Mesh Normal Vectors:A New Texture Feature for Single-View 3D Object Classification

1st Yufei Que
Soochow University
Soochow, China
20225246039@stu.suda.edu.cn

2nd Jie XIe
Soochow University
Soochow, China
xiejie@suda.edu.cn

3th Jin Zhang
Soochow University
Soochow, China
zhangjin1983@suda.edu.cn

4th Cheng Wu
Soochow University
Soochow, China
cwu@suda.edu.cn

***Abstract*—The depth information provided by LiDAR 3D point clouds provides help for object detection when the depth direction is occluded. However, large amounts of data from lidar point clouds requires unusual computing resources. Inspired by images and 3D reconstruction, we propose to use normal vectors of the triangular mesh as a new texture feature of point clouds. Our current work indicates that the texture features we created have a certain contribution in the classification task of single-view point clouds.**

***Index Terms*—Classification, Single-view, reconstruction, Normal vectors**

I. Introduction and Realted Work

Considerable progress has been made on point clouds classification tasks since the release of some public benchmark datasets such as ModelNet and ShapeNet. However, most of the existing public datasets provide complete object surfaces. That is, in order to obtain the 3D data and surface information, it is often necessary for multiple lidar devices to simultaneously collect data and complete the fine registration task. However, reality often deviates from the perfect assumption. Especially in autonomous driving, it's difficult for a single vehicle LiDAR to collect complete object surface information. The returned point cloud information tends to be part of the complete surface of the object. Therefore, object detection and classification based on single-view point clouds are tasks in autonomous driving.

Based on the traditional point clouds feature extraction network like PointNet and PointNet++, point clouds information is focused on three-dimensional coordinates. However, it fail to describe the shape of an object accurately and resist noise interference. Semantic ambiguity is one of the main causes of false detections. Literature [1] adds reflection intensity, which improves the performance of discrimination. However, intensity is usually determined by reflection or echo signal collected by sensors, and the range of the value may vary greatly among different sensors and environmental conditions. In some cases, intensity may contribute less to classification tasks, or even introduce redundant information. This means that high-dimensional feature spaces increases the complexity of training and inference. Furthermore, the representation of point clouds is affected by poses, viewpoint and scale. Literature [2] uses the symmetry of the learning problem to improve statistical efficiency, and proposes a method of equivariant convolution. When the input data is rotated or translated, the output of is also rotated or translated accordingly. This provides convenience for point clouds spatial variation. On this basis, the literature [3] uses a network structure accompanied by supervised pose estimation. Specifically, the invariance of point clouds during translation and rotation is exploited to alleviate semantic ambiguity by learning the pose under normalized geometry.

II. Create texture features from point clouds

Conventional point-based 3D detection networks including 3D point cloud coordinate information highly rely on computing resources. To this end, we propose to explore a new low-dimensional feature for the classification task of single-view LiDAR 3D point clouds. Object detection in the image domain is already a mature technique. We try to find similar features to 3D algorithms from 2D object detection or semantic segmentation. Image texture refers to the change rules of grayscale in image. Local texture is represented by gray distribution and surrounding spatial neighborhoods. Different degrees of repetition of local textures form global textures. The texture information of two-dimensional image collected by a single optical camera has the ability to intuitively and effectively reflect the roughness and boundary gradient. In contrast to images, point clouds do not have well-defined pixels or sub-regions. Therefore, it is generally believed that point clouds extracted from LiDAR do not have texture features in themselves. In previous studies, intensity was considered to extract texture of point clouds. However, intensity is mainly related to the distance traveled by light, the reflectivity of the surface and the angle of incidence. In general, smoother surfaces reflect more light and therefore have a higher reflection intensity; rougher surfaces reflect light in more directions and therefore have a lower reflection intensity. However, in our experiments, equipment from different manufacturers is affected by the transmission power of the line beam. Under the same data acquisition conditions, the reflection intensity of the same object sometimes has differences in orders of magnitude. It is difficult to convert sampling data from different manufacturers to a unified standard. In practical application scenarios, the feasibility of extracting point clouds texture using only intensity information needs to be further studied. Therefore,

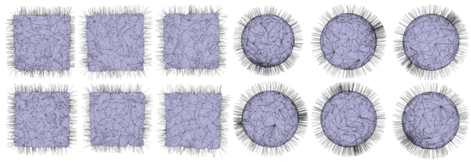

Fig. 1. triangle mesh normal vector

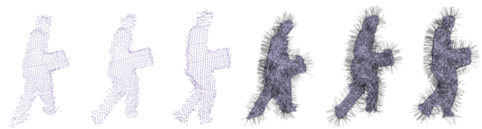

Fig. 2. Single-view original point cloud and triangular mesh normal vector

we hope to create a new texture from LiDAR point clouds for 3D object classification tasks. This new texture needs to have clear intra-class identity and inter-class dissimilarity.

3D reconstruction is another hot research field at present. The 3D reconstruction of an object generally includes object structure reconstruction and surface reconstruction, which correspond to the geometric features of the object and image texture features respectively. Geometric features can also be used to describe the appearance of objects to a certain extent. Inspired by 3D reconstruction techniques, we try to combine 3D reconstruction with 3D object detection. We aim to create information that can reflect geometric surfaces from the properties of points (such as coordinates, normal vectors, intensities, etc.).

Polygon meshes are widely used in graphics and modeling to simulate the surfaces of complex objects such as buildings, vehicles, etc. The normal vector direction of the polygon mesh can replace the local curvature information of the object. The surfaces of a sphere and a cube are spherical and planar, respectively. Even for information from a single view, there is a clear difference in the local curvature of the sphere and cube surfaces. The purpose of our current work is to demonstrate that the normal vector feature of polygon meshes has the ability to distinguish spheres from cubes.

III. PRELIMINARY EXPERIMENT AND CONCLUSION

First, verify that the normal vector feature of the polygon mesh has the ability to distinguish a sphere from a cube. We use software to randomly generate points on the surface of spheres and cubes to simulate point cloud information collected by LiDAR. When reconstructing the object surface, the directed point cloud is triangulated using a traditional greedy triangulation algorithm(Fig 1). Replace local surfaces with triangular mesh planes and local curvatures with triangular mesh normal directions. The classification task is completed by building a Fully Connected Neural Network, and using BCE loss and SGD optimizer. We conduct experiments on 1000 cubes and 1000 spheres, divide the training set, validation set and test set according to the ratio of 6:2:2, and the accuracy rate on the test set is 94.5%.

After initially confirming that the normal vector direction can be used as the texture feature of the LiDAR point clouds, we use the equipment of three LiDAR manufacturers to collect data respectively to construct a single-view dataset. In order to combine the strength and curvature information, the normal vector is proposed as the carrier of the feature for follow-up work on the strength. Take the average intensity of the three points in each triangle mesh as the length of normal vector. This means that each normal vector has curvature and strength information. The trigonometric value of the angle between the normal vector and the depth direction is used to characterize the orientation dimension.

Due to the characteristics of LiDAR data collection, the number of triangles is usually unequal. Zero padding is a traditional way to make eigenvectors equal in length. In order to avoid the number of zero padding becoming a new feature, we only select data with a close number of points. The correct rate of more than 90% is obtained on test set. Our experiments at this stage show that the normal vector of a single-view LiDAR point clouds has the ability to distinguish planes from curved surfaces on the currently constructed dataset at least. It's indicated that mesh normal vectors be used as a new texture to a certain extent.

In real scene of autonomous driving, avoiding living organisms is the basic goal. Our method can now distinguish between a person and a cube, and between a person and a sphere. At present, it's verified that normal vectors can complete the classification task of two types of objects with completely different shapes. When a person holds a cube(Fig 2), since the number of point clouds of the cube is much smaller than that of a person, the entire point clouds are classified as people. In addition, whether our proposed algorithm is robust and generalizable requires more experiments to verify.

Later research will focus on object recognition, especially human recognition. However, people are different in height, fat and thin, and point clouds of different people have different local curvatures after 3D reconstruction. The same person may make multiple movements in a short period of time, and the change of movements makes the local curvature of the same person different after 3D reconstruction. Whether false detections occur similarly to other object surfaces remains to be studied. In addition, curvature features are sensitive to noise and deformation, which may lead to false or missed detections. Based on literature [1] [2] [3], 3D reconstruction after converting the point clouds to a standard pose may achieve better results.

REFERENCES

[1] W. Dai, S. Chen, Z. Huang, Y. Xu, and D. Kong, "Lidar intensity completion: Fully exploiting the message from lidar sensors," Sensors, vol. 22, no. 19, p. 7533, 2022.

[2] T. Cohen et al., "Equivariant convolutional networks," Ph.D. dissertation, Taco Cohen, 2021.

[3] Z. Xu, K. Liu, K. Chen, C. Ding, Y. Wang, and K. Jia, "Classification of single-view object point clouds," Pattern Recognition, vol. 135, p. 109137, 2023.