Segmentation and Visualization of Airborne Sparse Forest LiDAR Data using Region Growing Approach

Swati R. Maurya
P.G. Department of Computer Science,
S.N.D.T. Women's University
Mumbai-400049, India

Ganesh M. Magar, PhD
P.G. Department of Computer Science,
S.N.D.T. Women's University
Mumbai-400049, India

ABSTRACT

Point clouds data acquired from airborne LiDAR point cloud data sources have great ability to provide vital structural information about geospatial objects. Identification, segmentation and visualization of airborne LiDAR point cloud data is interesting but considerably challenging problem. Processing LiDAR point cloud data can reveal several interesting properties of the geospatial objects. However the complexity of segmentation and visualization varies across different structures of geospatial objects viz. building structures, tree crowns, canopies etc. It is relatively complex compared to manmade objects. This paper explores and establishes the usefulness of variant of region growing algorithm for segmentation and 3-D visualization of sparse forest obtained from airborne LiDAR data stored as point clouds. The role of point normal and curvature in progressive region growing process is highlighted. The algorithm implicitly segments the ground from non-ground objects on the forest surface and enables improved visualization in 3-D space. The proposed approach achieves accuracy as high as 95.67% in segmentation of forest region and exhibits significant visual performance when compared to standard visualization platform.

General Terms

Data Mining, Geographical Information Analysis. Pattern Recognition

Keywords

LiDAR, 3-D Visualization, Region Growing, Segmentation, Forest Visualization.

1. INTRODUCTION

We Point cloud data has emerged as a vital form of available data formats for analytical processing and representation of geospatial objects and phenomenon in the geospatial environment. Every physical object structure on the surface of the earth can be described with three or more dimension viz. x, y, z, intensity, normal and color information etc. The geospatial visualization of earth objects and features can be improved using such information. To visually understand any objects and its positioning in the 3D space, the surface points play the vital role. These surface points can be stored as point cloud data [1] and used for analysis, reconstruction and visualization [2]. Collectively visualizing each surface point cloud data in the volumetric space provides better perceived object structure and hence potentially provides better perceived visual experience.

The point cloud data can be acquired from multiple sources including range sensors, kinetic devices, aerial or airborne scanning systems or mobile acquisition methods. However, the point cloud data acquired from these sources result into large files that contain raw location and surface specific properties of each points in 3-D space. Segmenting [3], [4] and relating points to the object surface is a challenging task as they lack any form of linkage between them. Specific algorithm for segmentation [6] are required to associate specific points with specific objects and also segregate them from neighboring ones. Although point cloud data for natural as well as manmade objects are stored in same format but segmentation of 3-D point data [7] and thereby rendering them is difficult for natural sites such as terrain and forests.

Several point cloud formats have been proposed and used for LiDAR data representation and storage and is continually being improved. While earlier versions of the data formats were more focused about three dimensional space coordinates, later versions of these formats added capabilities to include specific surface properties such as point normal, intensity, color etc. Popular point cloud data formats includes- polygon file format(ply), format native to the CAD software (stl), geometry defined storage format (obj), the ISO standard for XML-based format (x3d), formats popularly libraries such as Point Cloud Library(pcl) and - public file format for the interchange of 3-dimensional point cloud data (las). Of these point cloud formats .las and .pcd are among those which support ASCII and binary formats. Any point cloud processing algorithm that is used either for analytics or visualization would require large number of computations and thereby require substantial time and memory space.

It is of immediate interest for visual analysts to partition/segment forest land parcels between ground and non-ground regions for several reasons including planning, tracking, ecological and geographical studies etc. Although several commercial products for point cloud processing is available but their performance on natural objects such as forest trees is limited. Advanced algorithms such as nearest neighbor and K-d tree can be used for clustering and classification of region of interest. Point cloud data handling libraries such as PCL, and PDAL can be helpful for such task. Advanced clustering techniques such as K-d-tree based region growing approach can be more efficient in identifying clusters of forest trees and barren land parcels. It can also be useful for identifying outlier data among the point clouds. This paper evaluates the application of region growing algorithm for segmentation, classification and visualization of ground and non-ground objects viz. trees from forest LiDAR data.

2. RELATED WORK

This decade has seen growing interest of researchers and scientist working with data for 3-D geospatial analysis as opposed to traditional 2-D planar geospatial analysis in past. With advancements in the storage and computing capabilities in recent years, several large and ambitious projects like National Lidar Dataset (United States) have been taken to capture and store and thereupon facilitate data to analyze the

three dimensional world around us. Open Topography, USGS Earth Explorer, NOAA Digital Coast and National Ecological Observatory Network – NEON are popular repositories of attempts to gather and store LiDAR data.

Segmentation of point clouds is essential in processing and visualization of 3D views both for man-made and nature made object structures. The usefulness of point clouds in visualization and understanding of geospatial structures have been highlighted in [6] which described the methods and algorithms relating to point cloud filtering Airborne laser scanning systems. A. Nurunnabi and others in [7] highlighted the difficulties with segmenting surface data in laser scanning point clouds and emphasized on the usefulness of PCA based method in dealing with noisy data. Detection of building structures and similar manmade geospatial objects have been described in [8],[9],[10],[11]. Anandakumar M. Ramiya et al. in their work [8] used point cloud library (PCL) to compute the histogram based methodology to segment building from other structures and achieved accuracy of 82% which was comparable to the performance of commercially available software like Terasolid. Shaohui Sun et al. in [9] demonstrated the graph cut based method to create building from the airborne LiDAR point clouds. G. Forlani in [10] also demonstrated the usefulness of segmentation by region growing and region topology.

Region Growing based techniques has been very successful in identifying object structures and hence contribute significantly to the visualization. It has been studied by Runzhen Huang and Kwan-Liu Ma in [12] which demonstrated improvement over trial-and-error process used in point cloud data exploration and visualization. Several attempts have also been made to study and analyze natural fields and establishments including forestry and hydrology using point clouds. The work by Li et al. in [13] explored use of applied segmentation algorithm to detect trees and obtained accuracies between 86% and 94%. Visualization of forest is also experimented by En Mi Lima et al. in [14] using VRML.

Point cloud segmentation using spectral clustering has been describe by Teng Ma in [15]. It uses graph-cut theory and k-nearest neighbor techniques. Surface visualization of geospatial structures from point cloud is equally important. Terrestrial Laser Scanning and processing of point cloud data has been described by F. Laefer et al. in [16] to generate strategies for solid models. Triangular mesh model reconstruction based on template has been explained by Liu Bin, et al. [18] in which they explained the importance of curvature detection in forming angular mesh from point clouds. Spatial information contained in the point cloud for each point is vital in segmentation and hence the reconstruction and visualization of object structures. It has been tested by Oytun Akman et al. in [19] using kernel density estimation method. Another study by Wen-xue Xu et al. in [20] showed experimental proof of the usefulness of graph theory and region growing for segmentation terrestrial point clouds.

The study of the research initiatives in past two decades suggests growing need of segmenting and visualization of point clouds for improved understanding and exploration of 3D geospatial world around us. Although point cloud data can be acquired from various sources, fundamentally the complexity is due to large number of points in the point cloud and must be address accordingly. Region growing theory, graph based algorithms and neighborhood function can prove to be useful in reconstruction of arbitrary natural scenes and visualizing them thereafter.

3. EXPERIMETAL TEST DATA

This decade has seen growing interest of researchers and scientist working with data for 3-D geospatial analysis as opposed to traditional 2-D planar geospatial analysis in past. With advancements in the storage and computing capabilities in recent years, several large and ambitious projects like National Lidar Dataset (United States) have been taken to capture and store and thereupon facilitate data to analyze the three dimensional world around us. Open Topography, USGS Earth Explorer, NOAA Digital Coast and National Ecological Observatory Network – NEON are popular repositories of attempts to gather and store LiDAR data.

Fig 1: Aerial view of the Test Site of Californian Forest

4. PREPARING POINT CLOUD DATA FOR VISUALIZATION

Point cloud visualization requires understanding the format specification of the concerned representation. This paper uses PCL PCD file format for storing point cloud data for experimental analysis. Each PCD file is uniquely defined by its header which identifies and declares certain properties about the point cloud data stored in the file. For ease of use and interoperability, the header of a PCD is specified precisely and encoded in order in ASCII with each point on a new line. It has the several entries including version, fields, size, type, count, width, height, viewpoint and data respectively. With usual meaning of the header entries, FIELDS specify the name of the dimensions of the point, and POINTS specifies the total number of points in the given point cloud file. As already said, the point data can be in either ASCII or binary format, denoted by the DATA field.

5. PROCESSING POINT CLOUDS FOR VISUALIZATION

Before point cloud is rendered, the segmentation of ground from non-ground objects is performed. To achieve the objective, a variant of region growing algorithm [12] is used which inherently uses K-d clustering and nearest neighbor classification to associate points from the given point cloud to respective clusters of points which represent ground region and non-ground objects such as tree respectively.

5.1 K-d Tree Processing

K-d trees are very useful for range data and complements nearest neighbor searches for segmenting and classifying objects. With reference to point clouds, k-d trees would deal with three-dimensional data. Each level of a k-d tree procedure would split all its children along a specific dimension, using a hyperplane that is orthogonal to the corresponding axis. It is obvious that at the root of the tree, all children will be split based on the first dimension. Inherently, at each level down in the tree, it divides on the next dimension, returns to the first dimension once all others have been exhausted. The K-d tree approach is shown in figure 2.
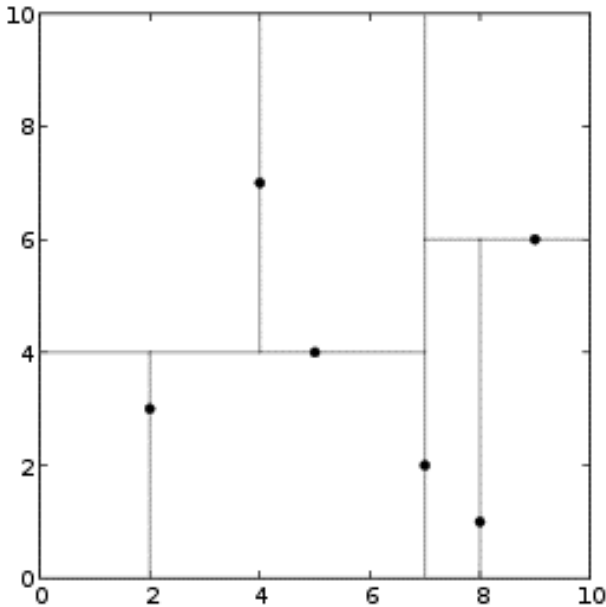

Fig 2: K-d Tree Processing

5.2 Nearest Neighbor Classifier

It is a non-parametric method used for classification of point cloud data. The input comprises of the k closest training samples in the feature space. The output is a class membership. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors. It is sensitive to the local structure of the data and hence influences membership of the points with object and surface structures.

6. MODEL FOR SPARSE FOREST VISUALIZATION

It is implied that sparse forest would encompass trees of various heights and crown size distributed over the forest land. The model proposed in this paper assumes that there is a sparse distribution of trees that are positioned with accountable distance between them. The model is capable of segmenting and clustering subset of point clouds into possible trees. Any two trees are separated from each other depending on the curvature threshold c_{th} and the angle threshold θ_{th} controlled further by the neighborhood measurement. The model assumes input in point cloud format is obtained from airborne sensors. The macro level sparse forest processing can be modeled as shown in figure 3.
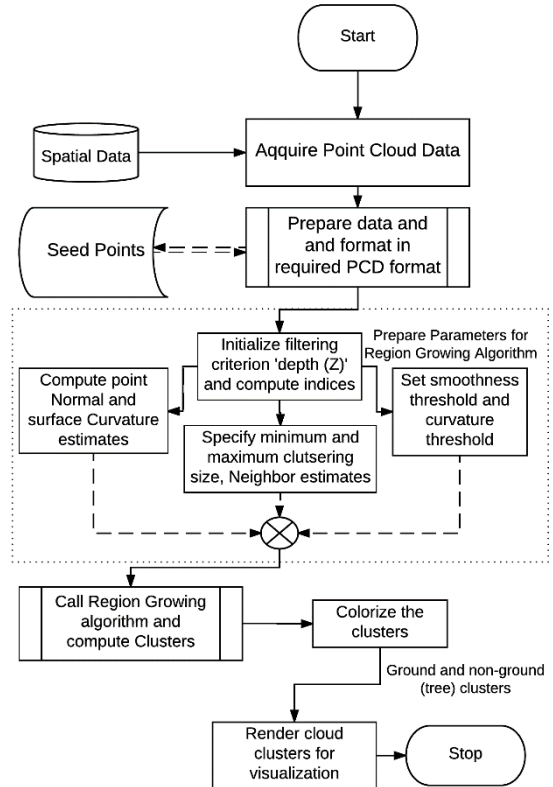

Fig 3: Model for Sparse forest Segmentation and Visualization

The clusters formed by the region growing algorithm correspond to the trees and ground region respectively. This process used search tree based K-d Tree algorithm together with k-nearest neighbor algorithm. These clusters are further colorized to explicitly develop visual sense of segregated trees in the scene. The ground level irrespective of the geospatial location attain same color. The details of the naïve procedure for segmentation and visualization of point cloud data are describe next.

7. REGION GROWING ALGORITHM FOR SEGMENTATION OF POINT CLOUDS

The region growing algorithm used for segmentation of forest LiDAR data used K-d Tree and nearest neighbor techniques. K-d Tree is used as a space-partitioning data structure for organizing point clouds in a k-dimensional space. Nearest neighbor is used for clustering.

Note that the ground point data differ from non-ground point data in terms of-

 a. Normal distribution along three axis

 b. Intensity variations

 c. Depth values (the z- coordinate)

 d. Color continuities

Segmentation of non-ground data points can be used to extract, study and visualize objects on ground.

7.1 Region Growth Theory in Point Cloud Analysis

The region growth theory [12] can be used for progressively determining and separating ground from non-ground 3-D data.

Since point clouds representing natural environmental features such as forest region show intermittent variations, hence a region growth based approach is more suited. The principles, input and output of this algorithm are as follows.

a. Principle: Merge the neighboring return points within the point cloud controlled by smoothness limit.

b. Input: Raw point cloud data storing LiDAR first returns.

c. Output: Set of collections of points. Each collection represents a cluster corresponding to homogeneous geographical feature such as land parcel or and crown of a tree with similar smoothness.

During clustering of points within point clouds, the comparison of the angles between the respective points normal are highly useful. The algorithm sorts the points by their curvature value (region begins its growth from the point that has the minimum curvature value). Point with the minimum curvature is located in the flat area growth from the flattest area which allows to reduce the total number of segments. Until there are unlabeled points in the cloud, algorithm picks up the point with minimum curvature value and starts the growth of the region.

7.2 Applying Region Growth Workflow to Sparse Forest Analysis

The algorithm inputs include the point cloud P, the normal N for the points in space, the function Ω to find the neighbors, the curvature threshold c_{th} and the angle threshold θ_{th}. The region growing algorithm assumes to start with an empty region list R and set of available point list A to contain all points in the given source sample point cloud. That is,

Region list $\{R\} \leftarrow \phi$

Available points list $\{A\} \leftarrow \{1,\ldots,|P|\}$

The generic workflow of the region growing algorithm can be outlined as follows.

Step 1. Pick a point P_{min} with minimum curvature in A. The picked point is added to the set S_c called seeds.

Step 2. For every seed point i, in the node list, the algorithm finds neighbor $\Omega\big(S_c(i)\big)$ points.

a. Every neighbor j is tested against the angle between its normal and normal of the current seed point. If the angle is less than threshold value θ_{th} i.e.

$P_j \in A$ and $cos\left(\frac{1}{|N\{S_c(i),\ N\{S_c(j)\}|}\right) < \theta_{th}$, then current point is added to the current region and removed from A

$R_c \leftarrow R_c \cup P_j$ and $\{A\} \leftarrow \{A\} \backslash P_j$

b. Every neighbor is tested for the curvature value. If the curvature is less than threshold value i.e. If $c(P_j) < c_{th}$ then this point is added to the seeds. i.e.

$S_c \leftarrow S_c \cup P_j$

c. Current seed is removed from the seeds.

Step 3. If the seeds set becomes empty this means that the algorithm has grown the region and the process is repeated from the beginning.

The application of region growing algorithm for segmenting point clouds generates clusters of points adhering belonging to the respective tree structures and are clearly visible. Precisely, the following results for the given input point cloud data is achieved.

a. Identification of trees and their geospatial distribution.

b. Rendering of tree crowns

c. Segregation and visualization of ground and non-ground 3D space visualization of distribution of surface points data

d. Identification of Shadow regions from 3D perspective.

The tree crown are described and rendered in 3D space and assumes no color information as these are clustered and separated in volumetric space by their curvature and angular thresholds. However each cluster representing individual tree can be colored separately for specific analytical results. The rendering of the final clustered point clouds gives 3-D perspective of the trees in the forest.

8. RESULTS VISUALIZATION AND COMPARISION

The experimental results in this paper establishes the fact that region growing algorithm can be used to identify and segment not only planar objects but also for nature made arbitrary structures such as trees. The method is equally suitable for irregular shaped objects and enables us to quantify memberships of 3-D points in space to clusters representing each object structure. It has successfully achieved ground and non-ground classification and generated clusters of individual trees. The visual performance of the proposed approach on sparse forest data is comparable to that of commercial tools. The visualization of the point cloud representing sparse forest before segmentation is shown in figure 4.
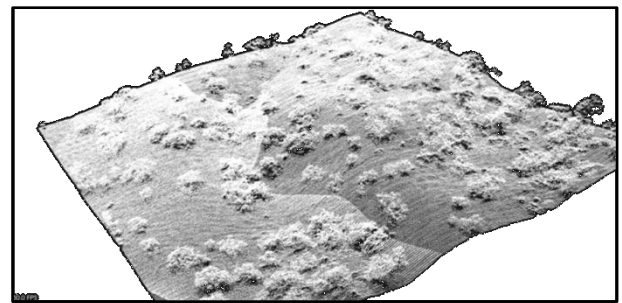

Fig 4: 3-D View for Unclassified and Unsegmented Lidar Data of Forest Woodland Area

Application of the region growing algorithm segments the point cloud into clusters viz. ground and non-ground point clouds. Non-ground point clusters are further clustered into sub clusters that represent individual regions trees in the forest. The result of applying region growing segmentation to the point cloud data is shown in figure 5.
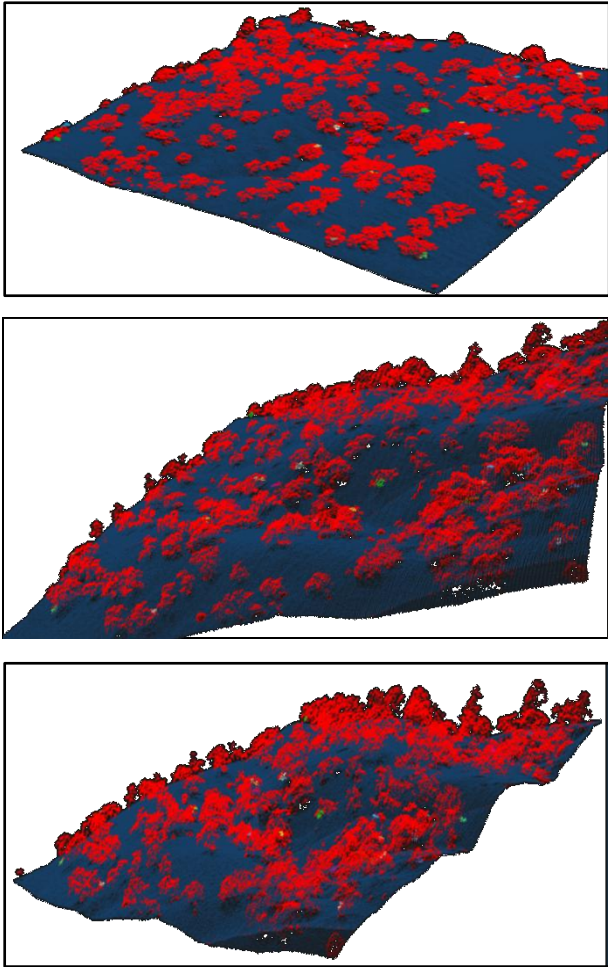
Fig. 5: Different 3-D Views for Segmented Forest Woodland Area

The results obtained with the proposed algorithm for separating ground from non-ground forest region is convincing and gives better visual performance compared to the visualization outputs for the same point cloud data with commercial alternative like ESRI's ArcScene. The visual outputs in ESRI's ArcScene is shown in figure 6.
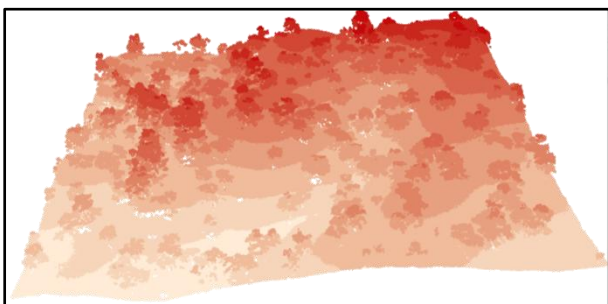

Fig. 6: 3-D Visualization using ArcScene

The proposed model obtained 269 clusters for the given sample point cloud with 567197 points in it. The number of points in each cluster vary according to membership to the respective segment of tree and also with the ground data based on curvature and angular association and neighborhoods. The ground cluster has encompasses 417540 points from the point cloud while the total number of points in non-ground clusters representing collection of trees is of which 125242 represents trees and 24415 denote points corresponding other non-

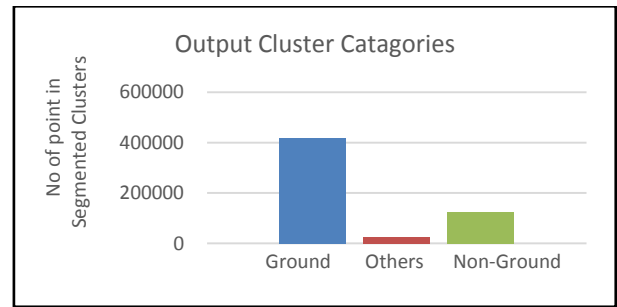ground features. The number of 3-D points in each cluster is shown in figure 7.

Fig 7: Cluster Categories for Point Cloud for test Californian Forest Region

The visual analysis of the region visualized in 3-D (see figure 6) and quantitative analysis (see figures 7, 8 and 9) suggest that given test area is sparsely populated with tree trees having major uncovered land parcel. The region growing model for clustering point cloud data was applied with minimum cluster size set to 50 and neighbor threshold set to 30. The largest crown identified has 288 points in the cluster. The region also has large number of smaller trees or shrubs with number of points as low as 50 enough to identify and visualize them.
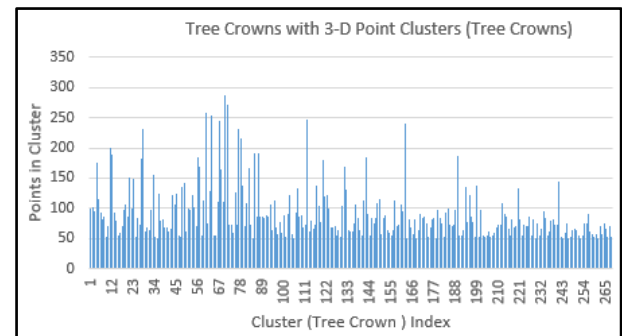

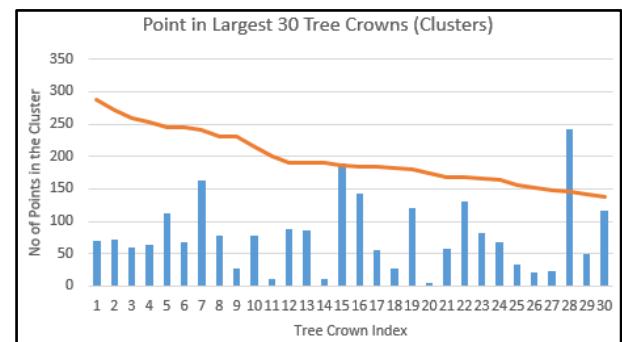**Fig 8: No of Clusters (Tree Crowns) and Points in Each Crown**

Fig 9: Quantitative Distribution of points in Largest 30 Tree Crowns

The analysis of the clusters obtained after applying the model suggested that of the 268 clusters of non-ground points, 99 clusters denoting trees are fairly distinct from each other. The summary of observations of applying the proposed model on the test region of Californian forest region is outlined in table 1.

Table 1: Results of Applying Region Growing Model for Sparse Tree Analysis

Feature Type	Average No of Points in the Cluster	Maximum No of point in the Cluster	Minimum No of points in Cluster
Land Surface Area	417540	417540	417540
Tree Crowns	91.10	288	50

9. CONCLUSION

Application of region growing algorithm on forest point cloud data achieved segmentation and visualization of ground and non-ground (tree) data with accuracy up to 95.67%. The performance of the proposed method is visually verified to be comparable with commercial software available around. The curvature estimates influence significantly the segmentation of geographical features and can be taken as an important parameter for classification and clustering natural objects. Region growing approach can be easily applied on large point clouds. The results also shows that non uniform distribution of points can be easily classified and partitioned which makes it suitable to study several physical phenomenon occurring on the earth surface including rough surfaces, tidal waves, etc. This theory can also be extended in future for simulating particle based visualization. In future the segmented clusters of objects can be processed further generate triangulation/mesh representing surfaces and simulating the environment. It can also be used for extraction of object structures from large point clouds and study surface properties.

10. REFERENCES

[1] Aloysius Wehr, , Uwe Lohr, "Airborne laser scanning—an introduction and overview", ISPRS Journal of Photogrammetry and Remote Sensing, Volume 54, Issues 2–3, July 1999, pp. 68–82

[2] Er Li, Xiaopeng Zhang, and Yanyun Chen. 2014. Sampling and surface reconstruction of large scale point cloud. In Proceedings of the 13th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry (VRCAI '14). ACM, New York, NY, USA, 35-41. DOI=http://dx.doi.org/10.1145/2670473.2670480

[3] A. Nurunnabi, D. Belton and G. West, "Robust Segmentation for Large Volumes of Laser Scanning Three-Dimensional Point Cloud Data," in IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 8, pp. 4790-4805, Aug. 2016. doi: 10.1109/TGRS.2016.2551546

[4] Geetha M and Rakendu R, "An improved method for segmentation of point cloud using Minimum Spanning Tree," Communications and Signal Processing (ICCSP), 2014 International Conference on, Melmaruvathur, 2014, pp. 833-837. doi: 10.1109/ICCSP.2014.6949960

[5] A. Nguyen and B. Le, "3D point cloud segmentation: A survey," 2013 6th IEEE Conference on Robotics, Automation and Mechatronics (RAM), Manila, 2013, pp. 225-230. doi: 10.1109/RAM.2013.6758588

[6] Peter Axelsson, "Processing of laser scanner data—algorithms and applications", ISPRS Journal of Photogrammetry and Remote Sensing, Volume 54, Issues 2–3, July 1999, pp. 138–147

[7] A. Nurunnabi, D. Belton and G. West, "Robust Segmentation in Laser Scanning 3D Point Cloud Data," Digital Image Computing Techniques and Applications (DICTA), 2012 International Conference on, Fremantle, WA, 2012, pp. 1-8. doi: 10.1109/DICTA.2012.6411672

[8] Anandakumar M. Ramiya, Rama Rao Nidamanuri, Ramakrishan Krishnan, "Segmentation based building detection approach `from LiDAR point cloud", The Egyptian Journal of Remote Sensing and Space Sciences, 2016, http://dx.doi.org/10.1016/j.ejrs.2016.04.001, pp.1-7

[9] S. Sun and C. Salvaggio, "Aerial 3D Building Detection and Modeling From Airborne LiDAR Point Clouds," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 6, no. 3, pp. 1440-1449, June 2013. doi: 10.1109/JSTARS.2013.2251457

[10] G. Forlani, C. Nardinocchi, M. Scaioni, P. Zingaretti, "Building reconstruction and visualization from LIDAR data", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV, Part 5/W12, 2007, pp. 151-156

[11] K. Zhang, J. Yan and S. C. Chen, "Automatic Construction of Building Footprints From Airborne LIDAR Data," in IEEE Transactions on Geoscience and Remote Sensing, vol. 44, no. 9, pp. 2523-2533, Sept. 2006. doi: 10.1109/TGRS.2006.874137

[12] Runzhen Huang, Kwan-Liu Ma, "RGVis: Region Growing Based Techniques for Volume Visualization", Proceeding PG '03 Proceedings of the 11th Pacific Conference on Computer Graphics and Applications, 2003, ISBN:0-7695-2028-6, Page 355

[13] Li, Wenkai; Guo, Qinghua; Jakubowski, Marek K.; Kelly, Maggi, "A New Method for Segmenting Individual Trees from the Lidar Point Cloud", Photogrammetric Engineering & Remote Sensing, Number 1 / January 2012, pp. 75-84(10), http://dx.doi.org/10.14358/PERS.78.1.75

[14] En-Mi Lima,Tsuyoshi Honjob, "Three-dimensional visualization forest of landscapes by VRML", ScienceDirect-Landscape and Urban Planning, Volume 63, Issue 3, 30 April 2003, Pages 175–186

[15] Teng Ma, Zhuangzhi Wu, Lu Feng, Pei Luo and Xiang Long, "Point cloud segmentation through spectral clustering," The 2nd International Conference on Information Science and Engineering, Hangzhou, China, 2010, pp. 1-4. doi: 10.1109/ICISE.2010.5690596

[16] F. Laefer, Debra; Truong-Hong, Linh; Fitzgerald, Michael, "Processing of Terrestrial Laser Scanning Point Cloud Data for Computational Modelling of Building Facades", Recent Patents on Computer Science, Volume 4, Number 1, January 2011, pp. 16-29(14)

[17] Johannes Otepka, Sajid Ghuffar, Christoph Waldhauser, Ronald Hochreiter, Norbert Pfeifer, "Georeferenced Point Clouds: A Survey of Features and Point Cloud Management", ISPRS International Journal of Geo-Information, 2013, 2(4), pp. 1038-1065; doi:10.3390/ijgi2041038

[18] LIU Bin, JIANG Kaiyong, "Triangular Mesh Model Reconstruction from Scan Point Clouds Based on Template", Tsinghua Science and Technology, June 2009, 14(S1): pp. 56-61

[19] O. Akman, N. Bayramoglu, A. A. Alatan and P. Jonker, "Utilization of spatial information for point cloud segmentation," 2010 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video, Tampere, 2010, pp. 1-4. doi:10.1109/3DTV.2010.5506339

[20] Wen-xue Xu , Zhi-zhong Kang ; Tao Jiang, "Segmentation approach for terrestrial point clouds based on the integration of graph theory and region growing," 2009 Joint Urban Remote Sensing Event, Shanghai, 2009, pp. 1-8.doi: 10.1109/URS.2009.5137519