

Automatic Extraction of Road Curb and Road Surface from Lidar Point Cloud Data

Maldeniya M R R, Perera G S N

Department of Remote Sensing & GIS, Faculty of Geomatics, Sabaragamuwa University of Sri Lanka.

DOI: <https://dx.doi.org/10.51244/IJRSI.2026.1304000199>

Received: 19 April 2026; Accepted: 24 April 2026; Published: 14 May 2026

ABSTRACT

Road feature extraction is crucial for wide range of geospatial applications such as road management, intelligent transportation and road safety evaluation. Due to the efficient vehicle-based on-road scanning opportunity, Mobile Laser Scanning (MLS) has become the most appropriate data acquisition system for road environments. Most of methods available for road feature extraction are classical approaches that do not mitigate the problems caused by the presence of outliers and occlusions. This study proposed an automated method for extraction and vectorization of road surface and curb, utilizing a grid-based segmentation and classification. The method begins with extracting the ground surface from the point cloud, where an elevation-based thresholding and Cloth Simulation Filtering (CSF) are applied to isolate terrain points. A region growing based segmentation algorithm is applied to identify the road surface and curb structures based on the elevation differences and surface normal orientations. To address main challenges in region growing, such as curb cuts and occlusions from parked vehicles, a local re-segmentation is developed. The local edge detection and gap bridging are applied for efficient region-growing. The extracted curb boundaries are then vectorized using the alpha-shape algorithms, ensuring a structured, GIS compatible representation. The developed algorithm was tested on multiple urban datasets, and reached a classification accuracy of 98% from the constructed confusion matrix. The geometric accuracy exceeded 85% for most curb islands, with some achieving 97.7% precision. The results validated the robustness and scalability of the developed method in urban environments, and provided a computationally efficient solution for automated road and curb extraction.

Keywords: Curb Cuts, Ground Points Filtering, Mobile Laser Scanning (MLS), Region Growing, Road Extraction.

INTRODUCTION

Roads are among the most significant and largest publicly owned national assets in all countries and it is considered as the backbone of the modern transportation infrastructure. Road networks are continuously and rapidly growing to handle increasing traffic loads day-by-day, due to the expansion of urbanization and economic activities. Today, understanding details and getting accurate information about the road surfaces and related components such as road curb, road divider and pavements, is a fundamental requirement of various applications. Roads play an important role in 3D (three dimensional) city analysis, urban planning and corridor mapping as well as autonomous driving, intelligent transportation, monitoring and reconstruction of road networks.

Numerous road extraction strategies have developed since 1990's for find efficient solutions for various challenges, and achieved various degrees of success for different road structures. Traditionally, road features were identified and measured by in-site measurements with manual field surveys. Although the measurements were very high in accuracy, they were very time consuming, labor intensive and impractical in various situations such as inaccessible areas. By the time, aerial photography was extensively adopted for feature extraction, as well as road feature mapping. Camera based methods which mainly aim to detect lane markings and road boundaries, were fast, low cost and operating power, but it was limited by poor illumination, occlusions from tall buildings and trees, bad weather conditions and difficulties in extracting features in complex urban environments. Radar systems have also been developed for road feature extraction and it can work in a high

range of detection and high reliability in bad weather conditions. However, due to the low angular resolution against the high distance resolution, it was difficult to detect road features and get measurements accurately.

In recent years, LiDAR (Light Detection and Ranging) based scanning systems have become very popular due to its efficiency and cost-effectivity. MLS systems provide high density 3D point clouds with millimeter level precision and the road surfaces and features are represented with very high in details. MLS recognizes complex road geometrics precisely, and enables use cases such as high-definition map generation, autonomous driving and road maintenance planning.

Curbstones are essential component of road networks which is important for defining the drivable areas. Various studies have been proposed for road and curb extraction using MLS data. Ibrahim and Lichti (2012) introduced a curb-based street floor extraction method, which includes 3D density-based ground segmentation, Gaussian filtering for curb detection and polygonal based street floor extraction, relying solely on 3D coordinates. Yadav et al. (2018) incorporated intensity information, using square gridding, PCA for planar surface detection and height differences to extract roads in complex urban environments. Some approaches have utilized MLS trajectory information for extracting the road features. Wang et al. (2012) applied hypothesis testing to determine road surfaces and boundaries. Sui et al. (2021) estimated the ground track from raw point clouds and detected boundaries via a pseudo-mileage spacing map. Mi et al (2022) focused on vectorized 3D road boundaries using supervoxel-based curb extraction and clustering algorithms for boundary fitting with high computational efficiency and fine border preservation.

This study aims to provide an automated approach for extracting road surface and curb from MLS data, addressing key challenges in complex urban environments such as noise, occlusions as well as computational inefficiencies. Most of the algorithms which utilizing region-growing based approaches, rely on the sudden elevation change between different road components, and fails where the presence of curb cuts and not well constructed curb. The proposed algorithm addresses these issues in urban environments and enhances the accuracy and efficiency in road feature extraction compared to the existing methods.

METHODOLOGY

Study Area

An MLS dataset, collected in 2008, in Enschede, The Netherlands was used for the analysis of this study. The MLS dataset was acquired with an Optech LYNX system, which has two 360 degrees rotating laser scanners with a 100 kHz measurement rate. The study area is shown in figure 01.



Figure 1: Study Area

METHODS AND METHODOLOGY

Dealing with large volumes of unorganized point data is one of the largest difficulties in 3D point cloud processing, with common applications resulting in billions of points being sampled. Analyzing point-by-point is not practical since it's computationally prohibited. The point cloud contains points in various densities in different locations. In most cases, exhaustive searches and local-calculations are required to analyze every point to evaluate their geometric properties. Therefore, the point cloud was ordered and stored into a 2D grid and each point within a particular cell was identified with the cell number.

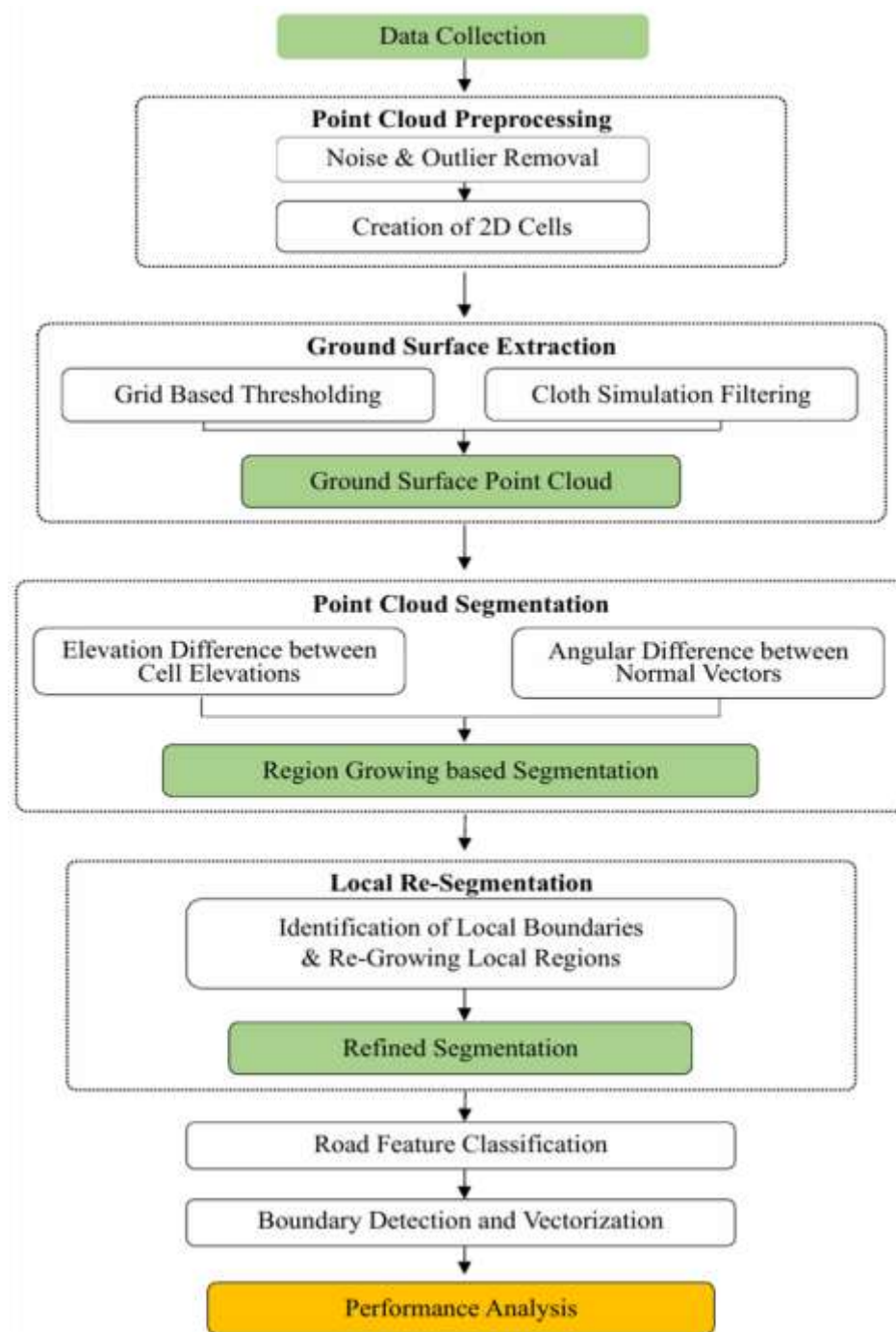


Figure 2: Methodology

Ground Surface Extraction

Typically, the road surfaces and road features are located in the ground level. Therefore, the points which belong to the ground surface were identified first and all non-ground points were removed. The ground surface was extracted by two steps approach. When considering grid cells, the lowest elevated points in each cell are

likely to be part of the ground surface. In each grid cell, the point with minimum Z value was identified and the points that are above a certain height threshold were filtered out and removed from the point cloud. Additionally, a neighborhood analysis was also performed by comparing the elevation of candidate cell with the average elevation of its direct neighbors, to ensure the consistency of the ground surface points. Then the Cloth Simulation Filtering (CSF) algorithm was utilized to further refine the ground surface.

Region Growing based Segmentation

Once the ground surface points are identified, these unstructured spatial data points were segmented into continuous group regions. For this, a region-growing algorithm was designed in this study with respect to the grid cells. In here, the elevation difference and the normal orientation of grid cells were analyzed within the neighbors. The principle of Least-Squares was applied to fit planar surfaces over the points within each cell and the surface normal was calculated using Singular Value Decomposition (SVD).

A neighborhood was defined for each grid cell with a 3x3 window. A region-growing with breadth-first algorithm was performed such that, starting from a particular unvisited cell, each region was grown by adding neighboring cells which meet the pre-defined criteria. Two thresholds were defined as one for the elevation difference and other for the angular difference between surface normal vectors. Each created segment was stored with a unique ID number.

Most of the places, the pavement is at a higher elevation than the road surface, but there are not well constructed curbs or curb cuts appeared on roads in reality. In those areas with little or no abrupt change in elevation between the road and the side pavements, the straightforward height-based segmentation will grow into both areas and fail to segment these two surfaces separately. With these reasons, the road surface and the pavements can be merged into a single segment in region growing. To overcome this issue with the curb cuts, a local re-segmentation procedure was developed that accurately separate the road and the pavement.

To identify the internal edges of the road regions, the cells were analyzed locally to find sharp elevation changes inside the road regions, and the cells that have a signification height difference from its neighbors were identified. Once these raw edge cells were identified, the nearby cells were grown to obtain separate boundary segments inside the dataset. Two corners of each edge segment were identified and stored. After identifying all edge segments and their corners, the azimuth of each edge corner should have to be calculated for estimating the direction of the corners. The azimuth was calculated by stepping back from a corner cell into each previous cell (up to 5 cells). Finally, the average azimuth of each edge corner was calculated to get a stable direction of corners.

After obtaining all the corner directions, for each corner, every other corner was checked whether its azimuth was differed by nearly 180 degrees (Since the two directions are opposite to each other) and it was between the specified distance threshold. When they meet the criteria, corner cells were bridged by a Bresenham line, marking all the intermediate cells in that line into the continuous boundary. Then the re-segmentation was done by treating the identified boundary cells as 'no-cross' barriers during the BFS segmentation. These cells broke the connectivity of the BFS and split into separate regions.

Road Feature Classification and Vectorization

Then the point cloud was classified into different road feature classes, the road surface, curb islands in the middle of the road and the side pavements. Identifying the road surface is a straight forward process because, in most cases in MLS LiDAR point clouds, the road is the largest continuous and smooth surface, therefore the road surface will be the largest continuous segment in the middle of the point cloud. After classifying the road surface, all the neighboring regions to the road were identified and separated from other segments as the road features and they were classified as curb islands which are surrounded by the road surface and pavements in which one side is adjacent to the road surface. Then the boundary cells of each segment were identified by analyzing each grid cell's neighborhood and determining the cells whether it is on the perimeter of a particular segment.

The vector transformation is important for transforming the identified raster boundary cells into a continuous, vector representations that defines the extent of each road features. Once the boundary cells were identified, their

centroids were computed and polygons were generated by using the alpha-shape algorithm. The alpha-shape is a generalized method of the convex hull that creates a tighter, and more natural boundary around a set of points. It is a flexible boundary representation that can create concave or convex polygons depending on the α value. A lower α value results in tighter and more detailed boundary around the points, while a α higher value results in smooth and more convex boundary. After the polygons were generated, they were stored in a geo-data frame. It helps in managing, analyzing and exporting these polygons as GIS-Compatible files.

Performance Analysis

The performance of the developed algorithm was evaluated for understanding its ability to correctly classify the road features. The accuracy of the extraction and classification was analyzed in two ways, by confusion matrix and by geometric accuracy. The point cloud was manually digitized using CloudCompare, to use as a reference dataset for evaluating the accuracy.

To analyze the classification accuracy, a three-class confusion matrix was constructed based on the classification from the analysis and a manually classified point cloud. In this classification, it mainly considered three classes for assessing the classification, road, curbs and non-classified. The confusion matrix was visualized as a heatmap and additionally standard metrics such as overall accuracy, precision, recall and F1 score were derived from the matrix.

To evaluate the geometric accuracy, each area of the generated shapefiles of curb islands was compared with the area of the manually digitized curb boundaries from the raw point cloud. By this comparison, a quantitative measure of how far the extraction process represents the real-world structures can be obtained. The curb islands were manually digitized and the area was computed by ArcGIS Software.

RESULTS AND DISCUSSION

Two subsets of the point cloud were used for the analysis and for testing the algorithm. Initial subset contained nearly 2.2 million points and contained essential road features like road surface, side pavement and curb islands ensuring diversity, shown in figure 3.

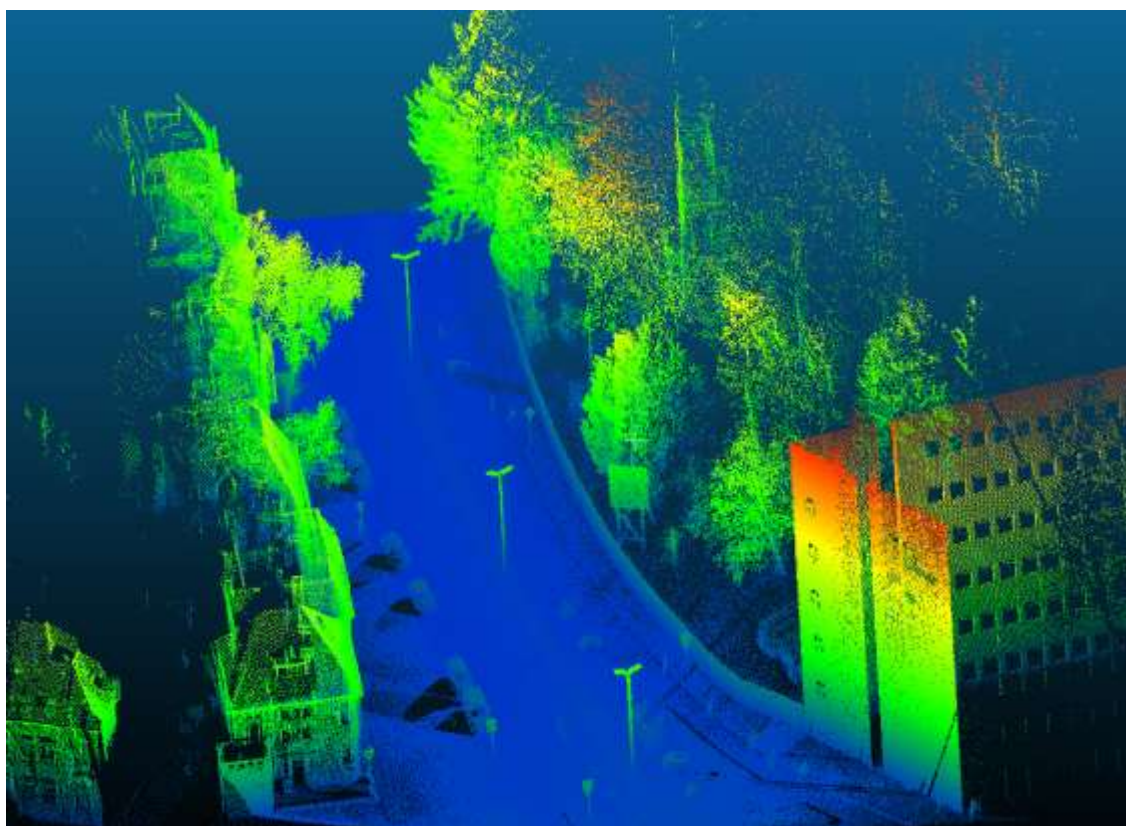


Figure 3: Point Cloud used for the Analysis

Data Storage

By dividing the point cloud into a 2D grid network, each grid cell was able to be examined independently, enabling localized analysis such as calculating statistical metrics like mean or elevations of cells and also the geometric properties of points within a cell can be identified. From this, not only individual cells, but the nearest neighborhood could also be compared and incorporated to analyze if common feature properties are existed across the cells. The cell size was selected as 20 centimeters for an optimal balance between preserving spatial details and maintaining computational efficiency. This resolution is better to capture important features such as curb edges while avoiding excessive data fragmentation. If the cell size is smaller, it would increase computational cost and sensitivity to noise due to sparse point distribution within cells and when it's larger, it would lead to over-generalization of the surface, potentially smoothing out critical features like curbs and reducing the accuracy of segmentation.

Ground Surface Extraction

Ground surface extraction significantly reduced the size of the dataset and complexity. It removed all the outliers which didn't relate to the road features by reducing the point count from 2.2 million to 1.6 million. The initial ground surface obtained from grid-based thresholding was further refined by CSF algorithm. Cloth Simulation is a highly effective tool for extracting ground surface by simulating a cloth draped over the inverted point cloud. The reason for not using the CSF algorithm alone for ground surface filtering was, with the presence of noise and elevated structures like vegetation and structures, certain limitations can be arisen when it is applied to the raw point cloud. These can affect to the initial cloth simulation by causing to misfit the actual terrain, especially in areas where the ground points are sparse and obscured by dense elevated objects. The combination of those two methods was effectively helped in reducing noise and improving the accuracy of ground point detection. The extracted ground surface was shown in figure 4.

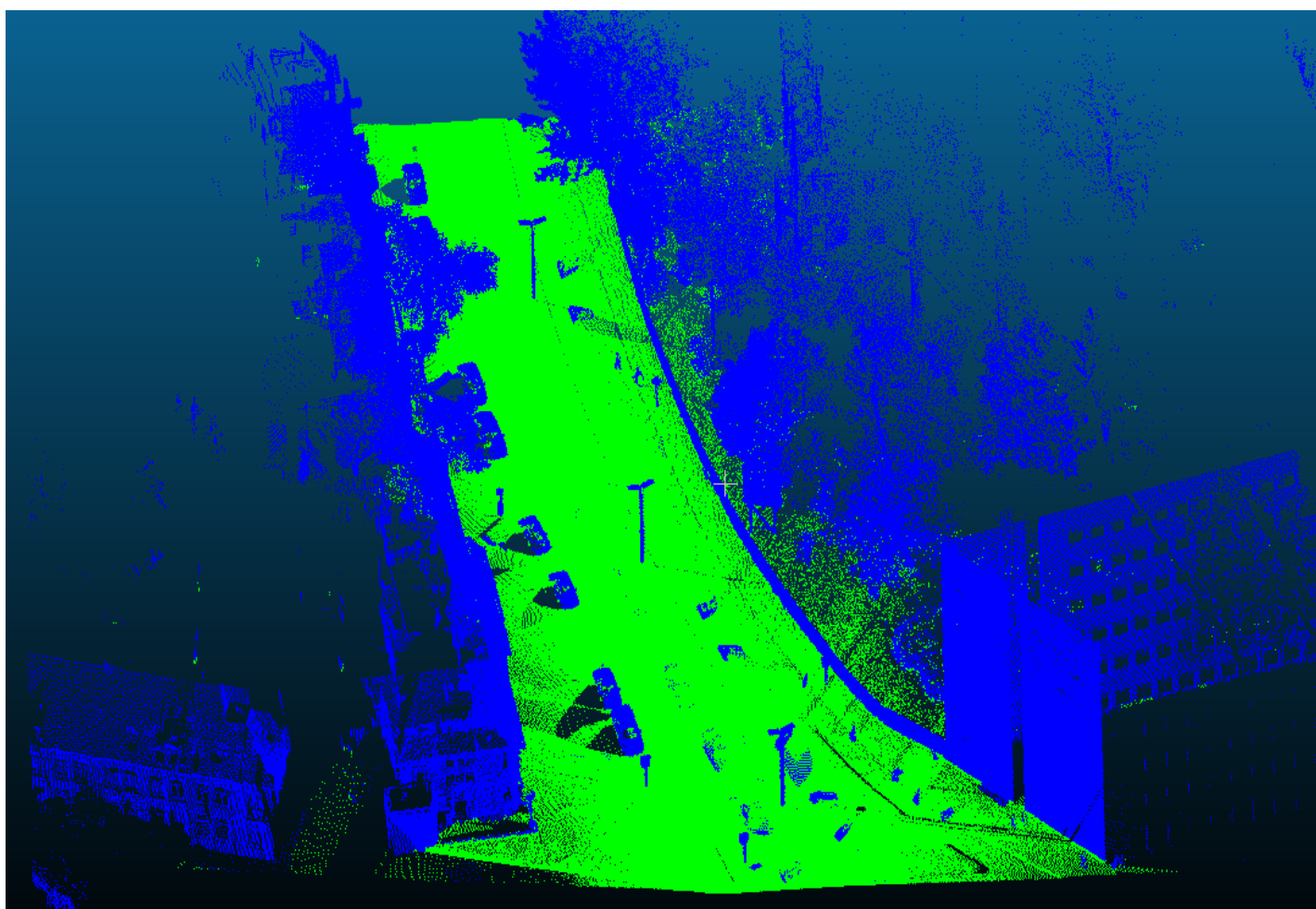


Figure 4: Extracted Ground Surface

The process of extracting ground points relied on several key parameters. These parameters were selected through empirical testing and by considering the geometric characteristics of road environments such as elevation smoothness and curb height variations. Table X summarizes the selected parameter values along with their roles and effects on the extraction process.

Table 1: Ground Filtering Parameter Selection Framework

Parameter	Value	Description	Effect of Lower Values	Effect of Higher Values
Cloth Resolution (CSF)	1.0	Controls the resolution of the simulated cloth surface	Sensitive to noise and small irregularities	May miss terrain details
Max Iterations (CSF)	500	Number of iterations for cloth simulation convergence	Incomplete filtering, unstable ground estimation	Higher computation time with minimal improvement
Classification Threshold (CSF)	1.0	Distance threshold for classifying ground vs non-ground points	May misclassify ground as non-ground	May include non-ground points as ground
Slope Smoothing (CSF)	True	Enables smoothing for sloped terrains	Without it, poor performance in sloped areas	Improves robustness but may slightly smooth sharp features

Point Cloud Segmentation

In segmentation, the regions with similar heights were grouped together, making it effective for flat surfaces like roads and curb. By considering geometric orientation of surfaces, the segmentation was made robust by incorporating normal vectors. In here, the results heavily depend on the thresholds, maximum elevation difference (0.01m) and maximum normal angle difference between neighbors (20°). For efficiently choose those parameters, the average curb height in the area was utilized and histograms were generated for those parameters considering a small segment of road features. Also, iterative testing was done by applying different threshold values and evaluating the segmentation results on the small subsets of the dataset.

The real-world problem which was described in the previous chapter, where the segmentation went inaccurate, due to curb cuts and poor-constructed curbs, was solved in this re-segmentation approach. When looking at the initial dataset, there is a curb cut in a particular location of one side of the road and due to the absence of elevation difference in this place, the region growing passes through this location and merged the road surface and the pavement into one segment. These merged regions were separated by a local re-segmentation procedure using the existing parts of the curb.

The proposed solution was to identify the existing parts of the curb, to merge them into a single continuous boundary line and to re-segment the region using the boundary as a barrier for BFS. This approach was successfully addressed the problem of curb cuts by creating a virtual boundary for segmentation. In this process, the existing road edge is important for estimating the absent boundary in the areas with curb cuts. Connecting the nearby separate boundary segments couldn't always represent the correct road boundary. Therefore, to find the accurate road boundary, the directional analysis of each boundary segment was very important. This prevented connecting boundary segments into incorrect other edges in different directions, even they are very close in distance. Then, the gap cells were filled by a Bresenham line, and obtained the accurate continuous road boundary. The method will be challenging when the actual boundary is continuously missed in a long curve. However, it is robust for short curb cuts along the road boundary. Figure 5 shows the existing boundary gap due to the curb cut and bridging the gap between them.

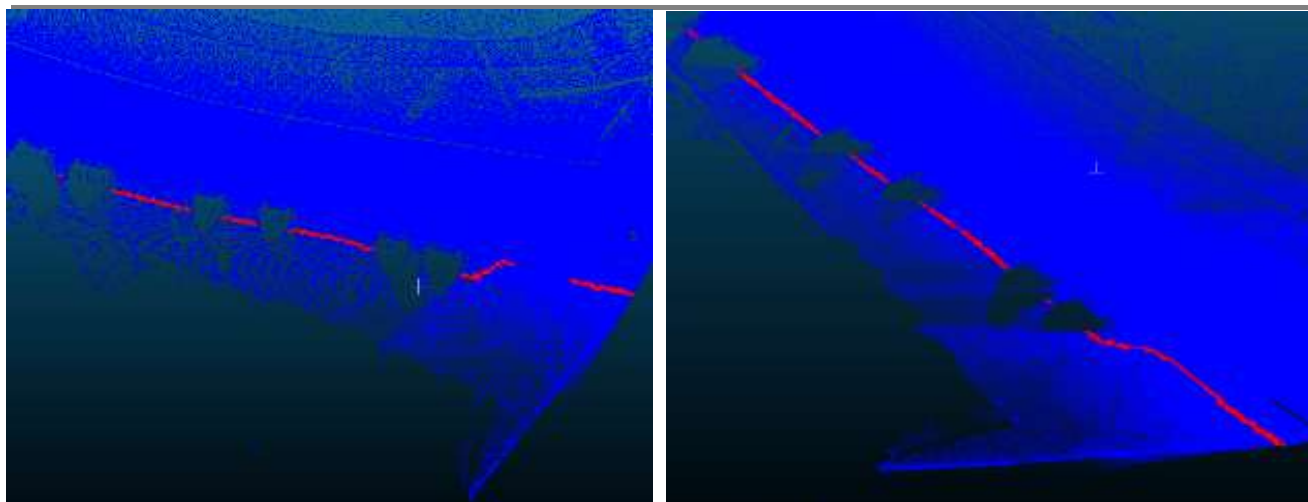


Figure 5: Filled Gap due to the Curb Cut

Boundary Identification and Vectorization

Identification of boundary cells of the road surface and curbs, played a crucial role in defining a clear distinction between the urban features. Boundary is important for defining the edges of curbs and road surface with high precision, defining the drivable area on the road and obstacles on the road. Finally, it was essential for vectorizing the boundaries, which allowed converting the rasterized boundaries into polygonal representations for GIS applications.

After developing and applying the methodology for the initial dataset, the process was tested on a more complex dataset, which contains more urban features and increased structural variations. The dataset included more curb islands and curved road edges. The objective was to evaluate the robustness of the developed algorithm and its ability to classify road features. The methodology performed exceptionally well in delineating road surface, curb islands and pavement features and the re-segmentation process accurately addressed the missing curb boundaries as well to precisely extract the road surface (Figure 6).

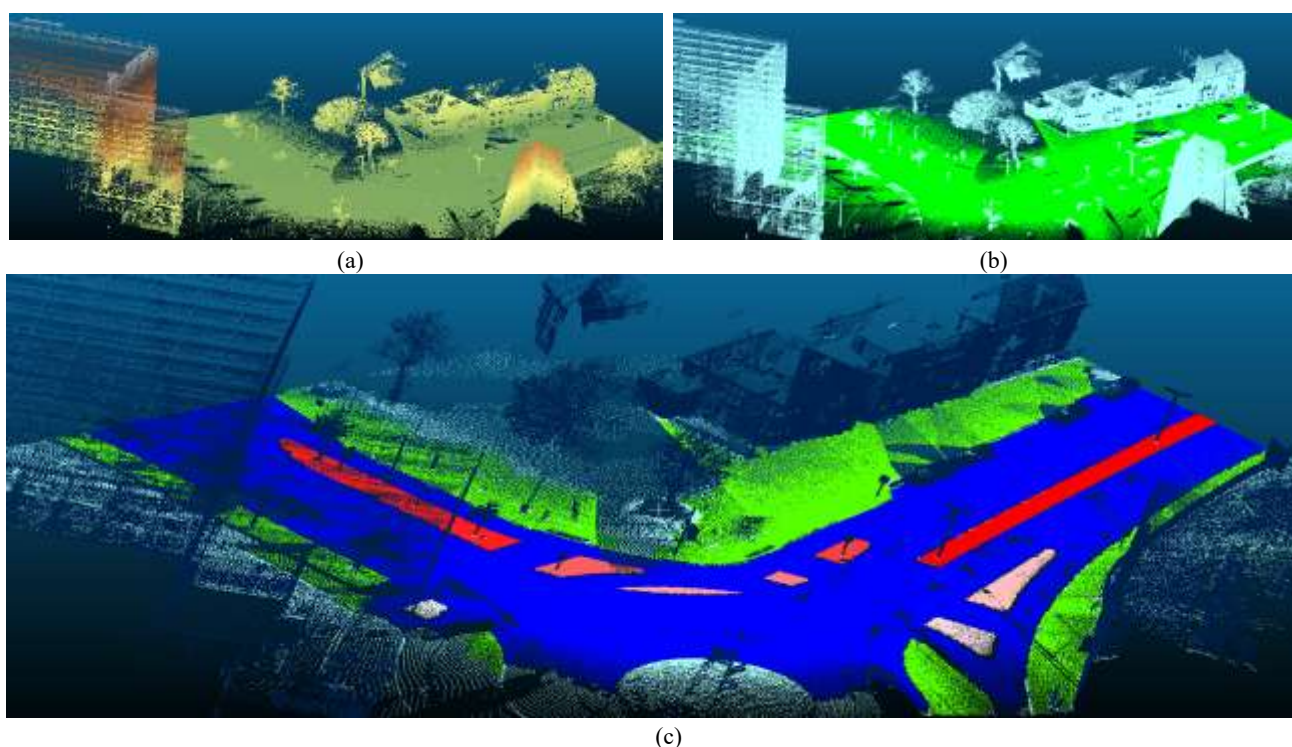


Figure 6: (a) shows the Second Subset used for Testing, (b) shows the extracted Ground Surface and (c) shows the Results of Classification.

Analyzing in raster format was efficient for processing and reducing the computational cost. It simplified the operations such as segmentation, neighborhood analysis and classification. However, to use these in practical applications, it is essential to convert these extracted features into continuous and GIS-compatible geometric representations. The Alpha-shape algorithm was a powerful and efficient techniques to convert the discrete boundary to vectorized polygons. It creates concave and tightly fitting boundaries that better preserves the true geometry of the features, unlike the traditional convex full methods that can overestimate the boundaries.

Performance Analysis

Finally, the confusion matrix showed that the model resulted excellent performance with an overall accuracy of 98%, effectively distinguishing road surfaces and curbs.

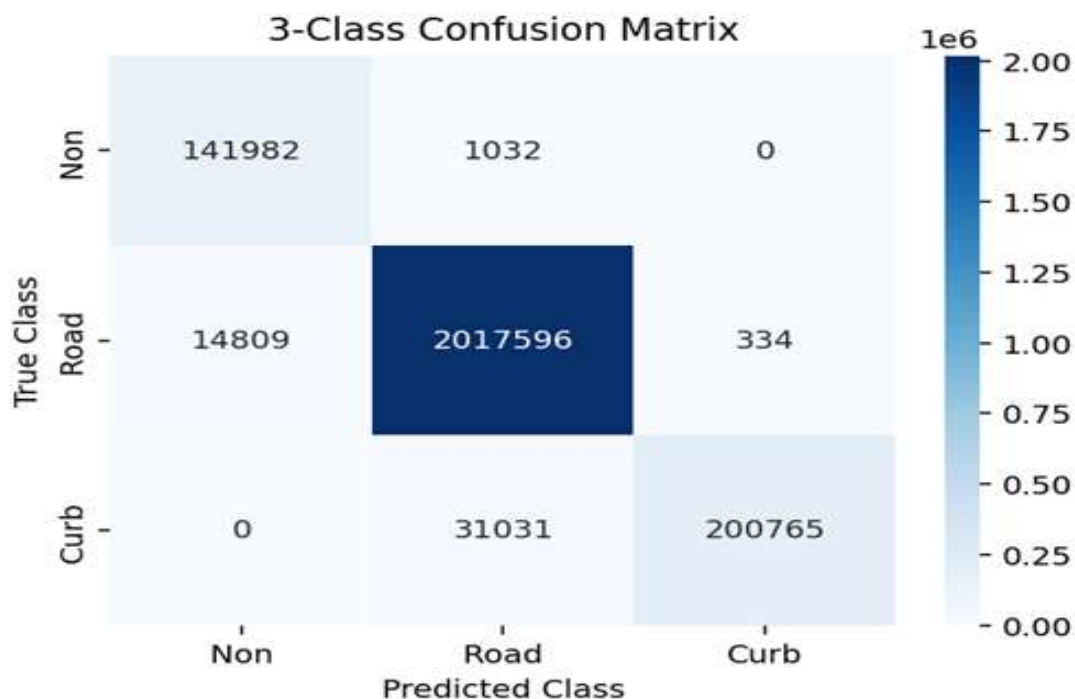


Figure 7: Results of the Confusion Matrix

Evaluating the geometric accuracy of the developed algorithm by computing area-based accuracy for curb islands, provides insights into the precision, as well as the reliability of the proposed method in real world applications. The accuracy analysis showed high reliability, with most curb islands achieved over 85% accuracy, and the highest reached 97.7%. There were several factors that affected the accuracy such as the data density of the particular area, occlusions and shadows due to the running and parked vehicles, and pedestrians, the complexity of the curb shapes, and parameter tuning. One exhibited the lowest accuracy of 65% due to the shadowing of complete opposite side of the data capture, and the complex shape.

Table 2: Area of Curb Islands

ID	Extracted Area (m ²)	Actual Area (m ²)	Accuracy
1	66.8025	70.4619	94.8
2	21.8362	24.3125	89.8
3	14.5012	17.1516	84.5
4	9.3262	10.2357	91.1
5	4.3762	4.9633	88.2
6	5.4562	8.3975	65.0
7	6.885	8.0076	86.0
8	15.5812	16.6621	93.5
9	55.485	56.7917	97.7

CONCLUSION AND FUTURE WORK

The study successfully developed an automated algorithm for extracting the road surface and road curb from MLS point cloud data. Initial ground extraction proved efficient for removing outliers and the algorithm achieved high precision with the implementation of the region-growing based segmentation based on elevation differences and normal vector orientation. The key challenge due to the presence of curb cuts was addressed effectively by a refined segmentation process. The results of assessing the geometric accuracy showed the most curb islands achieved an accuracy over 85% with some reaching 97.7% demonstrating the effectiveness of the developed approach. Also, the confusion matrix resulted 98% of excellent overall accuracy of the classification, ensuring the precise performance of the developed approach.

Processing all the steps for the above point cloud with over 3 million points took approximately 18 minutes on an Intel Core i7 (8th generation) laptop with 8565U CPU and 16 GB RAM, and the processing time can be further reduced when using a higher performance computer.

Future work should aim to reconstruct the final road surface boundary by filling gaps due to parked vehicles and occlusions, and provide the full road surface as the vectorized output. An adaptive curb height thresholding is better to automatically adjust the segmentation criteria based on the local height variations rather than using a fixed threshold. Also, optimizing the algorithm for real-time processing would enhance its applicability in autonomous navigation and traffic monitoring.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Data Availability

The data used in this study are available from the corresponding author upon reasonable request.

REFERENCES

1. Bartels, M., & Wei, H. (2010). Threshold-free object and ground point separation in LIDAR data. *Pattern Recognition Letters*, 31(10), 1089–1099.
2. Boyko, A., & Funkhouser, T. (n.d.). Extracting roads from dense point clouds in large scale urban environment.
3. Caltagirone, L., Bellone, M., Svensson, L., & Wahde, M. (2018). LIDAR-Camera Fusion for Road Detection Using Fully Convolutional Neural Networks.
4. Cartwright, W., Gartner, G., Meng, L., & Peterson, M. P. (n.d.). *Advances in GIScience: Proceedings of the 12th AGILE Conference (Lecture Notes in Geoinformation and Cartography)*.
5. Clode, S., Kootsookos, P., & Rottensteiner, F. (2004). The Automatic Extraction of Roads from LIDAR data.
6. Hinz, S., & Baumgartner, A. (2003). Automatic extraction of urban road networks from multi-view aerial imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(1–2), 83–98.
7. Honma, R., Date, H., & Kanai, S. (2020). Extraction of road edges from MLS point clouds using bend angle of scanlines. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 43(B2), 1091–1097.
8. Hu, X., Li, Y., Shan, J., Zhang, J., & Zhang, Y. (2014). Road centerline extraction in complex urban scenes from LiDAR data based on multiple features. *IEEE Transactions on Geoscience and Remote Sensing*, 52(11), 7448–7456.
9. Huber, R., & Lang, K. (n.d.). Road Extraction from High-Resolution Airborne SAR using Operator Fusion.
10. Ibrahim, S., & Lichti, D. (n.d.). CURB-BASED STREET FLOOR EXTRACTION FROM MOBILE TERRESTRIAL LIDAR POINT CLOUD. www.ambercore.com
11. IEEE Staff, (2010). 2010 IEEE Intelligent Vehicles Symposium. IEEE.

12. Kulathunga, G. P., Fedorenko, R., Klimchik, A., & Prathap, G. (n.d.). Ground and Non-Ground Separation Filter for UAV Lidar Point Cloud.
13. Li, Y., Hu, X., Guan, H., & Liu, P. (2016). An efficient method for automatic road extraction based on multiple features from lidar data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 41, 289–293.
14. Liu, L., & Lim, S. (2016). A framework of road extraction from airborne lidar data and aerial imagery. *Journal of Spatial Science*, 61(2), 263–281.
15. Ma, H., Ma, H., Zhang, L., Liu, K., & Luo, W. (2022). Extracting Urban Road Footprints from Airborne LiDAR Point Clouds with PointNet++ and Two-Step Post-Processing. *Remote Sensing*, 14(3).
16. Ma, L. (n.d.). Road Information Extraction from Mobile LiDAR Point Clouds using Deep Neural Networks.
17. Ma, L., Li, Y., Li, J., Wang, C., Wang, R., & Chapman, M. A. (2018). Mobile laser scanned point-clouds for road object detection and extraction: A review. In *Remote Sensing* (Vol. 10, Issue 10). MDPI AG.
18. Meng, X., Currit, N., & Zhao, K. (2010). Ground filtering algorithms for airborne LiDAR data: A review of critical issues. In *Remote Sensing* (Vol. 2, Issue 3, pp. 833–860).
19. Mi, X., Yang, B., Dong, Z., Chen, C., & Gu, J. (2022). Automated 3D Road Boundary Extraction and Vectorization Using MLS Point Clouds. *IEEE Transactions on Intelligent Transportation Systems*, 23(6), 5287–5297.
20. Siegemund, J., Franke, U., & Förstner, W. (2011). A temporal filter approach for detection and reconstruction of curbs and road surfaces based on conditional random fields. *IEEE Intelligent Vehicles Symposium, Proceedings*, 637–642.
21. Sui, L., Zhu, J., Zhong, M., Wang, X., & Kang, J. (2021). Extraction of road boundary from MLS data using laser scanner ground trajectory. *Open Geosciences*, 13(1), 690–704.
22. Wang, G., Wu, J., He, R., & Yang, S. (2019). A point cloud-based robust road curb detection and tracking method. *IEEE Access*, 7, 24611–24625.
23. Wang, H., Cai, Z., Luo, H., Wang, C., Li, P., Yang, W., Ren, S., & Li, J. (2012). Automatic Road Extraction from Mobile Laser Scanning Data.
24. Yadav, M. (2021). A multi-constraint combined method for road extraction from airborne laser scanning data. *Measurement: Journal of the International Measurement Confederation*, 186.
25. Yadav, M., Lohani, B., & Singh, A. K. (2018). Road surface detection from mobile lidar data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4(5), 95–101.
26. Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., & Yan, G. (2016). An easy-to-use airborne LiDAR data filtering method based on cloth simulation. *Remote Sensing*, 8(6).