

## RASTER VS. POINT CLOUD LIDAR DATA CLASSIFICATION

N. El-Ashmawy<sup>ab\*</sup>, A. Shaker<sup>a</sup>

<sup>a</sup> Ryerson University, Civil Engineering Department, Toronto, Canada - (nagwa.elashmawy, ahmed.shaker)@ryerson.ca

<sup>b</sup> Survey Research Institute, National Water Research Center, Cairo, Egypt

### Commission VII, WG IV

**KEY WORDS:** LiDAR, intensity data, Land Cover classification, Point Clouds.

### ABSTRACT:

Airborne Laser Scanning systems with light detection and ranging (LiDAR) technology is one of the fast and accurate 3D point data acquisition techniques. Generating accurate digital terrain and/or surface models (DTM/DSM) is the main application of collecting LiDAR range data. Recently, LiDAR range and intensity data have been used for land cover classification applications. Data range and Intensity, (strength of the backscattered signals measured by the LiDAR systems), are affected by the flying height, the ground elevation, scanning angle and the physical characteristics of the objects surface. These effects may lead to uneven distribution of point cloud or some gaps that may affect the classification process. Researchers have investigated the conversion of LiDAR range point data to raster image for terrain modelling. Interpolation techniques have been used to achieve the best representation of surfaces, and to fill the gaps between the LiDAR footprints. Interpolation methods are also investigated to generate LiDAR range and intensity image data for land cover classification applications. In this paper, different approach has been followed to classifying the LiDAR data (range and intensity) for land cover mapping. The methodology relies on the classification of the point cloud data based on their range and intensity and then converted the classified points into raster image. The gaps in the data are filled based on the classes of the nearest neighbour. Land cover maps are produced using two approaches using: a) the conventional raster image data based on point interpolation; and b) the proposed point data classification. A study area covering an urban district in Burnaby, British Columbia, Canada, is selected to compare the results of the two approaches. Five different land cover classes can be distinguished in that area: buildings, roads and parking areas, trees, low vegetation (grass), and bare soil. The results show that an improvement of around 10% in the classification results can be achieved by using the proposed approach.

### 1. INTRODUCTION

Airborne Laser Scanning system (ALS) acquires high accurate 3D point data of the ground surface using the LiDAR technology. The 3D coordinates of the point clouds are determined based on both the LiDAR range data, and the position and orientation data, (Wehr, 1999). The integrated GPS and IMU units system on board of ALS systems are used to determine the position and orientation of the aircraft at the moment of sending and receiving the Laser signals. The reflecting surface characteristics (such as the surface material and the inclination angle) influence the backscattered energy received by the Laser scanner, and hence the LiDAR intensity data (Hug and Wehr, 1997).

The distribution of the collected point clouds depends on the scanning pattern of the ALS, (Wehr, 1999). The ALS missions collect data in strips from different flight lines, with side overlap of around 30%, (Rentsch and Krzystek, 2009). The different characteristics of the flight lines lead to having different distribution of point clouds in elevation and position, and gaps appear between the LiDAR footprints. Additionally, the characteristics of the features on the ground affect distribution of the point clouds.

Commonly for producing land cover maps from LiDAR data, the point data are converted into raster format and the gaps between the points are estimated using one of the interpolation

techniques. Several interpolation techniques such as inverse distance weighting, median filtering and Kriging are investigated. In Song et al, 2002, different interpolation techniques were investigated to check their effect on the classification results. It was found that the Kriging interpolation technique based on the LiDAR intensity data produce high separability between the land cover materials of interest; trees, grass, asphalt and roofs, (Song et al 2002).

When the point data are resampled into a grid space, the 3D data are converted into 2D grid, causing losses in the details, (Bao et al., 2008). Therefore, this paper proposed a classification approach where the classification algorithm is applied on the LiDAR data in the original point format. Afterward, the classified point data are converted into raster format to produce a land cover map. This approach is expected to produce more accurate classification results. That is because the classification is conducted based on the original values of the point data ( $x, y, z$ , and  $I$ ).

This paper aims at comparing the classification results using two different approaches. Firstly the common approach of converting the point data into raster format then classify the raster data using one of the common supervised classification algorithms. The Kriging interpolation technique is used for converting the intensity and elevation point data into intensity image and DSM. The Maximum Likelihood classifier (MLC) is used for the land cover classification. The second approach is

\*Corresponding Author: Nagwa El-Ashmawy, [nagwa.elashmawy@ryerson.ca](mailto:nagwa.elashmawy@ryerson.ca)

conducting the supervised classification algorithm (MLC) on the point data (range and intensity) then the classified points are converted into raster data with filling the gaps between the points.

The paper is divided into five sections. It starts with the introduction in Section 1, and then a description of the methodology is comprised in the second section. Section 3 covers the study area and the data used in this research. Section 4 includes the results of the experimental work and further analysis. The paper is concluded by a summary of the work.

## 2. METHODOLOGY

The methodology used to achieve the research goal consists of two approaches for land cover classification of LiDAR data. The first approach is classifying the data in raster format, and the second is the classification of the point data.

### 2.1 Classification of Raster Data

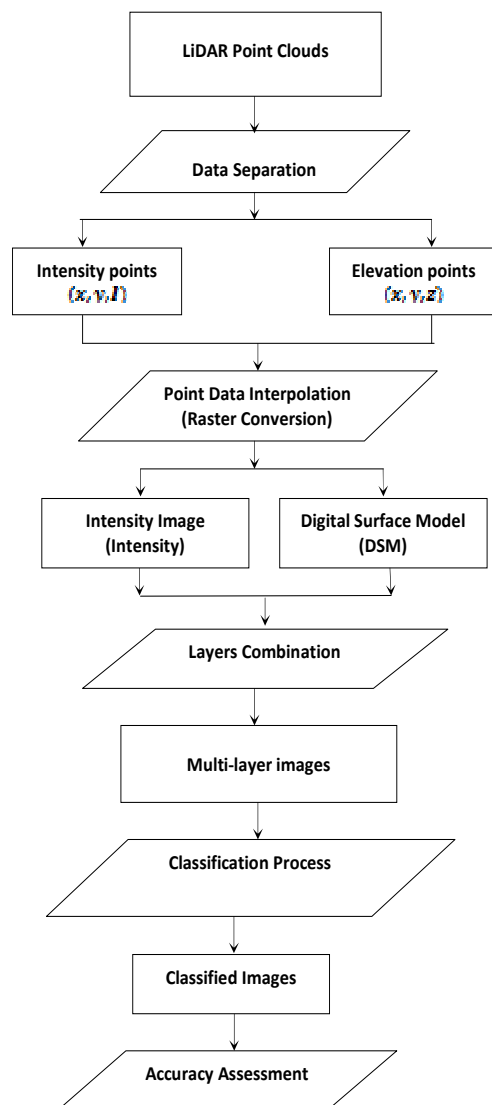


Figure 1. Flow Chart of the LiDAR Data Classification in Raster Format

The LiDAR data are provided in ASCII format as x, y, z, and I values of the point clouds. The flow chart of the first approach

is illustrated in Figure 1, where the first step is separating the intensity data from the elevation data into x, y, I and x, y, z point data files. Then each file is converted into raster format by using an interpolation technique producing an intensity image and a Digital Surface Model (DSM) files. Then the raster intensity image and the DSM are classified using one of the common supervised classification algorithms to the distinguished land cover classes.

The Kriging interpolation technique was used to convert the point data into raster format, the gaps between the LiDAR point clouds were covered by interpolating the point data surrounding these gaps. The advantage of Kriging interpolation technique over other techniques is that it depends on the geo-statistical analysis of the data. The Maximum Likelihood Classifier was used to classify the raster data into five distinguished land cover classes. The evaluation of the classification results were conducted using the confusion matrix approach.

### 2.2 Classification of Point Cloud Data

The proposed approach is classifying the LiDAR point data into the distinguished classes while they are in the point cloud format to avoid losing the information when the 3D point data are converted into 2D grid. The procedure of this approach is illustrated in Figure 2. The procedure consists of five steps. It starts with Defining the grid space that covers the whole area. This grid space is comparable to the raster file of the first approach. Afterwards, the original point data are classified based on the considered attributes in the second step. The third step is resampling the classified points into the predefined grid space. In this step the grid points that are coincide with the resampled points will have the same classes as the resampled points, and the other points will remain unclassified. In Step four, the unclassified grid points will assigned to the distinguished classes using a filling gaps technique that is defined for this work. The last step is the accuracy assessment where the classification results are evaluated.

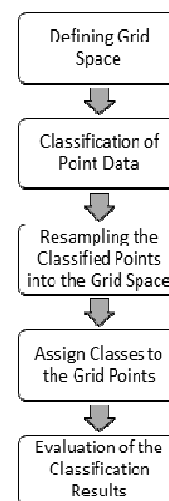


Figure 2. LiDAR Data Classification in Point Cloud Format

The grid space was defined to have a space between the grid points equal to the pixel size of the raster file in the first approach. For the classification of the original point data, the common Maximum Likelihood Classifier (supervised classification technique) was used to classify the point data into

the distinguished classes based on the intensity and elevation attribute values. To resample the classified point data into the grid space, the majority class of the classified points within an area covered the space between the grid points was selected to be assigned to the resampled points. To assign classes to the unclassified grid points that were not coincided with the resampled points, each unclassified grid point would assigned to the majority class of the surrounding eight neighbours of this point. This step was conducted in an iterative process. This process started by the unclassified points that were adjacent to the classified ones. This process was continued until no gaps (unclassified points were remaining). The final step was the evaluation of the classification results where the confusion matrix approach were used based on same reference points used with the first approach for more reliable comparison.

### 3. STUDY AREA AND DATA SETS

#### 3.1 Study Area

Two study areas were selected to test the classification approaches. The study areas were urban district areas at Burnaby, British Columbia, Canada (122°59'W, 49°15'N)., Figures (3-a and 3-b). These areas contained various land cover types.

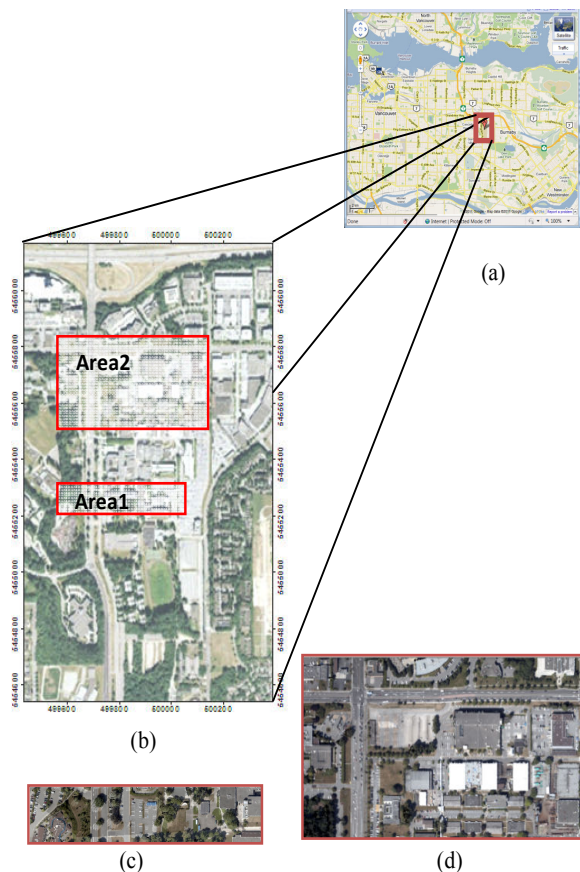


Figure 3. The Location of the Study areas as appeared in Google Map (a), and the Ortho-photo of the study area (b), Study area 1 (c), and Study area 2 (d)

A small area of around 360 m width x 85 m height was clipped to apply the two classification approaches (Figure 3-c). This study area had complex land cover types comparing to its size, where different types of roof surfaces and different ground elevations were found within the small area. A larger area that contained same land cover types was selected to verify the

outcomes of the first study area. The second area was around 600 m width x 350 m height, (See Figure 3-d).

#### 3.2 Data Sets

The LiDAR data were acquired during a mission that conducted on July 17, 2009 at local time 14:55 by Leica ALS50 sensor to cover the British Columbia Institute of Technology (BIT). McElhanney Consulting Services Ltd, BC, Canada provided the authors with the real LiDAR data. The LiDAR data were acquired in six different strips. The flying height of the LiDAR mission was around 540 m. The ALS50 sensor was operating in 1.064  $\mu\text{m}$  wavelength, 0.33 mrad beam divergence and 83 kHz pulse repetition frequency. The point density of the acquired LiDAR data was around 4-5 point/m<sup>2</sup>. The two study areas were clipped out of two different strips for resulting more reliable conclusions.

The provided data were geometrically corrected and calibrated as part of the SII P-IV # 72 GEOIDE project. The accuracy of the x, y, and z coordinates of the LiDAR point clouds were within 19 cm. As another task of the project, the data were radiometrically corrected to overcome the problem of inhomogeneity of the acquired data. Digital aerial photos were captured at the same time of the LiDAR mission. The aerial photos were ortho rectified, and used to collect ground truth for assessing the classification results.

### 4. RESULTS AND DISCUSSION

After examining the study area and the provided data sets, it was noticed that five different land cover types were distinguished. The distinguished classes were: Trees, Buildings, areas covered by grass, uncultivated areas covered by bare soil, and areas covered by asphalt (roads and parking lots). Additionally, both study areas contained same land cover types. The first study area (Area1) contained around 160,000 points which was much less than the second area (Area 2) which contained more than million points. Nevertheless, the Area 1 had large gaps between the LiDAR points especially within the areas covered by asphalt, but Area 2 contained smaller gaps. These varieties of the number of points and the size of the gaps will make the conclusions more reliable.

500 reference points were randomly selected to assess the classification results of Area 1 and 850 points for Area 2. The Confusion matrix approach was used to assess the classification results. The ground truth for the reference points was collected from the ortho-rectified aerial photos that were captured during the LiDAR mission.

#### 4.1 Classification Results of Raster Format

The results of the interpolation of the intensity and the elevation values are illustrated in Figures 4 and 5. The classification results of the Maximum Likelihood classifier for the raster data based on both the intensity and the elevation values are illustrated in Figure 6. The overall accuracy of the classification results were 44% for Area 1 and 50% for Area 2.



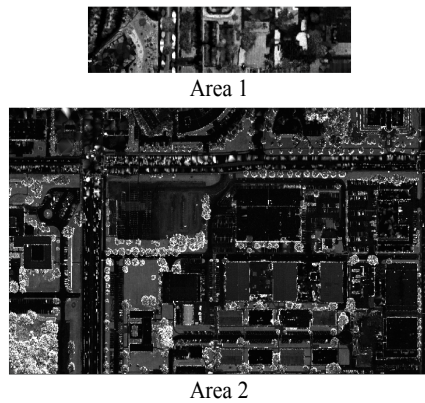


Figure 4. Intensity Surface of the Two Study Areas

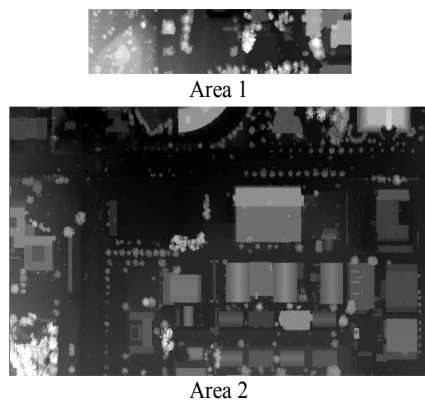


Figure 5. DSM of the Two Study Areas

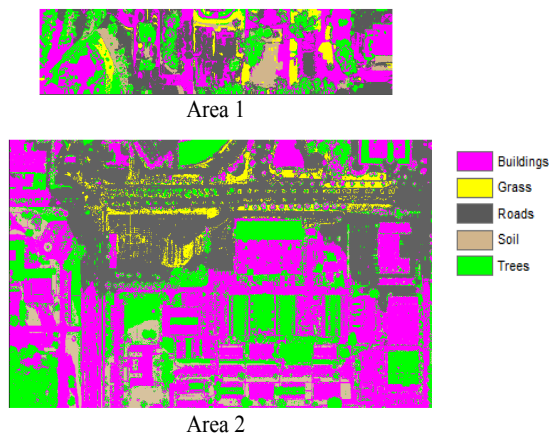


Figure 6. Classification Results of the Raster Format Data

#### 4.2 Classification Results of Point Format

The classification results of the LiDAR point clouds using the Maximum Likelihood classifier based on the intensity and the elevation values are illustrated in Figure 7 and 8, where Figure 8 shows the classification results of the whole grid points after filling the gaps between the original point data. The same reference points were used to assess the classification results of the second approach. The overall accuracy of the classification results were 53% for Area 1 and 65% for Area 2.

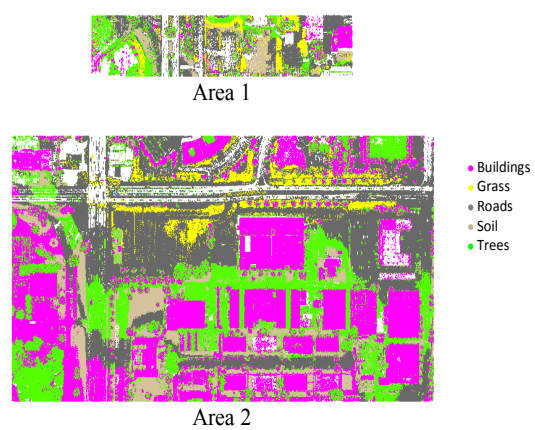


Figure 7. Classification Results of the Point Format Data

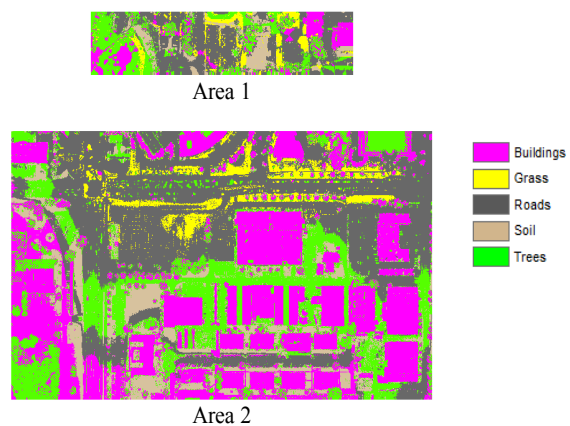


Figure 8. Classification Results of the Point Format Data after Filling the Gaps

The accuracy obtained using the second classification approach (Classification of Point format data) is higher than the one obtained by applying the classification of the raster format approach. However, it is expected that changing the method of filling the gaps may improve the final classification results. Further researches are underway to improve the accuracy of the classification.

#### 5. CONCLUSIONS

LiDAR intensity data have been examined to be used for Land-Cover classification. Two classification approaches were investigated and their results were compared. The common classification of raster data was compared to a proposed approach, where the point data were classified based on the considered attributes, and then the classified points were converted into raster file and/or grid points. The grid points were assigned to classes based on their locations. The grid points that were coincide with the original points after resampling to the grid space had classes same as the majority classes of the original points within small areas equivalent to the pixel size. The remaining grid points were assigned to the classes based on the majority of their eight adjacent neighbours. From the results obtained, it can be concluded that the proposed approach can be used successfully to improve the classification results of LiDAR point data the proposed accuracy improves the classification results by around 10%. Further research work is underway to further improve the classification accuracy.

## ACKNOWLEDGEMENTS

This research work was supported by the Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the GEOIDE Canadian Network of Excellence, Strategic Investment Initiative (SII) project SII P-IV # 72. The authors would like to thank McElhanney Consulting Services Ltd, BC, Canada for providing the real LiDAR and image dataset.

## REFERENCES

- Bao, Y., Li, G., Cao, C., Li, X., Zhang, H., He, Q., Bai, L., and Chang, C., 2008. Classification of LiDAR Point Cloud and Generation of DTM from LiDAR Height and Intensity Data in Forested Area. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Beijing, China, Vol. XXXVII, Part B3b.
- Hug, Ch., and Wehr, A., 1997. Detecting and Identifying Topographic Objects in Imaging Laser Altimeter Data. In: *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 32, Part 3–4W2, pp. 19–26.
- Rentsch, M., and Krzystek, P., 2009. Precise Quality Control of LiDAR Strips. *Proceedings ASPRS 2009 Annual Conference*, Baltimore, MD, 11 p, 9-13 March 2009.
- Song, J., Han, S., Yu, K., and Kim, Y., 2002. Assessing the Possibility of Land-Cover Classification Using LiDAR Intensity Data. *ISPRS Commission III, Symposium*, Graz, Austria. P.B-259ff.
- Wehr, A. and Lohr, U., 1999. Airborne Laser Scanning—An Introduction and Overview. *ISPRS Journal of Photogrammetry and Remote Sensing* 54:68–82.