Parameter analysis of disc fitting method

Gábor Nagy*
*Óbuda University, Alba Regia Technical Faculty, Institute of Geoinformatics
nagy.gabor@amk.uni-obuda.hu

***Abstract*—The Disc Fitting method is an point cloud processing algorithm which can calculate the elevation of the ground surface in a position. This result depends on two parameters: an radius (R) and an quantile (q), because the Fitting Disc method selects the points of the cloud that are nearest the examined position that R, and fit a plane (a disk) to this area whose three sectors contain q quantile of the points under the plane. This article describes some study about the accuracy of the Disc Fitting method with different R and q parameters.**

***Index Terms*—point cloud, LiDAR, Digital Elevation Model**

I. Introduction

The airborne LiDAR technology is a very productive method of spatial data acquisition. The created point cloud is a huge data set which contains millions (or some cases billions) of points and provides a very detailed representation of the terrain, but each point means a small bit of the information: the laser beam reflected from something in these points.

The elevation models describes the terrain surface. The most of the points of a LiDAR point cloud are located over the ground surface on different natural and artificial objects, for example trees and bushes. There are varied processing method to create digital elevation models, for example [3], [9] . This article studies the Disc Fitting method [7].

II. The Disc Fitting method

The Disc Fitting method fits a plane to a point cloud and a horizontal position so, that the points located nearest than an R radius to the position are selected, and this circle is divided to three sectors. The q portion (for example 50% in $q = 0.5$) of the points are under the fitted plane (the part of this plane in the R radius circle is a disk) in each sector.

All sectors have an control point in the center of these sectors (the horizontal coordinates are fixed), and the elevation of the plane in these points are defined the plane. The parameters of the plane can be estimated by an iterative algorithm. The initial values are the q quantiles of the elevations of the points in the sector. The iteration steps modify the elevation of the control point of the sector so, that the q quantile of the points will be under the plane (the two other control point do not change). These iteration steps are repeated around the three sector until the difference between the before and the after elevation of the sector's center is lower than a limit.

The three parameters of the equation of the plane (H_0, a and b in the $H = H_0 + a\Delta X + b\Delta Y$) can be determined from the three elevations of the sector's centers (and backwards). The plane has an elevation in the inspected position (H_0) and two slope (a and b) in the two horizontal axes. The Disc Fitting

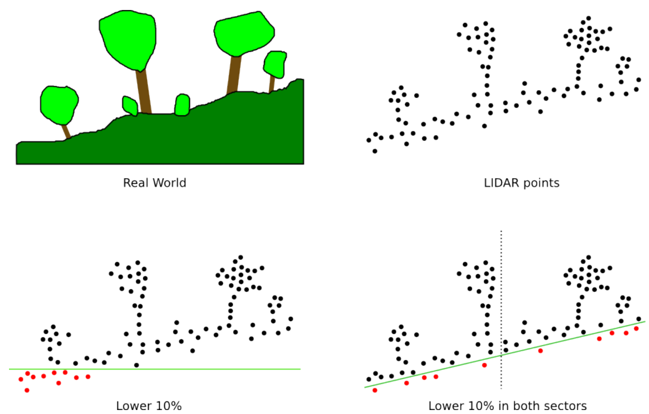

Figure 1. The principle of the method in two dimension. In this case, the area is divided to two sectors, because the regression line has two parameters. The 10% of the points ($q = 0.5$) are located under the regression line in all sectors.

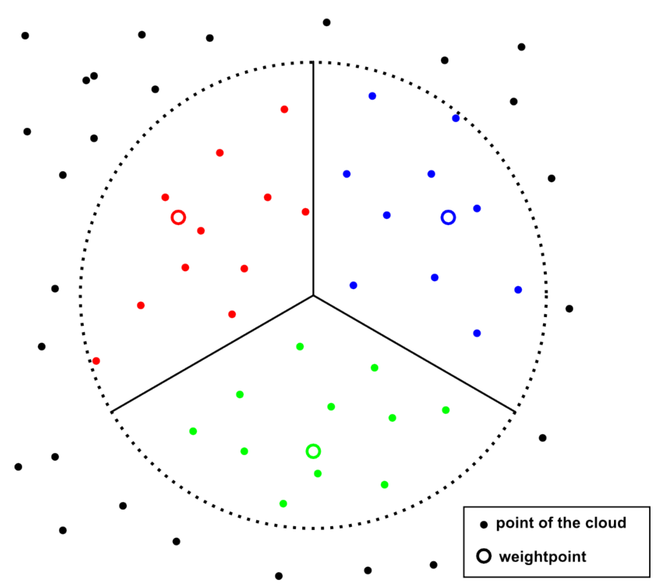

Figure 2. The sectors of the method in three dimension. Disc Fitting is applied in the R radius of an inspected horizontal positions, the regression plane is calculated from the points of this area. The circle area is divided to three sectors denoted by different colors. The q portion of the points are located under the regression line in all sectors.

method can determine an elevation in any horizontal position, and the slope (and aspect) of the terrain on this position.

Disc Fitting method can be computed in any horizontal position. If these positions are the points of a grid, an GRID type digital elevation model will be made. This calculating process is well applicable with parallel computing and distributed computing, because the calculation is independent in each grid point. The method can work without any filtering process (neither before nor after).

The mathematical background of the Disc Fitting method is an quantile linear regression. This linear regression can be generalized to any dimension. In an N dimension space, the regression has $N-1$ independent and 1 dependent variables (in case of Disc Fitting, the two independent variables are the horizontal coordinates, and the dependent variable is the elevation), and the area will be divided to $N+1$ sectors, because the N dimension plane has $N+1$ parameters. [6]

III. The effect of the parameters

Disc fitting method has two important parameters: the radius of the disc (R), and the quantile of the linear regression (q). In wooden areas a bit of the points are located on the terrain surface, therefore the method needs very low q value. The very low q value needs many points, and many points need large R (radius) value. The large R means large disk, and the terrain surface is modeled by a plane in this large area, but the terrain surface is not sure a plane in a large area, the small details of the surface may be lost.

IV. The study area

I wanted to compare the Disc Fitting method with different R and q parameters and other LiDAR point cloud processing methods. I surveyed the ground surface in 195 points. This study area located in Iszkaszentgyörgy, near Székesfehérvár. I have an airborne LiDAR point cloud from this area which was made in May of 2008.

The survey made with total station and GNSS technology in summer of 2015. The elevations of the surveyed points was accepted the elevation of the ground surface in the horizontal positions of the points. The elevation was calculated from the point cloud by different methods: the Disc Fitting methods with various parameters and the TopoSys Filled Digital Terrain Model (FDTM) which was created by the LiDAR surveying company.

V. The result by different R and q parameters

The differences between the elevation of the surveyed ground surface and the surface of the examined elevation model are calculated in each test points in each models. The distribution of the differences can be illustrated in a violin plots [1]. (Figure)

The violin plot is a good tool for illustrating the distributions, but one diagram can show only some R and q parameter-pairs. I created other plots for showing various R and q pairs. Each parameters generated by geometric series. The first element of the R sequence is $a = 1\,m$ and the common ratio

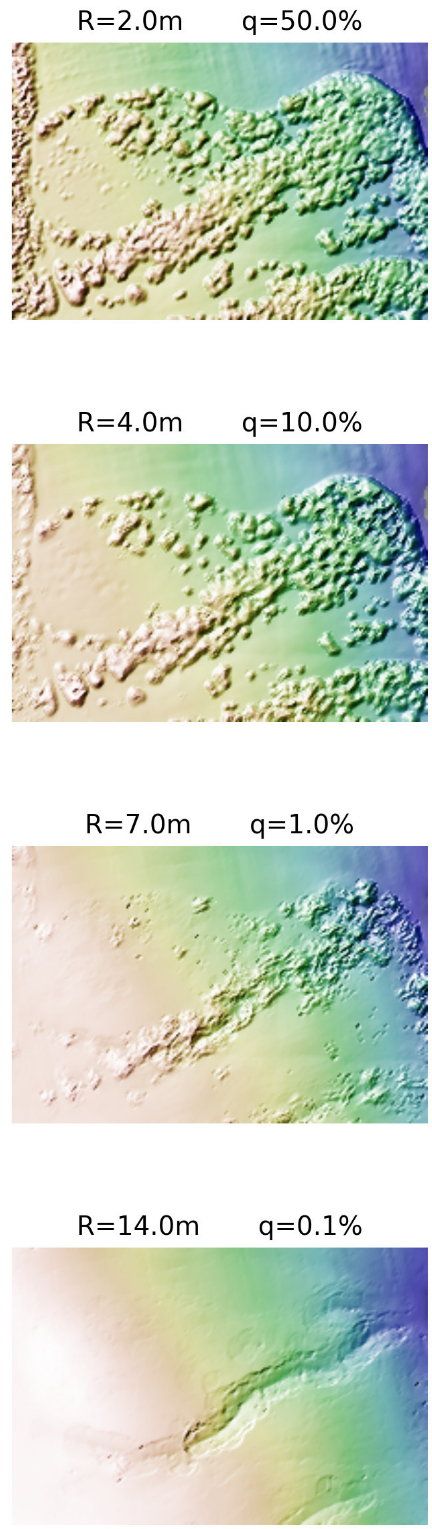

Figure 3. The processed surfaces with different R and q values. The $R = 2,\ 4,\ 7,\ 14\,m$ and $q = 0.5,\ 0.1,\ 0.01,\ 0.001$ (from the top to the bottom).

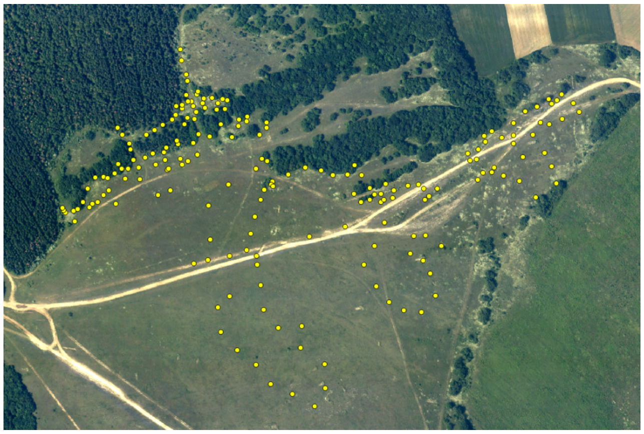

Figure 4. The surveyed points in the test area.

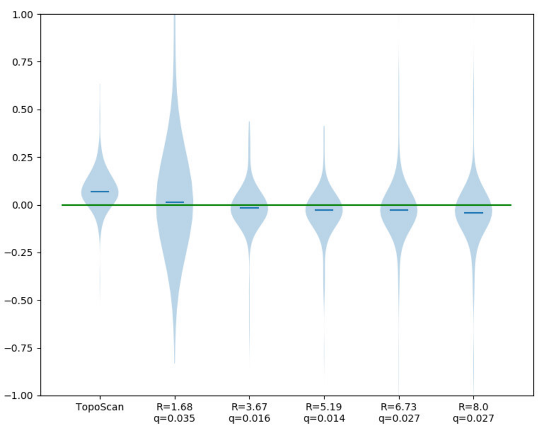

Figure 5. The distribution of the differences with some R and q values. (The first graph is the TopoSys FDTM.) This violin plot made by Matplotlib [2], [5].

is $r = \sqrt[8]{2}$. The both of first element and common ratio are $a = r = 0.875$ in the q sequence.

A Python [8], [4] program calculate the average of absolute values of differences in each R and q pair. These values are shown by a color scale in the Figure.

VI. The thickness of the point cloud

The thickness of the point cloud is defined by the difference of the elevation with two different quantiles. In this article the thickness of the point cloud is defined as the Disc Fitting elevations with $q = 0.875$ and $q = 0.118$. The $R = 4\,m$ in both of these cases.

The test area may be divided to two about the same size set: the thin areas where the thickness is less than $0.4\,m$, and the thick areas where the thickness is greater than $0.4\,m$. The plot of the Figure 6. can be created separately in the thin and thick areas (Figure 7). The optimal parameter pairs are different the thin and the thick areas.

The Figure 5. also can be created separately. The difference of the distributions can be seen in the Figure 8.

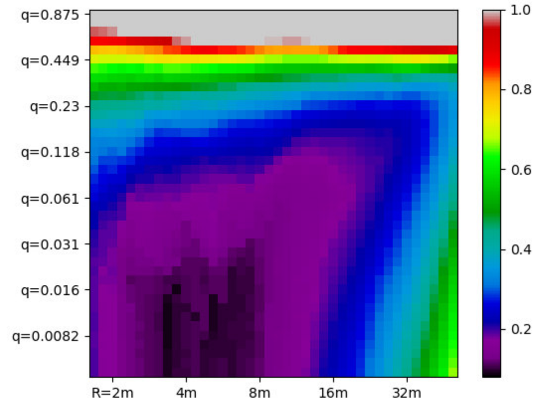

Figure 6. Averages of the absolute values of the differences in different R and q values. The plot made by Matplotlib [2], [5].

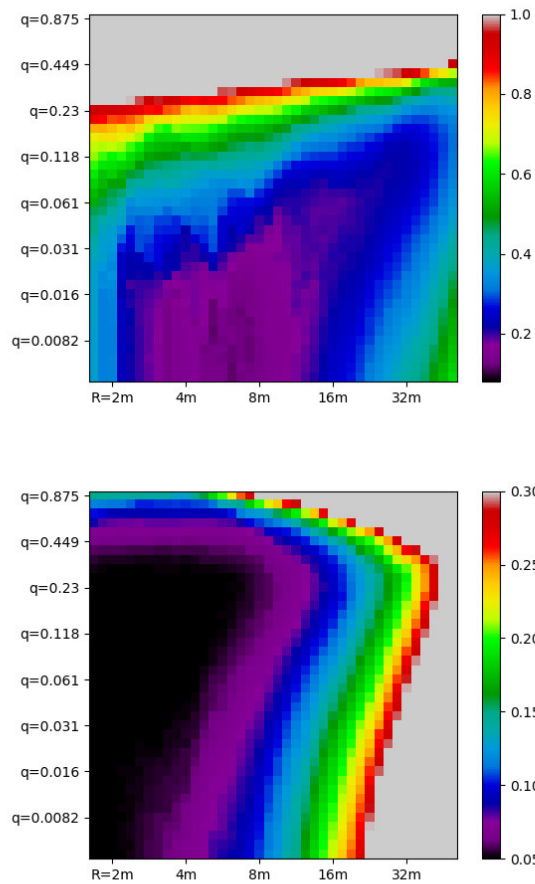

Figure 7. Averages of the absolute values of the differences in different R and q values in the thick (up) and the thin (down) areas. The plot made by Matplotlib [2], [5].

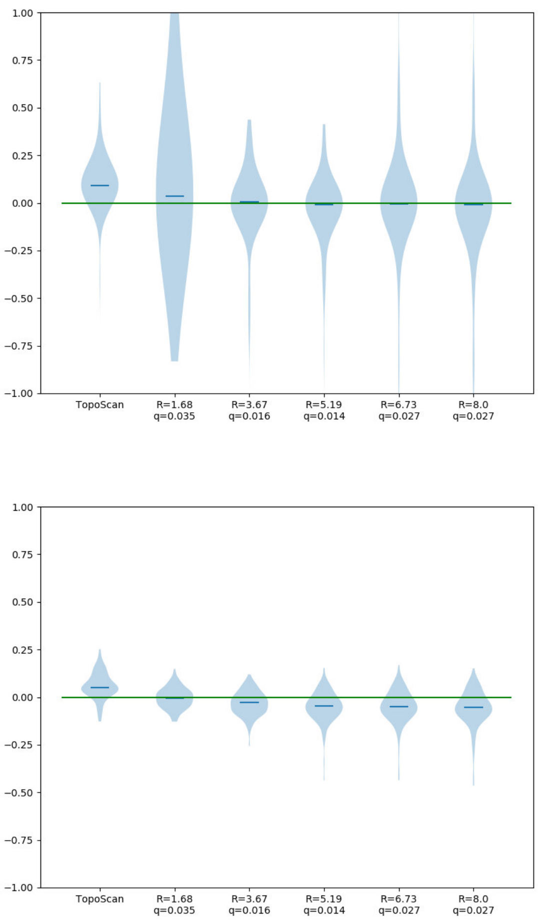

Figure 8. The distribution of the differences with some R and q values in the thick (up) and the thin (down) areas. (The first graphs are the TopoSys FDTM.) This violin plots made by Matplotlib [2], [5].

VII. Using variable R and q parameters

As we have seen before, different parameters (R and q) given good results in different point cloud thickness. A two-step method can be applied: the firs step calculates the thickness (from two Disc Fitting), and the second step calculates the elevation of the terrain surface with parameters depended by the thickness.

In my research I use a simple method, but this simple method provide good results. The thickness is calculated between $q = 0.875$ and $q = 0.118$ with $R = 4\,m$. In the thin areas (where the thickness is less than 40 centimeters) the $R = 1.68\,m$ and $q = 0.035$. In the thick areas (where the thickness is greater than 40 centimeters) the $R = 5.19$ and $q = 0.14$.

The violin plots are created with this thickness depended parameters in the Figure 9.

VIII. Conclusion

The Disc Fitting method can create good quality digital elevation models from LiDAR point clouds, if we use suitable R and q parameters. An area may be combined different parts, different parameters may be suitable in these parts.

This topic need more research in further test areas and sample data. More test points can provide more sophisticated models for the ideal R and q values.

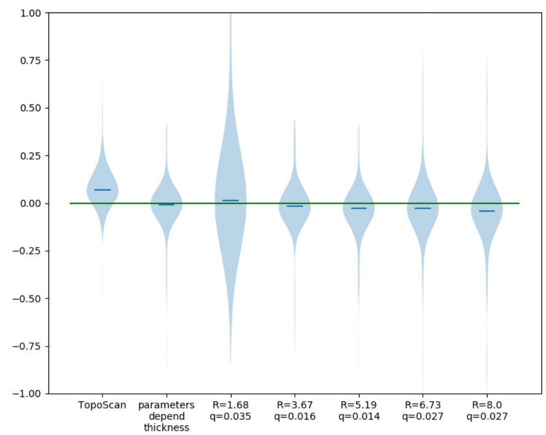

Figure 9. The distribution of the differences with some R and q values. The second graph is the distribution of the thickness depended method. (The first graph is the TopoSys FDTM.) This violin plot made by Matplotlib [2], [5].

References

[1] Jerry L Hintze and Ray D Nelson. Violin plots: a box plot-density trace synergism. *The American Statistician*, 52(2):181–184, 1998.

[2] J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing In Science & Engineering*, 9(3):90–95, 2007.

[3] Xuelian Meng, Nate Currit, and Kaiguang Zhao. Ground filtering algorithms for airborne LiDAR data: A review of critical issues. *Remote Sensing*, 2(3):833–860, 2010.

[4] K Jarrod Millman and Michael Aivazis. Python for scientists and engineers. *Computing in Science & Engineering*, 13(2):9–12, 2011.

[5] G. Nagy. Matplotlib in the geospatial analyzes. In Gábor Tamás Orosz, editor, *12th International Symposium on Applied Informatics and Related Areas (AIS 2017)*, pages 90–93. Óbudai Egyetem, 2017.

[6] Gábor Nagy. Sector based linear regression, a new robust method for the multiple linear regression. *Acta Cybernetica*, 23(4):1017–1038, 2018.

[7] Gábor Nagy, Tamás Jancsó, and Chongcheng Chen. The fitting disc method, a new robust algorithm of the point cloud processing. *ACTA POLYTECHNICA HUNGARICA*, 14(6):59–73, 2017.

[8] Travis E Oliphant. Python for scientific computing. *Computing in Science & Engineering*, 9(3), 2007.

[9] George Sithole and George Vosselman. Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. *{ISPRS} Journal of Photogrammetry and Remote Sensing*, 59(1-2):85–101, 2004. Advanced Techniques for Analysis of Geo-spatial Data.