CLASSIFICATION OF LiDAR POINT CLOUDS BASED ON THE 2D CELLS

Kottage Saumya Denuwan Jayasundara (1), Sanka Perea (2), Nalani Perera (2)

[1] G I S Branch, Survey Department, 150, Bernard Soysa Mawatha, Colombo 00500 , Sri Lanka.
[2] Sabaragamuwa University of Sri Lanka, P.O. Box 02, Belihuloya, 70140, Sri Lanka.
Email: saumyadenuwan.sab@gmail.com; sanka.perera@gmail.com; nalani.ha@gmail.com

KEY WORDS: LIDAR

ABSTRACT:

LiDAR (Light Detection And Ranging) is a modern active remote sensing technology which can collect accurate three dimensional data efficiently and effectively. Accurate digital terrain models are required for many applications such as topographic mapping, disaster management, constriction projects. But LiDAR 3D point cloud contains broth terrain and non-terrain points. So classifying points in to ground and above ground is the primary and important step of LiDAR data processing. This process is generally called as LiDAR filtering and this is very important because accuracy of final product is depending on accuracy of filtering. Generally classifying points into terrain and non-terrain is the most time consuming and complex step. If ground undulation is high, classification of point cloud is very challenging. Most common methods that are used to classify point cloud are semi-automated or manual which is time consuming and costly. To overcome that 2D cell based automated method to classify point in to terrain and non-terrain is proposed by this research. By this proposed method points are filtered only by using geometry of the points without using number of returns or intensity of return.

For the minimum boundary box area, 2D gridded cells are created. Then minimum Z value for each cell is found which represents the cell (generally ground is represent by lowers Z values) or use “NAN” (Not-a-Number) string for empty cells.

To distinguish ground and non-ground, continuity of surface is used. Generally ground is continuous surface. So if there is discontinuity, height difference between adjacent cells must greater than tolerance value. Height difference (dZ) is calculated along rows, columns and diagonals. If height difference is greater than tolerance value, large Z value cell is considered as non-terrain cell and replace cell value as “NAN”. Then interpolation is carried out to remove empty cells if the height difference between beside “NAN” value cells is smaller than tolerance. For define object boundaries first iteration is done. Then cell values which are inside object boundaries are classified. After cell classification is done, point which are inside the cells classified in to terrain and non-terrain classes by using cell classification outcome. If height difference between minimum Z value of the cell and other point of cell is less than tolerance point classification is same as cell classification and if not point classification is opposite as cell classification.

Five samples of Hermanni study area was used to assess the accuracy of the developed algorithm and accuracy was over 93%.

INTRODUCTION

Background of the study

Remote sensing is a method of obtaining information about an object or phenomenon, without making direct contact with it. In present, numerous active and passive remote sensing techniques are used to acquire data for various fields. Airborne laser scanning (ALS) also called as LiDAR (Light Detection and Ranging), is one of the active remote sensing techniques used for acquiring accurate three dimensional (3D) point data since the mid-1990s (Amolins, 2016). ALS sensor platform is consisting with laser scanner for distance calculation (measure the time delay between an emitted and received pulse), GPS/GNSS for provide position of the aircraft and IMU for angular measurements (pitch, roll and yaw) of aircraft.

At present, LiDAR technology has become one of the major techniques which provide high point density with high elevation accuracy (decimeter level) over large area in a very short time (Wehr et al, 2010). Airborne laser scanner data is a powerful tool for generating DTM, DSM than other data sources such as photogrammetry and Synthetic Aperture Radar (SAR) since this technology made high dense, more accurate, and reliable 3D point clouds possible, which provides direct 3D information about the features. Also, LiDAR provides high accurate point clouds with multiple returns and intensity values which represent surface characteristics. LiDAR point data is used in various fields such as geography, land surveying, mapping, forestry, most earth science disciplines, disaster management, project planning, telecommunication etc. At present, advanced LiDAR sensors are capable of capturing not only 3D (X, Y and Z) coordinates but also number of returns, return number, intensity of returns, pulse density (Tomljenovic et al, 2016).

Justification

A LiDAR point cloud includes terrain points and non-terrain points. For generating digital terrain models (DTMs) or extracting different object points such as vegetation, building, road from massive unstructured and dense point clouds with different local densities, a point cloud must be separated or classified in to terrain and non-terrain points first. Therefore, it is the most important and essential step of any LiDAR application. This process is generally called as filtering.

Problem statement

If manual method is used to classify points in to terrain and non-terrain classes it would be a costly, time consuming and difficult process because LiDAR provides high density point clouds (Yastikli and Cetin, 2016). So an automated method for classifying LiDAR points in to terrain and non-terrain is required. Different methods to classify terrain points from ALS point clouds have been proposed. Progressive triangular irregular network (TIN)-based method, surface-based methods, slope-based methods and morphological methods, are some of interesting methods that can use to classify a LiDAR point cloud into terrain and non-terrain classes (Xiangyun and Yuan, 2016). Although these methods work well, they still have several limitations, especially when classifying area with complex ground profile (Blaszczak-Bak and Sobieraj, 2013). Therefore, to develop an automated process to classify LiDAR point clouds in to terrain and non-terrain classes is still a challenging topic, for large areas with various terrain types (Alsubaie et al, 2014). Thus, this research work aims to propose a novel automatic method to classify terrain points from ALS point clouds more accurately and efficiently. For this purpose, geometric characteristics of the terrain such as smoothness, continuity and lowest elevations can be used. Further, this research propose a 2D cell based processing chain instead of using very complex data structures which take a long processing time.

Study Area

Study area is located on center of Finland's southern capital, Helsinki city. It is urban area with multi storage buildings, vegetation and different topography.

MATERIALS AND METHODS

LiDAR Surveying

LiDAR is high accurate three dimension (3D) data collecting method which can used for many fields and applications such as Update digital elevation models (NED), Glacial Monitoring, Detecting Faults and measuring uplift, Forest Inventory, Shoreline and Beach Volume Changes, Bathymetric Surveying (SHOALS), Landslide Risk Analysis, Habitat Mapping, Subsidence Issues, Telecom Planning, Urban Development and ect. Because of that LiDAR is one of most popular data collection method in present. In LiDAR technique optical signal is used and two way travels time of laser pulse is measure to calculate distance between receiver and surface. While aircraft is moving forward ground is swept by LiDAR pulses (10-150 kHz pulse rate) and the precise time of reflected pulse is recorded. Slant range to surface can calculate by using speed of light and travel time as below.

Slant range = (speed of light * two way travels time of laser pulse) / 2

Position and orientation of the sensor is known and the XYZ coordinate of the reflective surface can be calculated.

When LiDAR pulse is emitted from scanner, pulse can reflect back from number of objects in between scanner and surface because pulse not a single beam but a footprint. Because of that for a single pulse there can be number of returns. Generally first return is upper vegetation and last return is terrain or solid object. In modern LiDAR systems not only return number but also intensity value of the return is recorded which can use to identify surface characteristic.

Generally LiDAR survey provides very dense point data with lots of information such as X,Y,Z coordinates, number of returns, return number, intensity of return, scan angle, scan direction , edge of flight line, gps time and etc.

LiDAR Point Cloud Classification

LiDAR point cloud classification or filtering has been popular research topic for over 30 years and in the beginning ground filtering was carried out by image-matching and object-recognition algorithms which were basic image-processing methods (Weidner and Förstner, 1995). With the development of LiDAR techniques and spatial resolution of LiDAR data it has become popular and reliable data source for large-scale applications.

There are many filtering algorithms developed in present and according to their principle they can be divided in to main four methods as slope-based, morphology-based, surface-based and segmentation-based filters methods (Sithole and Vosselman, 2004).

Data

Data was collected for "Joint Building Extraction test of the EuroSDR" project and captured by TopEye and topoSys laser scanners (Perera, 2007). Only Hermanni test site data was used from project data for this study which consists of multi storage buildings, vegetation and slope landscape. Data in ASCII format was used for this research and output also an ASCII file. Characteristics test site data given as follows.

- Data acquired end of June 2002.
- Flying altitude 200m.
- TopEye laser scanner.
- 7000Hertz pulse frequency.
- 20 degrees (+ or -) field of view.
- 7-9 points per one square meter on average.
- 130m swath width.
- Two pulses mode.Application Architecture

Software

For develop classification algorithm Matlab (matrix laboratory) 2016 (R2016a(9.0.0.341360)) software was used. It is an integration of high-performance programming language, multi-paradigm arithmetical computing environment and visualization software which is user friendly. Matlab allows matrix manipulation which was very useful for this study. General uses of Matlab software are;

- Math and computation
- Image processing
- Algorithm development
- Modeling, simulation, and prototyping
- Data analysis, exploration, and visualization
- Scientific and engineering graphics
- Application development, including Graphical User Interface building

But for LiDAR point cloud viewing process, Fugro Viewer software has used because viewing LiDAR point cloud using Matlab takes extensive time (generally LiDAR data set is a large data set). Fugro Viewer is user friendly software which allows 2d and 3D visualization and supports different data formats.

Methods

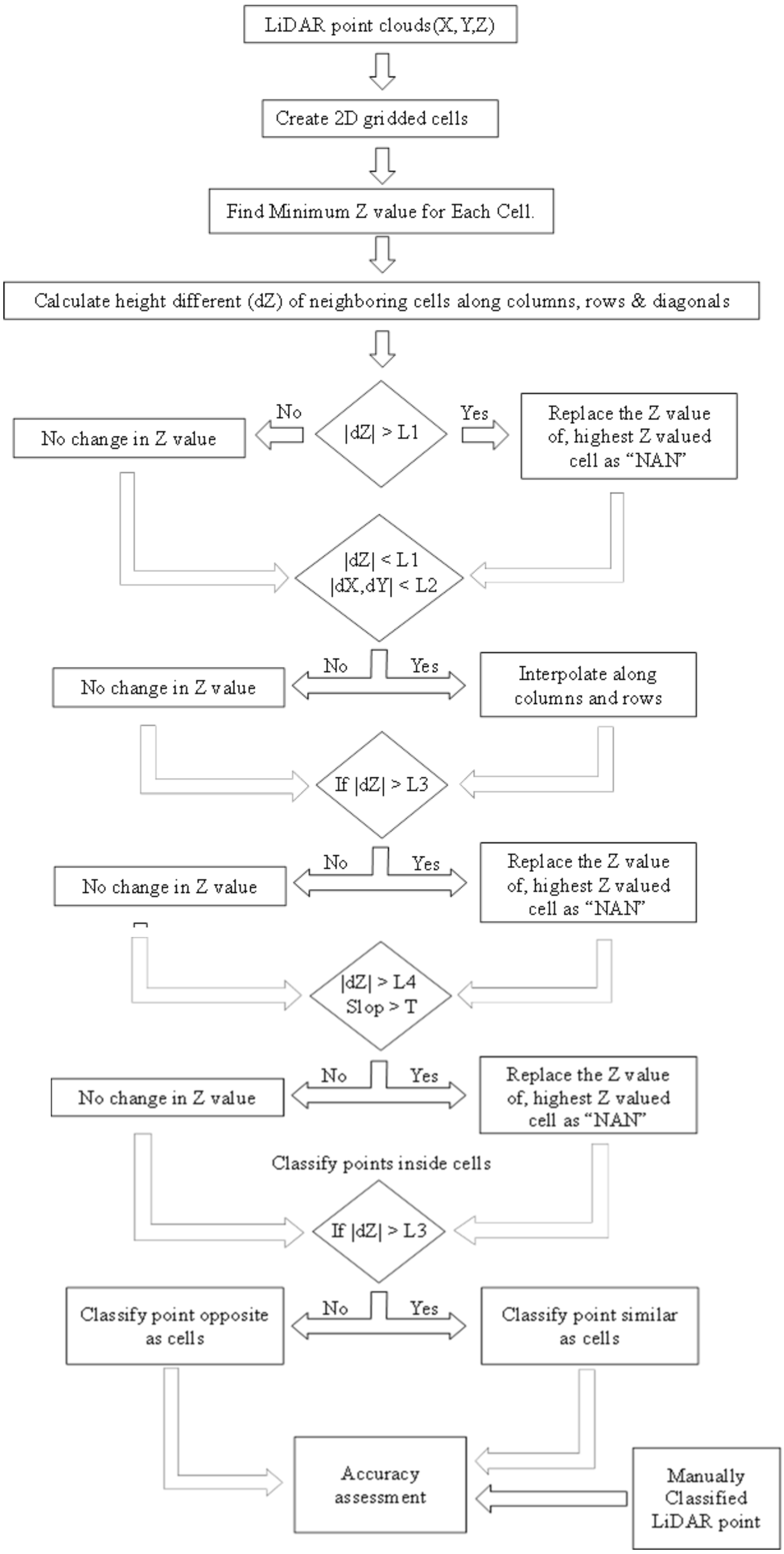


Figure 1: Methodology

Load point data in to matlab

LiDAR point cloud is loaded to matlab and minimum, maximum of X, Y values are found to define minimum boundary box area for loaded LiDAR points. Then 2D gridded cells are created for minimum boundary box area which is defined in above step. Structure array is created to insert points which are inside each cell. Cell size is defined as required for the processes.

Minimum height value matrix

For each cell minimum height (Z) value is found from the points inside the cell and stored in 2D matrix which represents the 2D cell structure. If the cell is empty 'NAN' string is stored in respective location on 2D cell matrix. Created minimum Z value 2D matrix is used for cell classification and it is same as raster image (DEM). 'NAN' (Not a Number) string is used because it will not effect on arithmetic calculation done in next steps.

Object boundary creation

First height different between adjacent cells are calculated along X, Y and diagonal direction respectively if broth of them are not equal to 'NAN'. Then if calculated height difference (dz) is greater than tolerance, Z value of highest cell is replaced by 'NAN', else Z value will be same.

Remove unnecessary points and give value to empty cells by interpolation

In this step height difference (dz) is calculated beside 'NAN' cells along X and Y directions. If 'dz' and distance is lower than tolerance 'NAN' is replaced by interpolated height (Z) value. This step is used to create continuous surface and remove unnecessary cells.

Object boundary creation form interpolated surface

Interpolated continuous surface is used to create object boundaries. Same algorithm as Object boundary creation is applied along X, Y and diagonal direction to determine object boundaries and mark boundary cell as 'NAN'.

Classifying cells using object boundaries

After classifying object boundary cells, object cells which are inside identified object boundary must be recognized. For that height different and slope is used. Classification is carried out along X and Y direction respectively. Firstly 'NAN' value cell (object boundary cell) is identified. Height difference (dz) and slope (dxs or dys), beside 'NAN' value cells are calculated along X and Y directions. If the conditions ($|dz|$ > tolerance1 and slope > tolerance2) are satisfied, cell values are replaced as 'NAN' from first 'NAN' value cell to direction of highest Z value cell until next 'NAN' value is found. This process is done along both X and Y direction.

Classifying points using cell classification

After cell classification is done, points which are inside each cells are classified. Point classification is done by calculating height deference (dz) between Z values of each point with minimum Z value of cell. If height difference ($|dz|$) is lower than tolerance, point classification is marked same as cell classification. If not points are classified as opposite of cell classification.

Analysis of Data

LiDAR point cloud is classified following above steps and five sample site of study area which was manually classified, is used to check the accuracy of the algorithm. Accuracy assessment was done by same algorithm. By this algorithm point cloud is labeled in to two groups, terrain and non-terrain classes. So there are two types of errors.

- Type 1 error: terrain points are classified as non-terrain points (Perera, 2007).
- Type 2 error: non-terrain points are classified as terrain points (Perera, 2007).

RESULTS AND DISCUSSION

Results

Five sample sites (which are manually classified) of study area are used to assess the accuracy of the algorithm. These sample areas are consisting with different terrain types, objects and trees. Results of the algorithm are as follows

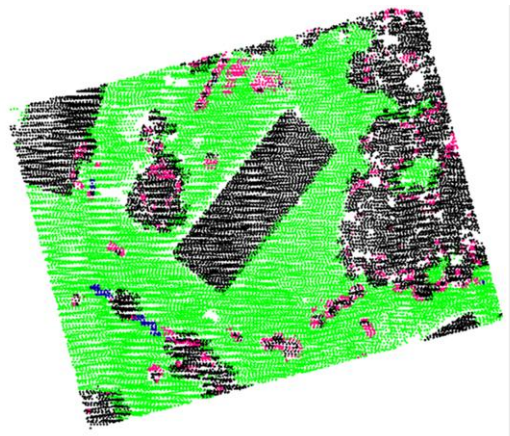


Figure 2: Classification of sample site 1

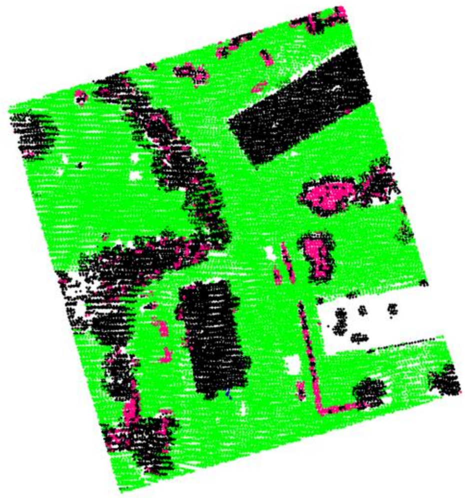


Figure 3: Classification of sample site 2

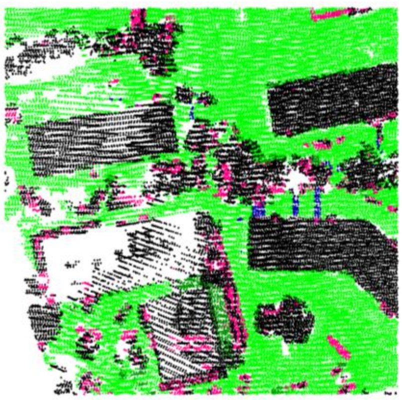


Figure 4: Classification of sample site 3

Figure 5: Classification of sample site 4

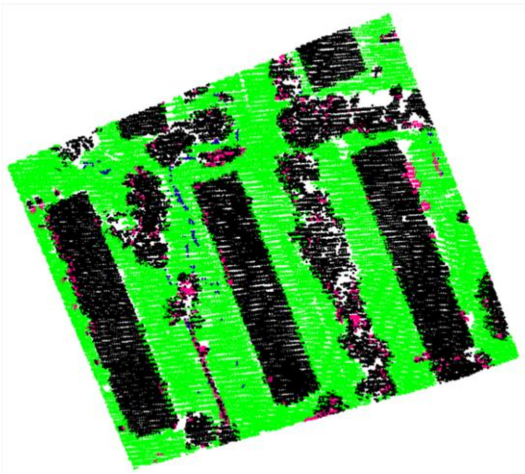
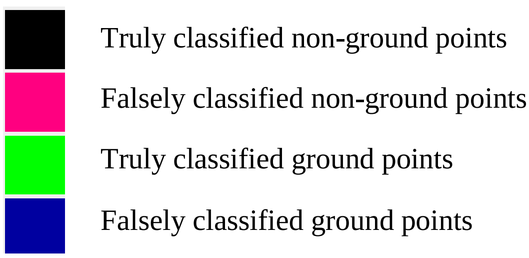


Figure 6: Classification of sample site 5



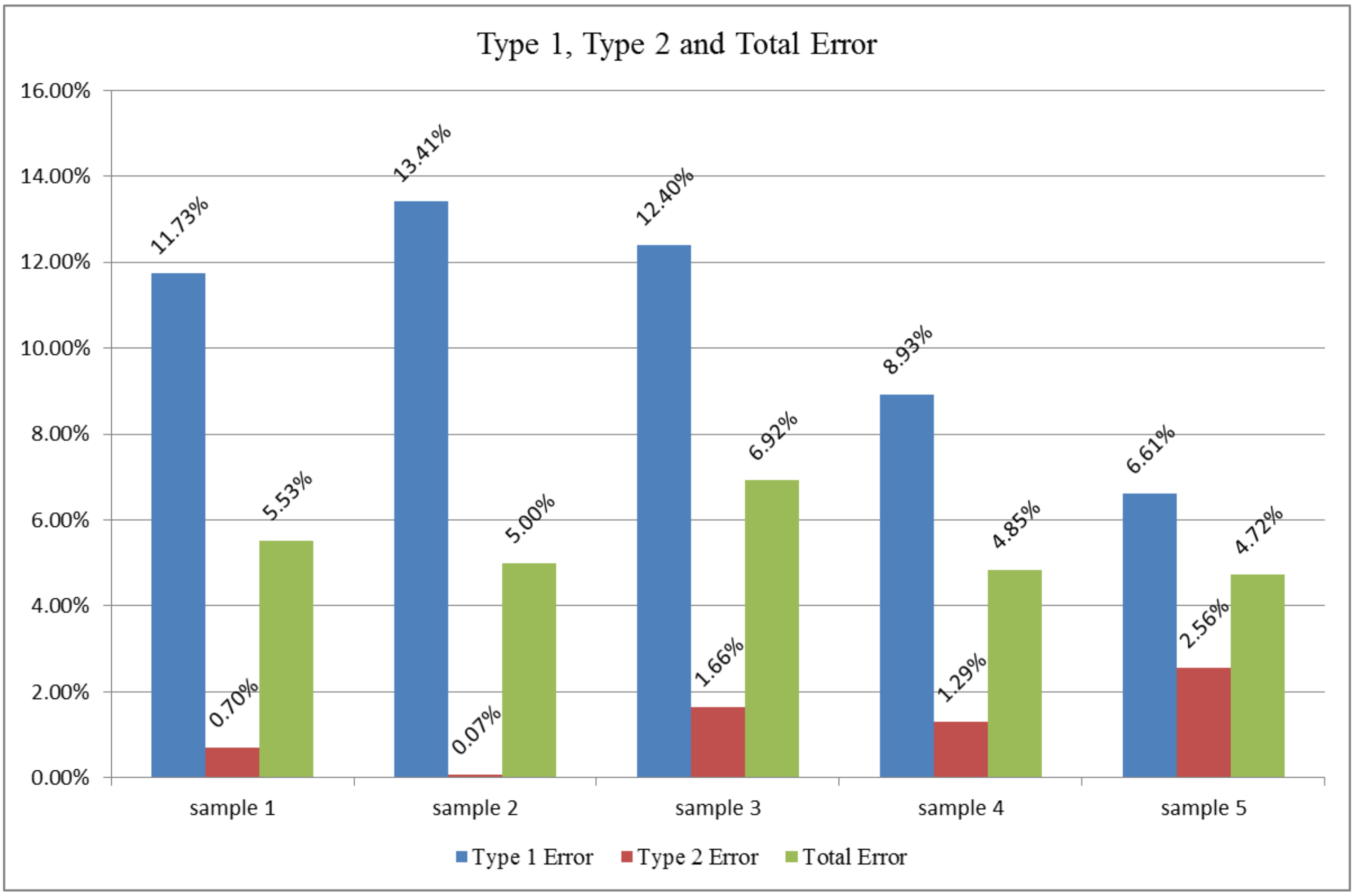


Chart 1: Type 1, type 2 and total error for same parameters

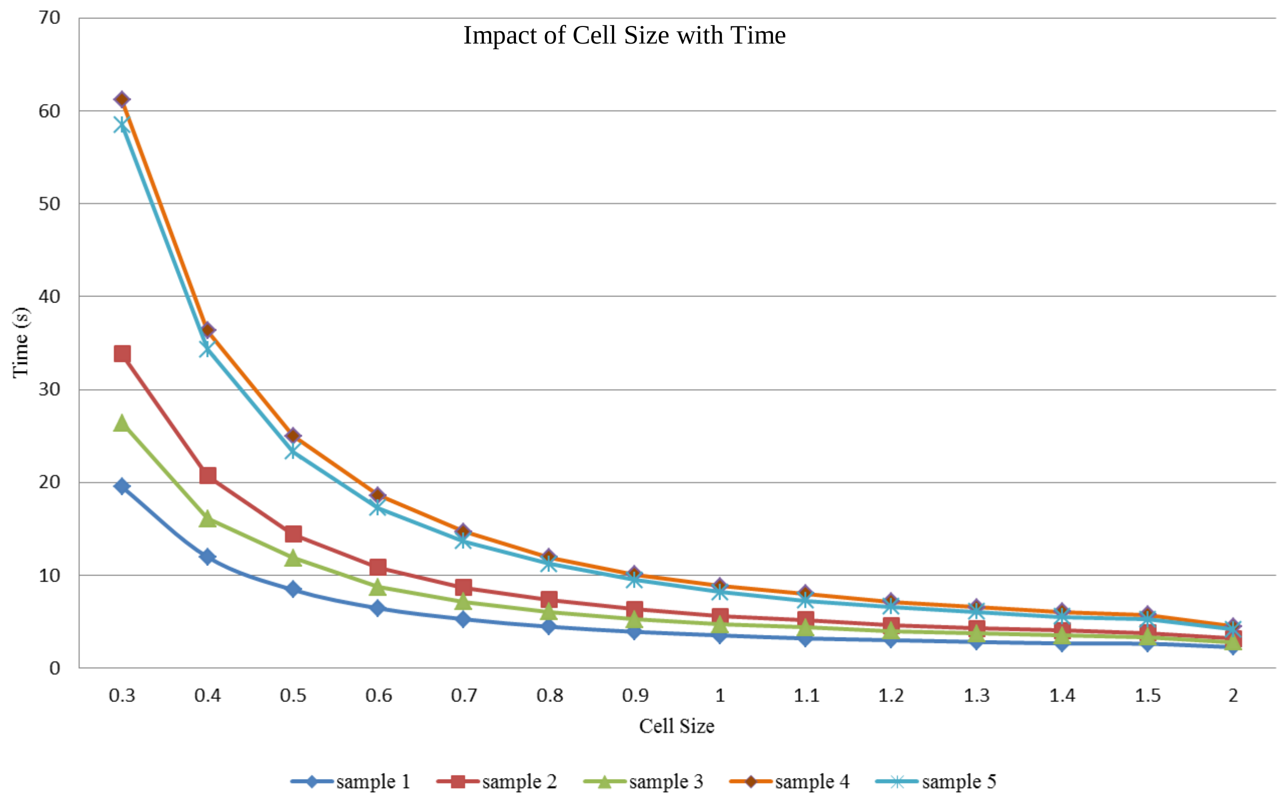


Chart 2: Impact of Cell Size with Time

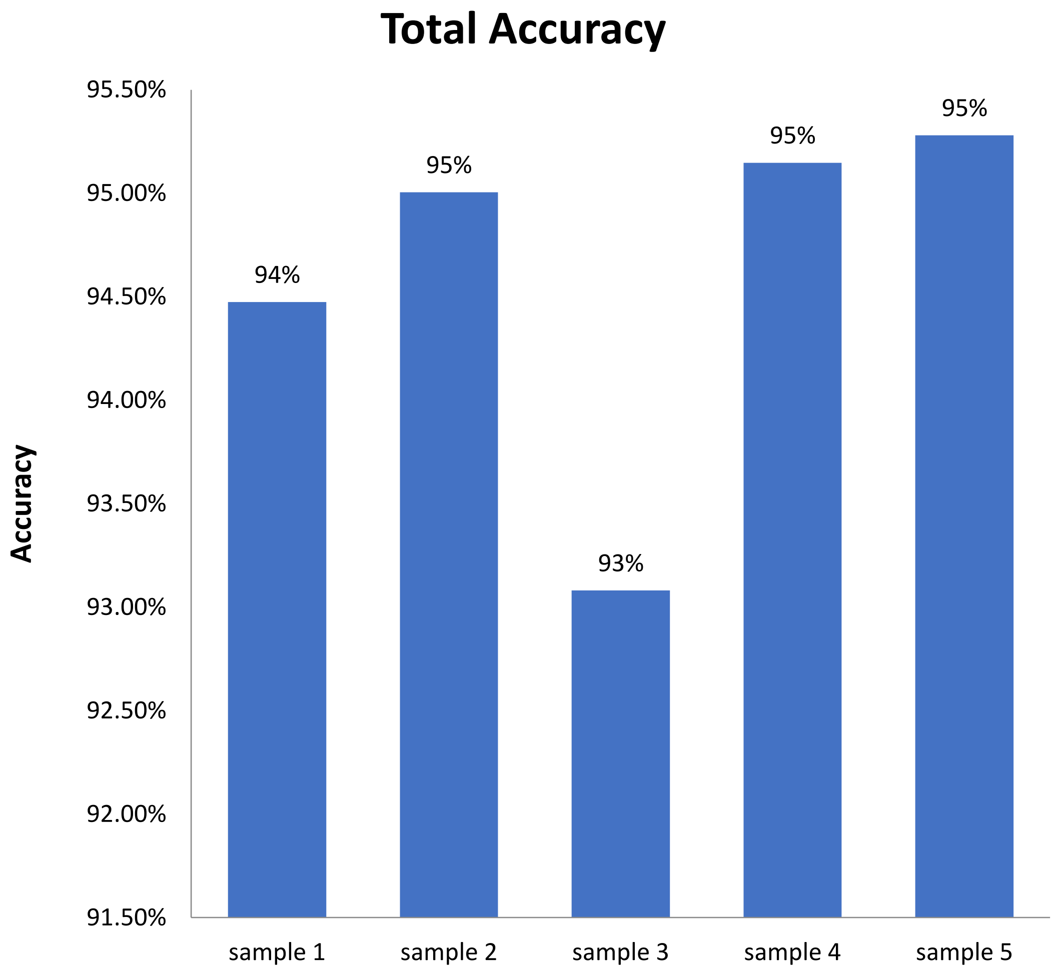


Chart 3: Total accuracy for same parameters

Discussion

In this section, outcome of the study and complications which occurred during development of the algorithm is discussed.

When finding minimum height value for each cell, there could be more than one minimum height value points in same cell. If so, when calling for X and Y coordinates of minimum height value point, algorithm could not deliver an answer. To overcome that average value of X and Y coordinates are used.

With the cell size, number of point contains in a cell is changing. When using smaller cell size, numbers of empty cells (no point inside cell) are rising. To overcome that interpolation process was carried out to assign value to empty cells. But interpolation process could change real surface characteristics. So before applying interpolation, object boundary was extracted to protect surface characteristics.

Over 90% accuracy is achieved by the Developed algorithm. But type 1 error (terrain points are classified as non-terrain points) in all samples output, are higher than type 2 error. Type 2 error (non-terrain points are classified as terrain points) is not a significant value when compared with type 1 error. Most of the points which misclassified as non-terrain points are representing vegetation points. Total time for classify study area (consist of 410498 points) is around 200 seconds with 1 meter cell size which is highly efficient.

CONCLUSION AND RECOMMENDATIONS

Conclusion

LiDAR point cloud filtering is popular research topic and there are many algorithms developed to carryout point classification. But most of them fails because lack of ability to classify different terrain types (steep terrain areas, complex terrain areas) and low efficiency due to complex computation. Generally LiDAR data consists with large number of point data which is very difficult to handle. Manual methods to filter LiDAR point cloud are accurate but time consuming. When developing algorithm to classify LiDAR point data in to terrain and non-terrain classes, developer must focus on above mansion complications to get better results.

Developed automated algorithm can classify LiDAR point data in to terrain and non-terrain, effectively, efficiently and accurately. Time for the classification process is inversely proposal to cell size and with large cell size accuracy is low. To get accurate output, parameters must be optimized and for that trial and error method is used. Finest tolerance slope angle which used to identify objects is around 30 degree for study area. But slope angle is depending on the steepness of the terrain.

Recommendations

Instead of using only X, Y & Z data, can use intensity value, multiple returns and X, Y and Z based method to improve the accuracy of the classification. But time for filtering process will be greater than simple X, Y and Z based methods.

In this study all conditions are applied along X, Y and diagonal directions separately. When classifying cells in single direction at a time, algorithm will ignore the influence of other neighboring cells for classification. So eight neighbor (3 by 3 kernels) comparison method will improve the accuracy of classification than individual comparison along X, Y and diagonal directions.

In this method cell classification is very important, because points which are inside particular cells are classified by using cell classification.

REFERENCES

Amolins, K. (2016). Design of a semi-automated lidar point classification framework. Department of Geodesy and Geomatics Engineering, University of New Brunswick.

Axelsson, P. (2000). DEM generation from laser scanner data using adaptive TIN models. International Archives of Photogrammetry and Remote Sensing.

Arachchige, N. & Maas, H.(2012). Automatic building facade detection in mobile laser scanner point clouds. DGPF, Germany

Belton, D. & Bae, K.(2010). Automating post-processing of terrestrial laser scanning point clouds for road feature surveys. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII, Part 5.Commission V Symposium, Newcastle upon Tyne, UK.

Blaszczak-Bak, W. & Sobieraj, A. (2013). Impact of optimization of ALS point cloud on classification. Technical Sciences, Abbrev: Techn. Sc., No 16(2), Y 2013.

Chen, Q., Wang, H., Zhan, H., Sun, M. & Liu, X. (2016). A Point Cloud Filtering Approach to Generating DTMs for Steep Mountainous Areas and Adjacent Residential Areas. International Journal of Remote Sensing.

Favorskaya, M.N & Jain,L.C.(2017). Handbook on Advances in Remote Sensing and Geographic Information Systems, Intelligent Systems Reference Library 122, Springer International Publishing AG 2017

Lin, X., Zhang, J. (2014). Segmentation-based filtering of airborne LiDAR point clouds by progressive densification of terrain segments. International Journal of Remote Sensing.

Maguya, A.S., Junttila, V., Kauranne, T. (2014). Algorithm for extracting digital terrain models under forest canopy from airborne LiDAR data. International Journal of Remote Sensing.

Meng, X., Wang, L., Silván-Cárdenas, J.L., Currit, N.(2009). A multi-directional ground filtering algorithm for airborne LIDAR. International Society for Photogrammetry and Remote Sensing Journal of Photogrammetry and Remote Sensing.

Maas, H., Bienert, A., Scheller, S. & Keane, E.(2008). Automatic forest inventory parameter determination from terrestrial laser scanner data. International Journal of Remote Sensing.Vol. 29, No. 5.

Mongus, D., Zalik, B. (2014). Computationally efficient method for the generation of a digital terrain model from airborne LiDAR data using connected operators. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

Niemeyer, J., Rottensteiner, F., Soergel, U. (2012).Conditional random fields for lidar point cloud classification in complex urban areas. International Archives of Photogrammetry and Remote Sensing.

Shan, J., Aparajithan, S.(2005). Urban DEM generation from raw lidar data. Photogrammetric Engineering & Remote Sensing.

Sithole, G.,Vosselman, G. (2004). Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. International Society for Photogrammetry and Remote Sensing Journal of Photogrammetry and Remote Sensing.

Susaki, J. (2012). Adaptive slope filtering of airborne LiDAR data in urban areas for digital terrain model (DTM) generation. International Journal of Remote Sensing.

Tomljenovic, I., Tiede, D. & Blaschke, T. (2016). A building extraction approach for airborne laser scanner datautilizing the object based image analysis paradigm. International Journal of Applied Earth Observation and Geoinformation 52 (2016), 137–148.

Vosselman, G. (2000). Slope based filtering of laser altimetry data. International Archives of Photogrammetry and Remote Sensing.

Weidner, U., Förstner, W.(1995) Towards automatic building extraction from high-resolution digital elevation models. International Society for Photogrammetry and Remote Sensing Journal of Photogrammetry and Remote Sensing. 1995, 50, 38–49.

Wehr, A., Duzelovic, H. & Punz, C. (2010). Fusion of als point cloud data with high precision surveying data. In: Wagner W., Székely, B. (eds.): ISPRS TC VII Symposium – 100 Years ISPRS, Vienna, Austria, IAPRS, Vol. XXXVIII(Part 7B), 639-643.

Wu, B., Yu, B., Huang, C., Wu, Q. & Wu, J. (2016). Automated extraction of ground surface along urban roads from mobile laser scanning point clouds. International Journal of Remote Sensing Letters.

Yastikli, N. & Cetin, Z. (2016). Classification of Lidar data with point based classification methods. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B3, 2016 XXIII ISPRS Congress, Prague, Czech Republic. 441-445. doi:10.5194/isprsarchives-XLI-B3-441-2016.

Yan, M., Blaschke, T., Liu, Y., Wu, L. (2012). An object-based analysis filtering algorithm for airborne laser scanning. International Journal of Remote Sensing.