Enhanced Clustering Method using 3D Laser Range Data for an Autonomous Vehicle

KUK CHO, SEUNGHO BAEG, and SANGDEOK PARK
Korea Institute of Industrial Technology
Department of Robot Research & Business Development
1271-18, Sa-1dong, Sangruk-gu, Ansan
Republic of Korea
gogi33, shbaeg, sdpark@kitech.re.kr

Abstract: This paper describes an enhanced clustering method for 3D point cloud data which is acquired from a laser range scanning system. The proposed method overcomes the instinctive problems of the vertical or horizontal scanning system. The acquired 3D laser range data for an autonomous vehicle has a disadvantage in that the point cloud data from the system is not equally distributed based on distance. Moreover, the data information describes only the object’s shown surface. The proposed method uses essentially a radially bounded nearest neighbor approach and describes the definition of nearest neighbor points on the point cloud data from the scanning system. The method uses a Gaussian-based approach to find nearest neighbor points during the clustering process. In conclusion, the experimental results for the clustering and segmentation are shown with real data.

Key–Words: 3D point cloud data, clustering, segmentation, autonomous vehicle, object recognition.

1 Introduction

The current foremost research objective of autonomous vehicle technology is to support and assist human driving. This technology should become more reliable and accurate to achieve this goal. One important issue is the creation of a sophisticated perception algorithm which undertakes object recognition and environmental understanding in complex environments [1, 2, 3]. There have been many studies in which various sensors were assessed in terms of their environment recognition capabilities. Similarly, there are a number of devices that recognize objects as a means of determining human perception, including image sensors, ultrasonic sensors, RADAR sensors, 2D LRF sensors, and 3D LIDAR sensors. The laser sensors known as LRF (Laser Range Finder) and LADAR (Laser Detection and Ranging) are among the primary sensors used for object recognition in various situations. Laser sensor-based approaches include a feature-based method and a grid-based method. The feature-based method uses geometrical features such as lines, boxes and circles [4, 5]. The grid-based method use occupancy grid maps [6]. Traditional classification techniques need a considerable amount of information for success with two sequential processes: feature extraction and classification. The methods work using multi-layer perception (MLP), support vector machine (SVM), and adaptive boosting (Adaboost) [7, 8, 9, 10]. One approach uses a LRF sensor to process one scan layer for object detection. Point cloud data with a 3D scanning sensor provides reliable object information with multiple scan layers. Over the last decade, the laser range finders has leading indoor and outdoor mobile robotics to a new level of object detection and environment understanding. Localization, map building, obstacle avoidance, obstacle detection, classification and tracking are nowadays almost exclusively performed based on 2D and 3D range data. The amount of data obtained from either 3D sensors or a combination of 2D LRF and a scanning mechanism is generally several orders of magnitude laser than that obtainable from a simple 2D scan.

In this paper, an enhanced clustering method for 3D laser range data is proposed. We use an agglomerative nearest-neighbor clustering algorithm to segment the acquired point cloud data to separate a basic unit for further processing. The obtained clusters and distinctive segment may then be processed by additional classification algorithms. The structure of this paper is as follows: In section 2, the motivation of this work is briefly explained. Section 3 briefly touches on previous methods and explains the proposed method. Section 4 describes the experimental results using the proposed method in detail. Finally, section 5 concludes the paper and gives an outlook on future work.
2Motivation

The objective of this study is to develop an enhanced object recognition system using point cloud data from either a 3D laser range sensor or a 2D laser range finder with a scanning mechanism. Before the high-level processing steps such as classification and tracking, the data should be divided into the basic units to be processed. It involves a preprocessing step which includes clustering and segmentation, through the clustering and segmentation steps require improvement. In the proposed system, there are 3D laser range data which is acquired using a 3D sensor provided with 3D scanning data and a 3D combined sensor which includes a 2D LRF with a scanning mechanism. The 3D laser range sensor provides 3D point cloud data which describes the width, height and depth characteristics of the 3D coordinates of x, y, and z. The 3D laser range sensor extracts a more suitable ROI compared to the results of the LRF approach. Nevertheless, 3D laser range sensor data is associated with an irregular point distribution due to inherent problems with these types of sensors, which do not consider the distance and the shape. Even before the challenge of classifying the information from the laser range sensor, we are faced with the more fundamental problem of 3D segmentation. Thus, in this paper, the focus is on a method which segments 3D point cloud data robustly.

For an outdoor autonomous vehicle, 3D laser range data which is acquired by a laser range finder with a scanning mechanism or a 3D scanning sensor such as a Velodyne sensors (HDL-64e and HDL-32e). For the further high-level processing, 3D point cloud data should be divided into a unit of data processing unit called a segment through a clustering process. There are several weaknesses. First, there are instinct problems related to scanning points based on the acquisition mechanism. For examples, for a composition with a LRF sensor and a scanning mechanism, the scanning frequency determines the vertical resolution. The vertical data will be sparse with a slow scanning frequency and dense with a rapid scanning frequency. Well-known scanning sensors by Velodyne also have similar problems stemming from their laser component assembly. Second, the vertical and horizontal resolutions are different. With the combination of the LRF and a scanning mechanism, the mechanism is scanned vertically or horizontally depending on the application [11, 12, 13]. The scanning points of the horizontal layer of the LRF have a dense distribution; i.e., the horizontal points are close by but the vertical points are far away. Additionally, we assume that all of the scanning points are mainly unrelated to the shape of the object. Therefore, the condition with three different coordinate conditions is not feasible.

Fig. 1 shows an example of a scanning scene. A scene is shown in which a car is scanned using a 3D scanner with either a 3D sensor or a 2D sensor with a tilting mechanism sensor. The yellow points indicate the reflections of the laser points, such as laser targets. Fig. 1(a) depicts the scanning points of the front of a car from a scanning sensor installed on the vehicle, and fig. 1(b) illustrates only the object’s scanning points. As shown in the figures, the horizontal distribution of the acquired points is dense but the vertical distribution of the points is sparse. Fig. 2 depicts an illustration of a human using point cloud data. As shown the figure, the horizontal resolution is dense but the vertical resolution is sparse. The measurement points of the nine layers describe the human. Here, the difference in the vertical and horizontal resolution depends on the distance to the object. However, the traditional Euclidean approach is not feasible for distinguishing nearest-neighbor points. This paper pro-
poses a method to overcome this certain condition. It will be described as an upcoming section.

3 The Proposed Method

This section briefly describes a previous clustering method which uses point cloud data which also introducing the proposed approach. Klasing et al. uses essentially a radially bounded nearest neighbor (RBNN) approach [14], while Cho et al. applies an efficient segmentation method (RBNN-2) to detect an object [15]. The former is a point-based description and the latter is a cluster-based description. The proposed algorithm is based of two previous methods. It suggests the definition of nearest-neighboring-points for laser range data for a 3D scanning system. The acquired point cloud data $P$ is represented with the ground points $P_g$ and object points $P_o$. The object points require the partitioning of one unit, such as a segment, for further processing. The clustering step works with only object points $P_o$ to detect objects. Previously, the four parameters needed for the clustering and segmentation processes are as follows: the radius range of neighbor points ($r_n$), the minimum number of neighboring points in the radius ($\nabla_r$), the minimum number of points ($\nabla_P$) and the number of layers in one cluster ($\nabla_{\ell}$). Two parameters, i.e., the radius range of the neighbor points ($r_n$) and the number of neighboring points in the radius ($N_n$), are effective when used to connect each point and remove noisy data. In addition, the RBNN-2 algorithm has two additional criteria for object segmentation: the minimum number of points ($\nabla_P$) and the minimum number of layers in one cluster ($\nabla_{\ell}$). Given $N$, the total number of point cloud data instances for one frame scene, let the $i$-number of points of point cloud data be $\mathcal{P}_i$, $i = 1, \ldots, N$. In this case, the measurement point of one frame consists of ground points and object points, $P = P_g \cup P_o$.

In this algorithm, $\mathcal{Y}^0_i$ is the searched neighbor non-clustered point set in the $\mathcal{P}_i$-centered radius ($r$). $\mathcal{Y}^-_i$ is the searched neighbor clustered point set, which is connected to the nearest neighbor point of the $\mathcal{P}_i$-centered radius. $\mathcal{Y}^+_i$ is the newly established cluster point set with the $\mathcal{P}_i$ point. $\mathcal{P}_i$ denotes the center points of the neighbor range. $\Psi$ denotes the neighboring points of the $\mathcal{P}_i$-centered $r_n$ radius. Algorithm 1 shows the pseudo-code with four parameters. The clusters do not comprise the full composition of object points; each cluster can be a partial object, an instance of miss-clustering, or an instance of miss-discrimination from the previous step such as a ground removal step. There are common situations in data processing. Additional criteria are needed for accurate segmentation to an object for further high-level processing, such as case 4 in algorithm 1. In addition, the use of the number of the neighboring points had the effect of removing noise data using $\nabla_P$. The definition of the nearest neighboring points can be found in one point. The Euclidean distance is the approach normally used to estimate the distances between points. The laser range points from the scanning system are not uniformly distributed. The 3D Gaussian distribution value shows the index of the nearest neighboring points.

Data: object points($P^o, r_n, \nabla^o_r, \nabla^o_P, \nabla^o_{\ell}$)

Result: clustered objects($\mathcal{Y}^+$)

Initialization parameter $j$;

for $i \leftarrow P^o$ do

if $\neg (P_i \in \mathcal{Y})$ then

$(\mathcal{Y}^0_i, \mathcal{Y}^-_i, \Psi) \leftarrow GetNeighbor(P^o, P_i, r_n)$;

if $\mathcal{Y}^-_i = \phi$ then

$\mathcal{Y}^+_j \leftarrow CreateNewCluster(\mathcal{Y}^0_i)$;

$j \leftarrow j + 1$; /* case 1 */

else if $\mathcal{Y}^-_i \neq \phi$ then

if $NumberOf(\mathcal{Y}^0_i) > \nabla^o_P$ then

$\mathcal{Y}^+_j \leftarrow \mathcal{Y}^+_j \cup \mathcal{Y}^0_i$;

/* case 2 */

forall the $\mathcal{Y}^-_i$ do

$\mathcal{Y}^+_j \leftarrow \mathcal{Y}^+_j \cup \mathcal{Y}^-_i$;

/* case 3 */

end

end

end

forall the $\mathcal{Y}^+_j \in Clusters$ do

if $\mathcal{Y}^+_j < \nabla^o_{\ell}$ or $\mathcal{Y}^+_j < \nabla^o_{\ell}$ then

delete $\mathcal{Y}^+_j$;

/* case 4 */

end

end

Algorithm 1: The proposed object clustering and segmentation procedure

Here, $R$ denotes the parameters of the three-dimensional Gaussian distribution. Each values shows the distribution of the coordinates.
\[ \chi = \sqrt{(p - m)^T R^{-1} (p - m)} < \gamma \quad (2) \]

In this equation, \( m \) is the center point, which is not included in the cluster, for in case 1 in algorithm 1. Also, \( p \) denotes the candidate neighboring points. \( \chi \) is the index which determines whether or not the point is a neighboring point. For a given point, if \( \chi \) is less than \( \gamma \), the point is determined with its nearest points.

A different effect arises with the Euclidean distance. For an example of a 2D plane example, there are two points \( P_1(2,5) \) and \( P_2(5,2) \) under the center point \( (5,5) \) with a 2-D gaussian distribution \([1 \ 0; 0 \ 5]\). The Euclidean distance is identical in the two cases, but the g-RBNN shows different results. These results are 3 and 1.3416, respectively.

Fig. 3 depicts the first step and the second step of the two clustering methods and shows the final clustering results with given the same data. Figs. 3(a)-3(d) and figs. 3(e)-3(h) show corresponding processing illustrations of each step. The gray circles depicts the boundary range of the nearest-neighbor. The red circles denote the center point which crates a cluster with the nearest-neighbor points within a certain range. The black circles show the same cluster. Figs. 3(a), 3(b) and 3(c) depicts the previous algorithm. Figs. 3(e), 3(f) and 3(g) shows the illustrations using the proposed nearest-neighbor points definition. Figs. 3(d) and 3(h) depicts the clustering results, and Fig. 3(h) shows the improved performance as compared to that in Fig. 3(d).

4 Experimental Results

The proposed method was evaluated on a collected dataset obtained from a 3D LIDAR sensor (Velodyne HDL-32E) mounted on an experimental all-terrain vehicle in a forest environment. The sensor has 32 vertical scanning lines and provides 3D range data with a horizontal scan of 360. This section presents the results from the ground-object discrimination and feature extractions that were performed. Furthermore, a classifier was implemented based on the acquired 3D data using a previously introduced algorithm. The test environment was based on the following scenario: in a forested environment, an unmanned ground vehicle (UGV) with an installed 3D LADAR system moves and distinguishes objects as humans, trees, or other objects. The test uses a system with a 3D LADAR sensor, GPS, INS, and an odometry encoder installed in the UGV. The test environment was based on an Intel Core i7 CPU x980 with 3.3GHz and 24GB RAM. The overall system algorithm was optimized using IPP. The 3D LADAR sensor used HDL-32E (Velodyne, Inc.) and the data update rate was 100ms. The performance of the first ground and object discrimination stage required 40-50ms. The second stage of object segmentation and feature extraction required 15-20ms. The overall system process managed real-time data processing within 60-70ms.

The proposed method is evaluated with the KITTI dataset [3]. The KITTI dataset was obtained from real-world traffic situations by a mobile vehicle with a variety of sensors, including a high-resolution stereo camera and a 3D laser scanner. The camera is calibrated with respect to the laser scanner and the images from the camera are synchronized to 3D point cloud data from the laser scanner. Each object segment is initially obtained from the point cloud data by the proposed g-RBNN clustering and segmentation method, after which its 3D bounding box is projected onto a 2D image plane to generate the ROI. Figs. 4(c) shows the object segments in the point cloud data and the projected ROIs in the 2D image. Figs. 4(h) shows the test samples obtained from the ROIs in the 2D images of the KITTI dataset. Assuming that we have a descriptor which describes the local shape of an object, the position of the block for the descriptor can vary by one or two pixels due to the segmentation result. These outcomes indicate that the proposed method provides a more stable descriptor to represent the object shape.

5 Conclusion

This study demonstrates an experiment in which clustering and segmentation are used as preprocessing techniques. In a forested environment, a 3D LADAR system installed into a UGV performed the clustering and segmentation while mobile. In the object segmentation step, both the Cartesian coordinate information and the polar coordinate information were utilized. The performance in such areas as those assessed here may depend on the 3D LADAR specifications. The results indicate that the proposed method is very effective and efficient for real-time processing in autonomous mobile vehicles. The proposed algorithm works with only four parameters, depending on the laser scanner and not on a specific data set. Furthermore, the algorithm is deterministic which means that running times and segmentation results are repeatable and do not require initialization. We have analyzed the factors influencing the runtime complexity of the g-RBNN method and have verified the analysis by a series of benchmark evaluations with synthetic data. Finally, the efficient performance of the algorithm on a real 3D laser point cloud acquired by a mo-
bile robot has been demonstrated. In future research, the currently analyzed laser range data may be used to estimate the amount of movements of the UGV designed to recognize the location of the UGV and to establish a leader-following system.

References:


Figure 3: Schematic diagram of the clustering method: (a-c) an illustration of the previous clustering approach at the first and second iterations, (d) the result with the previous clustering method, (e-g) an illustration of the proposed clustering approach at the first and second iterations, and (h) the result with the proposed clustering method.

Figure 4: The experimental results of the proposed clustering and segmentation methods: (a-b) the image, (c-d) measurement points acquired from the scanning system, (e-f) clustering and segmentation (g-h), and verification using a camera and a comparison of the LIDAR data.