Chapter 3

Basics of LiDAR Data Processing

Dong, Pinliang; Chen, Qi.
LiDAR Remote Sensing and Applications
CRC Press

3.1 INTRODUCTION

- Since global positioning system (GPS) uses the World Geodetic System of 1984 (WGS84) datum, the originally derived three-dimensional (3D) coordinates are also georeferenced to the WGS84 datum and its ellipsoid (this means that z is ellipsoidal elevation).

- A LiDAR vendor often needs to transform the original 3D coordinates of laser points from WGS84 to a new horizontal datum and/or a new vertical datum, especially upon the request of the end users.

- The processing and analysis methods for LiDAR data are usually application-specific (for example, multi-scale modeling for building extraction), and many new methods are being proposed.

3.1 INTRODUCTION

- However, two basic steps are usually needed in LiDAR data processing & analysis: (1) classification of laser points and (2) interpolation of discrete points into a continuous surface. For example, the generation of digital terrain model (DTM) requires the classification/extraction of ground returns and the interpolation of ground returns into raster or triangulated irregular network (TIN).

- Extracting ground returns from a point cloud is usually the first step of LiDAR data processing and the most important classification. Conventionally, this step is called “filtering” because the main driver of adopting airborne LiDAR data in its infancy stage (in the 1990s) was to generate DTM, which needs filtering out or removing non-ground returns.

- Although various interpolation methods exist in software, they have to be fast enough to process massive LiDAR points.

3.1 INTRODUCTION

- This chapter will introduce filtering, classification of non-ground returns, and spatial interpolation, respectively.

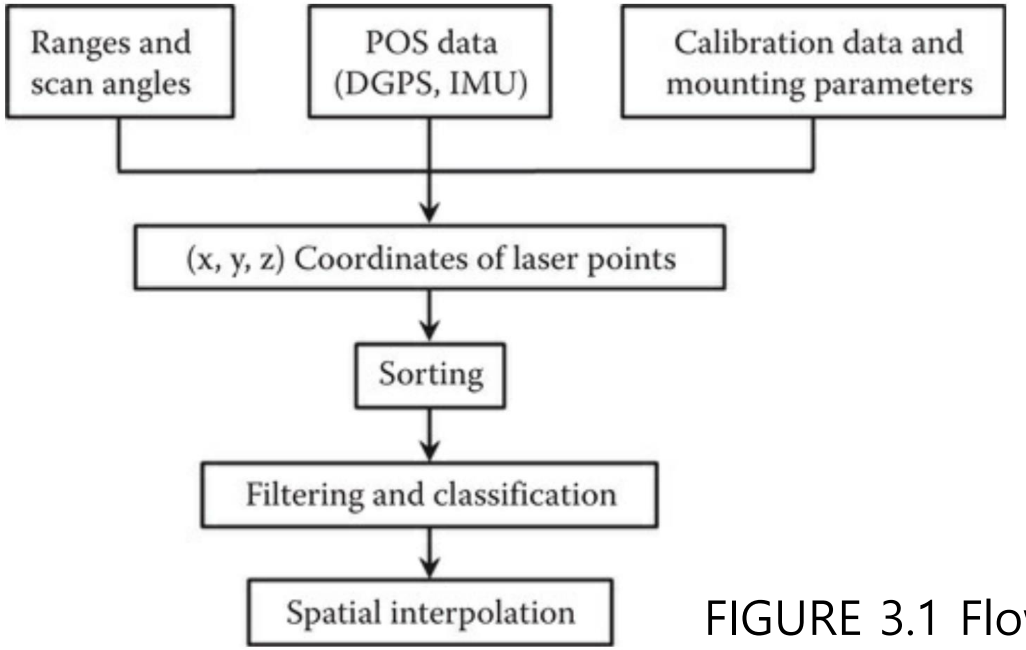


FIGURE 3.1 Flowchart of LiDAR point data processing.

3.2 FILTERING

- Filtering is used to remove non-ground LiDAR points so that bare-earth digital elevation models can be created from the remaining ground LiDAR points.

- Most filtering methods are unsupervised classifiers, which mean that users do not need to collect training data for ground and non-ground returns.

- The design of a filtering algorithm is usually based on two criteria: (1) the ground has the lowest elevation compared to the objects above it, and (2) elevation and slope change more slowly for bare earth than for DSM (Digital Surface Model).

- Among the large variety of filtering algorithms that have been developed, surface-based approaches are probably the most popular and effective ones and have been implemented in many commercial or free software.

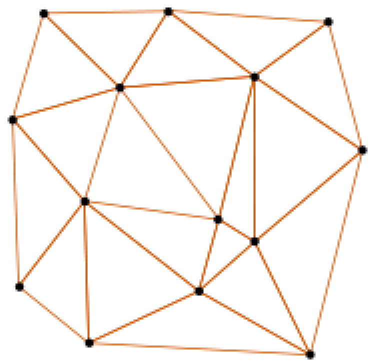
3.2 FILTERING

- Surface-based approaches usually start with an initial surface that approximates the bare earth and then generates another approximate surface of the bare earth that utilizes the information from the previous step. This process could be repeated iteratively until the next surface does not substantially differ from the previous one.

- Surface-based approaches can be based on either TIN (Triangulated Irregular Network, 불규칙삼각망) or raster.

3.2.1 TIN-Based Methods

- A TIN surface consists of a tessellated network of non-overlapping triangles, each of which is made of irregularly distributed points.



3.2 FILTERING

- TIN is well suited for constructing terrain surface from LiDAR points because: (1) LiDAR points are often irregularly distributed due to variations in scan angle, attitude (pitch, yaw, and roll) of the airplane, and overlaps between flight lines; (2) adding or removing points into TIN can be implemented locally without reconstructing the whole TIN; and (3) the speed of constructing a TIN is usually much faster than the grid-based interpolation methods.

- There are many different triangulation networks, but the Delaunay triangulation is the standard choice.

- The algorithm proposed by Axelsson (2000, Adaptive TIN) is probably the most famous TIN-based method, which works as follows: (1) the lowest points within a coarse grid are chosen as seed ground returns, which are used to construct an initial TIN. (2) points are added into the TIN if they are close to the triangular facet and the angles to their overlaying triangular nodes are small, and (3) the densification of the TIN continues until no more ground returns can be added.

3.2 FILTERING

- The main idea of Axelsson (2000) is to start with a small set of ground returns and then iteratively add the remaining the ground returns. Different from such an “addition” strategy, ground returns can also be extracted via “subtraction”: start with all returns and then iteratively remove non-ground returns; the remaining points in the end of iteration are ground returns.

3.2.2 Raster-Based Methods

- An alternative data structure for terrain surface is raster. Various interpolation methods (also called interpolators) can be used to generate raster grids from points.

- Interpolators can be exact or inexact, depending on whether the interpolated surface will go through the points or not.

3.2 FILTERING

- The interpolators that were often used are kriging, thin-plate spline (TPS), inverse distance weighting (IDW), and natural neighbors.

- Similar to TIN-based approaches, raster-based filtering algorithms can be designed based on either an “addition” or “subtraction” strategy.

- As an example of the “additive” method, Chen et al. (2013) developed a multiresolution hierarchical classification (MHC) algorithm for separating ground from non-ground LiDAR point clouds based on point residuals from the interpolated raster surface. The MHC algorithm uses three levels of hierarchy from coarse to high resolutions, and the surface is iteratively interpolated towards the ground using TPS at each level, until no ground points are classified. The classified ground points are then used to update the ground surface in the next iteration.

3.2 FILTERING

- Kraus and Pfeifer (1998) were among the first to develop a “subtractive” raster filtering algorithm. Using simple kriging (also called linear interpolator in their paper), they first fitted a surface using all returns. Then, points with large positive residuals are removed and the rest are assigned with weights according to their residuals: points with large negative residuals are more likely to be ground returns and therefore assigned with larger weights. The surfaces are iteratively refitted with remaining points with weights until convergence.

- One of the main challenges in interpolation from points to raster is the demanding computation involved because (1) a LiDAR file typically has several million points per square kilometer, and (2) interpolation usually involves the use of points within each point’s local neighborhood, which requires extra computation resources for indexing and searching.

3.3 CLASSIFICATION OF NON-GROUND POINTS

- After separating ground and non-ground LiDAR points through filtering, non-ground points can be further classified into buildings, roads, vegetation and other classes.

- Various methods have been proposed for LiDAR point classification, including unsupervised classification, bayesian networks, decision trees, and support vector machines. In addition, LiDAR intensity data and multispectral image data can be combined with LiDAR point data for classification. A review of some of the LiDAR point classification methods can be found in the work of Yan et al. (2015).

- Figure 3.3 is a classification map of LiDAR points, and Figure 3.4 is a profile of LiDAR points derived from P-P′ in Figure 3.3.

3.3 CLASSIFICATION OF NON-GROUND POINTS

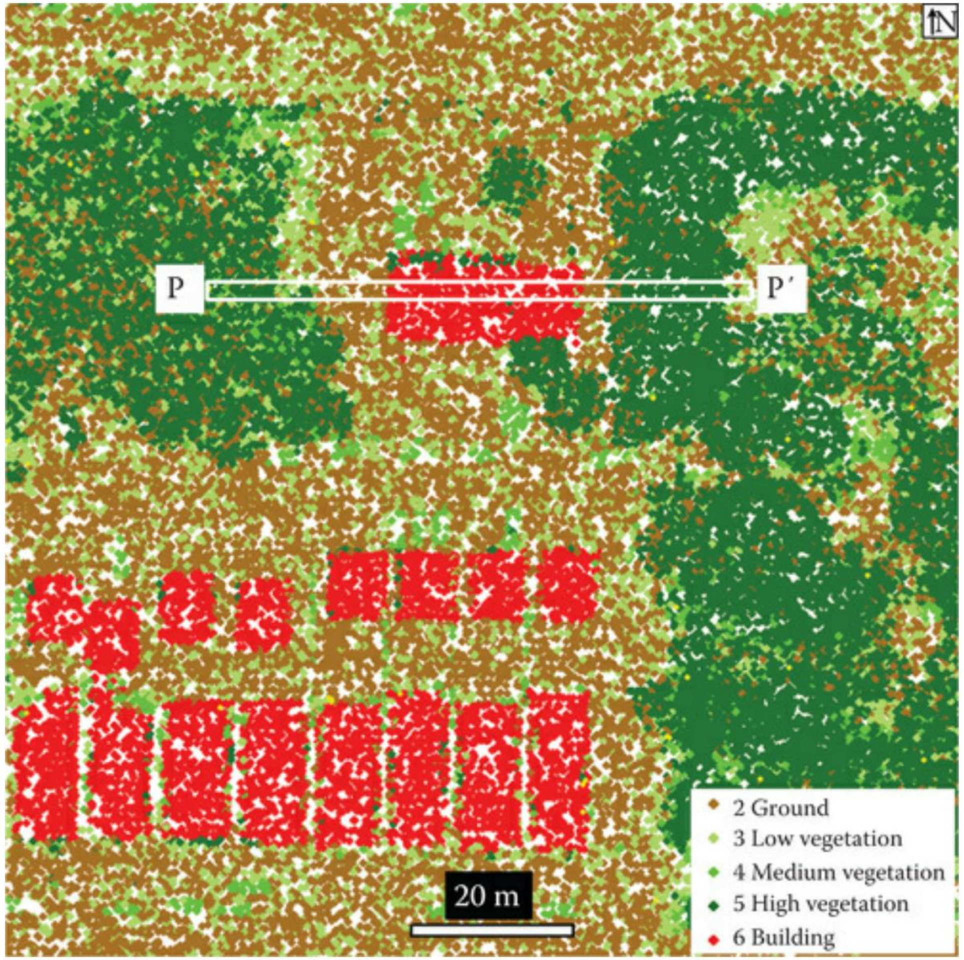


FIGURE 3.3 Classification of LiDAR points.

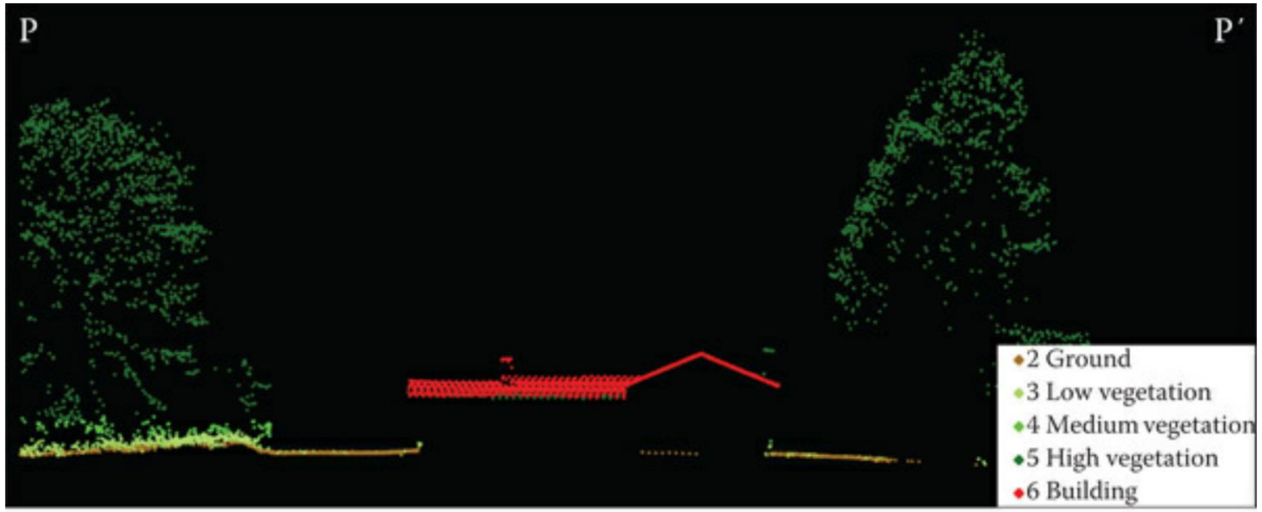


FIGURE 3.4 Profile of LiDAR points derived from P-P′ in Figure 3.3.

3.3 CLASSIFICATION OF NON-GROUND POINTS

➢ IDW (Inverse Distance Weighted) interpolation

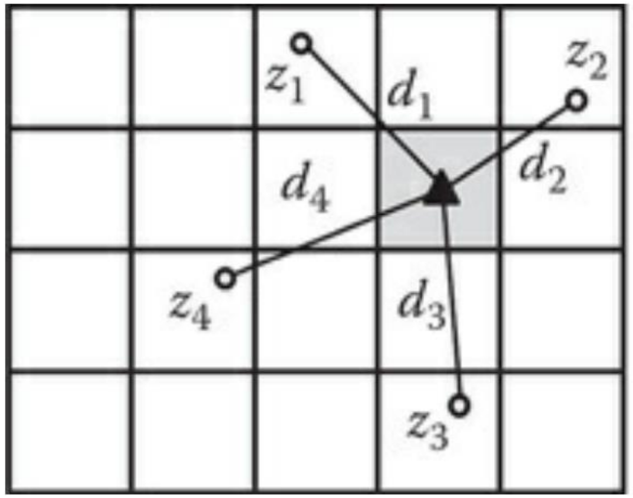
- IDW explicitly relies on Tobler's First Law of Geography: "Everything is related to everything else, but near things are more related than distant things.":

$$Z_j = \begin{cases} \dfrac{\sum_{i=1}^{n}\left(\dfrac{Z_i}{d_i^p}\right)}{\sum_{i=1}^{n}\left(\dfrac{1}{d_i^p}\right)}, & \text{if } d_i \neq 0 \text{ for all } i \\ Z_i & \text{if } d_i = 0 \text{ for some } i \end{cases}$$

- where *Zi* is the observation at the *i*th point, *Zj* is the interpolated value at output location *j*, *di* is the distance between the ith input point and the output location *j*, and *p* is the power of distance. If *di* = 0 for some *i*, the observation *Zi* is used as the output. If *p* = 1, a simple linear distance is used. A faster rate of distance decay may be obtained if *p* > 1, but a common practice is to use *p* = 1 or *p* = 2.

3.3 CLASSIFICATION OF NON-GROUND POINTS

- Figure 3.5 shows how to calculate the interpolated value Z_j at the output cell location (triangle) from four observations (z_1, z_2, z_3, and z_4) using IDW interpolation with p = 2.



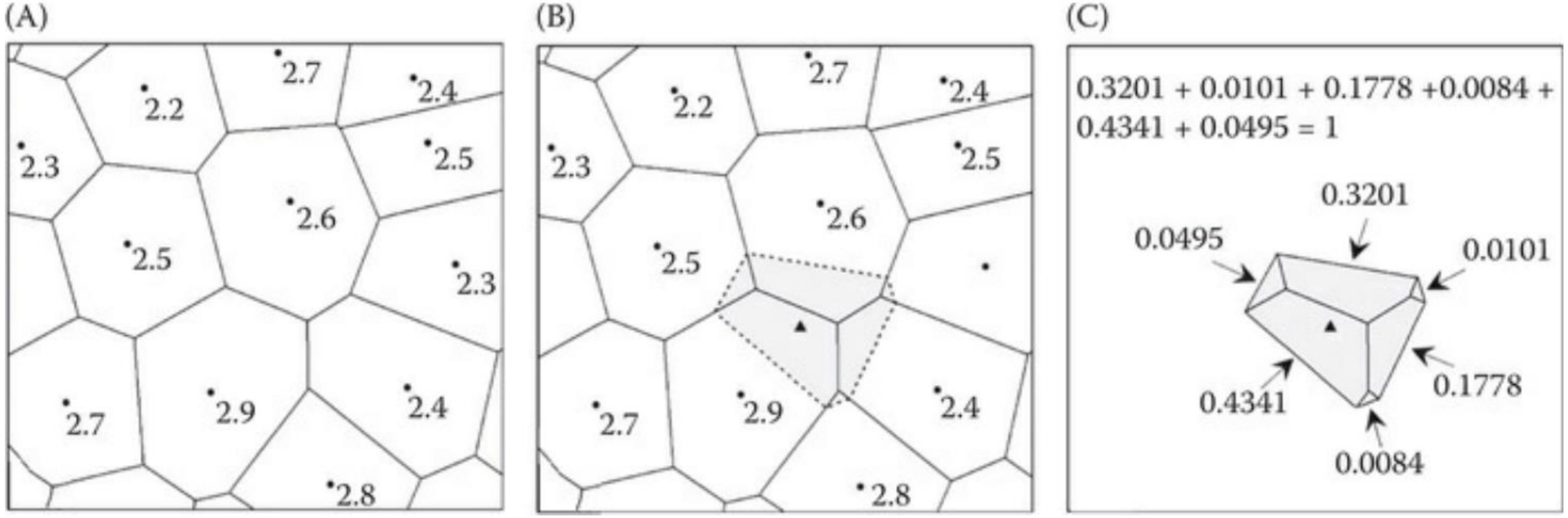
$$Z_j = \frac{\frac{z_1}{d_1^2} + \frac{z_2}{d_2^2} + \frac{z_3}{d_3^2} + \frac{z_4}{d_4^2}}{\frac{1}{d_1^2} + \frac{1}{d_2^2} + \frac{1}{d_3^2} + \frac{1}{d_4^2}}$$

FIGURE 3.5 Inverse distance weighting (IDW) interpolation

3.3 CLASSIFICATION OF NON-GROUND POINTS

➢ Natural neighbor interpolation

- Natural neighbor interpolation, developed by Sibson (1981), uses weights for each of the input points based on their area of influence, which is determined by Voronoi (Thiessen) polygons around each input point: $g(x, y) = \sum_{i=1}^{n} w_i f(x_i, y_i)$



$Z_j = 0.3201 \times 2.6 + 0.0101 \times 2.3 + 0.1778 \times 2.4 + 0.0084 \times 2.8 + 0.4341 \times 2.9 + 0.0495 \times 2.5 = 2.7$

FIGURE 3.6 Natural neighbor interpolation. (A) Voronoi diagram for input points; (B) Output location (triangle symbol); (C) Weights for natural neighbors of the output location.

Projects

- PROJECT 3.1: CREATING DTM, DSM, AND DHM FROM LIDAR DATA IN INDIANAPOLIS, IN, USA
- PROJECT 3.2: GENERATING A TERRAIN DATASET USING LiDAR DATA FROM ST. ALBANS, VT, USA