

基于激光雷达点云的3D目标检测方法综述

秦静, 王伟滨, 邹启杰, 汪祖民, 季长清

引用本文

秦静, 王伟滨, 邹启杰, 汪祖民, 季长清 [基于激光雷达点云的3D目标检测方法综述](#) [J]. 计算机科学, 2023, 50(6A): 220400214-7.

QIN Jing, WANG Weibin, ZOU Qijie, WANG Zumin, JI Changqing. [Review of 3D Target Detection Methods Based on LiDAR Point Clouds](#) [J]. Computer Science, 2023, 50(6A): 220400214-7.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于多传感器的室内建图导航系统的设计](#)

Design of Indoor Mapping and Navigation System Based on Multi-sensor

计算机科学, 2023, 50(6A): 220300218-8. <https://doi.org/10.11896/jsjcx.220300218>

[基于改进Yolov4-tiny的轻量型目标检测算法](#)

Lightweight Target Detection Algorithm Based on Improved Yolov4-tiny

计算机科学, 2023, 50(6A): 220700006-7. <https://doi.org/10.11896/jsjcx.220700006>

[CT影像阶段化目标检测方法研究](#)

Study on Phased Target Detection in CT Image

计算机科学, 2023, 50(6A): 220200063-10. <https://doi.org/10.11896/jsjcx.220200063>

[基于区域注意力机制和多尺度特征融合的输电线路螺栓缺陷检测](#)

Defect Detection of Transmission Line Bolt Based on Region Attention Mechanism and Multi-scale Feature Fusion

计算机科学, 2023, 50(6A): 220200096-7. <https://doi.org/10.11896/jsjcx.220200096>

[基于注意力机制最大化重叠的单目标跟踪算法](#)

Maximum Overlap Single Target Tracking Algorithm Based on Attention Mechanism

计算机科学, 2023, 50(6A): 220400023-5. <https://doi.org/10.11896/jsjcx.220400023>

基于激光雷达点云的 3D 目标检测方法综述

秦 静¹ 王伟滨² 邹启杰² 汪祖民² 季长清³

1 大连大学软件工程学院 辽宁 大连 116622

2 大连大学信息工程学院 辽宁 大连 116622

3 大连大学物理科学与技术学院 辽宁 大连 116622

(qinjing@dlu.edu.cn)

摘 要 近年来,利用激光雷达点云进行 3D 目标检测是计算机视觉领域的一个研究热点,并在自动驾驶领域备受关注。3D 相比于 2D 而言,结合了深度信息,更能体现出现实世界的特征,以有效解决如路径规划、运动预测、目标检测等方面的实际问题。介绍了 3D 目标检测的发展背景,概述了基于激光雷达点云数据的 3D 目标检测框架的流程,比较了几种常见的包含点云信息的数据集,并对主要研究方法进行分类。结合了自动驾驶的应用场景,对不同方法的性能和局限性进行了分析和比较。最后,总结了现阶段的技术难点,并展望了该领域未来的发展前景。

关键词: 目标检测;点云;计算机视觉;激光雷达;多模态融合

中图法分类号 TP391.41

Review of 3D Target Detection Methods Based on LiDAR Point Clouds

QIN Jing¹, WANG Weibin², ZOU Qijie², WANG Zumin² and JI Changqing³

1 College of Software Engineering, Dalian University, Dalian, Liaoning 116622, China

2 College of Information Engineering, Dalian University, Dalian, Liaoning 116622, China

3 College of Physical Science and Technology, Dalian University, Dalian, Liaoning 116622, China

Abstract In recent years, 3D target detection using LiDAR point cloud is a research hotspot in the field of computer vision and has attracted much attention in the field of autonomous driving. Compared with 2D, 3D combines depth information to better reflect the characteristics of the real world, to effectively solve practical problems such as path planning, motion prediction, target detection, and other aspects. This paper introduces the development background of 3D target detection, summarizes the flow of 3D target detection framework based on LiDAR point cloud data, compares several common data sets containing point cloud information, and classifies the main research methods. The performance and limitations of different methods are analyzed and compared. Finally, the current technical difficulties are summarized and the future development prospects of this field are forecasted.

Keywords Target detection, Point cloud, Computer vision, LiDAR, Multimodal fusion

1 引言

目标检测是计算机视觉领域中的核心问题,目的是在图像中找出人类感兴趣的物体,并进行识别和标注。目前较为成熟的方法是基于图像的 2D 目标检测,具有代表性的方法有 SSD^[1], YOLO^[2] 和 Faster-RCNN^[3] 等,这些方法被广泛应用于人脸识别^[4]、工业检测^[5] 等领域。但由于缺少距离信息,无法充分考虑 3D 空间环境影响,所以在无人驾驶^[6]、目标跟踪^[7] 等新兴领域,2D 目标检测应用有诸多局限性。基于 3D 数据的目标检测逐渐被人们所重视^[8],并成为该领域的研究热点。

2D 与 3D 目标检测的主要区别如图 1 所示。环境感知系统可利用立体摄像机或者激光雷达传感器获取环境的深度信息,算法根据感知到的空间信息计算并得到包含物体长宽高信息的预测框。

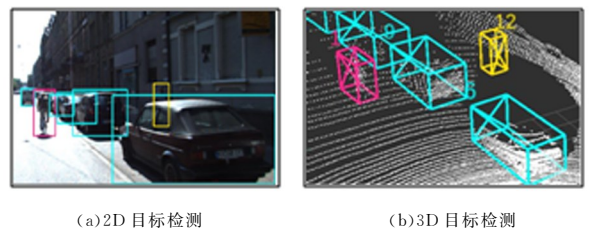


图 1 2D 和 3D 目标检测
Fig. 1 2D and 3D target detection

3D 目标检测方法目前最主要的应用场景是自动驾驶^[9]。环境感知系统是自动驾驶的重要组成部分,相当于人的眼睛^[10]。自动驾驶车辆的环境感知系统一般搭载了高精度相机和激光雷达等传感器,其中相机采集色彩鲜艳、纹理特征强、高密度的图像信息,激光雷达采集具有丰富空间信息的点云。将二者提供的数据结合,能够帮助系统理解周围现实

基金项目:国家自然科学基金青年科学基金项目(62002038)

This work was supported by the National Science Foundation of China(62002038).

通信作者:汪祖民(wangzumin@dlu.edu.cn)

环境,并作出合适判断^[11]。

激光雷达作为环境感知传感器的一部分,有着不可替代的作用,是目前主流传感器之一。激光雷达能够主动发射激光,有效地采集周围的环境信息,具有精确度高、不易受光照影响等优点。因此,利用激光雷达点云准确完成3D目标识别任务成为近年来的研究热点。本文主要就是对该技术的整体流程、数据获取、主要方法等进行深入的分析、对比和总结。

2 总体框架与研究难点

2.1 总体框架

目前基于激光点云的3D目标检测主流框架由点云特征表达、2D或3D的骨干网、检测头3部分组成,如图2所示。

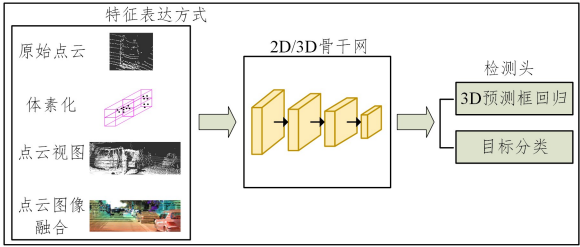


图2 基于点云的目标检测框架

Fig. 2 Object detection framework based on point cloud

首先,点云特征表达按不同的处理方式,分为了基于原始点云、体素化、点云视图及点云图像融合4种。其次,为了在较短的时间内提取出高质量的点云特征,需要设计骨干网进行提取特征,其中3D骨干网使用3D卷积。最后,使用检测头输出计算后目标的3D预测框以及目标类型。

2.2 研究难点

目前关于3D目标检测研究的主要难点有以下方面。

(1)特征提取困难。点云数据具有稀疏、不规则等特性,如何合理地划分点云集合,进而提取特征,是利用点云时考虑的重点和难点。目前处理点云的主流方法是将点云体素化成网格,再输入到卷积神经网络提取特征。基于体素的方法要人为规定体素大小,体素的大小影响捕捉到的特征细节,确立合理的体素大小是极为重要的问题。

(2)点云数据获取成本高,数据量少。激光雷达价格昂贵,获取点云的数据成本较高,虽然有一些开源数据集能够提供数据支持,但是难以满足在不同的天气状况、室外场景、白天或者黑夜下精确检测出3D环境中目标的条件。基于机器学习的方法需要在不同场景下采集大量数据供模型训练,因此需要花费大量时间去收集和标注数据。

(3)数据融合方法复杂。激光雷达在检测大型物体方面有着良好的性能,但是不能够提供颜色、纹理等信息,在识别小型物体上较为困难,而RGB图像能够提供丰富的视觉特征和密集的环境信息,因此将点云和图像进行融合能够得到更好的检测结果。多传感器信息融合需要考虑传感器校准^[12]、时间同步^[13]、特征融合^[14]处理等复杂问题。由于图像数据相比点云有着更高的分辨率,要将稀疏的点云数据与密集的图像数据结合也是一项难点。

3 点云数据的获取与处理

3.1 获取方法

点云主要通过激光雷达扫描获取。激光雷达一般被搭载在车辆顶部,主动向周围发射的脉冲激光触碰到物体后发生反射,系统接受单元负责接受反射信号生成点云^[15]。

除激光雷达外,还可使用相机获取。主要方式有两种:一种是通过双目立体摄像机拍摄到两幅不同位置的图像信息,利用位置偏差计算生成点云^[16];另一种是通过深度相机,利用红外激光设备采集得到深度信息生成点云^[17]。相机获取点云的优点是设备价格便宜,缺点是检测精度低,视野窄,存在盲区,容易受光照影响等。

点云主要以三维坐标的形式存储,此外还可能存储反射强度、时间戳等信息。

3.2 预处理方法

点云数据具有海量性、分布不均匀性、散乱性、噪声多样性以及坐标系随意性等特点^[18],因此需要对采集到的点云进行预处理。点云的预处理方式主要有以下几个步骤,处理效果如图3所示。

(1)去除离群点。设备采集点云时受到环境因素、物体材质或者设备本身的影响,采集的点云中包括许多散列和孤立的点,离群点云会干扰后续处理和判断,因此需要进行滤波操作将这些点云去除^[19]。

(2)数据压缩。点云数据量十分庞大,一帧点云数量就可达上万个点。如果不对点云进行压缩,则后续处理会消耗大量的时间,降低算法效率。因此要对点云进行简化和压缩。

(3)坐标变换。点云采集系统由惯性测量单位、激光雷达、相机等设备组成。不同的设备分布在载体的不同位置,从而造成系统中包含多个坐标系。因此需要将采集到的数据进行坐标变换,将其映射到一个固定的坐标系下。



图3 点云数据处理

Fig. 3 Point cloud data processing

对采集后的点云数据可以使用点云库(Point Cloud Library,PCL)^[20]进行预处理。PCL是一个2013年发布的开源C++库,集中了多种点云处理的算法模块,如点云滤波、关键点配准及可视化等模块,支持多种操作平台。

3.3 常用数据集

目前有许多公开的数据集可以使用,数据集中包含经过数据清理的点云信息。本节将介绍几种常见的3D目标检测数据集。

2012年提出的KITTI^[21]是被广泛使用的经典数据集。由德国卡尔斯鲁厄理工学院和丰田美国技术研究院联合创建,对数据中的汽车、货车、卡车、行人、骑行的人等对象提供了标注。该数据集在立体视觉、光流估计、视差估计、3D物体检测和3D跟踪等任务中评估计算机视觉技术的性能。

nuScenes^[22]数据集是自动驾驶公司nuTonomy建立的大规模数据集,除激光雷达外,还有毫米波雷达数据。nuScenes

由 1 000 个场景组成,每个场景 20 s 长,对 3D 边界框进行完整的标注,并包含了一些其他数据集所没有的标注,如行人姿势或车辆状态等。

Alphabet 旗下的自动驾驶公司 Waymo 开源了自动驾驶数据库 Waymo open dataset^[23],这是迄今为止最大、最多样

化的自动驾驶数据集。收集信息的车辆配备了先进的传感器,涵盖了不同的天气条件,白天、夜晚不同的时间段,以及多样的驾驶场景,为 3D 目标检测提供了高质量、大规模的数据集。

数据集相关信息及获取地址如表 1 所列。

表 1 3D 目标检测的数据集
Table 1 3D target detection dataset

数据集	发表年份	采集设备数量			下载地址
		摄像头	激光雷达	毫米波雷达	
KITTI	2012	4	1	—	www.cvlibs.net/datasets/kitti/index.php
NuScenes	2019	6	1	5	www.nuscenes.org/download
Waymo	2019	5	5	—	www.waymo.com/open

4 点云的 3D 目标检测方法

根据特征处理的不同,基于点云的 3D 目标检测方法可分为:基于原始点云、基于体素、基于多视图以及基于融合的方法。

4.1 基于原始点云的方法

由于点云的稀疏性和不规则性,神经网络通常不能直接处理点云数据。如果能够直接计算并处理点云,就能够最大程度地利用点云信息训练出高鲁棒性的模型。

2017 年 Qi 等^[24]开创性地提出能够直接处理点云数据的神经网络 PointNet,成功解决了该问题,并将应用于多个场景,如图 4 所示。但该网络无法获得局部特征,导致了在复杂场景不能够很好地识别。后经过改进,其又提出了 PointNet++^[25],它构建了点的层次分组,递归地将 PointNet 应用于输入点集,能够从上下文学习中学习局部特征。为了解决非均匀点的采样问题,利用两个集合抽象层来聚合多尺度信息,提升了结果的准确性。PointNet++ 的提出,搭建出了高效的点集特征学习网络,为点云处理提供了优质算法。

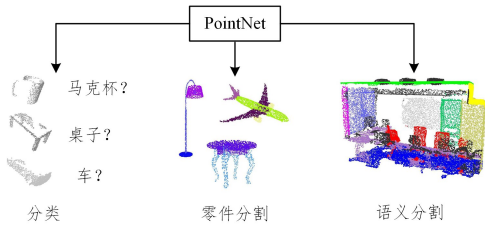


图 4 PointNet 的应用场景

Fig. 4 Application scenarios of PointNet

此外,常见的对 PointNet++ 改进的网络包括:PointRCNN^[26]将点云分割成前景点和背景点,从前景点云中生成边界框提议,利用池化后的点云特征以及分割掩码来细化坐标,以更好地学习局部空间特征;3DSSD^[27]是去掉了大计算量的特征传播层和细化模块来提高运行效率,相比于双阶段方法,少了提出候选区域的步骤,在运行速度上比双阶段方法 PointRCNN 快了 2 倍以上,具有较高的实时性。

4.2 基于体素的方法

体素是体积元素的简称。在 2D 图像中,将组成图像的小方格称为像素;类似地,在 3D 空间中划分成的立方体被称为体素。

VoxelNet^[28]是体素化方式的开篇之作,将点云划分为等间距的体素,并通过体素特征编码层(Voxel Feature Enco-

ding, VFE)将每个体素内的点转换为矢量,叠加多个 VFE 层以学习复杂的特征。由于使用了高计算代价的 3D 卷积运算,导致检测速度缓慢。为了减少计算量,SECOND^[29]网络采用了稀疏卷积的方法代替了 3D 卷积,经典结构如图 5 所示。PointPillars^[30]进一步改进网络,将点云体素化为柱子,柱子是一个在 z 方向上具有无限空间范围的体素,利用 PointNet 来学习在柱中组织的点云的表示,提高了网络适应性,速度达到了 62 帧每秒。

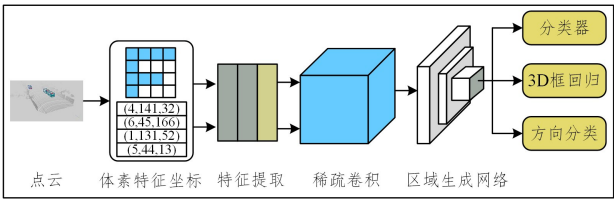


图 5 SECOND 网络结构

Fig. 5 SECOND network structure

此外,为了解决体素化导致的信息丢失问题,HVNet^[31]融合多种尺度的 VFE,使用两种体素的混合解决了不同尺度的体素提取特征的问题。Part-A² net^[32]根据不同的场景,可选择无锚策略或基于锚策略,具有更好的适应性。

为了提高检测速度,涌现了一批单阶段检测方法(Single-Stage Detector, SSD)。SA-SSD^[33]通过两个辅助任务来丰富骨干网的特征,具有更好的定位效果,通过能量模型^[34]可进一步提升性能。CIA-SSD^[35]通过降低空间分辨率的方式减少计算量。SE-SSD^[36]利用软目标、硬目标以及制定的约束共同优化模型,提高了车辆检测性能。

针对远处目标点云稀疏问题,SIENet^[37]使用空间信息增强模块将稀疏的目标点云补全。Hotspots^[38]将体素中的每个点的特征集合起来,形成热点表示,鼓励网络从最具区别性的特征中学习,对稀疏点云具有良好的鲁棒性和有效性。

为了比较上述算法性能,对不同体素方法在 KITTI 测试集的检测精度进行了测试,如表 2 所列。SE-SSD 在检测汽车方面有着最高的检测精度,而且将速度控制在了 30ms 左右,体现出了在单阶段方法上使用双模型训练的优越性。Part-A² net 与 SIENet 为双阶段方法,可以明显看出双阶段方法比较耗时,但其提升了对小物体检测的准度,这两种方法分别在检测行人和检测骑行的人这两类目标上体现出了最佳的成绩。PointPillars 在实时性方面表现优秀,将点云体素化成柱子是加快检测速度的较好方案。

表 2 不同体素方法在 KITTI 测试集的检测精度

Table 2 Detection accuracy of different voxel methods on KITTI test set

方法	参考文献	汽车/%			行人/%			骑行的人/%			每帧 用时/ms
		简单	中等	困难	简单	中等	困难	简单	中等	困难	
VoxelNet	[28]	77.47	65.11	57.73	39.48	33.69	31.51	61.22	48.36	44.37	225
SECOND	[29]	83.13	73.66	66.20	51.07	42.56	37.29	70.51	53.85	46.90	50
PointPillars	[30]	79.05	74.99	68.30	52.08	43.53	41.49	75.78	59.07	52.92	16
Part-A ² net	[32]	85.94	77.86	72.00	54.49	44.50	42.36	78.58	62.73	57.74	71
SA-SSD	[33]	88.75	79.79	74.16	—	—	—	—	—	—	40
CIA-SSD	[35]	89.59	80.28	72.87	—	—	—	—	—	—	31
SE-SSD	[36]	91.49	82.54	77.15	—	—	—	—	—	—	30
SIENet	[37]	88.22	81.71	77.22	—	—	—	83.00	67.61	60.09	161
Hotspots	[38]	87.60	78.31	73.34	53.10	45.37	41.47	82.59	65.95	59.00	40

4.3 基于多视图的方法

基于视图的方法不直接处理点云数据,而是将点云经过不同的投影方式压缩维度,其核心思路如图 6 所示。将 3D 点云转为 2D 的鸟瞰图(Bird-Eye View, BEV)或者前视图(Perspective View, PV),然后将投影输入到卷积神经网络中进行相应处理。不同视图能够表现点云的不同特点,在 BEV 下避免了物体间的遮挡影响,丢失了高度特征,但保留了位置与尺寸特征。在 PV 下,点云图与图像类似,点云密集能够方便获取语义信息,但是存在遮挡问题。

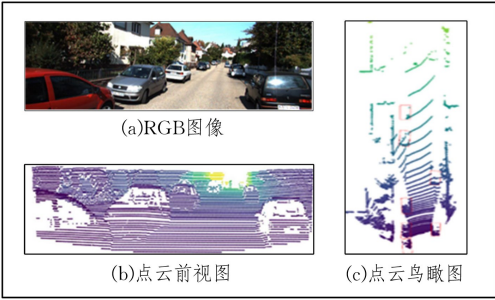


图 6 不同视角的点云图像

Fig. 6 Point cloud images from different perspectives

BANet^[39]对物体的局部边界相关性进行建模,为每个提议提供完整的边界,改善了投影后边界信息丢失的问题。Li 等^[40]将点云通过 VFE 提取 3D 体素特征,并投影到 BEV 上作为地图特征,再将点特征和地图特征与提议对齐,以准确感知目标位置。

为了充分利用视图信息,H² 3D R-CNN^[41]通过多视图融合块将两个视图特征进行融合,生成特征,利用双视图的互补形成 3D 表示。X-view^[42]融合了不同坐标空间的特征,通过转换 FV 坐标的原点和扩展 FV 的数量突破以自我中心约束来提升性能。

4.4 图像与点云融合的方法

只以点云作为输入的方法存在难以检测远方物体或偏小物体的缺陷,而 RGB 图像代表真实世界的投影,将二者的特征进行融合,能够提高小物体的检测的精度。这类方法主要分为特征融合类和以 2D 检测为驱动的形式。

MV3D^[43]融合了 RGB 图像和点云的 PV, BEV 两种视图特征,提高了检测的稳定性,但不适于检测小对象实例。针对该不足,AVOD^[44]舍弃了 PV 视图,通过在高分辨率特征图上进行多特征融合,提高了检测精度。3D-CV^[45]在多个阶段实现了特征融合,使用插值投影校正空间偏移,改善了坐标

转换导致的特征丢失问题。CLOCs^[46]将单个 2D 和 3D 预测框转换为一组联合预测框,大幅度抑制了计算过程中导致的预测框丢失问题。

VMVS^[47]使用图像对点云着色并绘制物体的多个视点来学习细节特征。类似的方法还有 PointPainting^[48],该方法利用图像分割后生成的掩码对应到点云完成着色,增强点云信息。在此基础上,SemanticVoxels^[49]将点云进一步体素化成柱子,提升了模型的泛化能力。这几种方法利用图像语义来增强点云,改善了点云语义信息较少的问题。

F-PointNets^[50]是以 2D 检测为驱动的代表,将 RGB 图像通过 2D 检测器得到 2D 预测框,然后从框中挤出目标的视锥体,如图 7 所示。在视锥体中,使用 PointNet 等网络对目标点云进行处理和检测。针对该方案,后续又提出了多种改良策略。F-ConvNet^[51]将视锥体轴以同样的步长滑动,形成一组视锥体,利用得到的视锥对局部点进行分组。F-PointPillars^[52]对每个 2D 检测创建掩码来表示属于该对象的可能性,通过构建对象掩码改善目标的定位,虽然性能有所提升,但是检测速度下降到了 17 帧每秒。

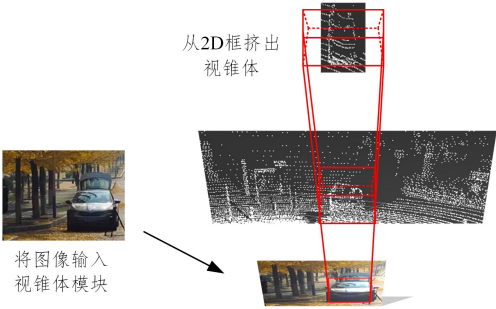


图 7 2D 引导的 3D 目标检测

Fig. 7 2D guided 3D target detection

不同融合方法在 KITTI 测试集的检测精度如表 3 所列。由该表可知,特征融合方法检测的速度较慢,这是由于融合方法需要考虑到激光雷达数据和图像数据分别来源于不同传感器,在处理时需要考虑到传感器校准同步,在特征融合时通常结构较为复杂。

综上所述,对点云采用不同的特征处理方法有着不同的优点与缺点,总结如表 4 所列。为了尽可能地提取点云的特征,现在的方法更趋向于采用多处理方式结合的模式,例如结合点和体素的处理方式,如 PV-RCNN^[53]和 P2V-RCNN^[54]等。

表 3 不同融合方法在 KITTI 测试集的检测精度

Table 3 Detection accuracy of different fusion methods on KITTI test set

方法	参考 文献	汽车/%			行人/%			骑行的人/%			每帧 用时/ms
		简单	中等	困难	简单	中等	困难	简单	中等	困难	
3D-CVF	[45]	89.20	80.05	73.11	—	—	—	—	—	—	75
AVOD	[44]	73.59	65.78	58.38	38.28	31.51	26.98	60.11	44.90	38.80	80
CLOCs	[46]	88.94	80.67	77.15	—	—	—	—	—	—	100
F-PointNet	[50]	81.20	70.39	62.19	51.21	44.89	40.23	71.96	56.77	50.39	200
VMVS	[47]	—	—	—	53.98	45.01	41.72	—	—	—	245
F-ConvNet	[51]	85.88	76.51	68.08	52.37	45.61	41.49	79.58	64.68	57.03	—
F-PointPillars	[52]	—	—	—	51.22	42.89	39.28	—	—	—	70
SemanticVoxels	[49]	—	—	—	50.90	42.19	39.52	—	—	—	—

表 4 不同特征处理方式对比

Table 4 Comparison of different feature processing methods

类型	优点	缺点
基于原始点的方法	能够保留点云的全部信息，不会造成信息损失	由于点云稀疏，不规则特性导致处理时间较长
基于体素的方法	将点云规则化，方便特征提取，处理速度快	体素化会不可避免地丢失原始点云的一部分信息，可能造成较大影响
基于视图的方法	体素化的另一种形式，通过压缩维度加快处理	
基于融合的方法	能够提供丰富的细节特征，弥补点云的不足	特征处理难度大，要解决特征融合及时间同步等问题

4 总结和展望

在对近年来相关方法研究分析的基础上,本文对其存在的共性问题及改进思路总结如下。

(1)数据单一。利用单一的传感器难以实现对周围环境全方位无死角的感知。传感器类型多样,例如将声音或者红外图像纳入参考,让机器得到更多有关环境的信息,就能够让模型适应更加恶劣的环境,从而做出更准确的判断。

(2)难以兼顾精确度和实时性。不同的任务对精确度和实时性的侧重点不同,例如自动驾驶领域对系统的实时性和准确性有着很高要求。如何兼顾精确度和实时性是 3D 目标检测任务中的一项重点。

(3)难以分辨细小的物体。不管是何种方法,越小的物体,识别的准确度越低,行人往往被忽略,这是现有方法的通病。小物体所含有的信息较少,点云也较为稀疏,从而导致方法无法获取足够多的信息来识别,多数据融合的方案可以作为改进的主要方法。

3D 目标检测还处于起步阶段,仍存在着许多的缺点和不足,并且受到计算机及感知系统性能的限制。随着科技的不断发展,硬件性能的大幅提高,及相关设备价格的下调,3D 目标检测会得到更好的发展,更加方便人类生活,为人类带来更大的社会及经济效益。

参 考 文 献

[1] LIU W,ANGUELOV D,ERHAN D,et al. Ssd:Single shot multibox detector[C]// European Conference on Computer Vision. 2016:21-37.

[2] REDMON J,DIVVALA S,GIRSHICK R,et al. You only look once:Unified,real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016:779-788.

[3] REN S,HE K,GIRSHICK R,et al. Faster r-cnn:Towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing systems,2015,28.

[4] MASI I,WU Y,HASSNER T,et al. Deep face recognition:A survey[C]//2018 31st SIBGRAPI Conference on Graphics,Patterns and mages(SIBGRAPI). IEEE,2018:471-478.

[5] HUANG R,GU J,SUN X,et al. A rapid recognition method for electronic components based on the improved YOLO-V3 network[J]. Electronics,2019,8(8):825.

[6] YURTSEVER E,LAMBERT J,CARBALLO A,et al. A survey of autonomous driving:Common practices and emerging technologies[J]. IEEE access,2020,8:58443-58469.

[7] LIANG M,YANG B,ZENG W,et al. Pnpnet:End-to-end perception and prediction with tracking in the loop[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:11553-11562.

[8] XIAO Y Q,YANG H M. Research on Application of Object Detection Algorithm in Traffic Scene[J]. Computer Engineering and Applications,2021,57(6):30-41.

[9] BRESSON G,ALSAYED Z,YU L,et al. Simultaneous localization and mapping:A survey of current trends in autonomous-driving[J]. IEEE Transactions on Intelligent Vehicles, 2017, 2(3):194-220.

[10] BOJARSKI M,DEL TESTA D,DWORAKOWSKI D,et al. End to end learning for self-driving cars[J]. arXiv:1604. 07316, 2016.

[11] HECHT J. Lidar for self-driving cars[J]. Optics and Photonics News,2018,29(1):26-33.

[12] PARK Y,YUN S,WON C S,et al. Calibration between color camera and 3D LIDAR instruments with a polygonal planar board[J]. Sensors,2014,14(3):5333-5353.

[13] SOMMER H,KHANNA R,GILITSCHENSKI I,et al. A low-cost system for high-rate,high-accuracy temporal calibration for LIDARs and cameras[C]// 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS). 2017:2219-2226.

[14] CUI Y,CHEN R,CHU W,et al. Deep learning for image and point cloud fusion in autonomous driving:A review[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(2):722-739.

[15] CHAZETTE P,TOTEMS J,HESPEL L,et al. Principle and physics of the LiDAR measurement[M]//Optical Remote Sensing of Land Surface. Elsevier,2016:201-247.

[16] HE H,WANG H,SUN L. Research on 3D point-cloud registration technology based on Kinect V2 sensor[C]//2018 Chinese Control And Decision Conference(CCDC). IEEE, 2018: 1264-1268.

[17] XU Y,JOHN V,MITA S,et al. 3D point cloud map based vehicle localization using stereo camera[C]// 2017 IEEE Intelligent

- Vehicles Symposium(IV). IEEE, 2017:487-492.
- [18] LI G Y, LI M L, WANG L, et al. A Review of Preprocessing of Laser-scanned Point Clouds Data[J]. Bulletin of Surveying and Mapping, 2015(11):1-3.
 - [19] XU Z, ZHANG Z, ZHONG R, et al. Content-sensitive multilevel point cluster construction for ALS point cloud classification[J]. Remote Sensing, 2019, 11(3):342.
 - [20] RUSU R B, COUSINS S. 3d is here: Point cloud library(pcl) [C]//2011 IEEE International Conference on Robotics and Automation. 2011:1-4.
 - [21] GEIGER A, LENZ P, URTASUN R. Are we ready for autonomous driving? the kitti vision benchmark suite[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition. 2012:3354-3361.
 - [22] CAESAR H, BANKITI V, LANG A H, et al. nuscenes: A multimodal dataset for autonomous driving[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:11621-11631.
 - [23] SUN P, KRETZSCHMAR H, DOTIWALLA X, et al. Scalability in perception for autonomous driving: Waymo open dataset [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:2446-2454.
 - [24] QI C R, SU H, MO K, et al. Pointnet: Deep learning on point sets for 3d classification and segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017:652-660.
 - [25] QI C R, YI L, SU H, et al. Pointnet++: Deep hierarchical feature learning on point sets in a metric space[C]//Conference and Workshop on Neural Information Processing Systems. 2017.
 - [26] SHI S, WANG X, LI H. Pointnet: 3d object proposal generation and detection from point cloud[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019:770-779.
 - [27] YANG Z, SUN Y, LIU S, et al. 3dssd: Point-based 3d single stage object detector[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:11040-11048.
 - [28] ZHOU Y, TUZEL O. Voxelnet: End-to-end learning for point cloud based 3d object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:4490-4499.
 - [29] YAN Y, MAO Y, LI B. Second: Sparsely embedded convolutional detection[J]. Sensors, 2018, 18(10):3337.
 - [30] LANG A H, VORA S, CAESAR H, et al. Pointpillars: Fast encoders for object detection from point clouds[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019:12697-12705.
 - [31] YE M, XU S, CAO T. Hynet: Hybrid voxel network for lidar based 3d object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:1631-1640.
 - [32] SHI S, WANG Z, SHI J, et al. From points to parts: 3d object detection from point cloud with part-aware and part-aggregation network[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43(8):2647-2664.
 - [33] HE C, ZENG H, HUANG J, et al. Structure aware single-stage 3d object detection from point cloud[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:11873-11882.
 - [34] GUSTAFSSON F K, DANELLJAN M, SCHON T B. Accurate 3D object detection using energy-based models[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021:2855-2864.
 - [35] ZHENG W, TANG W, CHEN S, et al. CIA-SSD: Confident IoU-aware single-stage object detector from point cloud[J]. arXiv: 2012.03015, 2020.
 - [36] ZHENG W, TANG W, JIANG L, et al. SE-SSD: Self-ensembling single-stage object detector from point cloud[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021:14494-14503.
 - [37] LI Z, YAO Y, QUAN Z, et al. Sienet: spatial information enhancement network for 3d object detection from point cloud[J]. arXiv:2103.15396, 2021.
 - [38] CHEN Q, SUN L, WANG Z, et al. Object as hotspots: An anchor-free 3d object detection approach via firing of hotspots [C]//European Conference on Computer Vision. 2020:68-84.
 - [39] QIAN R, LAI X, LI X. Boundary-aware 3d object detection from point clouds[J]. arXiv:2104.10330, 2021.
 - [40] LI J, DAI H, SHAO L, et al. From voxel to point: IoU-guided 3d object detection for point cloud with voxel-to-point decoder [C]//Proceedings of the 29th ACM International Conference on Multimedia. 2021:4622-4631.
 - [41] DENG J, ZHOU W, ZHANG Y, et al. From Multi-View to Hol-low-3D: Hallucinated Hollow-3D R-CNN for 3D Object Detection[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(12):4722-4734.
 - [42] XIE L, XU G, CAI D, et al. X-view: Non-egocentric Multi-View 3D Object Detector[J]. arXiv:2103.13001, 2021.
 - [43] CHEN X, MA H, WAN J, et al. Multi-view 3d object detection network for autonomous driving[C]//Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. 2017:1907-1915.
 - [44] KU J, MOZIFIAN M, LEE J, et al. Joint 3d proposal generation and object detection from view aggregation[C]//2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2018:1-8.
 - [45] YOO J H, KIM Y, KIM J, et al. 3d-cvf: Generating joint camera and lidar features using cross-view spatial feature fusion for 3d object detection[C]//European Conference on Computer Vision. 2020:720-736.
 - [46] PANG S, MORRIS D, RADHA H. CLOCs: Camera-LiDAR object candidates fusion for 3D object detection[C]//2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2020:10386-10393.
 - [47] KU J, PON A D, WALSH S, et al. Improving 3d object detection for pedestrians with virtual multi-view synthesis orientation estimation[C]//2019 IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS). 2019:3459-3466.
 - [48] VORA S, LANG A H, HELOU B, et al. Pointpainting: Sequential fusion for 3d object detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.

nition. 2020;4604-4612.

[49] FEI J, CHEN W, HEIDENREICH P, et al. SemanticVoxels: Sequential fusion for 3D pedestrian detection using LiDAR point cloud and semantic segmentation[C]//2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems(MFI). 2020;185-190.

[50] QI C R, LIU W, WU C, et al. Frustum pointnets for 3d object detection from rgb-d data[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018;918-927.

[51] WANG Z, JIA K. Frustum convnet: Sliding frustums to aggregate local point-wise features for amodal 3d object detection [C]//2019 IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS). 2019;1742-1749.

[52] PAIGWAR A, SIERRA-GONZALEZ D, ERKENT Ö, et al. Frustum-pointpillars: A multi-stage approach for 3d object detection using rgb camera and lidar [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021; 2926-2933.

[53] SHI S, GUO C, JIANG L, et al. Pv-rcnn: Point-voxel feature set abstraction for 3d object detection [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020;10529-10538.

[54] LI J, SUN Y, LUO S, et al. P2V-RCNN: Point to Voxel Feature Learning for 3D Object Detection from Point Clouds[J]. IEEE Access, 2021, 9: 98249-98260.



QIN Jing, born in 1981, Ph.D, associate professor, is a member of China Computer Federation. Her main research interests include signal processing and big data analysis.



WANG Zumin, born in 1975, Ph.D, professor, is a member of China Computer Federation. His main research interests include smart cities and internet of things.