LiDAR Ground Segmentation and Modeling for Mobile Robots in Unstructured Terrain

1st David Skuddis
SysLab
Diehl Defence GmbH & Co. KG
Ueberlingen, Germany
david.skuddis@diehl-defence.com

2nd Juergen Koeberle
SysLab
Diehl Defence GmbH & Co. KG
Ueberlingen, Germany
juergen.koeberle@diehl-defence.com

3rd Norbert Haala
Institute for Photogrammetry
University of Stuttgart
Stuttgart, Germany
norbert.haala@ifp.uni-stuttgart.de

2022 8th International Conference on Automation, Robotics and Applications (ICARA) | 978-1-6654-8383-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICARA55094.2022.9738519

***Abstract*—Ground segmentation and modeling for point clouds generated by mobile laser scanners is an important processing step in the environment perception of mobile robots. Many methods that work well in urban areas do not work correctly in unstructured environments such as meadows or forests. We present a new method that works reliably in both structured urban areas and unstructured outdoor environments. Overall we design a three steps pipeline. First individual scattered ground base points are identified using purely geometric criteria. Based on the ground base points, the ground is then modeled as a triangle mesh in a second step. The triangle mesh is used to provide a relative height and a local slope. The effectiveness and robustness of the proposed method are confirmed and it shows promising results even under difficult conditions.**

***Index Terms*—LiDAR, Mobile Robots, Outdoor, Segmentation**

I. Introduction

Mobile robots are playing an increasingly important role in both private and industrial environments. A key capability is their perception and correct interpretation of the robot environment. The latter becomes increasingly difficult the more complex and unstructured the robot's operational area is. For robots with higher autonomy functions, such as automatic obstacle avoidance, it is important to determine the relative height of each object in the vehicle environment. The relative height can then be used to decide to what extent this object influences trajectory planning. In addition, it is important for further processing steps to obtain not only the relative height of the points or objects, but also a reliable estimate of the gradients in the robot's surroundings. The presented approach can serve with both.

II. Related Work

The task of modeling the ground based on a single LiDAR scan, or separating ground points from non-ground points, has already been investigated using a wide variety of methods. In many cases, the overall goal is to identify drivable areas for autonomous robots in real time.

In [1] the environment is modeled as a two-dimensional horizontal grid. For each cell of the grid, the vertical difference between the highest and lowest point in the cell is determined. If the difference is above a threshold, the cell is considered occupied. All points whose vertical difference from the lowest point in the cell is above the threshold can be seen as non-ground points. All others could be considered as ground points. In the case of slopes, however, this can lead to errors, depending on the selected grid size and point density. In [2], a graph is formed based on the local range image neighborhood. Local normal vectors are then computed for the nodes of the graph. Ground points are then identified using a local convexity criterion. In [3] and [4], the vehicle environment is divided into two-dimensional segments based on azimuth. For each segment, lines are then determined that describe the course of the ground level. Based on this, a relative height is assigned to all points, which is then used to divide them into ground points and non-ground points. In [5], Zermas et al. present a robust approach in which ground points are identified iteratively. In this approach points near the initial selected seed points whose elevation difference is within a threshold are used to iteratively model the ground locally. In [6] a efficient approach is presented where ground points are determined in the range image representation. For all points the slope angles of the connection line between the respective point and the neighbor above it in the range image are calculated. Assuming that there are ground points in the lowest scan line, the ground points are then identified by a breadth-first search. A limit value for the maximum allowable slope change between adjacent ground points is used as a connectivity condition. This approach will be used later as a comparison.

In current works, deep learning based approaches are increasingly used, e.g. in [7].

While all the methods mentioned so far are intended for real-time application of mobile laser scanners, there are also methods that originate specifically from the airborne laser scanning (ALS) field. Since there is no real-time requirement for processing ALS point clouds, more computationally intensive methods are also used here. In [8] different methods from the ALS area are compared. In the context of this work, the method in [9], which originates from the ALS area, is also used in a processing step.

Most of the methods presented are intended for use in structured environments such as urban areas. In comparison, our approach is able to reliably determine the ground level for each point even in unstructured terrain such as high meadows.

III. Notation and Definitions

Since different data representations are used, the required terminology is briefly explained. A point cloud is a list of x, y and z coordinates. The dimension of a point cloud with N points is therefore $N \times 3$. A structured point cloud is a data representation where the x, y, and z coordinates of the points are stored in a three-dimensional array. The dimension of a structured point cloud results from the number of scan rings n_r and the horizontal resolution n_h for rotating LiDARs. The dimension is therefore $n_r \times n_h \times 3$. A range image is a representation that corresponds to the structured point cloud, where instead of the three-dimensional coordinates in the grid, the distance of the corresponding point from the sensor is stored. The dimension corresponds to $n_r \times n_h$. In the following, a single row from the range image is considered as scan line S. It is therefore a vector that contains the distance measurements of a single scan ring of the LiDAR sensors with n_h elements.

IV. Methodology

In the presented approach a structured point cloud is expected as input. The result of the approach is a triangle mesh that represents the ground. Each triangle is described by a plane equation, which represents the relative ground level and the slope. The method consists of three steps. In the first step, individual scattered ground base points are identified. In this context, the approach of Vosselman [9] is applied. Then a triangle mesh is created based on the ground base points. In a final step, the mesh is used to assign a relative height to each point of the point cloud.

The procedure consists of the following steps:

1) Identify *Maximum Points* of all scan lines pointing towards the ground (cf. section IV-A1 and IV-A2).
2) Check the *Maximum Local Slope* [9] criterion for all *Maximum Points* and remove all points that do not meet it (cf. section IV-A3). Consider the result as ground base points.
3) Create a 2D triangle mesh using the ground base points (cf. section IV-B).
4) Calculate the normal vector for each triangle of the mesh and remove triangles whose slope exceeds a threshold value (cf. section IV-B).
5) Assign to each point of the point cloud the triangle in which it lies in the horizontal plane or which is closest to it.
6) Calculate the relative height for each point by determining the distance to the plane of the associated triangle (cf. section IV-C).

In the following, the processing steps are presented in more detail.

A. Ground Base Points Identification

To identify ground base points, three geometric criteria are defined. These are called *Beam Elevation Angle Criterion*, *Maximum Criterion* and *Maximum Local Slope Criterion*. If a point in the point cloud meets all three criteria, it is considered a ground base point. During implementation, the criteria are checked one after the other. The next criterion is only checked for the set of points for which the previous criterion(s) are true. Fig. 1a and 1b show the selection of ground base points for an exemplary scene.

1) Beam Elevation Angle Criterion: The *Beam Elevation Angle Criterion* is simple yet effective. Only points whose elevation angle in the sensor coordinate system is less than zero degrees, which means that the beam points towards the ground, are considered as ground points. This criterion is used to roughly filter the points and assumes that the sensor is approximately parallel to the ground plane.

2) Maximum Criterion: The *Maximum Criterion* is a morphological criterion and refers to the local neighborhood of a point along its scan line. To be considered as a maximum point, the distance of the point from the sensor must be maximum within a window of d_N along its scan line. The aim here is subsampling, while keeping relevant ground points. Points on small and filigree objects tend to be filtered out. The determination of the points that satisfy the *Maximum Criterion* is done using single scan lines from the range image representation of a given structured point cloud. The points that meet the *Maximum Criterion* are called maximum points P_{max} below. In Algorithm 1, an efficient algorithm is described that approximately identifies the points that satisfy the criterion. In Fig. 1a the maximum points for an exemplary outdoor scene are shown.

Algorithm 1: Algorithm for identifying maximum points

Input: Scan line ranges S, sensors horizontal angle increment $\Delta\alpha$, minimum distance between maximum points d_N
Output: Scan line indices of maximum points I_{MP}

$I_{MP} = \emptyset$
while $s_{max} \neq 0$ **do**
  [s_{max}, i] = max(S) *% find max element and index*
  $I_{MP} = I_{MP} \cup i$
  % Convert distance d_N into index difference Δi
  Δi = round($\frac{d_N}{s_{max} \cdot \tan(\Delta\alpha)}$)
  % Set neighbor points of i in S to zero
  $S(i - \Delta i : i + \Delta i) = 0$
end while
return I_{MP}

3) Maximum Local Slope Criterion: The Maximum Local Slope (MLS) criterion was introduced by [9] and is applied in the processing of airborne LiDAR point clouds. Due to the high computational costs and the sparsity of the point clouds, it is not used so far in online applications with mobile laser scanners. By filtering the point cloud through the previously presented criteria, the MLS criterion only needs to be checked for a small subset of the original points. This enables the online application.

For a given candidate point $p_c = [x_c, y_c, z_c]^T$ of the point cloud P, the MLS criterion is true if the slope between the

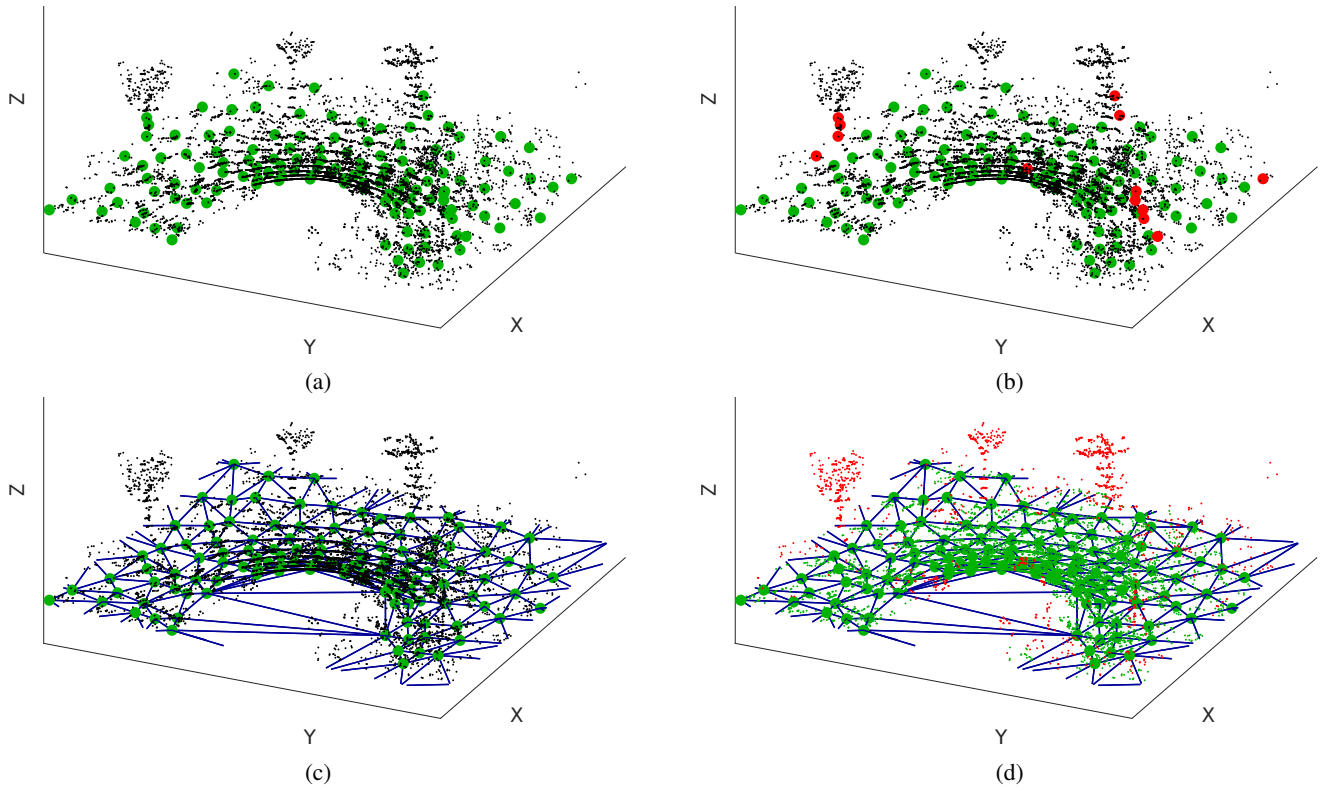

Fig. 1: Ground base points identification and ground modeling process for an exemplary outdoor scene. (a) Points that meet the beam elevation angle criterion and the maximum criterion are marked in green (b) Maximum points that meet the MLS criterion are marked in green, others in red otherwise. (c) Generated triangle mesh (blue) based on the ground base points (green). (d) Segmented point cloud. Red points indicate a relative height above 0.5 m and green points below.

current point and all other points is smaller than a boundary value m_{max}. This condition can be formulated as follows:

$$\forall p_j \in P : \frac{z_c - z_j}{\sqrt{(x_c - x_j)^2 + (y_c - y_j)^2}} \leq m_{max} \quad (1)$$

As suggested in [9], it is not necessary to calculate the slope between the current point and all other points. It is enough to check the points in the local area around the current point. In our case, points with a maximum horizontal distance of d_{hor} are checked. Fig. 1b shows the maximum points (green) that meet the MLS criterion. These points are considered as ground base points.

B. Mesh Generation

Through the previously mentioned steps, scattered ground base points have been identified. To segment the ground points and identify potential obstacle points, it is necessary to assign a relative height above ground to each point in the point cloud.

A naive approach would be to identify the nearest ground base point for each point and use the difference in z coordinates as the elevation. But this would lead to small errors depending on the ground point density in areas with slopes. A better result can be obtained if interpolation is used between the ground base points. To implement this interpolation, a two-dimensional triangle mesh is generated using the x and y coordinates of the ground points. The mesh is calculated using *Delaunay Triangulation* based on [10]. For each triangle of the mesh, a local plane is defined. Let p_1, p_2 and p_3 be the points of a triangle, then the normal vector $N_t = [n_{t,1}, n_{t,2}, n_{t,3}]^T$ and the slope m_t of the triangle are obtained as follows:

$$N_t = \frac{1}{\|(p_2 - p_1) \times (p_3 - p_1)\|} \cdot (p_2 - p_1) \times (p_3 - p_1) \quad (2)$$

$$m_t = \frac{\sqrt{n_{t,1}^2 + n_{t,2}^2}}{n_{t,3}} \quad (3)$$

If the magnitude of the slope of a triangle exceeds a limit value $|m_t| > m_{max}$, the triangle and the associated plane are removed. Fig. 1c shows the triangle mesh for an exemplary scene.

C. Calculation of Relative Height and Ground Segmentation

To calculate the relative height, for each point the triangle is selected in which it lies when projected into the horizontal plane or whose center point is closest to it. Then the distance to the plane associated with the triangle is calculated. This distance is considered as the relative height of the point. Figure 1d shows the relative height. Points with a relative height lower than 0.5 m are shown in green, other points red.

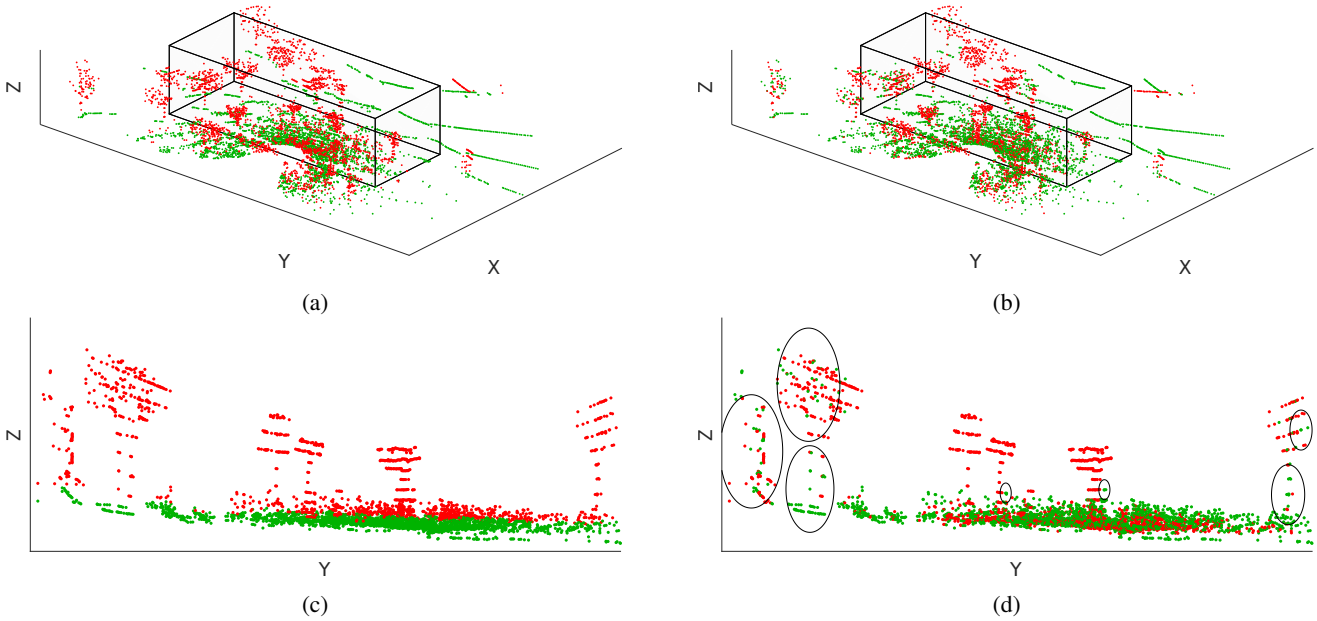

Fig. 2: Qualitative comparison based on an outdoor scene on a meadow with trees between the proposed approach and a state of the art approach [6]. Ground points are marked in green and others in red. The left column shows the result of the proposed approach and the right column shows that of [6]. Black ellipses indicate places where the comparative approach is incorrect. (a) and (b) show the scene from a bird's eye view and (c) and (d) show the area marked in (a) and (b) as a projection in the Y-Z plane.

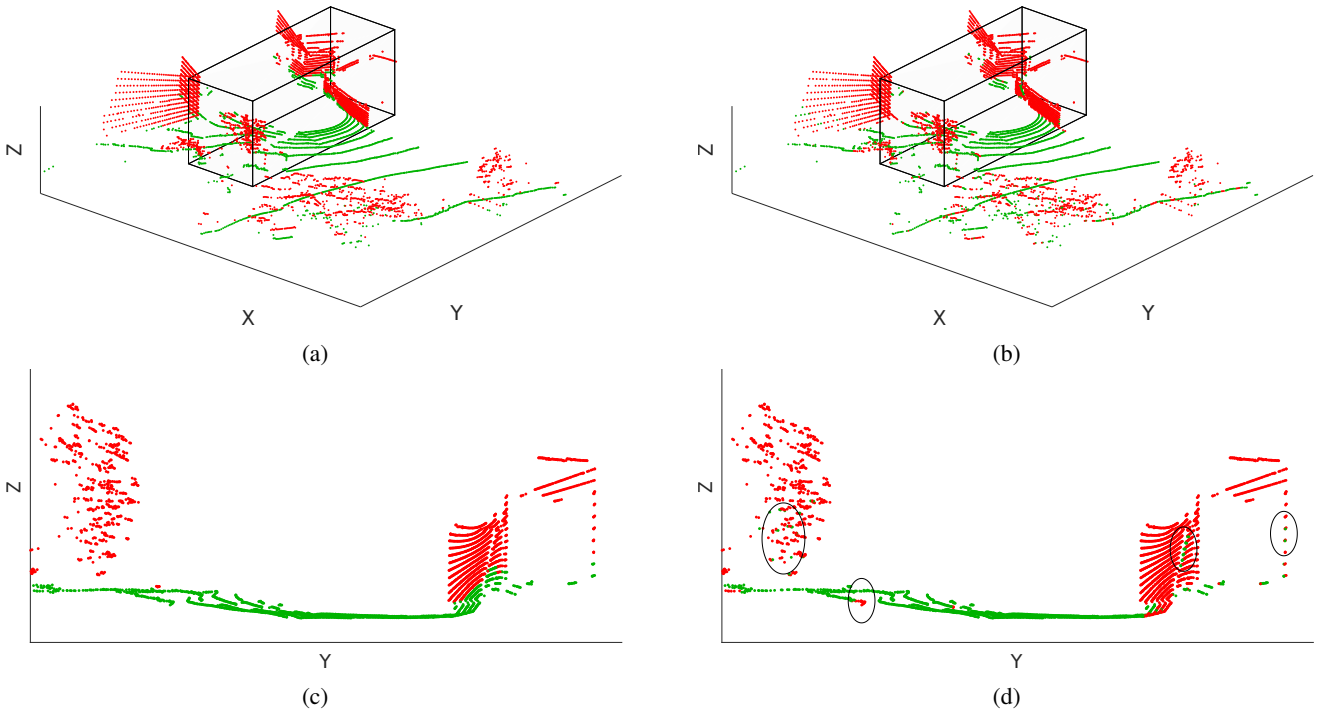

Fig. 3: Qualitative comparison based on an urban scene between the proposed approach and a state of the art approach [6]. Ground points are marked in green and others in red. The left column shows the result of the proposed approach and the right column shows that of [6]. Black ellipses indicate places where the comparative approach is incorrect. (a) and (b) show the scene from a bird's eye view and (c) and (d) show the area marked in (a) and (b) as a projection in the Y-Z plane.

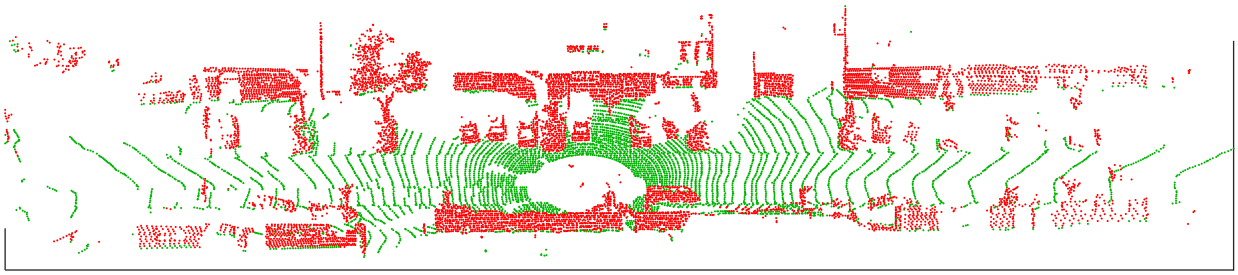

Fig. 4: Segmented exemplary scene from SemanticKITTI dataset.

V. Results

The presented method was qualitatively tested using recorded data from the outdoor area. The data was collected with Velodyne's Puck LiDAR sensor. The data sets included high meadows, bushes, slopes and trees. In experiments, reliable ground segmentation could be observed even under very difficult conditions. In figures 2 and 3, the proposed approach is compared to a state of the art procedure [6] based on two scenes. In this qualitative comparison, the comparison procedure showed errors in segmentation, especially in the scene in the meadow. Here, ground points were incorrectly classified as non ground points and more frequently points of higher objects were classified as ground points (Fig. 2d). But even in the urban scene, individual points were incorrectly classified, while the proposed approach provides a comprehensible result (Fig. 3d).

To obtain a quantitative assessment of the approach, the method was applied to sequence 00 to 10 of the SemanticKITTI dataset [11], [12]. The method only segments non-ground points from ground points but does not distinguish between different ground types. Thus the different classes of the SemanticKITTI dataset (*road*, *sidewalk*, *parking*, *terrain*, *lane-marking* and *other-ground*) were considered as a single *ground* class. All other classes were considered as *non-ground*. An overall accuracy of 94.5 % was achieved. More detailed results are presented in Table I. In Fig. 4 an exemplary scene is shown.

TABLE I: Evaluation on the SemanticKITTI dataset

Metric	*ground*	*non-ground*
Intersection over Union (Jaccard index)	88.8 %	90.3 %
Recall	96.8 %	92.7 %
Precision	91.5 %	97.3 %
Overall accuracy	**94.5 %**	

VI. Conclusion

We presented a new method for mobile LiDAR sensors that can be used to reliably model the ground and segment ground points in real time in both urban and outdoor areas. The performance in outdoor areas was evaluated in experiments with recorded data. In order to demonstrate the advantages of the method, a qualitative comparison was made with a state-of-the-art method [6] on the basis of two scenes (cf. Fig. 2 and 3). A clear superiority of the presented method in the outdoor area could be shown. The SemanticKITTI dataset was used to successfully demonstrate reliable operation even in urban areas. In future work, the method will be combined with classifiers for vegetation detection. The goal is to obtain a representation of the vehicle environment that is suitable for path planning of mobile robots in unstructured outdoor area.

References

[1] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, P. Fong, J. Gale, M. Halpenny, G. Hoffmann, K. Lau, C. Oakley, M. Palatucci, V. Pratt, P. Stang, S. Strohband, C. Dupont, L. E. Jendrossek, C. Koelen, C. Markey, C. Rummel, J. van Niekerk, E. Jensen, P. Alessandrini, G. Bradski, B. Davies, S. Ettinger, A. Kaehler, A. Nefian, and P. Mahoney, "Stanley: The robot that won the DARPA Grand Challenge," *Journal of Field Robotics*, vol. 23, no. 9, pp. 661–692, 2006.

[2] F. Moosmann, O. Pink, and C. Stiller, "Segmentation of 3D lidar data in non-flat urban environments using a local convexity criterion," *IEEE Intelligent Vehicles Symposium, Proceedings*, pp. 215–220, 2009.

[3] D. Steinhauser, O. Ruepp, and D. Burschka, "Motion segmentation and scene classification from 3D LIDAR data," *IEEE Intelligent Vehicles Symposium, Proceedings*, pp. 398–403, 2008.

[4] M. Himmelsbach, F. V. Hundelshausen, and H. J. Wuensche, "Fast segmentation of 3D point clouds for ground vehicles," *IEEE Intelligent Vehicles Symposium, Proceedings*, pp. 560–565, 2010.

[5] D. Zermas, I. Izzat, and N. Papanikolopoulos, "Fast segmentation of 3D point clouds: A paradigm on LiDAR data for autonomous vehicle applications," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 5067–5073, 2017.

[6] I. Bogoslavskyi and C. Stachniss, "Efficient Online Segmentation for Sparse 3D Laser Scans," *Photogrammetrie, Fernerkundung, Geoinformation*, vol. 85, no. 1, pp. 41–52, 2017.

[7] M. Velas, M. Spanel, M. Hradis, and A. Herout, "CNN for very fast ground segmentation in velodyne LiDAR data," *18th IEEE International Conference on Autonomous Robot Systems and Competitions, ICARSC 2018*, pp. 97–103, 2018.

[8] K. Zhang and D. Whitman, "Comparison of three algorithms for filtering airborne lidar data," *Photogrammetric Engineering and Remote Sensing*, vol. 71, no. 3, pp. 313–324, 2005.

[9] G. Vosselman, "Slope based filtering of laser altimetry data," *International archives of photogrammetry and remote sensing*, vol. 33, no. B3/2; PART 3, pp. 935–942, 2000.

[10] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa, "The Quickhull Algorithm for Convex Hulls," *ACM Transactions on Mathematical Software*, vol. 22, no. 4, pp. 469–483, 1996.

[11] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The kitti dataset," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013.

[12] J. Behley, M. Garbade, A. Milioto, J. Quenzel, S. Behnke, C. Stachniss, and J. Gall, "Semantickitti: A dataset for semantic scene understanding of lidar sequences," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 9297–9307, 2019.