaklines using a LoG operator (Brügelmann, 2000). For example, Fig. 6a and b show the ground points of a terrain area and the corresponding range image. Fig. 6c shows the extracted breaklines from the terrain area. The points neighboring the extracted breaklines are labeled as the candidate points with a buffering operator (Fig. 6d), where the width of the buffering operator ranges from 2 m to 5 m. Next, the candidate points are split into a set of overlapping blocks. The length W_{len} of the block is specified as about 10 m and the ratio of overlapping ranges from 30% to 50%. Fig. 6e illustrates the overlapping rectangle blocks dotted in different colors. The breaklines in each block are extracted according to the following steps.

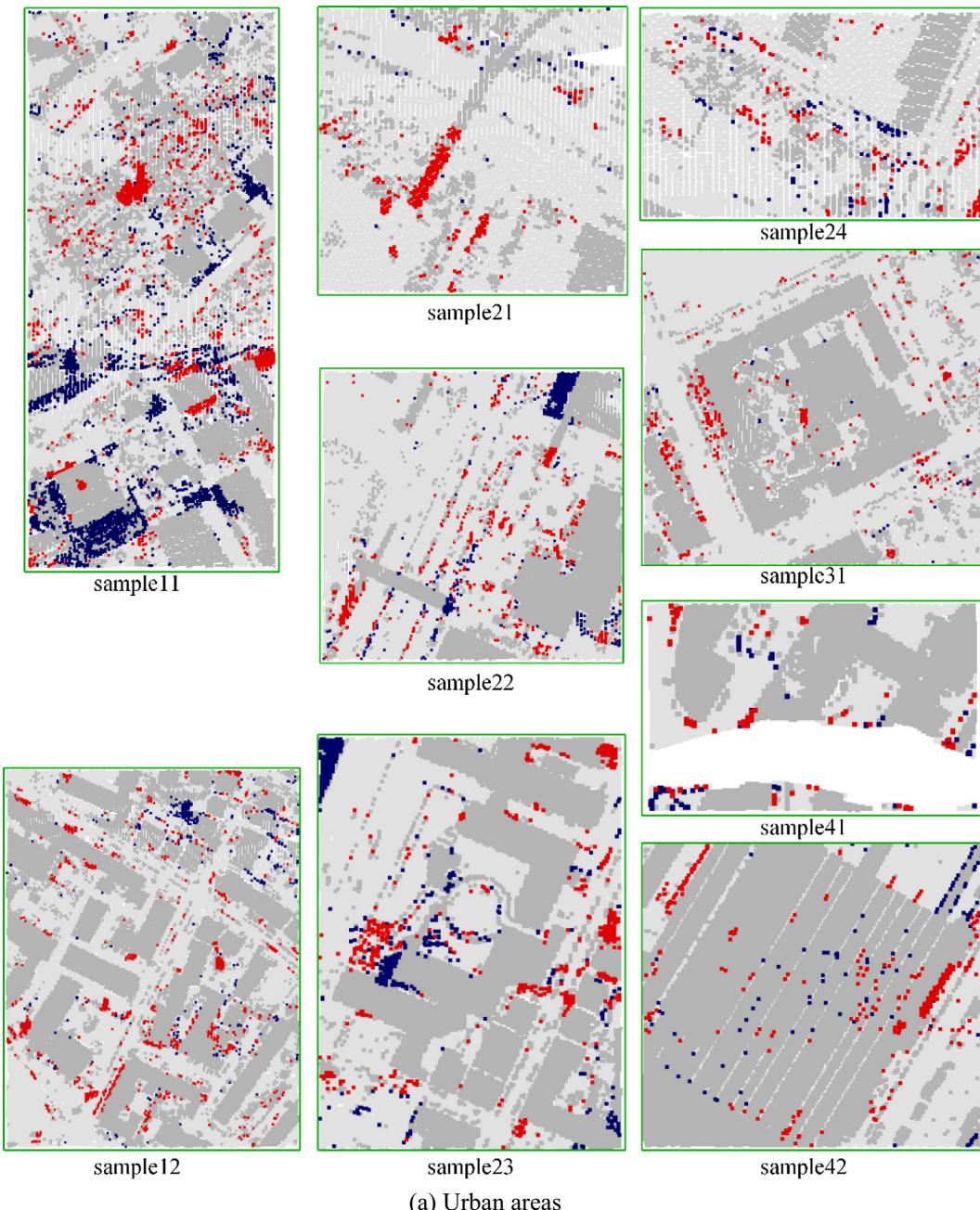


Fig. 7. The filtering results of 15 sample datasets.

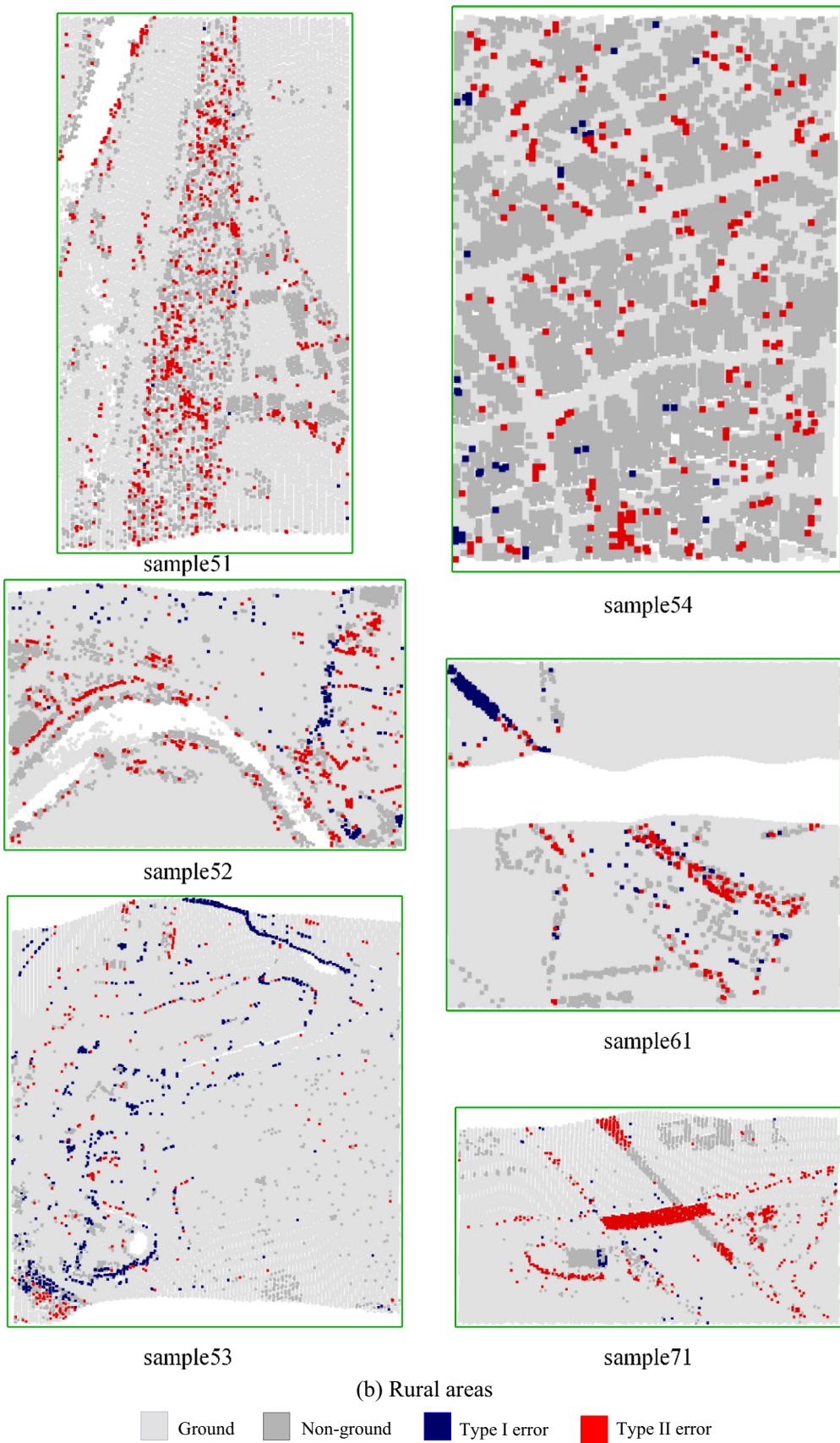
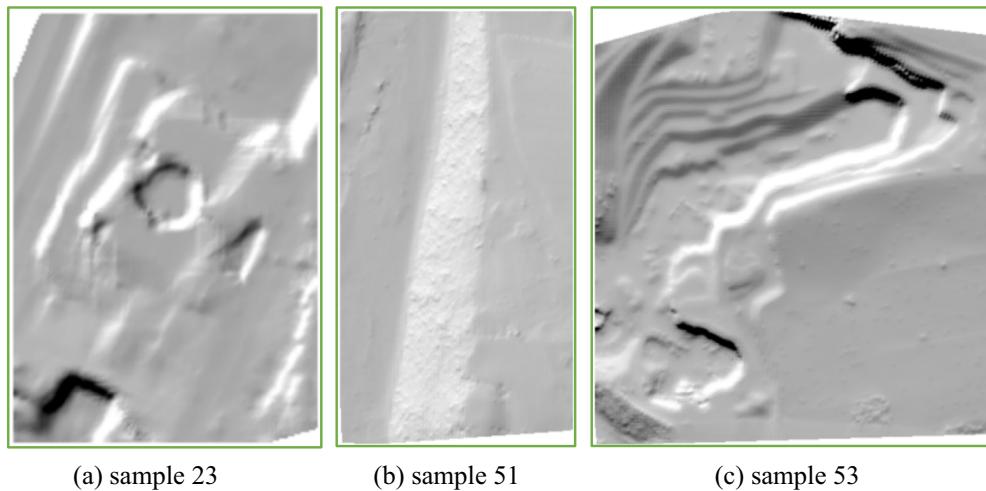
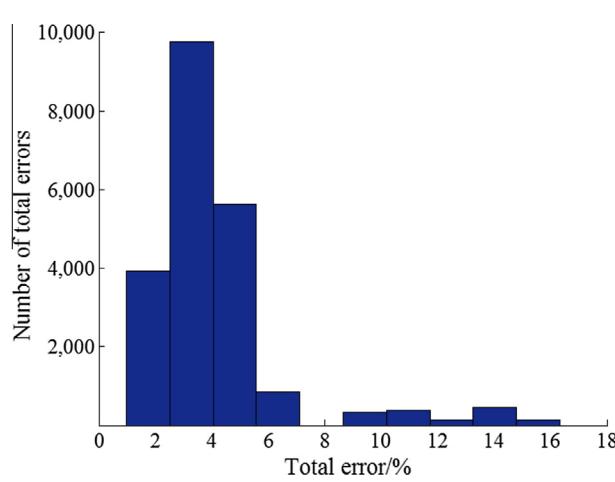
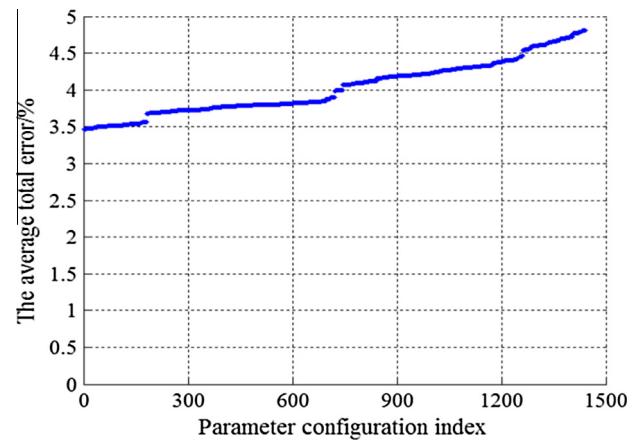


Fig. 7 (continued)

**Fig. 8.** The generated DEMs of three sample datasets.**Table 2**

Performance evaluation and comparison. The symbol “–” denotes that no value is available, and the values of the proposed method are highlighted.

Sample	<i>T.I</i> (%)	<i>T.II</i> (%)	<i>T.E</i> (%)	<i>k</i> (%)					
	Proposed method	Mongus (2012)	Pingel (2013)	Hu (2014)	Proposed method	Mongus (2012)	Pingel (2013)	Hu (2014)	Proposed method
11	11.69	8.93	11.01	8.64	8.40	10.52	–	82.40	82.78
12	1.88	3.54	5.17	3.10	2.99	2.68	–	93.80	94.02
21	0.59	10.42	1.98	1.88	1.92	2.76	–	94.43	94.26
22	3.66	6.83	6.56	3.40	3.50	4.65	–	92.07	91.76
23	3.81	5.24	5.83	6.48	4.75	4.48	–	87.02	90.47
24	1.68	8.02	7.98	4.19	4.12	3.40	–	89.49	89.52
31	0.28	3.11	3.34	2.48	3.26	1.58	–	95.00	93.41
41	2.25	2.73	3.71	10.79	6.27	2.49	–	78.41	87.47
42	0.93	1.39	5.72	2.93	1.21	1.26	–	93.07	97.10
51	0.09	16.04	2.59	3.00	2.78	3.49	–	90.74	91.49
52	1.38	16.22	7.11	4.17	3.18	2.92	–	78.80	83.69
53	2.50	17.60	8.52	7.41	5.67	3.11	–	47.24	53.06
54	0.99	4.97	6.73	3.67	2.71	3.13	–	92.65	94.57
61	0.66	17.28	4.85	2.02	2.41	1.23	–	75.38	71.08
71	0.43	39.95	3.14	1.85	1.86	4.90	–	90.52	90.54
Mean	2.19	10.82	5.62	4.40	3.67	3.51	–	85.40	87.01
Median	1.38	8.02	5.72	3.40	3.18	3.11	–	90.52	90.54
Max	11.69	39.95	11.01	10.79	8.40	10.52	–	95.00	97.10
Min	0.09	1.39	1.98	1.85	1.21	1.23	–	47.24	53.06

**Fig. 9.** The histogram of all total errors.**Fig. 10.** The graph of the relationship between the average total error and parameter configuration index, and the average total errors are sorted from the lowest to the highest.

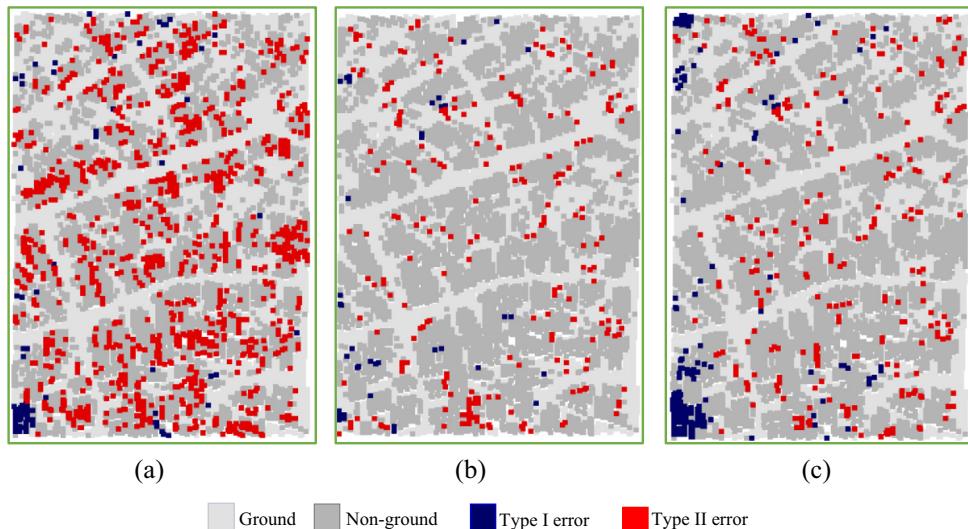


Fig. 11. The performance comparison of three filtering results with different point densities. (a) The filtering result without pseudo-grids. (b) The filtering result with pseudo-grids, where the grid size is 1 m. (c) The filtering result with pseudo-grids, where the grid size is 1.5 m.

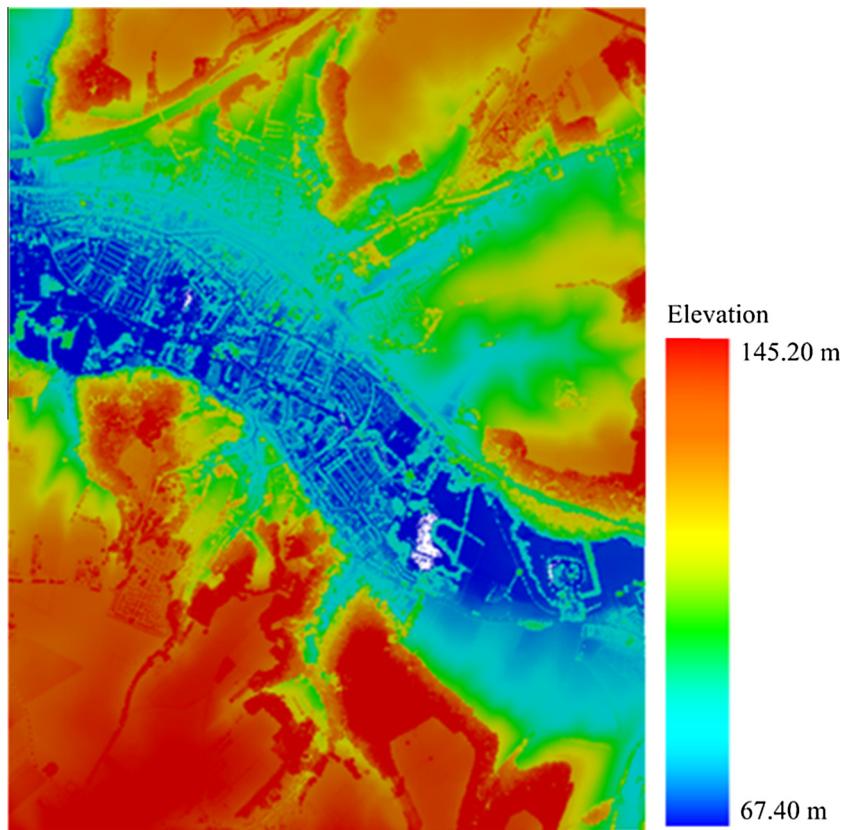


Fig. 12. Lidar point clouds from the city of Valkenburg in Netherlands.

Step 1: Calculate the normals for the candidate points of break-lines and classify the candidate points as different segments, as illustrated in Fig. 6f (each segment is dotted in one color).

Step 2: Extract the boundary of the candidate points in the block using the α -shape algorithm ([Edelsbrunner and Mücke, 1994](#)), and the boundary of each segment is also extracted ([Fig. 6f](#)).

Step 3: Remove boundary points of each segment that locate at the boundary of the block. The remaining boundary points of each segment are taken as the breaklines of this block (Fig. 6f).

In light of the above steps, the breaklines in all blocks are extracted, as showed in Fig. 6g. Then, the extracted breklines of

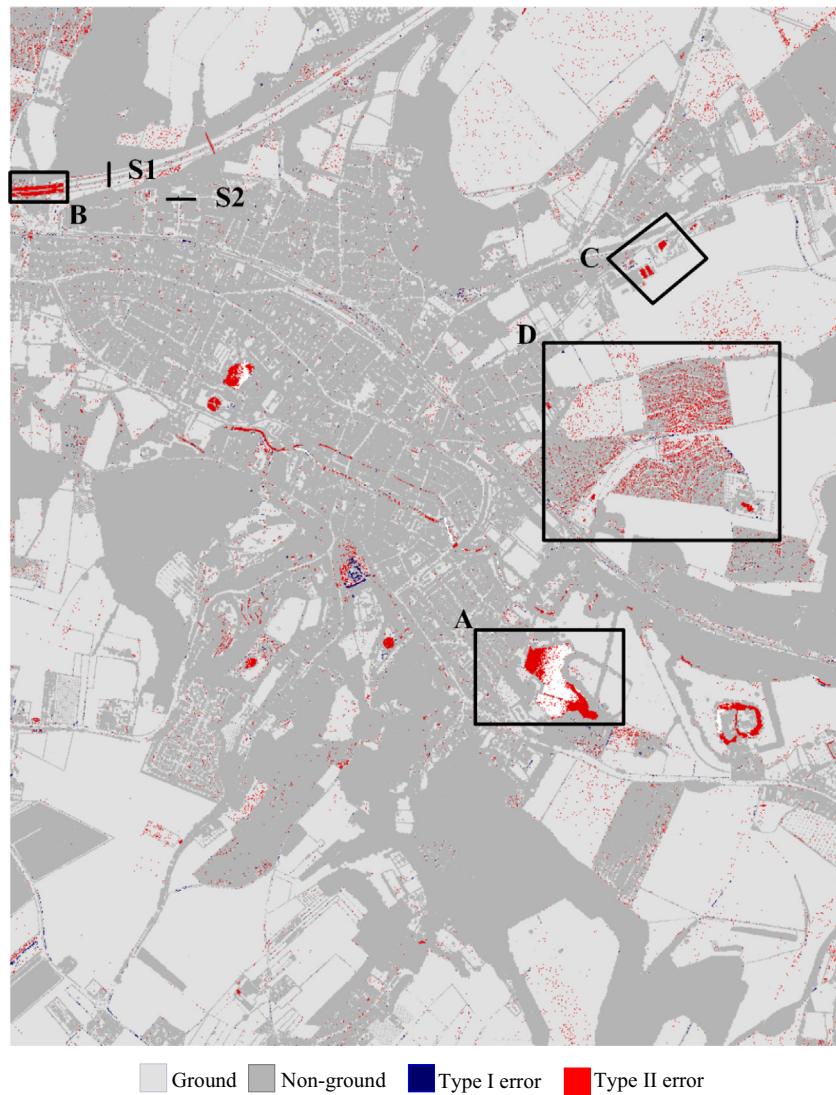


Fig. 13. The filtering result of the Valkenburg data.

Table 3

Performance comparison between the proposed method, the method described by Pingel et al. (2013), and the commercial software TerraScan®.

The proposed method (%)				Pingel (2013) (%)				TerraScan® (%)			
T.I	T.II	T.E	k	T.I	T.II	T.E	k	T.I	T.II	T.E	k
0.37	2.82	1.62	96.76	2.07	3.76	2.93	94.14	8.97	0.81	4.83	90.33

all blocks are linked if they share common regions to generate complete breaklines of the whole area (Fig. 6h). Finally, each complete breakline is fitted by the spline method (Fig. 6i).

3. Experimental results and analysis

3.1. Tests using ISPRS data

The dataset provided by ISPRS comprises 15 sample datasets, including diverse terrain types and various complex scenes (Sithole and Vosselman, 2004), and the point span is relatively large. For example, the point span in rural areas is about 2.0–3.5 m, causing major challenges during the filtering of non-ground points.

3.1.1. Performance evaluation and comparison

For the 15 sample datasets, the proposed method was applied with the same parameters values listed in Table 1. Fig. 7 shows the filtering results of urban and rural areas, the omission and commission points in each sample dataset were determined by comparing with the corresponding reference data. Fig. 8 illustrates the DEMs generated by the extracted ground points, showing good performances in preserving terrain features (e.g., cliffs). To quantitatively evaluate the performance of the proposed method, we calculated the type I error (T.I), the type II error (T.II), the total error (T.E), and the kappa coefficient (k) (Sithole and Vosselman, 2004). It can be seen from Table 2 that the mean total error and the kappa coefficient are 3.51% and 87.40%, the corresponding medians are 3.11% and 91.00%, and sample 61 has the lowest total error, sample 42 has the highest kappa coefficient. It also shows that sample 11

has the largest total error and sample 53 has the lowest kappa coefficient. The main reasons are that the quality of the provisional DEM in sample 11 is relatively low and few non-ground points in sample 53 contribute to the kappa coefficient. Besides, the proposed method achieves 80% of the type I errors less than 3.0%, indicating that it is robust in extracting ground points under various terrain types. For example, ground points in sample 53 have large elevation differences in local areas, the proposed method still achieves a good performance in terms of the type I error of 2.50%. Manual checking shows that most of the omission points (blue dots) are located on the vertical surface of the quarries in sample 53, as shown in Fig. 7b. However, 60% of the type II errors are greater than 5.0%, and type II errors are relatively larger than type I errors in most of sample datasets. The reasons are twofold. On the one hand, the number of non-ground points is generally less than that of ground points in these sample datasets, and few commission points could result in a large type II error. On the other hand, lots of commission points are caused by under-segmentation and low vegetation points. Fortunately, the commission points are easier to be found and edited than the omission points, because of the visible elevation differences with the surrounding ground points, such as the bridge points in sample 71 (Fig. 7b) and the building points in sample 11 (Fig. 7a). Another characteristic of the filtering results is that the average total errors in urban and rural areas have a minor difference, although the point densities in urban and rural areas have a large difference. The average total error of the 9 urban areas is 3.76%, and that of the 6 rural areas is 3.13%.

The performance of the proposed method was also compared with other methods (Mongus and Žalik, 2012; Pingel et al., 2013; Hu et al., 2014). In each method, the uniform parameters setting was used for the 15 sample datasets. The total errors and the kappa coefficients of the filtering results derived from these methods are listed in Table 2. The results demonstrate that the proposed method achieves the best performance for 8/15 sample datasets in terms of the total errors and the kappa coefficients. The 8 sample datasets achieving the best performance are sample 12, 23, 24, 31, 41, 52, 53, and 61. These sample datasets include terrain discontinuities, data holes, and varied point densities. However, the generated provisional DEMs have difficulties to describe the terrace terrain of sample 11 and 22, resulting in few ground segments omitted. In sample 21, 22 and 71, the attached object points are misclassified as ground points. In addition, low vegetation points and sparse point densities also affect the filtering quality, for example, the sample 51 and sample 54.

The experiment processed 15 sample datasets without tuning parameters, showing the robustness of the proposed method for different point densities, diverse terrain types and various complex scenes. The main reasons include three aspects. First, the proposed method generates the pseudo-grids to remedy the problems of the low and uneven point densities, resulting in a good estimation of the point normal for point cloud segmentation. Second, the proposed method classifies grid points into two sets and removes non-ground grid points by segment-based filtering and the multi-scale morphological filtering. Finally, the proposed method embeds an open operator and the segment-based filtering to decrease the maximum scale for relieving the classical cut-off problem and to improve the multi-scale morphological filtering.

3.1.2. Parameters values sensitivity analysis

To evaluate the robustness of the proposed method, we performed experiments with different parameters configurations, namely, W_{seed} ranges from 20 m to 28 m with an interval of 2 m, W_{scale} ranges from 5 m to 8 m with an interval of 1 m, h_1 ranges from 0.5 m to 0.8 m with an interval of 0.1 m, h_2 ranges from 0.2 m to 0.4 m with an interval of 0.1 m, h_3 ranges from 0.2 m to 0.4 m with an interval of 0.1 m, W_{grid} ranges from 1.0 m to 1.5 m with an interval of 0.5 m. The minimum area threshold for eliminating small segments is specified as 10 m^2 . Each sample dataset has 1440 configurations, resulting in $15 \times 1440 = 21,600$ results in total. The histogram of all total errors is illustrated in Fig. 9, showing that 92.50% of the total errors are between 1.00% and 7.00%. Simultaneously, the average total error of 15 sample datasets for each parameter configuration was calculated and illustrated in Fig. 10. It shows that the average total errors range from 3.47% to 4.81%. The result demonstrates that the proposed method is insensitive to the configuration of the parameters values in a reasonable range and thus achieves robust filtering results.

It is clear that low density of point clouds has an effect on point cloud segmentation, thus decreasing the quality of the filtering result. However, the generation of pseudo-grids could remedy the problem by interpolating some pseudo grid points. Sample 54 with the point span about 2–3.5 m was selected to check the benefits of pseudo-grids. Fig. 11a is the filtering result of the proposed method without the generation of pseudo-grids, and the total error is 13.08%. Fig. 11b is the filtering result of the proposed method with the generation of pseudo-grids (the grid size W_{grid} is 1.0 m), and the total error is 3.13%. Fig. 11c is the filtering result of the proposed method with the generation of pseudo-grids (the grid size W_{grid} is 1.5 m), and the total error is 5.35%. The comparison

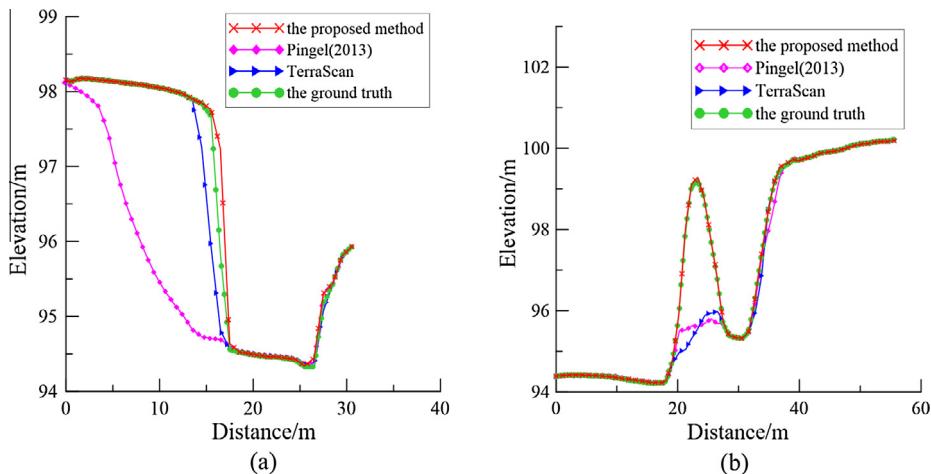


Fig. 14. The comparison between the cross-sections obtained by different methods and those of the ground truth in two special areas. (a) A cliff. (b) A small terrain protuberance area.

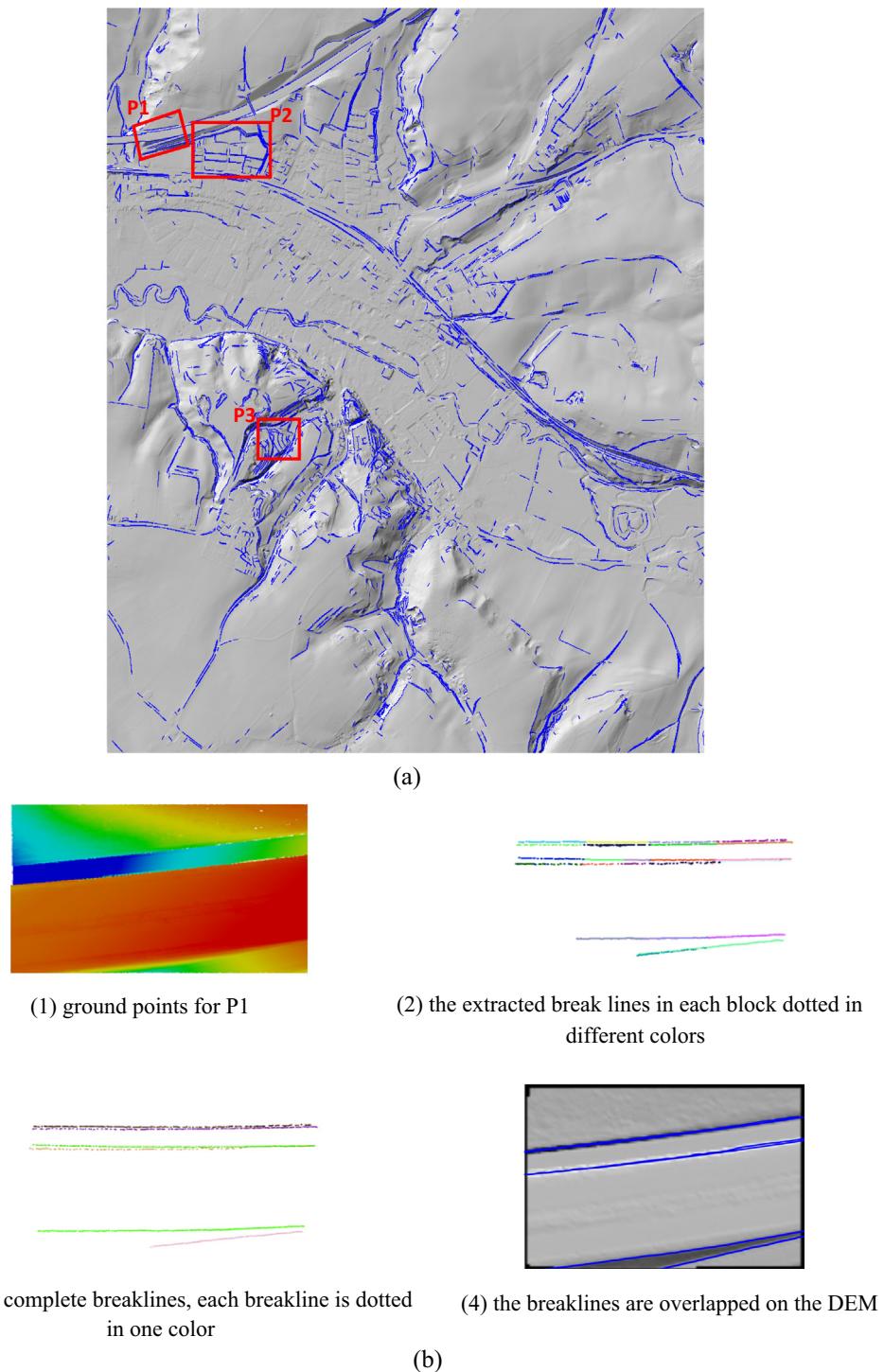


Fig. 15. Extracting breaklines and generating DEMs with breaklines: (a) DEMs generated with breaklines, (b) the process of extracting the breaklines from ground points in area P1.

Table 4

Quality evaluation of extracted breaklines for three areas.

Area	TP (m)	FN (m)	FP (m)	Completeness (%)	Correctness (%)	Quality (%)
P1	1182.94	112.26	23.42	91.33	98.06	89.71
P2	4750.80	556.56	129.19	89.51	97.35	87.39
P3	946.61	107.65	25.23	89.79	97.40	87.69
Mean				90.21	97.60	88.26

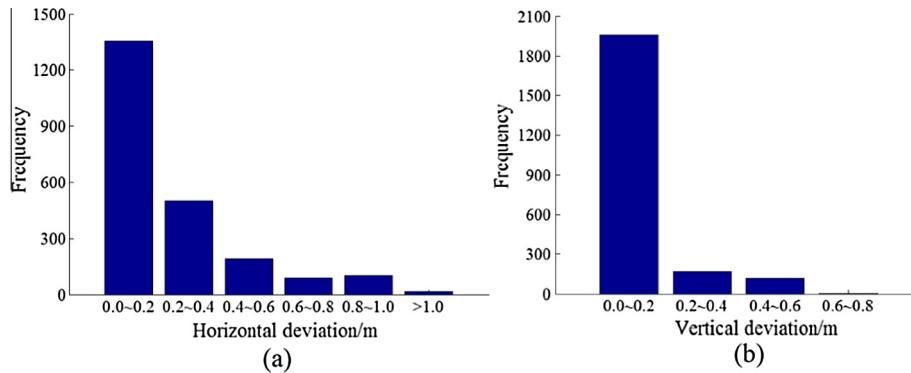


Fig. 16. The histograms of the horizontal and vertical deviations between the extracted breakline points and the corresponding reference points.

shows that proposed method improves the quality of the filtering result by the generation of pseudo-grids. It can also be seen that the increasing of grid size affects the quality of the filtering result.

3.2. Test using data from the Netherlands

The data were obtained from the city of Valkenburg in Netherlands with an area of 11.25 km^2 , 0.42 billion points in total, and elevations ranging from 67.40 m to 145.20 m, including the steep slope, the dense and low vegetation, rivers, and large buildings, as shown in Fig. 12. Fig. 13 illustrates the filtering result achieved by the proposed method, where $W_{seed} = 60 \text{ m}$ and the values of the other parameters are listed in Table 1. Manual checking demonstrates that the proposed method performs well in complex areas, and there are only a few type I error points. However, because of false ground seed points in water areas and under-segmentation, the commission points are mainly distributed in water areas and the attached objects (e.g. bridge and buildings), as indicated in the rectangles A, B, and C of Fig. 13. Besides, some vegetation points with low elevations are erroneously extracted, as shown in the rectangle D of Fig. 13.

We compared the filtering performance of the proposed method with the method described by Pingel et al. (2013) (<http://tpingel.org/code/smrf/smrf.html>) and the commercial software TerraSolid TerraScan®. The comparison results are listed in Table 3, showing that the proposed method performs better in terms of the total error and the kappa coefficient. Moreover, two cross-sections of the cliff S1 and the small terrain protuberance area S2 also show that the cross sections obtained by the proposed method approximate the ground truths better, as illustrated in Fig. 14.

Fig. 15 illustrates the extraction of breaklines and high quality DEMs generated with breaklines for the study area. To quantitatively evaluate the quality and positional accuracy of extracted breaklines, we selected three areas (P1, P2, P3) of Fig. 15a. Firstly, we manually digitized the reference breaklines on the DEMs, and measured the lengths of correctly extracted breaklines (TP), omission breaklines (FN) and commission breaklines (FP) based on the comparison of the extracted breaklines and the reference breaklines. Three indicators (Completeness, Correctness, and Quality) were calculated according to the method of Yang et al. (2013). Table 4 shows that the proposed method achieves the mean completeness of 90.21%, the mean correctness of 97.60% and the mean quality of 88.26%. Moreover, we manually extracted 2246 three dimensional breakline points from 45 breaklines as the reference points to evaluate the positional accuracy. The extracted breakline points include points of jump breaklines, crease breaklines and curvature breaklines. Then, the reference points were assigned to the nearest breakline points extracted by the proposed method. The horizontal deviation and the vertical deviation between each pair of the reference point and the extracted breakline point were

calculated. The mean values of the calculated horizontal and vertical deviations are 0.23 m and 0.10 m. And the histograms of the horizontal and vertical deviations are illustrated in Fig. 16. Fig. 16a shows that 82.50% of the horizontal deviations are less than 0.40 m, and 60% of the horizontal deviations are less than 0.20 m. Fig. 16b shows that 87.18% of the vertical deviations are less than 0.20 m. It demonstrates that the proposed method could extract breaklines with an acceptable quality and positional accuracy.

To evaluate the effects of breaklines on the quality of the DEMs generated, we compared the elevation differences between the DEMs with/without breaklines and the ground truths in three areas. Fig. 17 illustrates the comparisons of the 2D cross-sections in the three areas. It can be seen that the cross-sections of the DEMs with breaklines more closely approximate the ground truths. Although the proposed method extracts three types of breaklines from ground points, it can be found from Fig. 17a (the left rectangle in Fig. 17a) that topological errors between breaklines might still exist. On the other hand, it also shows that the extraction of jump breaklines may fail in the area with small elevation differences (the right rectangle in Fig. 17a). As far as the extraction of curvature breaklines and crease breaklines is concerned, the breaklines extraction may fail in the area with too gentle slope change (rectangle in Fig. 17b and 17c). The above problems could be corrected by manual editing.

Fig. 18 illustrates the elevation differences between all ground points and the interpolated DEMs with/without breaklines, showing that the interpolation errors are mainly distributed around the breaklines, and the interpolation errors in the DEMs without breaklines are larger than those of the DEMs with breaklines. It demonstrates that the proposed method performs better in generating high quality DEMs with breaklines.

4. Conclusion

In this study, we propose an automated method for extracting ground points and breaklines from airborne LiDAR point clouds. It generates grid points from airborne LiDAR point clouds and these grid points are further classified into two sets (i.e., a set of segments and one set of individual points), which are processed by segment-based filtering and multi-scale morphological filtering. The proposed method achieves a good performance by integrating two filtering methods and removing large objects before applying multi-scale morphological filtering, thereby facilitating the reliable extraction of ground points in various areas, such as urban areas and rural areas. Moreover, the proposed method simultaneously extracts multiple types of three dimensional breaklines from the ground points. To verify the validity and the robustness of the proposed method, 15 sample datasets provided by the ISPRS and one

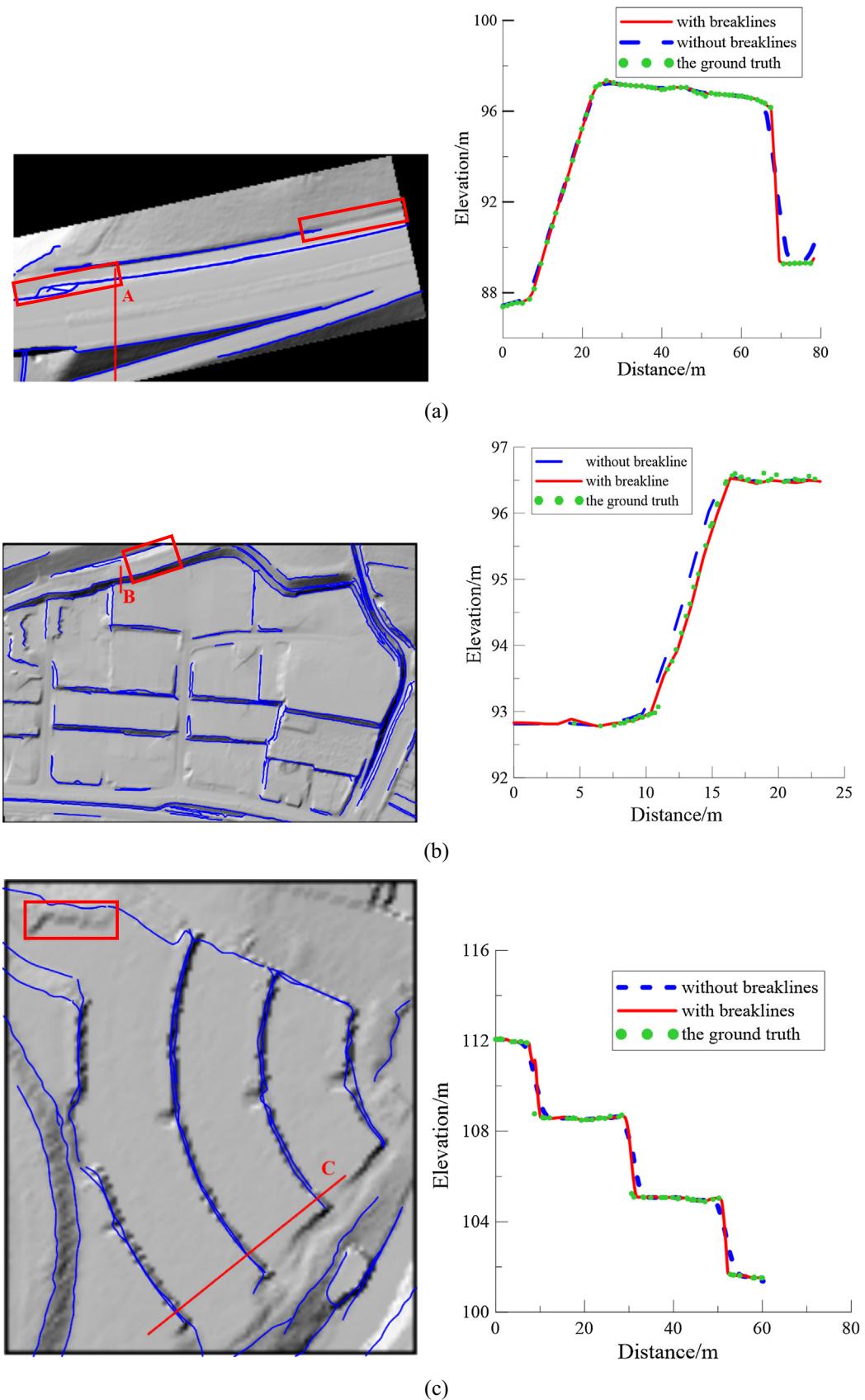


Fig. 17. 2D cross-sections comparisons between the DEMs with/without breaklines and the ground truths in three areas. (a) A road with jump and curvature breaklines. (b) A steep slope with crease breaklines. (c) A terraced terrain with jump breaklines.

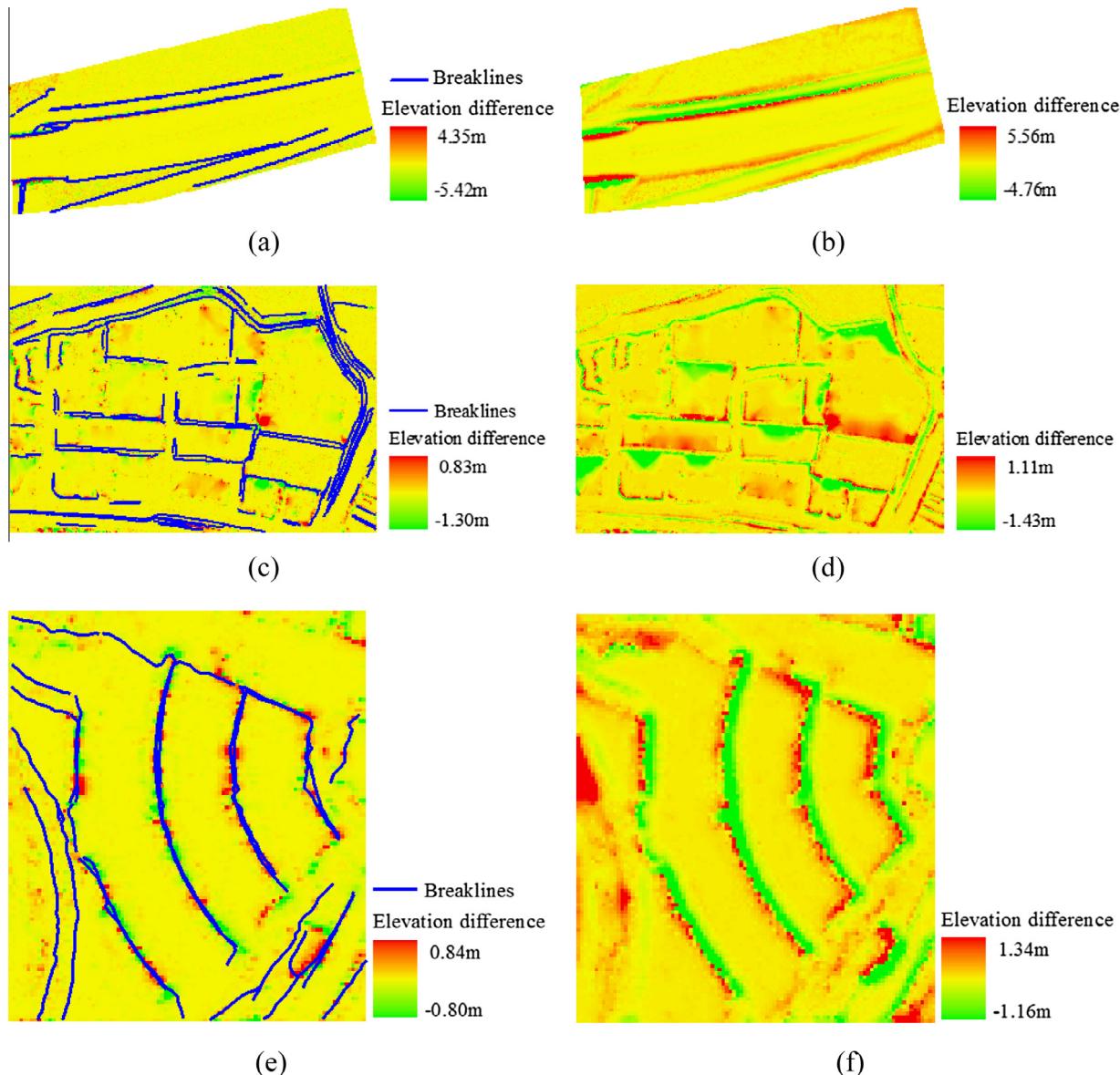


Fig. 18. The comparison of elevation differences between all ground points and the DEMs with/without breaklines. (a, c, e) The elevation differences between DEMs with breaklines and the ground truth in P1, P2 and P3 of Fig. 15a. (b, d, f) The elevation differences between DEMs without breaklines and the ground truths in P1, P2 and P3 of Fig. 15a.

large-scale dataset from the Valkenburg of Netherlands were selected to extract ground points and breaklines. Our comprehensive experiments and comparisons demonstrate that the proposed method exhibits a good performance in filtering non-ground points of the areas with various terrain types (e.g., rural areas and urban areas), thereby providing a good foundation for high quality DEMs generation while preserving the breaklines. More importantly, the proposed method can extract ground points from areas with various terrain types using uniform parameters values. However, the proposed method may erroneously classify the points of some objects (e.g., bridges) attached ground to be ground points, thereby resulting in commission points. This problem can be overcome by improving the quality of point cloud segmentation in the future research.

Acknowledgments

This study was jointly supported by the National Key Technology R&D Program (No. 2014BAL05B07), National Basic Research

Program of China (No. 2012CB725301), NSFC project (No. 41071268), and Public science and technology research funds projects of ocean (No. 2013418025). We also thank Prof. Jantien Stoter (TU Delft) for providing the data from the city of Valkenburg in Netherlands.

References

- Axelsson, P., 2000. DEM generation from laser scanner data using adaptive TIN models. *Int. Arch. Photogram. Remote Sens.* 33, 111–118.
- Briese, C., 2004. Three-dimensional modelling of breaklines from airborne laser scanning data. *Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci.* 35 (B3), 1097–1102.
- Briese, C., Pfeifer, N., 2008. Towards automatic feature line modelling from terrestrial laser scanner data. *Int. Arch. Photogram. Remote Sens.* 37, 463–468.
- Brügelmann, R., 2000. Automatic breakline detection from airborne laser range data. *Int. Arch. Photogram. Remote Sens.* 33, 109–116.
- Cho, W., Jwa, Y.S., Chang, H.J., Lee, S.H., 2004. Pseudo-grid based building extraction using airborne LiDAR data. *Int. Arch. Photogram. Remote Sens.* 35 (B3), 378–381.
- Chen, C., Li, Y., Li, W., Dai, H., 2013. A multiresolution hierarchical classification algorithm for filtering airborne LiDAR data. *ISPRS J. Photogram. Remote Sens.* 82, 1–9.

- Chen, Q., 2009. Improvement of the Edge-based Morphological (EM) method for lidar data filtering. *Int. J. Remote Sens.* 30, 1069–1074.
- Chen, Q., Gong, P., Baldocchi, D., Xie, G., 2007. Filtering airborne laser scanning data with morphological methods. *Photogramm. Eng. Remote Sens.* 73, 175–185.
- Cui, Z., Zhang, K., Zhang, C., Chen, S., 2013. A cluster-based morphological filter for geospatial data analysis. *ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data*, 1–7.
- Deng, S., Shi, W., 2013. Integration of different filter algorithms for improving the ground surface extraction from airborne LiDAR data. *Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci.* 1, 105–110.
- Edelsbrunner, H., Mücke, E., 1994. Three-dimensional alpha shapes. *ACM Trans. Graph. (TOG)* 13, 43–72.
- Evans, J., Hudak, A., 2007. A multiscale curvature algorithm for classifying discrete return lidar in forested environments. *IEEE Trans. Geosci. Remote Sens.* 45, 1029–1038.
- Filin, S., Pfeifer, N., 2006. Segmentation of airborne laser scanning data using a slope adaptive neighborhood. *ISPRS J. Photogram. Remote Sens.* 60, 71–80.
- Hu, H., Ding, Y., Zhu, Q., Wu, B., Lin, H., Du, Z., Zhang, Y., Zhang, Y., 2014. An adaptive surface filter for airborne laser scanning point clouds by means of regularization and bending energy. *ISPRS J. Photogram. Remote Sens.* 92, 98–111.
- Kilian, J., Haala, N., Englisch, M., 1996. Capture and evaluation of airborne laser scanner data. *Int. Arch. Photogram. Remote Sens.* 31, 383–388.
- Kraus, K., Pfeifer, N., 1998. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS J. Photogram. Remote Sens.* 53, 193–203.
- Kraus, K., Pfeifer, N., 2001. Advanced DTM generation from LiDAR data. *Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci.* 34 (3/W4), 23–30.
- Li, Y., Wu, H., Xu, H., An, R., Xu, J., He, Q., 2013. A gradient-constrained morphological filtering algorithm for airborne LiDAR. *Opt. Laser Technol.* 54, 288–296.
- Liu, X., 2008. Airborne LiDAR for DEM generation: some critical issues. *Prog. Phys. Geogr.* 32, 31–49.
- Meng, X., Currit, N., Zhao, K., 2010. Ground filtering algorithms for airborne LiDAR data: a review of critical issues. *Remote Sens.* 2, 833–860.
- Mongus, D., Lukač, N., Žalík, B., 2014. Ground and building extraction from LiDAR data based on differential morphological profiles and locally fitted surfaces. *ISPRS J. Photogram. Remote Sens.* 93, 145–156.
- Mongus, D., Žalík, B., 2012. Parameter-free ground filtering of LiDAR data for automatic DTM generation. *ISPRS J. Photogram. Remote Sens.* 67, 1–12.
- Pingel, T., Clarke, K., McBride, W., 2013. An improved simple morphological filter for the terrain classification of airborne LiDAR data. *ISPRS J. Photogram. Remote Sens.* 77, 21–30.
- Podobnikar, T., Vrečko, A., 2012. Digital elevation model from the best results of different filtering of a LiDAR point cloud. *Trans. GIS* 16, 603–617.
- Shan, J., Toth, C., 2008. *Topographic Laser Ranging and Scanning—Principles and Processing*. CRC Press Taylor & Francis, London.
- Shao, Y., Chen, L., 2008. Automated searching of ground points from airborne lidar data using a climbing and sliding method. *Photogram. Eng. Remote Sens.* 74, 625–635.
- Shen, J., Liu, J., Lin, X., Zhao, R., 2012. Object-based classification of airborne light detection and ranging point clouds in human settlements. *Sensor Lett.* 10, 221–229.
- Sithole, G., 2001. Filtering of laser altimetry data using a slope adaptive filter. *Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci.* 34, 203–210.
- Sithole, G., Vosselman, G., 2004. Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. *ISPRS J. Photogram. Remote Sens.* 59, 85–101.
- Sithole, G., Vosselman, G., 2005. Filtering of airborne laser scanner data based on segmented point clouds. *Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci.* 36, W19.
- Sohn, G., Dowman, I., 2002. Terrain surface reconstruction by the use of tetrahedron model with the MDL criterion. *Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci.* 34, 336–344.
- Tóvári, D., Pfeifer, N., 2005. Segmentation based robust interpolation – a new approach to laser data filtering. *IAPRS* 36, W19.
- Vosselman, G., 2000. Slope based filtering of laser altimetry data. *Int. Arch. Photogram. Remote Sens.* 33, 935–942.
- Vosselman, G., Maas, H., 2010. *Airborne and Terrestrial Laser Scanning*. CRC, Boca Raton.
- Xu, S., Vosselman, G., Oude Elberink, S., 2014. Multiple-entity based classification of airborne laser scanning data in urban areas. *ISPRS J. Photogram. Remote Sens.* 88, 1–15.
- Yan, M., Blaschke, T., Liu, Y., Wu, L., 2012. An object-based analysis filtering algorithm for airborne laser scanning. *Int. J. Remote Sens.* 33, 7099–7116.
- Yang, B., Fang, L., Li, J., 2013a. Semi-automated extraction and delineation of 3D roads of street scene from mobile laser scanning point clouds. *ISPRS J. Photogram. Remote Sens.* 79, 80–93.
- Yang, B., Wei, Z., Li, Q., Li, J., 2013b. Semiautomated building facade footprint extraction from mobile LiDAR point clouds. *IEEE Geosci. Remote Sens. Lett.* 10, 766–770.
- Zhang, J., Lin, X., 2013. Filtering airborne LiDAR data by embedding smoothness-constrained segmentation in progressive TIN densification. *ISPRS J. Photogram. Remote Sens.* 81, 44–59.
- Zhang, K., Chen, S., Whitman, D., Shyu, M., Yan, J., Zhang, C., 2003. A progressive morphological filter for removing nonground measurements from airborne LiDAR data. *IEEE Trans. Geosci. Remote Sens.* 41, 872–882.