TRA Vel: Traversable Ground and Above-Ground Object Segmentation Using Graph Representation of 3D LiDAR Scans

Minho Oh1,‡, Euing Jung1,‡, Hyungtae Lim1, Wonho Song1, Sumin Hu1, Eungchang Mason Lee1, Junghie Park2, Jaekyung Kim2, Jangwoo Lee2, and Hyun Myung1,∗, Senior Member, IEEE

Abstract—Perception of traversable regions and objects of interest from a 3D point cloud is one of the critical tasks in autonomous navigation. A ground vehicle needs to look for traversable terrains that are explorable by wheels. Then, to make safe navigation decisions, the segmentation of objects positioned on those terrains has to be followed up. However, over-segmentation and under-segmentation can negatively influence such navigation decisions. To that end, we propose TRAVEL, which performs traversable ground detection and object clustering simultaneously using the graph representation of a 3D point cloud. To segment the traversable ground, a point cloud is encoded into a graph structure, tri-grid field, which treats each tri-grid as a node. Then, the traversable regions are searched and redefined by examining local convexity and concavity of edges that connect nodes. On the other hand, our above-ground object segmentation employs a graph structure by representing a group of horizontally neighboring 3D points in a spherical-projection space as a node and vertical/horizontal relationship between nodes as an edge. Fully leveraging the node-edge structure, the above-ground segmentation ensures real-time operation and mitigates over-segmentation. Through experiments using simulations, urban scenes, and our own datasets, we have demonstrated that our proposed traversable ground segmentation algorithm outperforms other state-of-the-art methods in terms of the conventional metrics and that our newly proposed evaluation metrics are meaningful for assessing the above-ground segmentation. We will make the code and our own dataset available to public at
https://github.com/url-kaist/TRA Vel. 

Index Terms—Traversable ground segmentation; Object segmentation; Graph search; LiDAR; Autonomous navigation

I. INTRODUCTION

In recent years, there has been an increasing demand to perceive and represent surroundings in the robotics field. For autonomous navigation, robust object segmentation is required to identify meaningful objects—possibly to track and even avoid them in the subsequent autonomous tasks or utilize them in localization [1], [2] or mapping [3], [4]. In this paper, we explicitly explore segmentation of a point cloud captured by a 3D LiDAR sensor. A typical point cloud segmentation infers a class label of each data point, which is mostly tackled by learning-based approaches. Unfortunately, most of these approaches are supervised with ground-truth labels and computationally costly to process an immense number of data on a mobile GPU. These downsides have hindered a mobile robot platform from adopting the learning approaches into navigation in an unknown environment.

The focus of our work is primarily oriented towards safe navigation that is not subject to a specific environment. Rather than assigning class labels (e.g. vehicles, pedestrians, poles, etc.) to all data points, we detect a traversable area and spatially distinguish meaningful objects. To that end, we propose an efficient segmentation algorithm, named TRAVEL, that leverages a node-edge graph structure. TRAVEL consists of tri-grid field (TGF)-based traversable ground segmentation by terrain modeling and above-ground object segmentation by clustering, as shown in Fig. 1. We evaluate our proposed algorithm on CARLA [5], Semantic KITTI dataset [6], and our own rough terrain dataset, to underscore the following contributions:

- To the best of our knowledge, this research is the first to introduce TGF for ground segmentation. By leveraging TGF and geometrical discrepancy, our algorithm can effectively detect the traversability of terrains based on the proposed breadth-first search.
- We separately cluster object points horizontally and vertically with the following improvements:
  - We introduce a concept of skipped linkage and circular linkage to effectively handle over-segmentation.
  - Our binary search approach reduces the complexity of finding neighboring nodes to \(O(N \log(N))\), where \(N\) is the total number of nodes in a ring.
- We propose novel metrics, namely over-segmentation entropy and under-segmentation entropy. These metrics measure the distribution and uncertainty of labels.
that correspond to an object, being comprehensive and meaningful measures to assess segmentation performance.

The rest of the letter is organized as follows: Section [II] provides an overview of related works. Section [III] explains the proposed traversable ground segmentation algorithm; Section [IV] delineates the procedure of the above-ground segmentation; Section [V] describes the experiments and novel evaluation metrics; Section [VI] discusses the experimental results; and, finally, Section [VII] summarizes our contributions and explores future works.

II. RELATED WORKS

A. Ground Segmentation

To overcome the under-segmentation problem, a real-time ground segmentation that is robust to varying environments should come prior to object segmentation. The previous lines of research base themselves on RANSAC [7], [8] or ground plane fitting approaches using the principal component analysis (PCA) [9], [10]. Among them, Narksri et al. [8] and Lim et al. [10] proposed the slope-robust methods based on the ego-centric scan representation. On the other hand, Xue et al. [11] proposed an algorithm that examines density of scan points based on the grid-based scan representation. All of the above methods have solely focused on ground segmentation itself and have not examined the traversability, discerning an area on which a vehicle can stand and travel safely. Moosmann et al. [12] employed a point-wise graph search method to split the ground and vertical points considering continuity by local convexity. However, traversability is still not considered, and real-time operation is not guaranteed due to point-wise graph-search. To address this issue, our traversable ground segmentation explores the traversability considering all three attributes: slope, level, and continuity of a terrain.

B. Object Segmentation

Given the immense resources required to classify massive number of remaining scan points after the ground segmentation, above-ground object segmentation particularly focuses on real-time feasibility while keeping up reasonable accuracy. Bogoslavskiy et al. [13] employed a range image and calculated the difference in reflection angles of adjacent points to differentiate objects. Zermas et al. [9] proposed a novel method of using the ring structure of a 3D point cloud captured by a 3D LiDAR sensor, clustering points horizontally in a ring and vertically between rings in two steps. Burger et al. [14] proposed a mesh structure using loss functions that consist of cluster densities, slopes, distances, and angles; however, its three steps of horizontal, vertical, and fusion updates entail high computation complexity. Yang et al. [15] represented a set of horizontally neighboring points as a node and created a set graph. However, Yang et al. [15] did not consider the interaction between index-wise non-adjointing nodes, even when, due to occlusion, they can be possible candidates to be clustered geometrically. Establishing deficient interplay between scan points, the above segmentation methods are prone to over-segmentation, especially caused by occlusion, and under-segmentation, especially caused by ground segmentation failure. In addition, although the prior works measured computation times and qualitatively evaluated their performance, they lacked appropriate quantitative metrics to scrutinize over-segmentation and under-segmentation.

C. Learning-based Segmentation

Several works have tackled the segmentation task with learning-based methods. The research by Zhang et al. [16] directly identified objects from a point cloud. Wong et al. [17] proposed an instance segmentation method to recognize both known and unknown instance objects. Although these learning-based instance segmentation methods can accurately extract objects, they cannot recognize the ground or walls that are also important for autonomous navigation. Unlike the instance segmentation, others proposed the semantic segmentation method that can label the point-wise class [18]–[20]. These methods extract the ground without considering traversability. Also, since the extracted objects do not have instance labels, some separate post-processes (e.g., clustering) are necessary to be applied to autonomous navigation. Paigwar et al. [21] proposed the semantic segmentation with only two classes; ground and non-ground. It can process the ground estimation in real-time by considering the slope of terrains. Although learning-based methods work well in the learned environment, the performance cannot be guaranteed under the condition encountered for the first time [17].

III. TRAVERSABLE GROUND SEGMENTATION

TGF-based traversable ground segmentation, the first step of TRAVEL, mainly consists of three parts: node-wise terrain modeling on TGF, breadth-first traversable graph search (BTGS), and traversable ground modeling and segmentation.

A. Pre-processing for Traversability

The pre-processing step is essential for boosting segmentation performance especially in bumpy terrains. As shown in Fig. 2(a), a point cloud can be skewed by the motion of a mobile platform [2], [22]. To deskew a point cloud, we use an IMU pre-integration approach to fuse LiDAR and IMU data tightly. Also, we adopt a 6-DoF pose to alleviate the segmentation problem caused by a tilted platform, as shown in Fig. 2(b). The raw point cloud is rotated to the upright position by using the pitch and roll angles calculated from the pose. After these compensations, we can accurately estimate the terrain properties in the world coordinate [12].

Fig. 2. Effects by (a) point cloud deskewing and (b) attitude alignment. The black and green points represent point clouds before and after applying each compensation, respectively. (Best viewed in color)
B. Node-wise Terrain Model on TGF

Firstly, as shown in Fig. 3(a) a 3D point cloud is encoded into our graph representation, TGF, which is the pre-built grid with the corresponding grid points among a set of node initial ground points, tentative nodes, and unknown nodes. (Best viewed in color)

Each triangular node \( T_i \) contains a corresponding partial point cloud, \( P_i \). TGF consists of a set of nodes \( N = \{ T_i \mid i \in N \} \) and a set of edges \( E = \{ e_{ij} \mid i, j \in N \} \), where \( e_{ij} \) connects \( T_i \) and \( T_j \), and \( N \) is a set of node indices.

Accordingly, the PCA-based ground plane fitting approach estimates a planar model \( P_s \), initial ground points, and descending ordered eigenvalues \( \lambda_k \in \{1, 2, 3\} \) for \( P_s \) in each \( n_i^T \), where \( P_s \) consists of a normalized surface normal vector \( s_i \in \mathbb{R}^3 \), i.e. \( \| s_i \| = 1 \), and plane coefficient \( d_i \). A mean point \( m_i \in \mathbb{R}^3 \), which is obtained by averaging the initial ground points among \( P_s \), and weight \( w(T_i) \) for scoring the traversability of a corresponding tri-grid are also included in \( n_i^T \). They are denoted as follows:

\[
P_i^T \begin{bmatrix} m_i^1 \\ m_i^2 \\ m_i^3 \\ s_i \end{bmatrix} = \begin{bmatrix} [s_i^T] d_i \\ m_i \end{bmatrix} = 0,
\]

(1)

where \( \text{cohesion} = \lambda_1/\lambda_3 \), \( \text{planarity} = \lambda_2/\lambda_3 \), and \( \text{linearity} = \lambda_1/\lambda_2 \). A node that is less inclined than \( \theta^T \), more inclined than \( \theta^T \), or has fewer points than \( \sigma^T \), is classified into \( n_{T,s} \), \( n_{T,o} \), and unknown node \( n_{T,u} \). A node that is less inclined than \( \theta^T \), more inclined than \( \theta^T \), or has fewer points than \( \sigma^T \), is classified into \( n_{T,s} \), \( n_{T,o} \), and unknown node \( n_{T,u} \). A node that is less inclined than \( \theta^T \), more inclined than \( \theta^T \), or has fewer points than \( \sigma^T \), is classified into \( n_{T,s} \), \( n_{T,o} \), and unknown node \( n_{T,u} \).

C. Breadth-first Traversable Graph Search (B-TGS)

To search for a set of traversable nodes \( T^T \) among a set of \( n_{T,s} \), we adopt a breadth-first search approach into TGF, namely B-TGS. At first, \( n_{T,s} \) with the highest weight and the closest to the sensor is selected as a seed \( n_{T,s} \), and the adjacent nodes of the seed are considered as the neighbors, as illustrated in Fig. 5(b). Then, to determine the traversable node \( n_{T,s} \) among the neighbors, we calculate the traversability of \( e_{ij} \) that is the geometric relationship between \( n_{T,s} \) and \( n_{T,u} \).

Inspired by [12], the traversability of \( e_{ij} \), is determined by local convexity and concavity of itself, which are illustrated in Fig. 5(a). The following equation expresses whether an edge \( e_{ij} \) has acceptable local convexity and concavity:

\[
lcc(e_{ij}) = \begin{cases} 
true, & \text{if } (|s_i \cdot s_j| > 1 - \sin(|d_{ij}|/\epsilon_2)) \\
\land (|s_i \cdot d_{ij}| < ||d_{ij}|| \sin \epsilon_1) \\
\land (|s_i \cdot d_{ij}| < ||d_{ij}|| \sin \epsilon_1) \\
false, & \text{otherwise}
\end{cases}
\]

(2)

where \( d_{ij} = m_i - m_j \) is the displacement vector between two nodes, \( \epsilon_1 \) denotes the angle at which \( m_i \) (or \( m_j \)) may lie above \( P_j \) (or \( P_i \), and \( \epsilon_2 \) denotes the threshold angle for similarity. If \( lcc(e_{ij}) \) between the seed and other node is true, then the node belongs to \( T^T \) and becomes another seed for the following search steps. B-TGS continues until there are no more neighboring nodes. As shown in Fig. 6(a), at the end of B-TGS, the traversable set of \( e^T \) is stretched out by connecting the corresponding \( n_{T,s} \in T^T \subset N^T \).
D. TGF-wise Traversable Terrain Model Fitting (TTMF)

Finally, in traversable terrain model fitting (TTMF) process, by applying the weighted corner fitting to the triangular corners $c^T_{k\in\{1,2,3\},i} = (x^c_{k,i}, y^c_{k,i}, z^c_{k,i})$ of the nodes $n^T_{i} \in \mathcal{T}'$, the overall traversable ground on TGF is refined from $\mathcal{P}$ to $\mathcal{P}$ for each $n^T_{i}$. In the weighted corner fitting, all the corners are grouped into $\mathcal{C}'_{m} \in \mathcal{M} = \{c^w_{k,i,m}, x^c_{m} = x^c_{k,i,m}, y^c_{m} = y^c_{k,i,m}, z^c_{m} = z^c_{k,i,m}; \forall i \in N'_i\}$, where $\mathcal{M}$ is a set of grouped corner indices. Then, to make a node with a high weight have a more significant effect on the model of its neighboring nodes, the height of each corner in $\mathcal{C}'_{m}$ is updated to $\hat{z}_{c_{m}}$, i.e., $\hat{c}_{m} = (\hat{x}_{c_{m}}, \hat{y}_{c_{m}}, \hat{z}_{c_{m}})$ through the following weighted average step:

$$\hat{z}_{c_{m}} = \frac{\sum_{i} \hat{z}_{c_{m}}(x^c_{k,i,m} \cdot w^T(n^T_{i})/||c^w_{k,i,m} - m||_{xy})}{\sum_{i} w^T(n^T_{i})/||c^w_{k,i,m} - m||_{xy}}. \tag{3}$$

Furthermore, every $n^T \in \mathcal{N}'$ surrounded by three $\hat{c}_{m}$, is corrected to a terrain node, $\mathcal{N}'_{i}$ with the following updated elements:

$$\hat{P} = [\hat{s} \quad \hat{d}], \quad \hat{m} = (\hat{c}_{1} + \hat{c}_{2} + \hat{c}_{3})/3$$

$$\hat{d} = -\hat{s} \cdot \hat{m}, \quad \hat{s} = (\hat{c}_{2} - \hat{c}_{1})/||\hat{c}_{2} - \hat{c}_{1}|| \times (\hat{c}_{3} - \hat{c}_{1}). \tag{4}$$

As shown in Fig. 6(b), from $\hat{P}$ for each $n^T_{i}$, the point cloud is segmented as follows:

$$\text{label}(p_k) = \begin{cases} 
\text{Terrain}, & \text{if } p_k \cdot \hat{s}_i + \hat{d}_i < \epsilon_3 \\
\text{Obstacle}, & \text{otherwise} 
\end{cases} \tag{5}$$

for a point $p_k \in \hat{P}_i$, where $\epsilon_3$ denotes the point-to-plane distance threshold.

IV. ABOVE-GROUND OBJECT SEGMENTATION

The above-ground object segmentation first processes a spherical projection of a point cloud $\mathcal{P}$ that has points $\text{label}(p) = \text{Obstacle}$. Then, it executes the horizontal and vertical update iteratively on each row in the projection space for the efficient computation time.

A. Spherical Projection

A spherical projection [18] can be helpful in capturing spatial adjacency of points by their angles of reflection, making the upcoming horizontal and vertical update steps simpler. The spherical projection of a point cloud arranges points by their azimuth and elevation angles of reflection by mapping into a $\mathbb{R}^2$ space with width $w$ and height $h$. Fig. 7 shows an example of spherically projected points.

Fig. 7. A single scan is projected to a spherical projection space by the azimuth ($\phi$) and elevation ($\theta$) angles of points. Each of the filled containers indicates a measured point. Note that the containers in the second ring are colored to better elucidate the concept of a node and an edge in Fig. 8.

B. Graph Representation

Similar to [15], horizontally adjacent points that are separated by a distance below a threshold $T_{horz}$ form a node $n^C = (idx_x, idx_y, label)$, where $idx_x$, $idx_y$, and $label$ indicate its local index of a starting point, an ending point, and its label, respectively. Since points are organized in the horizontal order in terms of their azimuth angles, registering local start and end indices of a node can be useful for subsequent clustering. In brief, a set of directed edges $E^C = \{e^C_{i,j}| i, j \in N'_0, \theta_i, \theta_j \in T\}$, each of which connects two nodes, and a set of nodes $N^C = \{n^C_{i,k}| \theta \in T, k \in N'_0\}$ constitute our graph structure $G^C$ for clustering, where $T = \{0, \ldots, h - 1\}$ and $N'_0$ is a list of node indices in a ring $\theta$.

C. Horizontal Update

The horizontal update step creates a node by clustering neighboring points within a horizontal merge threshold $T_{horz}$ along the direction of the local index. After the node-edge construction through the horizontal update, the second ring in Fig. 7 can be compactly depicted as in Fig. 8.

By efficiently utilizing the compact representation, the circular linkage and skipped linkage tests can prevent over-segmentation due to occlusion. The circular linkage test checks whether the horizontal distance between the point in $idx_x$ of the last node $n^C_{y,n-1}$ and that in $idx_x$ of the first node $n^C_{y,0}$ in the same ring $\theta$, falls below $T_{horz}$, where $n$ is the total number of nodes in $\theta$. This linkage check is necessary for preventing separation of points from a 3D LiDAR. The skipped linkage test merges non-neighboring nodes as long as these two nodes are located within $T_{horz}$. Likewise, the
skipped linkage test plays its critical role under a frequently occurring situation as in Fig. 9(a).

D. Vertical Update

Once a ring finishes the horizontal update, it goes through the vertical update; these two procedures are repeatedly executed on each ring. The chief novelty of our vertical update procedure lies in adopting binary search in finding overlaps and computing an edge distance \( D_v \) between two nodes efficiently. As shown in Fig. 10(a), the search space is first extended by the extension window size \( T_{\text{ext}} \) so that more possible edges are constructed, reducing chances of vertical separation as in Fig. 9(b). Then, at the previous ring, binary search is used to find the lower- and upper-bound nodes that have index-wise overlaps with the node \( n_i^{C_0} \) at the current ring. This step can create inter-ring linkages from the current node to all nodes in between the lower- and upper-bounds.

Fig. 10(b) demonstrates how the vertical distance \( D_v \) of an inter-ring edge is computed. In a naive way, \( D_v(\mathbf{e}_i^{C_0}, \mathbf{e}_j^{C_0}) \) is defined by the distance between the points in \( n_i^{1.0} \) and \( n_j^{C_0} \). However, the order of computing the distances becomes critical if a large number of points were to be clustered into a node. Since scan points that are collected at the same local index in multiple rings are more likely measured from one object, our algorithm starts computing the point distance from the overlapping index. In this way, our algorithm can retrace computational costs by saving a considerable number of redundant calculations.

V. EXPERIMENTS

TABLE I: Parameter setting for TRAVEL. Units of \( r^T \), \( \epsilon_3 \), \( T_{\text{horz}} \), and \( T_{\text{vert}} \) are in unit of \( m \), and \( \theta^T \) is in the unit of degree(°).

| Param   | \( r^T \) | \( \theta^T \) | \( \sigma^T \) | \( \epsilon_1 \) | \( \epsilon_2 \) | \( \epsilon_3 \) | \( T_{\text{horz}} \) | \( T_{\text{skip}} \) | \( T_{\text{ring}} \) | \( T_{\text{vert}} \) | \( T_{\text{ext}} \) |
|---------|----------|-------------|--------------|----------------|----------------|----------------|---------------------|-------------------|-------------------|-------------------|-------------------|
| Value   | 8        | 30°         | 0.1          | 0.03           | 0.1            | 0.3            | 10                  | 5                 | 0.5               | 100               |                   |

A. Dataset

To evaluate our proposed algorithm, we use the simulation, public urban scene dataset, and our own dataset. Both conventional and newly proposed evaluation metrics are used for quantitative evaluation. Table 1 shows the parameters that are used throughout the evaluation.

1) CARLA Simulation: In CARLA [5] simulation, the points in the scan have semantic labels. The points with the labels, such as RoadLane, Road, SideWalk, Ground, and Terrain, are considered to be the ground-truth terrain. Also, object identification, given by CARLA simulation, is used as ground-truth cluster label for the evaluation of above-ground object segmentation. This dataset is designed to test the algorithms in an urban environment with tunnels, slopes, buildings, and vehicles.

2) Semantic KITTI Dataset: To evaluate the terrain segmentation performance of our proposed method against other ground segmentation algorithms, we experimented with the SemanticKITTI dataset [6], which presents an urban scene in real world. Note that the points that are labeled as Lane marking, Road, Parking, Sidewalk, Other ground, and Terrain are considered to be the ground-truth terrain points. The ground-truth cluster for each of the above-ground data points is decided by vanilla Euclidean clustering (EC). This vanilla EC can cluster the points with the same class label and object identification, provided by Semantic KITTI dataset, as one object. Note that, for example, a tree in the dataset has points at its trunk labeled as Trunk and at its leaves as Vegetation, but our algorithm can only segment the tree as a single object. Due to such limitation of the dataset, the evaluation of our algorithm will have some inevitable errors.

3) Rough Terrain Dataset: Unlike the other two datasets, our own rough terrain dataset introduces a challenging environment with bumpy terrains and low-lying obstacles. The dataset was acquired via a mobile robot that roams around in the forest and on sidewalks and roads. This robot platform, Husky from Clearpath Robotics, is equipped with a 3D LiDAR sensor of Ouster OS0-128 and an IMU sensor of Xsens MTI-300.

B. Evaluation Metrics

The traversable ground segmentation performance of TRAVEL is evaluated and compared with other ground segmentation algorithms through the conventional metrics: precision (P), recall (R), accuracy, and F1-score. On the other hand, the above-ground object segmentation is evaluated by our newly proposed metrics: over-segmentation entropy (OSE) and under-segmentation entropy (USE).

Yang et al. [15] was the first to adopt quantitative measures to evaluate the segmentation performance on a labeled dataset. However, the authors only compared the ratio of the largest number of clustered points to the total number of points in a labeled object. This criterion is similar to our OSE but does not take into account the effect of multi-instance prediction. In addition, the authors did not suggest under-segmentation evaluation of any kind. Held et al. [24] and Hu et al. [25] suggested over-segmentation and under-segmentation error metrics by counting the number of occurrences for each ground-truth object. However, they fail to capture the distribution of the multi-class labels in the metrics. On the contrary, our two measures are effectively designed so that they can reflect the performance of multi-class prediction. OSE measures the confusion caused by two
or more cluster labels in one ground-truth object, whereas USE measures the confusion caused by two or more ground-truth labels inside one prediction cluster. Suppose $N$ points are measured from an object and, out of $N$ points, $N_i$ points are clustered as label $i$ such that $\sum_i (N_i/N) = 1$. The OSE is calculated by

$$\text{OSE} = -\sum_i \left( \frac{N_i}{N} \right) \log \left( \frac{N_i}{N} \right).$$

Similarly, suppose $M$ points are clustered as a cluster and, among $M$ points, $M_i$ points have the ground-truth label $i$ such that $\sum_i (M_i/M) = 1$. The USE is calculated by

$$\text{USE} = -\sum_i \left( \frac{M_i}{M} \right) \log \left( \frac{M_i}{M} \right).$$

Both entropies increase as more conflicting labels exist in an object or in a cluster but reach 0 if only one label exists. Fig. 11 illustrates a simple example of OSE and USE. The validity of the two suggested metrics can be seen in Table II as the vanilla EC, which is close to the ground-truth, results in lower OSE and USE. However, as the number of points grows, EC is proven to be impractical in computation time [9].

![Fig. 11. The example illustration of over-segmentation entropy (OSE) and under-segmentation entropy (USE).](image)

**TABLE II: Average USE and OSE on the sample CARLA dataset**

| Dataset                      | Sample CARLA dataset (50 frames) | USE (%) | OSE (%) | Time (ms) |
|------------------------------|----------------------------------|---------|---------|-----------|
| Vanilla EC                   | 0.94                             | 4.64    | 7.438   |
| Proposed                     | 17.68                            | 22.84   | 46      |

VI. RESULTS AND DISCUSSION

Our proposed algorithm is evaluated in two separate categories: traversable ground segmentation and above-ground object segmentation.

A. Parameter Studies

We first shed light on the effect of some important parameters on the performance using sequence 07 of Semantic KITTI dataset. This dataset is carefully chosen for the studies since it has the most varying urban conditions so that we can generalize the parameters. Figs. 12 - 14 show the effect of trinodal resolution $r_T$, ring-wise search window size $T_{ring}$, and $T_{ext}$, respectively. The other parameters are kept constant as in Table II during the experiment of each parameter.

![Fig. 12. The effect of $r_T$ on (a) precision, recall, (b) $F_1$-score, accuracy, and (c) computation time.](image)

![Fig. 13. The effect of $T_{ring}$ on (a) the under-segmentation entropy, (b) the over-segmentation entropy, and (c) computation time.](image)

![Fig. 14. The effect of $T_{ext}$ on (a) the under-segmentation entropy, (b) the over-segmentation entropy, and (c) computation time.](image)

1) The effect of $r_T$: With a small resolution, the performance of ours worsens. This is because the number of points included in each node is insufficient for planar modeling. On the other hand, there is no significant change in the performance at a resolution of 6 m or higher. Considering the computation time as well, we found that a resolution of 8 meter would suffice.

2) The effect of $T_{ring}$: A large $T_{ring}$ allows a larger search space for the vertical linkages between nodes. Since the computation time increases with an increasing $T_{ring}$, we have to carefully determine the optimal value. From Fig. 13, we figured that the values within 3 to 5 are reasonable for $T_{ring}$.

3) The effect of $T_{ext}$: $T_{ext}$ helps the algorithm extract more overlapping node candidates to merge vertically. The larger $T_{ext}$ is, the less likely the vertical separation will occur. As shown in Fig. 14, thanks to the binary search explained in Section IV-D, the increased $T_{ext}$ does not significantly increase the computation burden. However, the performance of the segmentation saturates or worsens after 100. As a result, we chose 100 as an optimal value for $T_{ext}$.

B. Traversable Ground Segmentation

Table III shows that our proposed algorithm demonstrates the highest $F_1$-score and accuracy with the lowest perturbation and computation time in both datasets. Also, to further examine the robustness of the proposed algorithm, three other algorithms, which show high performance on both datasets, are compared on the bumpy terrain with a myriad of low-lying obstacles and trees as shown in Fig. 15.

Our ground segmentation considering the traversability allows not only to model the bumpy terrains but also to extract the low-lying obstacles such as low stairs and vegetation. On the other hand, since Ground Plane Fitting (GPF) [9]...
divides areas only in the moving direction, the estimated ground tends to converge to a local minimum on the uneven terrain. Patchwork [10] shows relatively decent segmentation results on ground modeling by taking into account the characteristics of LiDAR data distribution. Nevertheless, its lack of interaction between the bins causes the discontinuity of the planar models, thus making it difficult to extract the low-lying obstacles from the ground. As a result, our design of TGF, which focuses on the geometric relationship between terrain elements via a graph structure, assists in the accurate modeling of varying-sloped planes and makes our algorithm robust to both urban and wild environments.

TABLE III: Quantitative comparison for the traversable ground segmentation on CARLA simulator and whole sequences of the Semantic KITTI dataset. Units are ms for T (computation time) and %, otherwise. \(\mu\) and \(\sigma\) are the mean and stdv. of each metric, respectively.

| Metrics | Dataset | \(P\) | \(R\) | \(\mu\) | \(\mu\) | \(\sigma\) | \(\sigma\) | \(\mu\) | \(\mu\) |
|---------|---------|-------|-------|-------|-------|-------|-------|-------|-------|
| RANSAC | CARLA   | 88.3  | 92.1  | 89.0  | 20.0  | 88.5  | 21.9  | 67    |
| Zermas et al. | 97.5  | 92.7  | 94.3  | 10.3  | 95.0  | 7.9   | 20    |
| Narkisi et al. | 97.7  | 83.3  | 89.6  | 8.2   | 60.7  | 8.8   | 19    |
| Lim et al. | 94.2  | 95.6  | 94.8  | 3.9   | 93.6  | 3.8   | 28    |
| Proposed w/o TTMF | 95.8  | 97.0  | 96.3  | 2.8   | 93.4  | 2.9   | 14    |
| Proposed     | 96.6  | 97.2  | 96.7  | 2.7   | 95.9  | 2.5   | 14    |
| Proposed     | 96.6  | 97.2  | 96.7  | 2.7   | 95.9  | 2.5   | 14    |

C. Above-Ground Object Segmentation

Table IV shows that our proposed method demonstrates the lowest USE and OSE in both datasets. As illustrated in Fig. 16, the pole-like object, which is prone to vertical separation, is well clustered without over-segmentation by our inter-ring linkages. Also, the horizontal separations of the walls by occlusion are largely prevented by our skipping linkages. Note that TGS (ours) + Scan Line Run (SLR) [9] increases the performance significantly compared with GPF + SLR. This performance boost implies that our traversable ground segmentation outperforms GPF by a large margin.

TABLE IV: Quantitative comparison for the above-ground object segmentation on CARLA simulator and whole sequences of the Semantic KITTI dataset.

| Metrics | Dataset | USE | OSE | T |
|---------|---------|-----|-----|---|
|         |         | \(\mu\) | \(\mu\) | \(\sigma\) | \(\sigma\) | \(\mu\) | \(\mu\) |
|         | CARLA   |       |       |     |     |     |     |
| GPF + SLR | 67.01  | 64.31 | 315.01 | 2.5 |
| TGS + SLR | 37.59  | 28.28 | 230.48 | 8.9 |
| Proposed | 7.86   | 4.99  | 31.71  | 15.88 |
|         | Semantic KITTI |       |       |     |     |     |     |
| GPF + SLR | 117.08 | 127.40 | 301.16 | 114.29 |
| TGS + SLR | 45.33  | 22.14 | 228.53 | 86.13 |
| Proposed | 24.07  | 11.88 | 70.40  | 34.44 |

In conclusion, our algorithm notably outperforms the state-of-the-art algorithm at the cost of time complexity while still maintaining real-time performance by about 20-30 Hz.

VII. CONCLUSIONS

In this study, we propose a two-step segmentation of traversable ground and above-ground objects, TRA.VEL. Our node-edge representation of a point cloud allows accurate modeling of the ground and efficient searching of neighboring points. Our traversable ground segmentation outperforms the prior studies in terms of conventional evaluation metrics. Also, our object segmentation brings about less under-segmentation and over-segmentation, which are assessed using our newly proposed metrics. In essence, unlike learning-based methods, our algorithm cannot assign the same class label, for instance, to two spatially distant walls separated by a large object; \(T_{\text{skip}}\) mitigates separation of the walls caused by a thin object, yet \(T_{\text{base}}\) may limit the merging of them. Due to this limitation, our work is concerned more with real-time navigation by spotting relevant targets under unseen environments, than with classifying targets with specific class labels. As future works, we would like to apply the proposed algorithm to the navigation task, tracking the relevant objects to identify and remove dynamic motion from a scene.
Fig. 16. Qualitative comparisons on CARLA simulator, sequence 05 and 07 of Semantic KITTI datasets, and rough terrain dataset in the order from left to right column. The estimated ground points are shown in grey, and the segmented objects are represented by non-grey colors. TRAVEL overcomes the vertical separation problem on the pole as in the first column and the horizontal separation problem of the wall as in the third column.

REFERENCES

[1] Y. Pan, P. Xiao, Y. He, Z. Shao, and Z. Li, “MULLS: Versatile LiDAR SLAM via multi-metric linear least square,” in Proc. IEEE Int. Conf. on Robot. and Automat., 2021, pp. 11 633–11 640.

[2] C. Sung, S. Jeon, H. Lim, and H. Myung, “What if there was no revisit? Large-scale graph-based SLAM with traffic sign detection in an HD map using LiDAR inertial odometry,” Intell. Service Robot., pp. 1–10, 2021.

[3] M. Arora, L. Wiesmann, X. Chen, and C. Stachniss, “Mapping the static parts of dynamic scenes from 3D LiDAR point clouds exploiting ground segmentation,” in Proc. European Conf. on Mobile Robots (ECMR), 2021, pp. 1–6.

[4] D. Yoon, T. Tang, and T. Barfoot, “Mapless online detection of dynamic objects in 3D LiDAR,” in Proc. Conf. on Comput. and Robot Vis. (ICCV), 2019, pp. 113–120.

[5] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, “CARLA: An open urban driving simulator,” in Proc. Conf. on Robot Learning, 2017, pp. 1–16.

[6] 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 Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), 2019, pp. 9297–9307.

[7] M. A. Fischler and R. C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” Commun. ACM, vol. 24, no. 6, pp. 381–395, 1981.

[8] P. Narasri, E. Takeuchi, Y. Ninomiya, Y. Morales, N. Akai, and N. Kawaguchi, “A slope-robust cascaded ground segmentation in 3D point cloud for autonomous vehicles,” in Proc. IEEE Int. Conf. on Intell. Transport. Syst. (ITSC), 2018, pp. 497–504.

[9] D. Zermas, I. Izzat, and N. Papanikolopoulos, “Fast segmentation of 3D point clouds: A paradigm on LiDAR data for autonomous vehicle applications,” in Proc. IEEE Int. Conf. Robot. Automat., 2017, pp. 5067–5073.

[10] H. Lim, M. Oh, and H. Myung, “Patchwork: Concentric zone-based region-wise ground segmentation with ground likelihood estimation using a 3D LiDAR sensor,” IEEE Robot. Automat. Lett., vol. 6, no. 4, pp. 6458–6465, 2021.

[11] H. Xue, H. Fu, R. Ren, J. Zhang, B. Liu, Y. Fan, and B. Dai, “LiDAR-based drivable region detection for autonomous driving,” in Proc. IEEE/RSJ Int. Conf. on Intell. Robots and Syst., 2021, pp. 1110–1116.

[12] F. Mooosmann, O. Pink, and C. Stiller, “Segmentation of 3D LiDAR data in non-flat urban environments using a local convexity criterion,” in Proc. IEEE Intell. Veh. Symp. (IVS), 2009, pp. 215–220.

[13] I. Bogoslavskyi and C. Stachniss, “Fast range image-based segmentation of sparse 3D laser scans for online operation,” in Proc. IEEE/RSJ Int. Conf. on Intell. Robots and Syst., 2016, pp. 163–169.

[14] P. Burger, B. Naujoks, and H. Wunsch, “Fast dual decomposition based mesh-graph clustering for point clouds,” in Proc. Int. Conf. on Intell. Transport. Syst. (ITSc), 2018, pp. 1129–1135.

[15] H. Yang, Z. Wang, L. Lin, H. Liang, W. Huang, and F. Xu, “Two-layer-graph clustering for real-time 3D LiDAR point cloud segmentation,” Appl. Sci., vol. 10, no. 23, p. 8534, 2020.

[16] F. Zhang, C. Guan, J. Fang, S. Bai, R. Yang, P. H. Torr, and V. Prisacariu, “Instance segmentation of lidar point clouds,” in Proc. IEEE Int. Conf. on Robot. and Automat., 2020, pp. 9448–9455.

[17] K. Wong, S. Wang, M. Ren, M. Liang, and R. Urtasun, “Identifying unknown instances for autonomous driving,” in Conference on Robot Learning, PMLR, 2020, pp. 384–393.

[18] A. Milioto, I. Vizzo, J. Behley, and C. Stachniss, “RangeNet++: Fast and accurate LiDAR semantic segmentation,” in Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst., 2019, pp. 4213–4220.

[19] J. Xu, R. Zhang, J. Dou, Y. Zha, J. Sun, and S. Pu, “Rpynet: A deep and efficient range-point-voxel fusion network for lidar point cloud segmentation,” in Proc. IEEE/CVF Int. Conf. Comput. Vis., 2021, pp. 16 024–16 033.

[20] K. Peng, J. Fei, K. Yang, A. Roitberg, J. Zhang, F. Bieder, P. Heidenreich, C. Stiller, and R. Stiefelhagen, “Mass: Multi-attentional semantic segmentation of lidar data for dense top-view understanding,” Proc. IEEE Transac. on Intell. Transport. Syst., 2022.

[21] A. Paigwar, Ö. Erkent, D. Sierra-Gonzalez, and C. Laugier, “Gndnet: Fast ground plane estimation and point cloud segmentation for autonomous vehicles,” in Proc. IEEE/RSJ Int. Conf. on Intell. Robots and Syst., 2020, pp. 2150–2156.

[22] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and D. Rus, “LIO-SAM: Tightly-coupled LiDAR inertial odometry via smoothing and mapping,” in Proc. IEEE/RSJ Int. Conf. on Intell. Robots and Syst., 2020, pp. 5135–5142.

[23] M. Weinmann, B. Jutzi, S. Hinz, and C. Mallet, “Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers,” ISPRS Journal of Photo. and Remote Sens., vol. 105, pp. 286–304, 2015.

[24] D. Held, D. Guilleo, B. Rebsamen, S. Thrun, and S. Savarese, “A probabilistic framework for real-time 3D segmentation using spatial, temporal, and semantic cues.” in Robotics: Science and Systems, vol. 12, 2016.

[25] P. Hu, D. Held, and D. Ramanan, “Learning to optimally segment point clouds,” IEEE Robot. Automat. Lett., vol. 5, no. 2, pp. 875–882, 2020.