A method for separation of the terrain and non-terrain from Vehicle-borne Laser Scanning Data

Jiangxia Wei and Ruofei Zhong
Capital Normal University Key Lab of 3D Information Acquisition and Application, MOE, Beijing 100048, China
E-mail: weijiangxia2011@163.com

Abstract. Half the points from vehicle-borne laser scanning data are terrain data. If you want to extract features such as trees, street lights and buildings, terrain points must be removed. Nowadays, either airborne or vehicle-borne laser data, are mostly used to set an elevation threshold based on the scanning line or POS data to determine whether the point is a terrain point or not, but the disadvantage is part of low buildings or other feature objects will be lost. If the study area has high differences in the horizontal or the forward direction, this method is not applicable. This paper investigates a new methodology to extract the terrain points, which has great significance for data reduction and classification. The procedure includes the following steps: 1) Pre-processing: to remove discrete points and abnormal points. 2) Divided all the points into grid, calculating the average value of the XY and the minimum value of the Z of all the points in the same grid as the central point of the grid. 3) Choose nearest six points which are close to the centre point to fitting the quadratic surface. 4) Compare the normal vector of the fitting surface of the grid to the normal vector of the 8-neighborhood, if the difference is too big, it will be smoothed. 5) Determine whether the point in the grid is on the surface, if the point belongs to the surface, it will be classified as terrain point. The results and evaluation have shown the effectiveness of the method and its potential in separation of the terrain of various areas.

1. Introduction

Recently years, many researches in laser data have been devoted to separating the terrain point from others, but the number of terrain points is very large and it’s not necessary for classification. As the vehicle moving in the cities is very flexible and the laser’s precision is very high, it is very suitable to obtain the information of near-surface and three-dimension urban modelling.

Nowadays most of the filtering algorithms for point cloud proposed are aimed to deal with the airborne point cloud. Here are some methods for the airborne point cloud separation: [1] proposed a mathematical morphological method which is based on the slope of the scan line and use the slope difference segmentation to filtering the point cloud and separating ground and discrete points. But the disadvantage is that a reasonable height difference threshold must be set, if terrain is changed, the threshold values need to be changed. Besides, the parameters will be repeatedly adjusted based on experience. [3] and [4] designed an adaptive filter TIN structure to solve this problem, the amount of data of the Vehicle-borne Laser data is far greater than the amount of data of the airborne laser data for the same three-dimensional scene, in this case, the data computation of the TIN is enormous, so this
method is not suitable for the removal of the terrain data rapidly and further classification and city modeling.

The methods above are very suitable for low-relief surface, but it is not very satisfactory for undulating terrain which may contain low vegetation. In order to effectively solve this problem, [5] proposed a method which at first use of the least squares principle to fit the surface of the terrain points, and then compare the distance between the points in the grid and the lowest point, if the distance is less than the height difference threshold value, the point will be considered as terrain ground. But the disadvantage of this approach is that it is only effective in flat area, if the study area is more complex, a large number of points near-ground will be misclassified.

The method proposed in this paper is a combination of the advantages and disadvantages of the previous methods of separation, in order to reduce computational complexity, the best and the minimum number of points will be found in the same grid to fit a plane ([6]), and then to iterate until the best plane is fitted. Because of the number of points in the grid is limited, the iteration will not be caught in infinite loop. Furthermore, the method proposed in this paper can make automatic separation of the terrain and non-terrain point. The next section will describes the principle of this method, the improved plane fitting principle, the way to select the best points to fit the best plane and the conditions of whether the point belong to terrain. In section three the set up and the results of a series of experiments are presented. The conclusions and discussion will be given in the last section.

2. Methodology

2.1. Pre-processing

The original data used in this paper has been integration and registration with POS data, and all of them have been converted into geodetic coordinate data. However, due to the instrument or other reasons, there are some vacant, outliers and discrete points. If these points directly be used for the following calculation process, the calculation result of the parameters of the fitting surface will be affected. The de-noising method is as below:

At first, statistical all the points in the same scan line to search for outliers in the scan line. Secondly, count the point density of every grid, and the height distribution characteristics including the highest point and the lowest point. If the height difference is very large and the density is very small, that is used to be determined whether the grid density is less than the threshold, otherwise these discrete points will be removed.

2.2. The improved Quadratic Surface Fitting algorithm

Quadratic surface equation is: $Z = a \cdot X^2 + b \cdot XY + C \cdot Y^2 + d \cdot X + e \cdot y + f$ (1)

To get the values of the parameters of the quadratic surface[7], it requires a minimum of six points. The method presented here is to iteratively select the best six points involved in the calculation for plane fitting.

The parameter calculation in this part is based on the least square principle. In addition, the paper will also simplify the calculation process of traditional least square principle.

The surface equation can be transformed into:

$$a \cdot X^2 + b \cdot XY + C \cdot Y^2 + d \cdot X + e \cdot y + f - Z = 0$$ (2)

For a series of n points in the same grid, \((X_i, Y_i, Z_i)\) \(i = 0,1,6 \ldots, n-1\). Use point \((X_i, Y_i, Z_i)\), \(i = 0,1,6, \ldots, n-1\) fit to calculate equation (1), the aim of the least square principle is to get the S in formula (2) is the smallest:

$$S = \sum_{i=0}^{n-1} (a_0 \cdot X^2 + a1 \cdot XY + a2 \cdot Y^2 + a3 \cdot X + a4 \cdot Y + a5 - Z)^2$$ (3)

If you want to make the S minimum, formula (2) should meet formula (3):
After the data is pre-processed, first of all, set the centre of the grid as the seed point, set the point with the value of Z is minimum belong to ground. But if the angle of surface in the grid is very large, the lowest point must be in the edge of grid. In order to avoid this kind of situation, we need to adjust the “centre point” to a real centre of grid. The adjustment method is given in Equation (8):

\[
\frac{\partial S}{\partial a_k} = 0, \quad k = 0,1,2,3,4,5
\]

Formula (4) can be expanded to formula (5):

\[
\begin{align*}
\sum_{i=0}^{n-1} 2(a_0 \cdot X^2 + a_1 \cdot XY + a_2 \cdot Y^2 + a_3 \cdot X + a_4 \cdot Y + a_5 - Z) \cdot X^2 &= 0 \\
\sum_{i=0}^{n-1} 2(a_0 \cdot X^2 + a_1 \cdot XY + a_2 \cdot Y^2 + a_3 \cdot X + a_4 \cdot Y + a_5 - Z) \cdot XY &= 0 \\
\sum_{i=0}^{n-1} 2(a_0 \cdot X^2 + a_1 \cdot XY + a_2 \cdot Y^2 + a_3 \cdot X + a_4 \cdot Y + a_5 - Z) \cdot Y^2 &= 0 \\
\sum_{i=0}^{n-1} 2(a_0 \cdot X^2 + a_1 \cdot XY + a_2 \cdot Y^2 + a_3 \cdot X + a_4 \cdot Y + a_5 - Z) \cdot X &= 0 \\
\sum_{i=0}^{n-1} 2(a_0 \cdot X^2 + a_1 \cdot XY + a_2 \cdot Y^2 + a_3 \cdot X + a_4 \cdot Y + a_5 - Z) \cdot 1 &= 0
\end{align*}
\]

Further transform can get the formula (6):

\[
\begin{align*}
a_0 \cdot \sum_{i=0}^{n-1} X^4 Y^0 + a_1 \cdot \sum_{i=0}^{n-1} X^3 Y^1 + a_2 \cdot \sum_{i=0}^{n-1} X^2 Y^2 + a_3 \cdot \sum_{i=0}^{n-1} X^1 Y^3 + a_4 \cdot \sum_{i=0}^{n-1} X^0 Y^4 = \sum_{i=0}^{n-1} X^2 Z \\
a_0 \cdot \sum_{i=0}^{n-1} X^3 Y^1 + a_1 \cdot \sum_{i=0}^{n-1} X^2 Y^2 + a_2 \cdot \sum_{i=0}^{n-1} X^1 Y^3 + a_3 \cdot \sum_{i=0}^{n-1} X^0 Y^4 + a_4 \cdot \sum_{i=0}^{n-1} X^1 Y^2 = \sum_{i=0}^{n-1} XYZ \\
a_0 \cdot \sum_{i=0}^{n-1} X^2 Y^2 + a_1 \cdot \sum_{i=0}^{n-1} X^1 Y^3 + a_2 \cdot \sum_{i=0}^{n-1} X^0 Y^4 + a_3 \cdot \sum_{i=0}^{n-1} X^2 Y^1 + a_4 \cdot \sum_{i=0}^{n-1} Y^3 = \sum_{i=0}^{n-1} Y^2 Z \\
a_0 \cdot \sum_{i=0}^{n-1} X^3 Y^0 + a_1 \cdot \sum_{i=0}^{n-1} X^2 Y^1 + a_2 \cdot \sum_{i=0}^{n-1} X^1 Y^2 + a_3 \cdot \sum_{i=0}^{n-1} X^0 Y^3 + a_4 \cdot \sum_{i=0}^{n-1} X^1 Y^0 = \sum_{i=0}^{n-1} XZ \\
a_0 \cdot \sum_{i=0}^{n-1} X^2 Y^1 + a_1 \cdot \sum_{i=0}^{n-1} X^1 Y^2 + a_2 \cdot \sum_{i=0}^{n-1} X^0 Y^3 + a_3 \cdot \sum_{i=0}^{n-1} X^1 Y^1 + a_4 \cdot \sum_{i=0}^{n-1} Y^2 = \sum_{i=0}^{n-1} YZ
\end{align*}
\]

Finally, parameters can be directly expressed as formula (7):

\[
\begin{pmatrix}
a_0 \\
a_1 \\
a_2 \\
a_3 \\
a_4
\end{pmatrix} = \begin{pmatrix}
\sum_{i=0}^{n-1} X^4 & \cdots & \sum_{i=0}^{n-1} X^2 \cdot Y \\
\vdots & \ddots & \vdots \\
\sum_{i=0}^{n-1} X^2 \cdot Y & \cdots & \sum_{i=0}^{n-1} Y^2
\end{pmatrix}^{-1} \begin{pmatrix}
\sum_{i=0}^{n-1} X^2 \cdot Z \\
\sum_{i=0}^{n-1} XYZ \\
\sum_{i=0}^{n-1} Y^2 \cdot Z \\
\sum_{i=0}^{n-1} XZ \\
a \sum_{i=0}^{n-1} YZ
\end{pmatrix}
\]

If the number of points is only six, relevant parameters of the equation (1) will be obtained directly by calculate formula (7).

2.3. The choice of the centre point and the proximity point

2.3.1. The choice of the centre point

After the data is pre-processed, first of all, set he centre of the grid as the seed point, set the point with the value of Z is minimum belong to ground. But if the angle of surface in the grid is very large, the lowest point must be in the edge of grid. In order to avoid this kind of situation, we need to adjust the “centre point” to a real centre of grid. The adjustment method is given in Equation (8):
\[ x_0 = \frac{\sum_{i=0}^{n-1} X_i}{n} \]
\[ y_0 = \frac{\sum_{i=0}^{n-1} Y_i}{n} \]
\[ z_0 = \min_{i=1 \rightarrow n} (Z_i) \]

Where \( n \) is the total number of the points in the same grid, \((x_0, y_0, z_0)\) is the centre point of grid, \((X_i, Y_i, Z_i), i=1,2 \ldots n\), is the point in grid.

2.3.2. The choice of the best proximity points
In order to get the best surface which representing the distribution of grid, you’ld better choose the nearest six points away from the centre point, the way to select the nearest points is as figure 1:

\begin{itemize}
  \item a. Count all the distance \((H_1)\) between all the points in grid and the centre point in ascending order.
  \item b. Choose the nearest six points away from the centre point, in accordance with formula (6), the vector normal of the inner surface can be obtained by the curved surface parameters. Then we will calculate the angle between the vector normal of this grid and the eight adjacent grid’s vector normal, if the average value of the angles is greater than the threshold \( K \), the optimum points will be re-selected, it will go to step c, otherwise the iteration termination.
  \item c. Calculate the distance between optimal points and the surface, remove the farthest point, and then select the point with the minimum distance \( W \) which is close to the centre point and the surface. The following is the method of calculation of \( W \):
  \[ w = \frac{H_1 + H_2}{2} \]
  Where \( H_1 \) is the distance between point and the centre point, and the \( H_2 \) is the distance between point and the surface. The proportion of it is 1:1.
  \item d. Go to step b
\end{itemize}

Figure 5. The selection process of the best points
In the worst situation, which means all of the points satisfy the distance condition, the best surface is still not able be fitted. We will directly select the lowest point in grid as the terrain point, and use the height difference threshold to determine whether or not the points is belong to the terrain.

2.4. Determination of the ground

By using the method above, we got the lowest surface of each grid, but still can’t accurately know whether or not they belong to the terrain. If there is a car or a human being passing by, the terrain points behind it will not be obtained. This step is required to determine whether the surface we get in grid belongs to terrain.

It is widely accepted that terrain is flat to gentle transformation([8]), so if there are two surfaces adjacent to each other and the angle between them is less than a threshold \( \theta \), they must belong to the same plane.

The terrain surface must satisfy the following two conditions:

**Condition One**: Compare the distance in the normal vector ([9]) between the centre point and the centre point of the adjacent grid, the distance difference \( D_i \) (i=1,2,…,8) must be less than threshold \( D \).

**Condition Two**: Compare the angle between the normal vector \( n \) of the surface in grid and the eight adjacent grid’s vector normal \( n_i \), i=1,2,…,8, all of the angle difference \( \phi_i \) ,i=1,2,…,8 must be less than threshold.

If the surface can satisfy the two conditions above at the same time, the points belongs to the surface can be considered as terrain point (figure 2).

![Diagram](image)

**Figure 2.** The conditions which the terrain points must meet

![Image](image)

The original data

![Image](image)

The data without terrain

**Figure 3.** The result of this study
3. Result
The experimental data of this article is mainly located in the urban area of Beijing, and the suburbs, point spacing of 0.02m points to 300,000. Figure 3 is the result of this experiment, the terrain and non-terrain which contains vegetation and other coverings are represented. Figure 2 is the separation results of the method in this paper in other regions. From the result above, we can find that the method in this paper is very suitable for the vehicle-borne laser scanning data. As the set of constraints, many small objects near the ground such as a moving car human beings will be not separated as terrain points, these is a need for further processing of these near-ground data.

4. Conclusions and Discussion
This paper presented a simplified method for the separation of the terrain and non-terrain point from Vehicle-borne laser scanning data. First of all, all of the points will be vertical projected into grid, the optimum six points will be selected through iteration, and it will be used to fit the quadratic surface parameters for the normal vector, and then the normal vector will be compared with the normal vector of the eight adjacent grids, the angle between the normal vectors will be used to determine whether the point on the surface belongs terrain.

There are multiple iterations in this study, so we have to repeatedly comparison and adjustment process with the eight adjacent grids. But we only select the best six points for calculation which means it has greatly reduce the number of iterations and the computational. From the experimental results, we can find that the accuracy of surface parameters is very close to the reality, the method proposed in this paper is very suitable for automatic separation of the high precision and high density vehicle-borne laser data.

Acknowledgments
The work of this paper is supported by the open fund from State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University (NO:11P02) and the National Science and technology support program (NO: 2012BAH34B01)

References
[1] Vosselman G, 1997. Slope based filtering of laser altimetry data[C] International Archives of Photogrammetry and Remote Sensing 32(3-4w2) 19-26
[2] Vosselman G, Gorte, B.G.H., Sithole, G. and Rabbani, T. 2004. Recognising structure in laser scanning point clouds. The international Archives of Photogrammetry Remote Sensing and Spatial Information Science Vol. XXXVI, 2004 part 8/W2 Freiburg Germany,33-38
[3] Qiang Yunjuan Wu Kan, 2011. The exclusion method for removing noise in ground point clouds based on the 3D TIN. Engineering of Surveying and Mapping, 20(2), 43-45
[4] Hofmann, A., 2004. Analysis of TIN-structure parameter spaces in airborne laser scanner data for 3-D building model generation. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 35(B3), pp. 302-307.
[5] Peng Zhi, Li Chuan-rong, 2011. Study on Filtering of Point Cloud Based on Least Squares Curve Fitting Method. Remote Sensing Technology and Application. 86-89
[6] Lalonde, J.F., 2005. Scale Selection for Classification of Point-sampled 3-D Surfaces. Fifth International Conference on 3-D Digital Imaging and Modeling,285-292
[7] Hu Jianwei, 2009. Mesh Segmentation Based on 3D Delaunay Triangulation Journal of Computer-Aided Design & Computer Graphics, 21(12):1730-1732
[8] Filin, S., Pfeifer, N., 2006. Segmentation of airborne laser scanning data using a slope adaptive neighborhood. ISPRS Journal of Photogrammetry and Remote Sensing, 60(2), pp. 71-80
[9] Niloy J, An Nquyen, 2003. Estimating surface normals in noisy point cloud data. Proceedings of the nineteenth annual symposium on Computational geometry, 322-328