doi:10.3311/BMEZalaZONE2022-004

3D Object Detection in LIDAR Point Cloud Based on Background Subtraction

Szabolcs Nagy[1a], András Rövid[1]

[1] Department of Automotive Technologies, Faculty of Transportation Engineering and Vehicle Engineering, Budapest University of Technology and Economics

[a] Corresponding author: nagy.szabolcs@kjk.bme.hu

Abstract

Autonomous vehicles have a key role in transportation systems of the future, but there are still many difficulties to overcome. Nowadays one of the most critical problems in autonomous driving is the precise and robust detection of traffic participants. This paper presents a LIDAR-based 3D object detection method. The algorithm uses HD Map to subtract the static background points from the LIDAR point cloud. The remaining points are grouped by clustering, then 3D boxes are fitted to the clusters. The object detection method presented in this paper was tested on real sensor data collected by a solid-state LIDAR on the highway module of the ZalaZONE proving ground. The results showed that the developed algorithm performs as intended in a highway scenario, detecting vehicles even more than 100 meters away from the sensor by a framerate of ~20FPS.

Keywords: *background subtraction, HD map, LIDAR point cloud, object detection*

1 Introduction

One of the greatest challenges for self-driving vehicles is environmental perception, because autonomous vehicles rely on information about their surroundings acquired from perception systems. Thus, it is crucial to detect the presence of traffic participants like vehicles, pedestrians, and other elements. LIDAR-based perception systems have been introduced due to the need for reliable and accurate measurement capabilities [1]. The LIDAR is an active sensor which operates by emitting laser and measuring the time for the reflected light to return. The range is estimated based on the elapsed time between the transmitted and received signals, resulting in a 3D point cloud representing the surrounding environment. Many high-level perception systems use LIDARs to complement the lack of depth information in 2D image data captured by a digital camera [2].

HD maps are highly accurate maps used for autonomous driving purposes, at centimeter-level. High-definition maps provide information about the surrounding environment, including map elements as road shape, markings, traffic signs while maintaining high accuracy, thus high-definition maps are becoming a key technology for autonomous driving systems, although they are typically employed for motion planning applications [3]. However, there are already object detection solutions [4] that extract geometric and semantic features from the HD maps to improve performance and robustness.

Lidar-based environment perception algorithms have gone through increasing development in recent years. Based on the approach, the most object detection algorithms can be divided into two main groups: classic machine vision-based methods such as [5] using occupancy grid for object segmentation and techniques based on machine learning as PointPillars [6] adopting 2D convolutional layers to learn point cloud features in order to generate 3D bounding boxes for different object classes.

The aim of this paper is to present a traditional object detection method that only uses 3D point cloud from LIDAR as sensor data and benefits from the information contained in a high-definition map. The algorithm uses HD Map to subtract the background points from the LIDAR point cloud, then the remaining points are projected to the road surface. The projected point cloud is grouped into clusters, then to the clustered points 3D boxes are fitted by constructing convex hulls and minimum-area rectangle.

The method has been tested on real sensor data collected by a solid-state LIDAR installed as infrastructure sensor on the highway module of the ZalaZONE proving ground, which HD map is also available. Another important consideration regarding the performance of the proposed object detector was its applicability in real time applications.

2 Methodology

2.1 Calibration

The measurements were taken at the highway platform of the ZalaZONE proving ground, where the LIDAR was installed as infrastructure sensor in the middle of a bridge over the highway, facing the lanes as shown by Fig. 1.

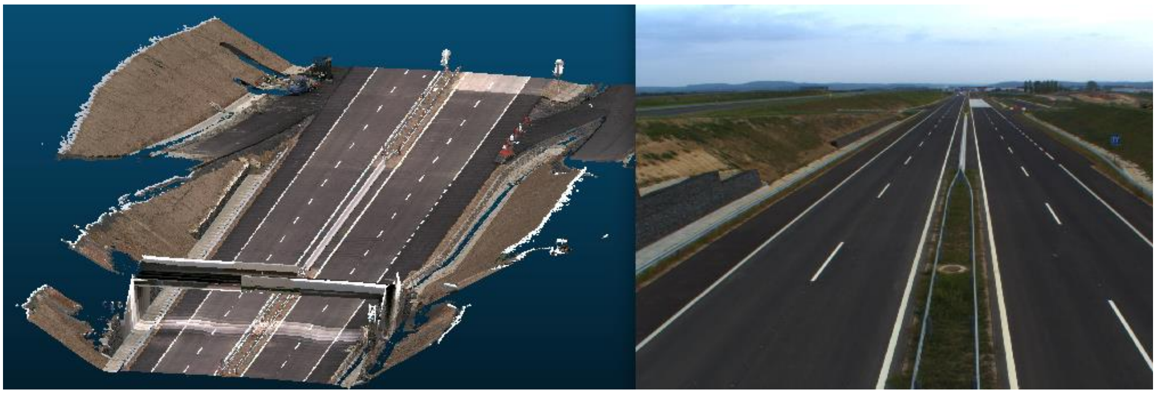

Fig. 1 *High-definition map of the highway platform (left). The LIDAR was installed in the middle facing the lanes*

Since only one side of the highway was used for the measurements, the not relevant parts of the HD map were removed manually, which resulted a 3D point cloud representing only the lanes being used. In order to reduce the complexity of the approach, the remaining point cloud was uniformly subsampled, followed by a transformation to a reference coordinate system which in our case was the UTM (Universal Transverse Mercator) frame. For estimating the rotation and translation parameters between the two coordinate systems the iterative closest point (ICP) algorithm [7] was applied.

2.2 Segmentation

Since the highway stretch used for the measurements can be approximated well as a plane, thus the equation of the plane is then calculated by using the least-squares method to solve the over-determined linear matrix equation. The segmentation task began with the removal of the LIDAR points that are on either side of the road. To achieve that, the point cloud was aligned with the XY plane of the UTM coordinate system. By projecting the transformed points to the XY-plane, we reduced the problem to 2D.

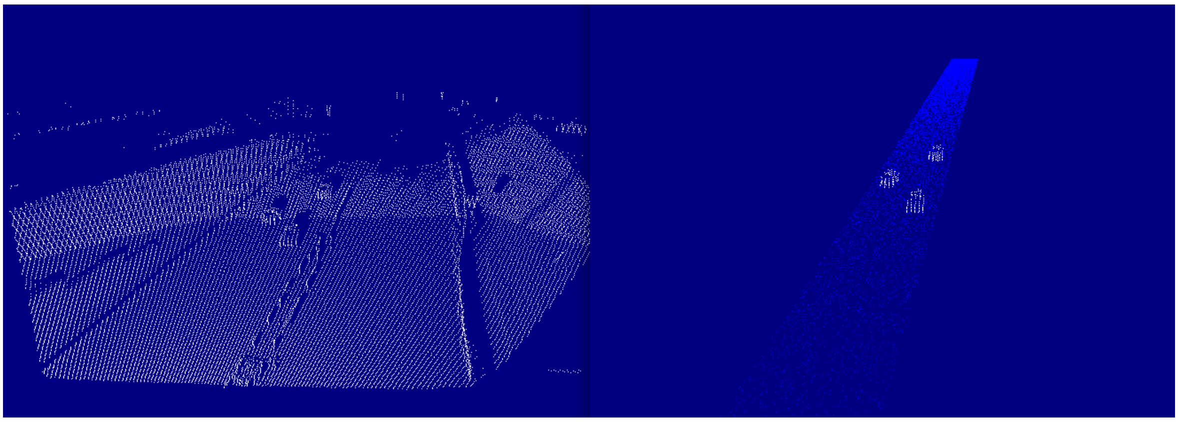

Fig. 2 *Original LIDAR point cloud (left) and the segmented point cloud with the plane from the HD map approximating the highway stretch (right)*

The remaining points were located either below or above the road surface. In the following step the perpendicular distances - of all kept 3D points - from the approximated plane were calculated. Then, based on the calculated distances, the points closer than a lower threshold and farther than an upper threshold were omitted. As Fig. 2. shows, the remaining point cloud contains only points that are reflected from objects on the highway stretch.

2.3 Fitting bounding boxes

To fit 3D bounding boxes to the objects, the point cloud is first clustered. Before the clustering the 3D point cloud was arranged into a k-d tree representation, then by using the Euclidean Cluster Extraction method of the Point Cloud Library [8] the clusters were determined, each representing an object. By setting the minimum cluster size to 3 points, the outliers were filtered out. The next step was to fit 3D bounding boxes to the clusters of points. For each cluster, a 3D box is fitted based on the minimum area rectangle (MAR), containing points that were projected to the plane representing the highway section. This was achieved by determining the convex hull of the projected points, thus one of the edges of the MAR will coincide with the corresponding edge of the convex hull. Then for each edge of the convex hull, the corresponding minimum area rectangle was determined. Among all the possible MARs, the rectangle with the smallest area is selected as the bounding box of the 2D point cloud as shown by Fig. 3. For each cluster, the point having the largest distance to the plane of the highway stretch will be considered as the height of the 3D bounding box.

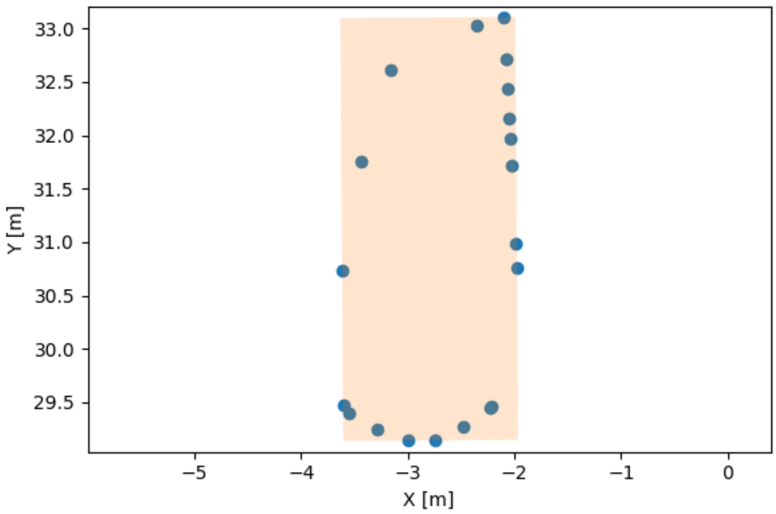

Fig. 3 *The minimum area rectangle fitted to the convex hull of a projected point cloud*

At longer ranges, where only few points are reflected from the objects, the shape and size of the bounding box fitted to the detected object may vary significantly. In order to reduce this effect, points from previous *k* frames (corresponding to the same object) have been aggregated.

3 Results

The 3D object detection algorithm presented in this paper was tested on real sensor data collected by a solid-state LIDAR on the highway module of the ZalaZONE proving ground. Several scenarios were examined with multiple vehicles driving at different speeds. As shown by Fig. 4, the algorithm can detect the moving vehicles as intended, while also maintaining the framerate of ~20 FPS. The first detections occurred at the distance of around 150 meters, but the size and orientation of the bounding box fitted to the detected object depends strongly on the number of points used, thus the presented method performs better in close range. Setting *k* to a small value (e.g. *k*=3) improved the bounding box fitting remarkably, on the other hand when taking too many frames into account the heading estimation degrades when the object changes direction. As future work, tracker will be incorporated into the algorithm in order to get better estimations on the orientation and position of the 3D bounding box. In addition, the number of points the vehicle is represented with, maybe used to emphasize the importance of clusters.

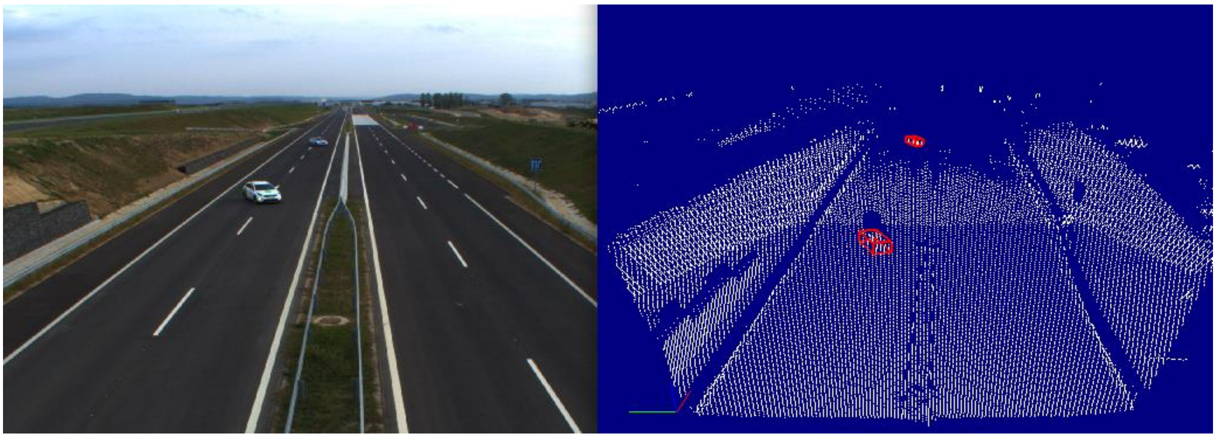

Fig. 4 *The 3D bounding boxes in the LIDAR point cloud. Vehicles were moving on a sine-trajectory*

4 Conclusion

The presented object detection algorithm can create 3D bounding boxes in LIDAR point cloud data by relying on high-definition maps as backbone for background subtraction. Although most object detection algorithms operate on a deep learning basis, this approach only uses classical methods. DL based solutions generally depend on supervised learning requiring a huge amount of labeled data. The trained network will perform as expected only when the setup of sensors doesn't change (height, angle). For example, LIDARs generate different patterns of objects from different positions, thus by changing the mounting height of the sensor may strongly affect the detection performance. This traditional solution can be applied for any LIDAR sensor setup, especially in a highway environment.

Acknowledgement

The research reported in this paper and carried out at the Budapest University of Technology and Economics has been supported by the National Research Development and Innovation Fund (TKP2020 National Challenges Subprogram, Grant No. BME-NC) based on the charter of bolster issued by the National Research Development and Innovation Office under the auspices of the Ministry for Innovation and Technology.

References

[1] Y. Li and J. Ibanez-Guzman (2020). Lidar for Autonomous Driving: The Principles, Challenges, and Trends for Automotive Lidar and Perception Systems, in IEEE Signal Processing Magazine, vol. 37, no. 4, pp. 50-61, doi: 10.1109/MSP.2020.2973615.

[2] Tihanyi V, et al. (2021). Motorway Measurement Campaign to Support R&D Activities in the Field of Automated Driving Technologies. Sensors. 21(6):2169. https://doi.org/10.3390/s21062169

[3] Seif, Heiko & Hu, Xiaolong. (2016). Autonomous Driving in the iCity—HD Maps as a Key Challenge of the Automotive Industry. Engineering. 2. 159-162. 10.1016/J.ENG.2016.02.010.

[4] Yang, B., Liang, M., & Urtasun, R. (2018). HDNET: Exploiting HD Maps for 3D Object Detection. ArXiv, abs/2012.11704.

[5] Himmelsbach, M & Müller, A & Luettel, Thorsten & Wuensche, Hans-Joachim. (2008). LIDAR-based 3D object perception.

[6] Lang, Alex & Vora, Sourabh & Caesar, Holger & Zhou, Lubing & Yang, Jiong & Beijbom, Oscar. (2019). PointPillars: Fast Encoders for Object Detection From Point Clouds. 12689-12697. 10.1109/CVPR.2019.01298.

[7] Bouaziz, S., Tagliasacchi, A., Pauly, M., (2013). Sparse iterative closest point. In Computer graphics forum, vol. 32, no. 5, pp. 113-123.Oxford, UK: Blackwell Publishing Ltd.

[8] R. B. Rusu and S. Cousins (2011). 3D is here: Point Cloud Library (PCL), 2011 IEEE International Conference on Robotics and Automation, pp. 1-4, doi: 10.1109/ICRA.2011.5980567.