

基于改进 RPN 的车载导航目标识别技术

史俊莉

(河南工业职业技术学院,河南 南阳 473000)

摘要:车载导航是无人驾驶的一项关键技术,而其主要技术难点是动态避障,即实时目标识别与反馈控制。为了能够实时获取车辆周围目标的运动状态,提出了一种基于三维点云数据快速体素化分析的识别算法。该算法以极坐标系替代直角坐标系,直接与车载激光雷达建立映射关系。再通过区域推荐网络技术中对区域框的限定计算完成可表征目标特征的准确定位。实验采用开源数据库 KITTI 中的点云数据进行验证,对比了动态目标(汽车、摩托车、行人)和静态目标(树木、楼宇)之间的点云特征。识别结果与两种常用的三维目标识别算法进行对比。结果显示本算法的最优平均精度为 88.45%,最优平均方向相似度为 93.27%,相比常用的 SECOND 算法对动态目标的识别具有更好的效果。验证了其可行性,其在车载导航目标识别领域具有一定的优势。

关键词:目标识别;点云数据;体素化;区域推荐网络

中图分类号:TN958.98 **文献标识码:**A **DOI:**10.3969/j.issn.1001-5078.2022.06.007

Target recognition technology for vehicle navigation based on improved RPN

SHI Jun-li

(Henan Polytechnic Institute, Nanyang 473000, China)

Abstract: Car navigation is a key technology of unmanned driving, and its main technical difficulty is dynamic obstacle avoidance, which means real-time target recognition and feedback control. In order to obtain the real-time movement status of the targets around the vehicle, a three-dimensional point-based algorithm is proposed to replace the rectangular coordinate system with a polar coordinate system, and directly establish a mapping relationship with the vehicle-mounted lidar. Then, the accurate positioning of the characterizable target features is completed through the limited calculation of the area frame in the area recommendation network technology. The experiment is validated using point cloud data in the open source database KITTI to compare the point cloud features between dynamic targets (cars, motorcycles, pedestrians) and static targets (trees, buildings). The recognition results are compared with two commonly used 3D target recognition algorithms. The results show that the optimal average accuracy of this algorithm is 88.45%, and the optimal average direction similarity is 93.27%. The proposed algorithm has a better effect on the recognition of dynamic targets compared with the commonly used SECOND algorithm. Its feasibility is verified, and it has certain advantages in the field of vehicle navigation target recognition.

Keywords: target recognition; point cloud data; voxelization; regional recommendation network

1 引言

当前无人驾驶技术已成为汽车导航领域的一个重要研究方向,因而车辆行驶过程中的目标识别与

避障技术成为了研究热点^[1]。依据不同的工作原理主要可以分为两大类,一类是基于图像识别技术,另一类是基于激光雷达技术。

基金项目:国家自然科学基金青年项目(No. 61803065);《控制测量技术》实践教学信息化项目(No. SXXM202004)资助。

作者简介:史俊莉(1984-),女,硕士,讲师,主要研究方向为精密测量与 GNSS 导航定位。E-mail:shijunlinuc@sina.com

收稿日期:2021-06-24;**修订日期:**2021-07-30

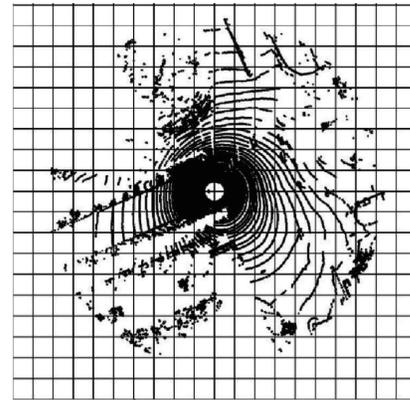
对于图像识别技术而言,常采用 RGB-D 图像识别技术,由此克服二维图像没有距离信息的问题^[2]。因为路上的车辆、行人等均是一个整体,故其在二维图像中的几何范围是一个动态变化的区域,同时该动态变化区域的 RGB 符合一定的变化范围,基于此结合其深度信息就能为车辆避障提供有效信息^[3]。它的优点是数据量小实时性好,缺点是对于行人、摩托车等相对小目标而言,其距离信息误差大,故其应用于高速公路等场合时十分适用。Engelcke^[4]提出了在 RGB-D 图像上使用滑动提取的方式获得目标边界框的算法,实现了利用深度变化值锁定目标的目的。Qi, C. R.^[5]将 F-PointNets 算法应用于特征面形识别,从而大幅简化了图像之间的中继过程,提高了识别效率。对于基于激光雷达技术而言,目标大小的影响可以忽略,都可以获取精确的距离信息,优点是精度高稳定性好,缺点是从点云数据集中提取目标时匹配准确度容易受训练样本影响,且三维点云数据运算量大。Chen, X.^[6]在其系列研究中对特征获取、边界提取等具有很好目标识别结果,唯一的缺点是运算量很大,实时性降低。González, A^[7-8]采用将三维数据向二维平面投影的方式,大幅提升了处理速度,但是尺寸较小的行人等目标的空间信息丢失严重,影响了其适用性。

由此基于体素化的三维目标识别方法由于其保留空间特征的同时,大幅降低了数据运算量,故得到了广泛关注。本文设计了一种基于极坐标系的体素表达方式,并采用改进的区域推荐网络(RPN, Region Proposal Network^[9-10])算法抑制了目标点云匹配中角度误差,从而提高了系统目标识别精度。

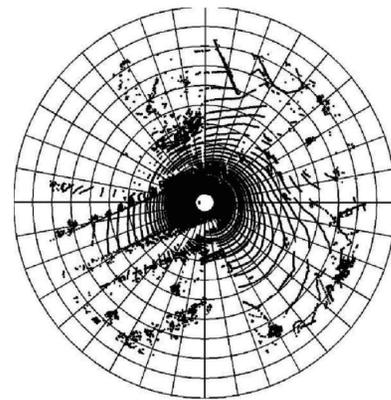
2 极坐标体素化模型构建

在 VoxelNet 中,点云数据的体素化如果采用不同的坐标系,则会有不同个划分方式,基于传统的直角坐标系的结果如图 1(a)所示,而基于极坐标系的结果如图 1(b)所示。由于车载激光雷达本身就是以一个激光雷达光源为原点的,故其采用极坐标系的体素化方法更有利于数据分类与提取。

图 1(a)是采用直角坐标系的体素化结果,一般这种体素化方式适用于具有扫描结构的系统中,可以沿着直角坐标系中的一个轴进行数据获取,并且与二维图像进行映射,从而达到较好的对应效果。而对于车载激光雷达而言,由于激光扫描方式是以车辆为圆心的,故其形成一个同心圆效果,所以采用极坐标系实现体素化更符合特点,也是抑制目标提取时角度误差的重要依据。



(a)直角坐标系



(b)极坐标系

图 1 体素化在不同坐标系下的表达形式

Fig. 1 The expression form of voxelization in different coordinate systems

假设极坐标系的三个自变量分别是 r , α 和 z 。则将直角坐标系中的任意点 (x, y, z) 变换至极坐标系时,有:

$$\begin{cases} r = \sqrt{x^2 + y^2} \\ \alpha = \arctan(y/x) \\ z = z \end{cases} \quad (1)$$

经过对原有点云数据进行体素化后,需要对点云点进行聚类。由于激光雷达在可探测角度范围内是均匀扫描的,所以基于极坐标系的体素分割比直角坐标系的更均匀。对获得的稀疏点云进行数据处理,设每个非空体素中的最大点数为 T ,并自动丢弃冗余点;如果点云小于 T 的,设置为零。假设每个体素的体素数为 K 、最大点数为 T ,则体素聚类的矩阵张量为 $4 \times K \times T$ 。

3 算法设计

3.1 点云特征选择

设非空体素的表达式为:

$$V = \{p_i = [r_i, \alpha_i, z_i, q_i]^T \in R^4\} \quad (2)$$

式中,其中 r , α 和 z 为极坐标系的三个分量; q 表示回波强度;下标 i 表示任意点; $i \in [1, t], i \in R^+$ 。首先,估计每个体素中所有点的平均值为 $V_m = (v_r, v_\alpha,$

v_z), 中心值为 $V_c = (c_r, c_\alpha, c_z)$ 。然后, 将平均距离和从每个点到体素中心的距离相加得到其特征。最后, 生成形式 $10 \times K \times T$ 的张量为:

$$P_i = [r_i, \alpha_i, z_i, q_i, r_i - v_r, \alpha_i - v_\alpha, z_i - v_z, r_i - c_r, \alpha_i - c_\alpha, z_i - c_z] \quad (3)$$

3.2 特征提取

体素特征提取一般通过体素特征编码 (Voxel Feature Encoding, VFE) 实现, VFE 层 (如图 2 所示) 以聚类体素为输入, 并使用全连接层, 批处理规范和 ReLU 激活函数提取特征。设 VFE 的输出节点数为 $C = 2 \times n, n \in Z^+$ 时, VFE 层获取对应于同一体素的 $10 \times T$ 个点, 并完全连接到 $C/2$ 的输出, 可获得尺寸为 $K \times C/2$ 的特征。然后使用最大池化获得每个体素的局部聚集特征尺寸 $K \times 1$, 最后将其复制到点对点特征中, 获得最终的体素特征 $K \times C$ 。

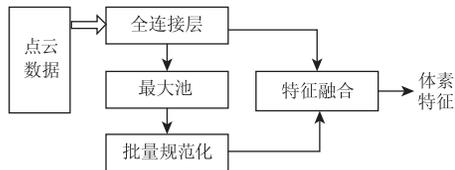


图2 VFE层结构示意图

Fig. 2 Schematic diagram of VFE layer structure

采用稀疏卷积提取空间特征, 在逐渐增加体素特征参数的过程中, 提高对测试区域中目标形状的描述。空间特征提取层可以提取有关 z 轴的信息, 并将稀疏的 3D 数据转换为密集的 2D 伪图像。空间特征提取层将从体素特征提取层获得的特征作为输入, 并根据体素网格中每个体素的坐标将非空体素转换为稀疏 4D 张量, 然后提取空间特征。

3.3 改进的 RPN 算法

由于卷积层中的卷积核具有固定的形式, 其数据限定域仍为正方形, 所以经过多次卷积后, 由极坐标形式获取的激光雷达数据的角度误差会不断被放大, 导致测试精度降低。RPN 算法通常用于确定特征图的局部是否存在要检测的目标, 故本文对传统 RPN 算法进行了改进, 通过在体素中引入阶跃参数 s 与填充参数 f , 对体素内点云数量变化与有效点云数进行评价。改进的 RPN 算法结构如图 3 所示。

在改进的 RPN 模块中, 将卷积输出的特征参数作为输入对应的输出检测框, 图中的 2D Conv 表示带有非线性激活函数的 2D 卷积, 然后再将所有不同形式输出数据进行反卷积, 最终完成数据特征的构成, 将三种数据特征整合后形成特征图。最终系

统采用特征图完成置信度的计算与回归图的生成。

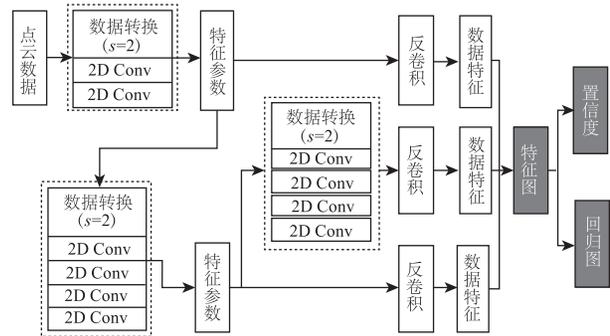


图3 改进的 RPN 算法结构图

Fig. 3 Improved RPN algorithm structure diagram

4 实验

4.1 数据来源与对比方法

为了验证算法的可行性, 激光雷达数据选用 KITTI 的开源数据。神经网络训练是通过映射模型参数的调整实现的。在深度学习网络模型训练过程中, 数据量对网络的收敛速度和泛化性能有很大影响。在训练过程中, 从数据集中随机提取 n 个地面目标, 并将其插入当前的训练数据中。采用该策略大大增加了每个帧点云中实际目标的数量, 降低了随机干扰。考虑到噪声对网络性能的影响, 使用了 VoxelNet 中的类似方法。对于每一帧都采用点云独立变换的方式完成, 而不是采用相同参数对所有点云同时变换。使用 Adam 优化器, 初始学习率设置为 0.0002, 并且在训练过程中使用学习率衰减策略, 共训练 100 代, 用时 10 h (硬件采用 i5 CPU 配合 NVIDIA RTX 2080 Ti 显卡)。数据分析结果采用平均精度和平均方向相似度来分析不同算法的测试结果。

4.2 测试结果分析

与用于点云数据目标框图选择常用的 SECOND 算法进行对比分析, 对包含汽车、摩托、行人、树木及建筑物的测试环境进行分析, 测试环境可见光图像、经本文算法获取的点云数据与目标识别框图如图 4 所示。

如图 4(a) 和 (c) 所示, 在可见光图像中车辆目标由于树荫的遮光效果, 使其清晰度很差, 但在点云数据中具有较清晰的回波点云特征, 结合点云特征通过本算法完成了测试目标的框选。如图 4(b) 和 (d) 所示, 当测试环境复杂时, 路面上除了汽车还存在摩托车、行人时, 根据不同的目标点云特征, 可以将不同类型目标进行区分。并且在目标点云识别图像中右侧出现的建筑物表面特征也十分清晰, 而左

侧的自行车堆呈现为杂散点云集合,类似建筑但弥散性更强。由此可见,采用本算法可以很好地完成对不同目标的区分,并为无人驾驶提供更准确的安全导航定位信息。

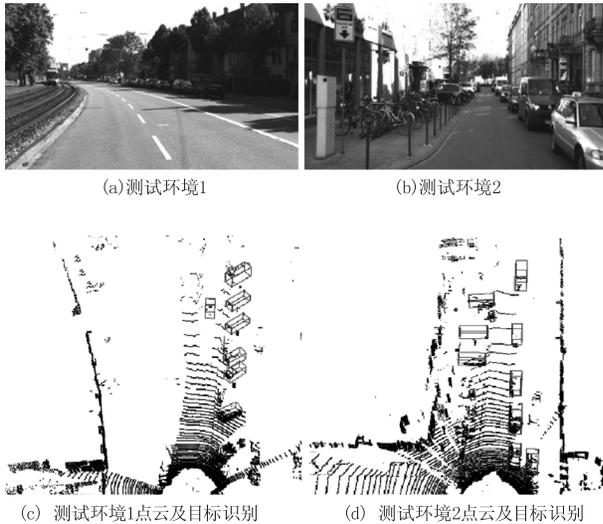


图 4 不同测试环境下目标识别点云结果

Fig. 4 Target recognition point cloud results in different test environments

4.3 测试数据对比

对同一组的点云数据进行比较,将本系统识别结果的平均精度 (Mean Average Precision, MAP) 与平均方向相似度 (Average Orientation Similarity, AOS) 与 SECOND 算法结果进行比较,两个参量的计算方程如下:

$$\begin{cases} \text{MAP} = \frac{1}{N} \sum_{i=1}^N p_i \\ \text{AOS} = \frac{1}{N} \sum_{r \in \langle 0, 0.1, \dots, 1 \rangle} \max_{r > r}(\bar{r}) \end{cases} \quad (4)$$

通过对测试区域中所有待测目标进行计算,可以得到每种目标的 MAP 和 AOS,对相同数据而言,两种不同算法的测试结果如表 1 所示。

表 1 算法测试结果对比

Tab. 1 Comparison of algorithm test results

测试目标	MAP		AOS	
	SECOND	本算法	SECOND	本算法
汽车	85.36	88.45	90.14	93.27
摩托车	75.41	79.32	88.35	91.54
行人	74.68	78.36	81.73	85.91
树木	69.55	70.57	75.36	75.82
平均	76.25	79.18	83.90	86.64

由表 1 的测试结果对比可知,对于汽车、摩托车

及行人目标而言,本算法的平均精度与平均方向相似度均优于常用的 SECOND 算法。其中,MAP 分别高出 3.09 %、3.82 % 和 3.68 %,AOS 分别高出 3.13 %、3.19 % 和 4.18 %。对于树木的点云区分中两种方法检测精度接近。分析认为,由于汽车、摩托车及行人均是运动目标,并且在不同测试站位对激光雷达几何面形是不同的,故 SECOND 算法中每个站位的几何映射都需要在直角坐标系中首先转化为相对激光雷达坐标系的点云,这个过程会造成角度误差。而采用本算法时,是基于极坐标系完成的激光雷达点云体素化的,这样即使是目标移动,测试面也是以激光雷达光源为原点的极坐标表达,所以其没有点云数据角度变换的过程,避免了测试过程中的角度误差。而树木测试数据由于不能运动,故其不同站位的点云集合具有连贯性,所以其点云数据匹配不需要坐标变换,仅需要点云数据映射即可,故两种方法检测的平均精度与平均方向相似度基本不发生改变。

4.4 算法处理速度对比

为了进一步验证改进算法的时效性,将本算法与 SECOND 算法同时对等量的点云数据进行处理,并且分别记录算法的收敛时间,测试点云数据量从 1000 个持续增加至 10000 个,算法收敛时间如图 5 所示。

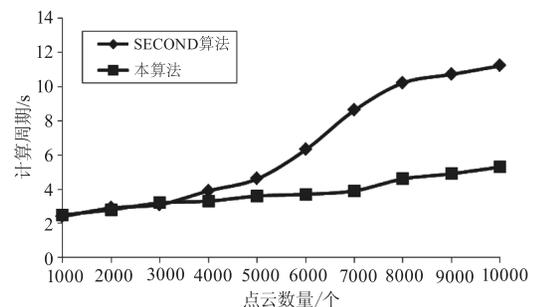


图 5 两种算法在不同点云数据量下的收敛时间

Fig. 5 The convergence time of the two algorithms under different point cloud data volumes

由图 5 可知,点云数据在 3000 个以下时,两种算法的处理速度基本一致,由于点云数据总量较小,所以在目标分类过程中是否采用更合理的点云分块设计并不能体现更好的优越性。但是,当点云数据超过 3000 个后,除了目标特征信息外的杂散点数据也逐渐增多,这时如果没有采用本文的极坐标体素化方式进行预处理,那么就会使后期目标识别过程中需要更多的相关计算与边界条件迭代才能达到预期分类精度。由此可见,本算

法的处理速度优于 SECOND 算法,尤其是在点云数据量较大的情况下。

5 结论

本文针对车载激光雷达三维点云数据中动态目标的识别问题,提出了一种基于极坐标系表达的体素化分析识别算法。在重新构建了体素化表征方式的基础上,通过设置阶跃参数与填充参数实现了对 RPN 算法的改进。实验对 KITTI 数据库中包含汽车、摩托车、行人与树木的点云样本进行测试,结果显示,本算法的平均精度与平均方向相似度均优于常用的 SECOND 算法,在点云目标的动态框选中具有更好的应用效果。

参考文献:

- [1] Giseok K, Jae-Soo C. Vision-based vehicle detection and inter-vehicle distance estimation for driver alarm system [J]. *Optical Review*, 2012, 25(6): 388 - 393.
- [2] Chavez-Garcia R O, Aycard O. Multiple sensor fusion and classification for moving object detection and tracking [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2016, 17(2): 525 - 534.
- [3] Li Chunmiao, Gao Jianbo, Yang Zhenyuan, et al. Target recognition algorithm of laser imaging radar based on feature value [J]. *Laser & Infrared*, 2015, (3): 335 - 338. (in Chinese)
李春苗,高剑波,杨镇源,等.基于特征值的激光成像雷达目标识别算法[J].*激光与红外*, 2015, (3): 335 - 338.
- [4] Engelage M, Rao D, Wang D Z, et al. Vote3deep: fast object detection in 3d point clouds using efficient convolutional neural networks [C]//International Conference on Robotics and Automation, 2017: 1355 - 1361.
- [5] Qi C R, Litany O, He K, et al. Deep hough voting for 3d object detection in point clouds [C]//2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), 2019: 9276 - 9285.
- [6] Chen X, Ma H, Wan J, et al. Multi-view 3d object detection network for autonomous driving [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, 2017: 6526 - 6534.
- [7] González A, Vázquez D, López A M, et al. On-board object detection: multicue, multimodal, and multiview random forest of local experts [C]//IEEE Trans. Cybern. 2017: 3980 - 3990.
- [8] González A, Villalonga G, Xu J, et al. Multiview random forest of local experts combining RGB and LIDAR data for pedestrian detection [C]//2015 IEEE Intelligent Vehicles Symposium, IV 2015, Seoul, South Korea, 2015: 356 - 361.
- [9] Wang Baofeng, Qi Zhiqian, Ma Guocheng, et al. Vehicle detection based on information fusion of radar and machine vision [J]. *Automotive Engineering*, 2015, 37(6): 674 - 678. (in Chinese)
王宝锋,齐志权,马国成,等.一种基于雷达和机器视觉信息融合的车辆识别方法[J].*汽车工程*, 2015, 37(6): 674 - 678.
- [10] Xia Xianzhao, Zhu Shixian, Zhou Yiyao, et al. LiDAR K-means clustering algorithm based on threshold [J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2020, 46(1): 115 - 121. (in Chinese)
夏显召,朱世贤,周意遥,等.基于阈值的激光雷达 K 均值聚类算法[J].*北京航空航天大学学报*, 2020, 46(1): 115 - 121.