Filtering of Airborne Lidar Point Clouds for Complex Cityscapes

JIANG Jingjue  ZHANG Zuxun  MING Ying

Abstract  A novel filtering algorithm for Lidar point clouds is presented, which can work well for complex cityscapes. Its main features are filtering based on raw Lidar point clouds without previous triangulation or rasterization. 3D topological relations among points are used to search edge points at the top of discontinuities, which are key information to recognize the bare earth points and building points. Experiment results show that the proposed algorithm can preserve discontinuous features in the bare earth and has no impact of size and shape of buildings.

Keywords  filtering; segmentation; laser scanning; Lidar; point clouds

CLC number  P237.3

Introduction

Airborne Lidar (light detection and ranging) can provide accurate and high sampling 3D geometry data of the earth’s surface. It has become the preferred method for the acquisition of digital elevation models (DTM)[1]. At present, there are two strategies for developing filtering algorithm. One is similar to image processing[2-5], the other is constructed by considering the local neighbor relations among points represented in the triangulated irregular network (TIN)[6,7]. When producing the DTM of complex cityscapes, most current filtering algorithms face the greatest challenges due to the multi-tier buildings, courtyards, plazas, stairways, ramp and discontinuities such as steep slopes and break-lines in the bare earth[8]. Usually, a lot of manual editing are needed to get accurate points on the bare earth. In practice, the time consumed by manual classification (including filtering) and quality control is approximately 60% to 80% of the total processing time[6]. Therefore, it is necessary to develop an accurate, efficient, robust and self-adaptive filtering algorithm.

Current filtering algorithms can be divided into two groups. The first is point-based method, which includes most of the filters. In these algorithms, each point is investigated and labeled as terrain or other objects according to its neighbors. However, it is difficult to confirm the class of a point when only considering the neighbor relations between points[2-7]. For example, points on the edge of the discontinuous bare earth and on the edge of building roofs cannot be distinguished if these points are treated individually. The second is segmentation-based method[1,9]. Point clouds are segmented first, and then these segments are classified based on their surrounding neighbors. This kind of filter considers more contexts. However, it is easy to make point clouds over-segmented.

A new approach is proposed for filtering raw air-
borne Lidar point clouds with two echoes. The algorithm can be performed without previous interpolation or triangulation. It holds the virtues of the two kinds of filters mentioned above. Edge points on the discontinuous surface are used to search the bare earth points and building points first. Then those bare earth points and building points are treated as seeds to cluster and recognize all other congeneres. Experiment results show that this filter has a good performance.

1 Filtering strategies

The bare earth in the real world includes some discontinuous features such as steep slopes and break-lines. Therefore, it is reasonable to treat the bare earth as a patchwork of piecewise continuous surfaces in concept. In this study, non-terrain objects will be classified as vegetation, buildings and uncertain objects when performing the filtering algorithm. Based on this, more detailed classification can be achieved for the different practical applications.

The classification of objects was guided by the following assumption in the study.

1) Bare earth. The local lowest point in a level neighborhood of a point on the edge of discontinuities is probably a bare earth point.

2) Vegetation. The height difference of the first-return and last-return data of vegetation is larger than a given threshold.

3) Building. Points at the top of discontinuities may be the edge of the roof. In a small local neighborhood, the height differences between a known building point and its neighbors are less than a given threshold.

4) Uncertain object. It includes cars, heap and uncertain patches in the courtyard. This kind of object cannot be subdivided further if only considering Lidar point clouds.

In fact, although the assumption is often valid, it does sometimes fail. The failure may exist in the following situations. For example, points on the lower roof of multi-tier buildings may be labeled as the bare earth points. In the bare earth with steep slope, points at the top of discontinuousness may be treated as the edge of the building roof. In addition, both vegetation and part of the edge of the building roof may produce several recordable reflections of one laser pulse. However, all these failures can be corrected during iteration of the solution.

2 Filtering algorithm based on Lidar point clouds

To obtain DTM, the proposed algorithm is performed based on the last return echoes, which are most probably reflected from the bare earth. The work flow is explained in the following six steps.

1) Removing outliers. First, \(k\)-nearest neighbors of each point are calculated based on the Euclidean distance between points. Then the least squares best fitting plane of a point is computed according to its neighbors. If the distance between a point and its fitting plane is greater than a given threshold, it is the outlier. The value of \(k\) must ensure that all \(k\)-nearest neighbors of a point come from the same surface. This method can deal with both high and low outliers. It can also remove a set of outliers which are close to each other.

2) Labeling vegetation candidates. The height difference \(h\) between the first and the last return of each point is calculated. If \(h\) is greater than a given threshold, it is labeled as a vegetation candidate. It is noticeable that some points on the edge of the building roofs may be marked as vegetation candidates.

3) Recognizing points on the edge of discontinuities. This is a key step for filtering. Usually, \(k\)-nearest neighbors of a point \(p\) can be used to define a small surface patch around \(p\) as shown in Fig.1(a). However, in some cases as shown in Fig.1(b), when a point is on the edge of the discontinuities, its \(k\)-nearest neighbors only cover half the surface patch. There

![Fig.1 k-nearest neighbors of a point](image-url)
fore, if we can evaluate that \( k \)-nearest neighbors in 3D space only cover half the area around the point \( p \), we can find that the point \( p \) lies on the edge of the discontinuities. Thus, we introduce an angle criterion to check the relationship between a point and its neighbors for edge point detection as shown in Fig.2.

At first, a least squares best fitting plane \( H \) is calculated according to a point \( p \) and its \( k \)-nearest neighbors, \( \{p_1, p_2, \ldots, p_k\} \). Then, all \( k \)-nearest neighbors of \( p \) are projected on to the plane \( H \). The projected neighbors are sorted as \( \{q_1, q_2, \ldots, q_k\} \) around \( q \), the projection of \( p \) in the plane \( H \). If the angle \( \angle q_i q_{i+1} q \) exceed a given threshold, \( p \) is detected as an edge point. Usually, the edge of the building roof is a straight line or two lines perpendicular to each other. So the angle threshold can be chosen a value greater than 90°.

![Fig.2 Relations between projected neighbors](image)

When \( k \)-nearest neighbors of a point lie on nearly one line (usually appearing on overlapping regions), the point will be signed as an edge point according to the above criterion. In Fig.3(a), those failure edge points in the seven horizontal curves are caused by uneven data distribution. Considering a fact, the value of height difference between a false edge point and the local lowest point among its horizontal neighborhood is small. Therefore, we can correct false edge points by examining the height differences. As shown in Fig.3(b), the false edge points in Fig.3(a) are removed correctly.

![Fig.3 Detection and correction of edge points in point clouds](image)

4) Classifying. According to the labeled edge points, the local lowest points in the horizontal neighborhood around the known edge points can be marked as the bare earth. The bare earth is defined as a piecewise continuous surface, so in a small neighborhood around a known bare earth point, the height differences between it and its neighbors have a small value. Therefore, the known bare earth points can be treated as seeds to cluster and recognize other bare earth points. If the points on the edge of discontinuities are not signed as the bare earth during clustering, they will be labeled as building seeds for recognizing other building points by the same clustering algorithm. To repeat the clustering until all edge points are labeled as the bare earth or buildings. The points with no sign will be labeled as one of uncertain object.

5) Correcting and refining. After performing the former four steps, Lidar point clouds have been roughly classified as the bare earth, vegetation, building and uncertain object. However, those bare earth points may include some multi-tier building points, small and low vegetation or other uncertain objects. Those multi-tier building points among the labeled bare earth points can be distinguished from the true bare earth by performing the 4th step. To recognize small and low objects, the bare earth should be defined as more flat. For example, in a small region, we can limit the value of height differences between the known bare earth points and their neighbors in a small range, for example 0.3 m. So the above bare earth points can be subdivided into the true bare earth and other uncertain objects. If those uncertain objects have been labeled as vegetation in Step 2, they are vegetation clearly.

6) Producing DTM. To generate DTM from the known bare earth points, we should fill the holes caused by the removal of non-terrain objects. There are a lot of methods for point clouds interpolation. We adopt the approach of weighted average interpolation which is common in practice and has high computing efficiency without resolving the equation set\(^{[10]}\). The height of a query point can be substituted by weighted average heights of its neighbors.

\[
Z_p = \frac{\sum_{i=1}^{n} w_i Z_i}{\sum_{i=1}^{n} w_i}
\]  

(1)
where $Z_p$ is the height of a query point $p$; $Z_i$ is the height of the $i$th neighbors; $n$ is the number of $d$ neighbors; $w_i$ is the weight of the $i$th neighbor. Usually, weight can be calculated by using a Gaussian function as shown in the following.

$$w_i = e^{-\frac{(p_i - p)^2}{h^2}}$$

where $p_i$ is the 3D coordinates of neighbors around point $p$; $h$ is a global parameter. The $h$ is smaller, the local influence on the height calculation for query point is stronger. Namely, neighbors close to the query point have greater influence on height calculation.

3 Experiment results and discussion

We have implemented our proposed algorithm in C++. Lidar point clouds of complex cityscapes are used to test the algorithm. The test data sets contain some typical features such as multi-tier and irregular shape buildings, courtyards, tunnel, plazas and steep slopes. The distribution of the test data set is uneven, especially in the strip overlaps, where the point spacing is reduced and $k$-nearest neighbors are almost arranged in one line as shown in Fig.4. The average level space between points is 0.67 m$^2$.

When filtering outlines, the number of neighbors was chosen as 16, i.e. $k=16$. That is, $k$-nearest neighbors can describe a region of 25 m$^2$. The distance threshold between a point and its fitting plane is 7 m. Because the earth surface and building roof are smooth relatively, the $k$-nearest neighbors of a point should be limited in a small region during clustering in our experiments, e.g. $k=20$. If the distance between a point and seed is less than 3 m, and the height difference between them is less than 0.5 m, the point will be marked as the same class as the seed.

Fig.5 shows the classification results. Fig.6 presents the bare earth removing non-terrain objects. Fig.7 shows the DTM produced by known bare earth points. Experiment results show that the algorithm presented in this paper is capable of recognizing complex non-terrain objects accurately, and at the same time preserving discontinuous features on the bare earth. The patch in the middle of courtyards sometimes belongs to the bare earth, while in other cases it is part of buildings. Because this patch is enclosed by objects and separated from other bare earth, the decision of whether enclosed patch is the bare earth is not always clear-cut. To eliminate bad influence caused by failure classification, these patches are defined as uncertain objects in this study.
4 Conclusions

This study presents a novel filtering algorithm performed on raw Lidar point clouds. Topological relations among points are used to search edge points on the discontinuities. These edge points as key information are used to recognize the bare earth points and points on the edge of building roofs, which are treated as seeds to cluster and separate the bare earth from other non-terrain objects. The algorithm can preserve discontinuous features in the bare earth and has no impact from the size and shape of buildings. It is fit for processing Lidar data from complex cityscapes due to its high automation and efficiency.

In the next work, we will investigate methods combining Lidar point clouds and other data sources to obtain accurate and reliable classification results for different applications.

References

[1] Sithole G, Vosselman G(2005) Filtering of airborne laser scanner data based on segmented point clouds[C]. Laser Scanning 2005, Enschede, the Netherlands

[2] Elmqvist M, Jungert E, Lantz F, et al.(2001) Terrain modelling and analysis using laser scanner data[C]. IAPRS, Annapolis, MD

[3] Roggero M(2001) Airborne laser scanning: clustering in raw data[C]. International Archives of Photogrammetry and Remote Sensing, Annapolis, MD

[4] Wack R, Wimmer A(2002) Digital terrain models from airborne laser scanner data: a grid based approach[C]. IAPRS, Graz, Austria

[5] Sithole G(2001) Filtering of laser altimetry data using a slope adaptive filter[C]. IAPRS, Annapolis, MD

[6] Sohn G, Dowman I(2002) Terrain surface reconstruction by the use of tetrahedron model with the MDL criterion[C]. IAPRS, Graz, Austria

[7] Axelsson P(2000) DEM generation from laser scanner data using adaptive TIN Models[C]. IAPRS, Amsterdam

[8] Sithole G, Vosselman G(2004) Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds[J]. ISPRS Journal of Photogrammetry & Remote Sensing, 59: 85-101

[9] Tóvári D, Pfeifer N(2005) Segmentation based robust Interpolation: a new approach to laser data filtering[C]. Laser Scanning 2005, Enschede, the Netherlands

[10] Zhang Zuxun, Zhang Jianqing(2001) Digital photogrammetry [M]. Wuhan: Wuhan University Press (in Chinese)