

三维激光雷达在无人车环境感知中的应用研究

张银¹, 任国全^{1*}, 程子阳¹, 孔国杰²

¹陆军工程大学石家庄校区车辆与电气工程系, 河北 石家庄 050003;

²中国人民解放军 63963 部队, 北京 100072

摘要 环境感知是无人驾驶的核心技术之一, 而利用三维激光雷达进行障碍物检测一直是国内外的研究热点。本文首先按照传感器的种类介绍了无人车障碍物检测方法的分类, 然后介绍了基于三维激光雷达进行障碍物检测的基本原理, 之后详细分析了基于三维激光雷达进行障碍物检测的传统方法。其中深度学习是二维图像目标检测及分类的重要方法, 在介绍三维激光雷达点云特点的同时分析了点云深度学习的挑战, 最后详细分析了三维点云深度学习在障碍物检测方面的研究现状以及发展趋势, 并且介绍了自动驾驶领域的 KITTI 数据集和 ApolloScape 数据集。

关键词 遥感; 无人驾驶; 三维激光雷达; 障碍物检测; 栅格网络; 深度学习

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP56.130001

Application Research of Three-Dimensional LiDAR in Unmanned Vehicle Environment Perception

Zhang Yin¹, Ren Guoquan^{1*}, Cheng Ziyang¹, Kong Guojie²

¹Department of Vehicle and Electrical Engineering, Shijiazhuang Campus of Army Engineering University, Shijiazhuang, Hebei 050003, China;

²The No. 63963rd Troop of PLA, Beijing 100072, China

Abstract The environmental perception of unmanned vehicles is a vital technology for automatic driving. The usage of three-dimensional (3D) LiDAR for obstacle detection becomes a popular research topic. In this paper, we first introduce the classification of obstacle detection methods for an unmanned vehicle according to different sensors. The basic principle of obstacle detection based on 3D LiDAR is then introduced in detail along with an analysis of the traditional method of obstacle detection using 3D LiDAR. Deep learning is an important method for two-dimensional object detection and classification. We analyze the characteristics of the 3D LiDAR point clouds and the challenges of deep learning for point clouds. Finally, we analyze the research status and development trend of deep learning in the point cloud obstacle detection application and introduce relevant datasets in the field of automatic driving, such as KITTI and ApolloScape.

Key words remote sensing; automatic driving; three-dimensional LiDAR; obstacle detection; grid network; deep learning

OCIS codes 280.3400; 140.6910

1 引言

环境感知是实现无人驾驶的核心技术之一, 其任务是通过车载传感器来识别周围的环境信息, 从而为无人驾驶汽车(无人车)的导航定位、路径规划

和决策控制提供信息支撑^[1-2]。对周围环境感知的目标主要包括静态障碍物(路面、路牙、植物、建筑等)和动态障碍物(车辆、行人、自行车等)。无人车障碍物检测是指通过分析传感器获取的道路环境信息, 判别车辆行驶过程中静态障碍物和动态障碍物

收稿日期: 2018-12-12; 修回日期: 2019-01-11; 录用日期: 2019-01-21

基金项目: 国防预研基金(9140A09031715JB34001)

* E-mail: rrrgggqq@163.com

的位置、大小、种类等。障碍物检测是无人车环境感知的关键步骤,随着无人驾驶技术的深入发展,三维激光雷达在障碍物检测方面的重要性也愈发突显,已应用于 Waymo 无人车和百度无人车之中。

目前,三维激光雷达在无人车环境感知中主要应用于车辆行驶过程中的障碍物检测。传统利用三维激光雷达进行障碍物检测的方法是先对采集的点云进行地面分割,再进行障碍物聚类。传统方法可分为基于栅格的方法、基于分类器的方法和基于模型的方法。传统方法的稳健性高,可以适应复杂的地形环境,但其计算量大,实时性较差。应用深度学习处理三维点云来进行障碍物检测的方法主要可分为直接法、投影法和特定网络法,该类检测方法的准确率较高、耗时少,但是在前期训练模型时,训练时间长,数据量大。

本文总结了基于三维激光雷达进行障碍物检测的研究进展,包括传统方法和深度学习方法,并分析了各类方法的优点以及存在的问题,同时对今后的研究方向也提出了建议。

2 障碍物检测方法分类

目前用于无人车环境感知的车载传感器主要包括被动型传感器(白光、红外摄像机等)和主动型传感器(毫米波雷达、激光雷达等)。而按照传感器种类的不同又可以分为基于摄像机的障碍物检测、基于毫米波雷达的障碍物检测和基于激光雷达的障碍物检测^[3-6]。

摄像机虽然能够采集障碍物的 RGB 信息(颜色、纹理等特征),实现对障碍物的检测及跟踪,但却很难得到障碍物的三维模型,且整个采集过程受光照、自然环境的影响较大,如在阴暗或者光线突变的环境中障碍物检测效果不佳。毫米波雷达技术是通过测量障碍物的距离、速度信息来判断是否为障碍物的,被广泛用于自动驾驶技术中的自适应巡航控制(ACC)、前向防撞报警(FCW)、盲点检测(BSD)、辅助停车(PA),以及辅助变道(LCA)等方面。目前常用车载毫米波雷达的波段有 24, 60, 77 GHz, 尽管其探测距离最远可达 200 m, 但发射的电磁波在遇到比其波长小的物体时,将会发生衍射现象,即波的大部分能流绕过物体继续向前方传播,而反射回来可供雷达接收的能量很少,因此无法探测行人等小目标,且由于其分辨率不高,无法实现对行人及周围小目标的三维建模。

激光雷达(LiDAR)作为一种主动型传感器,按照有无机械旋转部件可分为机械激光雷达和固态激光雷达,按照激光光束的数量可分为单线激光雷达和多线激光雷达。三维激光雷达(3D LiDAR)通过发射光束来探测障碍物的距离、角度、速度,以及反射强度等信息,得到具有非结构化、稀疏性、不均匀等特点的点云,进而生成障碍物的三维图像,如图 1 所示。利用激光雷达进行障碍物检测可以排除光线的干扰,具有测量精度高、方向性好、抗干扰能力强等优点。目前市场上除了老牌的美国 Velodyne(16 线、32 线及 64 线等)、德国的 IBEO 等激光雷达供应厂商外,还有国内的禾赛科技、镭神智能等激光雷达厂商。尽管市场上现有的用于自动驾驶的激光雷达价格普遍较高,但是随着自动驾驶技术商业化进程的不断推进,其价格必定会随之下降,同时,采用激光雷达进行障碍物检测的效果也必将会越来越好。未来车载激光雷达势必向着多线、固态式的方向发展。

对于无人驾驶场景,三维激光雷达在环境感知中主要用于障碍物检测、跟踪,以及定位构图。障碍物(汽车、行人及自行车等)检测只需通过物体外形轮廓就可以加以区分,三维激光雷达通过激光数据就能够轻易区分障碍物的外部轮廓,便于进行障碍物检测、分类、定位,以及动态障碍物跟踪。相比较而言,摄像头采集的 RGB 信息(颜色、纹理等特征)对于障碍物检测的意义就不是很大。因此,在无人车上使用三维激光雷达进行障碍物检测已经成为了趋势。

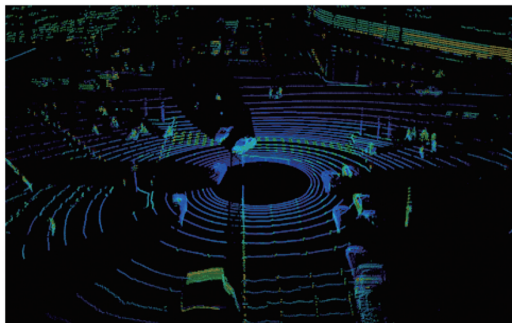


图 1 三维激光雷达点云图

Fig. 1 Point clouds of 3D LiDAR

3 基于三维激光雷达的障碍物检测研究现状

3.1 障碍物检测基本原理

障碍物检测是无人车环境感知中的关键环节,利用三维激光雷达进行障碍物检测具有较高的检测

精度和较好的环境适应性。其检测的基本原理如图2所示。1) 用已经标定好外参数的三维激光雷达对周围环境进行360°扫描,获取原始点云。2) 对原始点云进行预处理,排除因激光雷达安装、标定,以及行驶过程中的颠簸而造成雷达扫描角度和地面投影角度不一致的情况。3) 根据点云的高度、密度,以及反射强度等信息,对点云进行地面分割处理,分离出全部点云中的地面点和障碍物点。4) 对障碍物点进行聚类分析,提取相应的障碍物特征,对路面上的障碍物进行分类。其中地面分割的主要目的是分离出点云中的地面点和障碍物点,为下一步的障碍物检测等步骤奠定基础。地面分割的结果将直接影响到后续的障碍物检测、可通行区域提取,以及动态障碍物跟踪的效果。

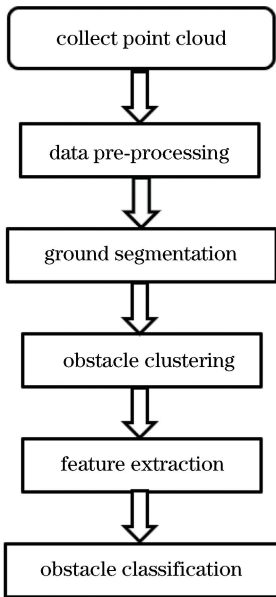


图2 障碍物检测基本原理

Fig. 2 Basic principle of obstacle detection

3.2 障碍物检测研究现状

3.2.1 基于栅格的方法

对于自动驾驶场景中的障碍物检测,传统的检测方式主要是基于栅格地图的方法:将标定过的三维激光雷达扫描得到的点云投影到2D或3D栅格中,按照扫描线遍历所有栅格,根据每个栅格内的密度、平均高度、高度方差、最大高度差及直方图等信息,进行聚类分析,判断对应的栅格是障碍物还是地面。栅格划分的本质是将非结构化、稀疏的点云进行结构化和信息压缩处理。作为最小计算单元的栅格,其大小是否合适将直接影响到检测的准确率。

基于栅格的方法最早是在2005年美国国防部高级研究计划局(DARPA)举办的无人车挑战赛中

得到了应用。斯坦福大学研制的Stanley无人车^[7]是将激光雷达扫描前方路面得到的点云投影到栅格中,根据相邻点间的高度差以及一阶马尔科夫模型进行快速地面分割,该方法在野外环境中能快速进行地面分割,但在城市道路环境中却容易漏检小体积障碍物、悬空障碍物(树冠、桥梁等)。

在2007年DARPA举办的城市无人车挑战赛中,斯坦福大学研制的Junior无人车^[8]利用激光雷达的扫描特性,采用了一种新的地面分割方法:一束激光雷达在平地上扫描一周,返回坐标点形成一个圆环,判断相邻光束落在平面上的距离差是否比落在障碍物上的距离差大。根据实际检测距离计算相邻圆环间期望的距离差,可以补偿车身倾斜的影响,如图3所示。

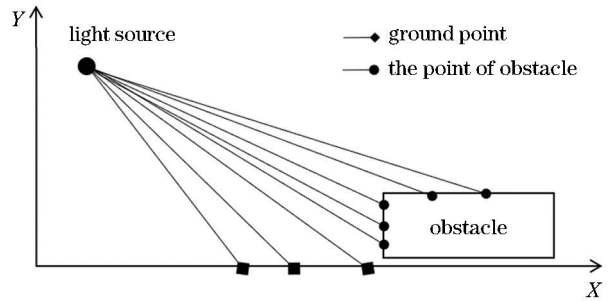


图3 三维激光雷达扫描特性

Fig. 3 3D LiDAR scanning characteristics

Douillard等^[9]则提出了一种平均高度和最大、最小高度相结合的地面分割方法:1) 按照点云图大小设置合适的二维网格划分,计算每个网格内点的平均高度,与相邻的4个网格做差,保留最大差值作为高度梯度,将差值小于一定阈值的点标记为地面点,并且将对点连成连通域,其中面积最大的作为参考地面。2) 将高度梯度大于阈值的点标记为障碍物点,并且将其连通域面积最大的标记为障碍物。但在实际检测中却存在障碍物点与地面点的平均高度小于阈值,而仍将障碍物点错误地标记成地面点的情况。因此需要再次以更小尺寸的网格重新划分先前标记为障碍物及其相邻的区域,用最大高度差筛选障碍物点,并将相邻的障碍物聚类,而后根据每个网格用体素(voxel)表示出高度的占据情况,将地面高度附近的体素标记为地面,并重新计算地面高度。该方法利用这种混合高度图(Hybrid Elevation Maps)能够提取地面,对障碍物点进行聚类分析。

3.2.2 基于分类器的方法

在基于栅格方法的基础上,为了消除使用单一特征进行检测的局限性,研究者们提出了基于分类

器的方法,即提取更多的特征(如:局部法向量),使用分类器^[10-12]进行训练,根据其提取的特征优化分类判断地面点和非地面点,得到了更好的障碍物检测效果。Samples等^[13]在Junior无人车中采用地面分割方法的基础上,通过统计每个网格的地面点和障碍物点的平均高度、方差、全局地面的先验高度信息和先验高度不确定度共6个特征,利用条件随机场(CRF)对网格地面高度进行优化,根据优化后的网格高度划分地面点和障碍物点。任国亭等^[14]提出了一种基于支持向量机(SVM)的可通行区域检测方法,该方法对使用支持向量机对多特征融合的样本集进行训练,利用训练好的模型对激光雷达数据进行检测,其结果明显优于传统的基于高度差的道路检测方法。McDaniel等^[15]将树干建模为圆柱或圆锥体,提取相应特征,利用支持向量机分类器将地面点和树干点进行分类。Spinello等^[16]提出了一种基于雷达点云的方向梯度直方图特征(HOG)进行检测物体位置的SVM分类器。Azim等^[17]则在三维激光雷达数据中采用矩形框表示障碍物,继而提出了一种基于长高比和宽高比两个特征分类的障碍物检测方法,该方法虽然检测效果较好,但却不能解决障碍物之间遮挡的问题。相比较而言,使用SVM分类器的方法,其检测效果要优于基于CRF优化的方法以及传统的基于高度差的方法,但其仍存在耗时长缺点,不适合进行实时检测。

3.2.3 基于模型的方法

对于野外起伏地形环境,相关研究人员提出了基于模型的方法,该方法主要采用分段平面拟合(RANSAC)、马尔科夫随机场(MRF)和高斯回归过程(GP-INSAC)等理论^[18-21],来达到良好的地面分割效果的目的。尽管大多数地面可以用平面进行表示,但由于实际路面存在起伏,因此可以将实际路面看成是由一个个小的平面拼接而形成的。

湛彤童^[5]针对野外起伏路面环境,提出了将极坐标栅格地图和一维高斯回归过程理论结合起来进行地面检测的方法,该方法可用于复杂的道路情况,计算量少,实时性好。吴倩倩等^[22]提出的方法利用高斯回归过程和马尔科夫随机场重建三维图像。Himmelsbach等^[23]利用分块直线拟合的方法进行地面分割,而后根据二维连通性对障碍物进行聚类分析,实验表明该方法能得到较好的障碍物检测效果。王肖等^[24]通过分块直线拟合算法剔除三维激光雷达数据中的地面点,根据计算机视觉算法对行人和车辆进行聚类分析,该方法有效地降低了过分

割和欠分割的错误率,提高了车辆和行人检测的准确率。Asvadi等^[25]在RANSAC方法和分段平面拟合算法进行地面分割的基础上,对基于栅格划分的三维点云模型进行判别分析,以检测静态和动态障碍物。

传统的基于三维激光雷达的障碍物检测方法主要包括地面分割、障碍物聚类两步操作。该类方法通常适用于单一的道路场景,在复杂道路场景下还需要结合多种理论方法进行障碍物检测,才能满足检测精度要求。在地面分割操作过程中,常常根据激光雷达的扫描特性、人工特性或决策规则的集合来进行地面点的提取,而障碍物的分类、聚类通常依据非线性拟合能力强以及小样本学习能力好的RANSAC、GP-INSACA、MRF等迭代优化算法来进行;另外,算法的随机初始化质量也会直接影响检测的速度和精度。

4 基于深度学习的三维点云障碍物检测

随着卷积神经网络(CNN)在ImageNet比赛^[26]中的广泛应用,深度学习(DL)又迎来了一波新的发展热潮。近年来,基于卷积神经网络架构的深度学习已被广泛应用在二维图像目标的检测及分类^[27]当中,其检测的准确率及耗时明显优于传统的基于机器学习的检测方法,使用基于卷积神经网络的深度学习已经成为图像分类任务的标准解决方案。

深度学习与传统模式识别方法的最大区别在于其是从大数据中自动学习特征,而非采用人工提取特征。基于激光雷达点云进行障碍物检测的方法是通过线下训练的卷积神经网络模型,学习点云的特征并预测障碍物的相关属性(比如前景物体概率、相对于物体中心的偏移量、物体高度等),进而根据这些属性进行障碍物检测。目前,越来越多的学者开始研究用深度学习的方法处理三维点云,以提高无人车环境感知中障碍物检测的效率(检测的时间更短且准确率更高)。

4.1 三维点云的特点

在无人驾驶技术领域,仅使用三维激光雷达就可以对行人、汽车、自行车,以及道路两旁的树木、路灯等障碍物进行检测,且检测效果较好。三维点云通过深度学习的网络结构就可以很好地学习到相关的空间几何特征。但由于三维点云具有稀疏性、无序性,以及信息表示的有限性等特点,导致在三维点

云上用深度学习的方法进行物体检测具有挑战性。

1) 点云具有稀疏性。在自动驾驶领域常用数据集 KITTI, 如果把原始的激光点云投影到对应的彩色图像中, 大约只有 3% 的像素才有对应的点云。点云的数量会因为传感器或采集等原因而发生变化。相比较而言, 在图像中, 像素的数量一般是一个定值。激光雷达采集的点云具有很强的稀疏性, 会导致基于点云的高层语义感知变得更加困难。另外, 三维点云在三维空间中是不规则分布的, 在空间和数量上可以看作是任意分布的, 所以需要先将其转化成结构化的形式(栅格化处理), 再用三维卷积神经网络(3D CNN)来处理数据。

2) 点云具有无序性。受到三维激光雷达的安装位置以及坐标系的影响, 同一障碍物利用不同的扫描设备或者从不同的方向扫描, 得到的三维点云排列顺序千差万别。在几何上, 点的顺序不影响其在底层矩阵结构中的表示方式, 例如相同的点云可以由两个完全不同的矩阵表示。这就导致三维点云很难直接通过端到端的模型处理。

3) 三维点云表示的信息有限。由于三维物体的自遮挡, 三维激光雷达只能扫描到三维物体的一部分, 这就导致点云信息只能表示被测物体的部分几何信息, 从而给利用三维激光雷达点云进行物体检测带来了更多的困难。

4.2 三维点云深度学习研究现状

近年来, 越来越多的学者开始研究用深度学习

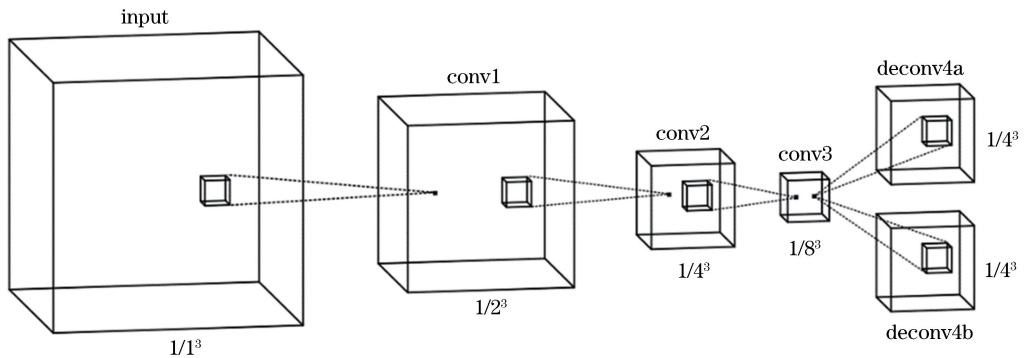


图4 基于3D FCN的车辆检测网络结构

Fig. 4 Vehicle detection network structure based on 3D FCN

4.2.2 投影法

投影法: 1) 将三维点云投影到某个平面上, 比如鸟瞰图(BEV)^[32-36]、前视图(FV)^[37]。2) 进行特征提取及检测。鸟瞰图包含有前视图, 但不具有空间深度信息, 采用前视图的方法一般会补充深度信息(幅度)特征, 以弥补障碍物检测过程中信息不足的缺陷。虽然在投影过程中会存在某一维度的信息

方法处理三维点云。无人车在环境感知中对于障碍物的检测方法可以总结为3类: 直接法、投影法和特定网络法。

4.2.1 直接法

直接法: 在三维点云上直接使用3D CNN进行物体检测。由于点云非常稀疏, 并且3D CNN需要在3个维度上操作, 整个检测、定位的过程分外耗时, 典型的方法有: 3D FCN^[28-29]、Vote3Deep^[30]。2017年, Li^[29]提出了一种基于3D全卷积神经网络(3D FCN)来进行车辆检测的方法, 该方法仅采用三维激光雷达数据作为输入, 将三维激光雷达扫描的区域分割为立体栅格空间, 生成简化信息的4D图像。基于3D FCN的车辆检测网络结构如图4所示(引用于Li的3D FCN^[29]中), 其流程为: 1) 通过三次卷积运算(conv1, conv2, conv3)对特征图以 $1/2^3$ 步长进行下采样, 每一层的激活函数均为Relu函数, 同时在每个卷积层上都对输出结果进行归一化。2) 利用相同步长的反卷积运算进行上采样, 分别从deconv4a和deconv4b层输出目标分类图和目标边界框图。Engelcke等^[30]提出了一种利用以特征为中心的投票构造的卷积神经网络架构(Vote3Deep)的方法, 该方法能够在三维点云上直接使用稀疏卷积层对障碍物进行检测, 平均检测精度可达40%, 但其仍存在检测速度慢的问题。Maturana等^[31]则将体素占用网格表示与3D CNN相结合, 提出了一种可用于三维点云的实时高效的目标检测方法。

丢失的情况, 但就自动驾驶而言, 所有的障碍物几乎都位于同一平面, 因此高度信息的部分丢失不会影响到检测结果。鸟瞰图虽然能够很好地表示车辆等较大目标的外观形状, 但对于行人、路灯等细长类型障碍物的检测效果较差。

Yang等^[32]利用通过三维点云得到的以高度信息和反射率为通道的鸟瞰图来进行障碍物的检测与

定位。Zeng 等^[33]则以栅格内点的最大、最小,以及平均高度信息将三维点云投影为三通道的鸟瞰图,使用基于 Faster R-CNN 改进的 R-FCN 网络进行快速障碍物检测,该方法(RT3D)在投影为鸟瞰图的过程中高度信息丢失较多,导致检测精度下降,但其检测速度较快。Wu 等^[34]构建了一个基于卷积神经网络和条件随机场的端到端检测模型(SqueezeNet):1)将三维点云通过球面投影得到前视图。2)使用基于 SqueezeNet 的卷积神经网络对输入图像进行特征提取与分割。3)将 CRF 作为 RNN 层对分割结果进一步优化。Caltagirone 等^[35]提出了一种基于完全卷积神经网络的道路检测方法,即将非结构化点云生成编码平均高度和密度等几种基本统计数据的鸟瞰图。将道路检测简化为单一尺度问题,从而通过简单快速的完全卷积神经网络来完成处理运算。此外,研究人员还将道路检测问题使用完全卷积神经网络构建为点云鸟瞰图图像中的像素方式语义分割任务。Beltrán 等^[36]基于三维点云的鸟瞰图提出了 BirdNet 网络,不同于其他利用鸟瞰图进行检测的方法,BirdNet 网络先进行 2D 边框回归再进行 3D 边框回归。BirdNet 网络首先将三维点云以高度、强度和密度为通道投影得到鸟瞰图,针对不同的 LiDAR 在相同的场景下点云密度不相同,且在同一网络下是不通用的问题,将点云密度进行归一化处理,然后使用 Faster R-CNN 在鸟瞰图上进行二维物体检测,最后结合检测结果和地面分割结果,离线进行三维物体检测。Minemura 等^[37]提出了一种新的三维物体检测网络(LMNet),该网络是目前最快的三维物体检测网络,但其检测精度较低。LMNet 网络通过三维点云的柱面投影得到前视图,使用带有扩张卷积

(Dilated Convolution)扩大感受野的 FCN 结构进行三维物体检测。为了弥补三维信息投影过程中的损失,在 LMNet 网络中还加入了反射强度、幅度,以及点的三维坐标等特征。

4.2.3 特定网络法

特定网络法:针对点云提出特定的深度学习网络结构,实现对三维物体的检测。比如 PointNet^[38]、PointNet++^[39]、VoxelNet^[40]等。利用这些网络模型的主要目的是为了高效地从三维点云中学习到更有效的空间几何表示。对于三维点云深度学习方法的研究中存在的困难,斯坦福大学的研究者在 2016 年提出了一种直接处理无序三维点云的深度神经网络—PointNet,具体网络架构如图 5 所示(引用于 Charles 等^[38]的 PointNet 中)。PointNet 是首个可以直接处理原始三维点云的网络结构,其处理流程为:1)按照一定的规则输入含有 n 个点的三维点云,由微型网络 T-Net 生成一个变换矩阵,将输入的三维点云数据对齐处理,再通过一个共享参数的双层感知机模型[$\text{mlp}(64, 64)$]进行每个点 64 维特征提取(局部特征)。2)第二个 T-Net 网络生成特征空间变换矩阵对齐输入的特征,再利用三层感知机模型[$\text{mlp}(64, 128, 1024)$]进行特征提取,将特征维度升到 1024 维。3)在特征空间上使用最大池化(max pool)提取三维点云的全局特征(1×1024),再利用一个 $\text{mlp}(512, 256, k)$ 进行融合,通过训练分类器实现逐点分类,网络输出对应 k 个分类类别的概率。对于分割问题,结合局部点特征和全局特征,训练网络实现所有 n 个点对 m 个语义类别的概率,即输出尺寸为 $n \times m$ 。该网络结构的特点是:1)通过 T-Net 微型网络得到了两个空间变换矩阵(3D 空

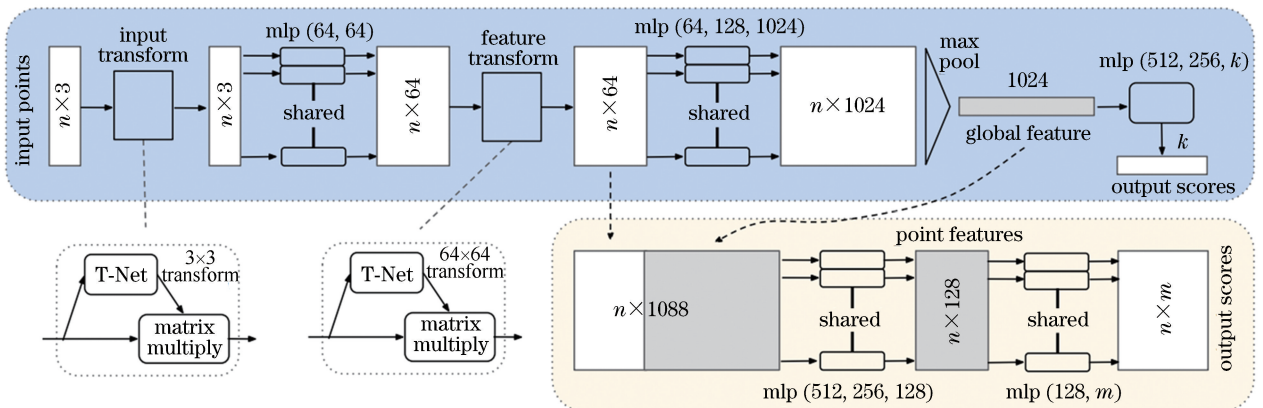


图 5 PointNet 网络架构

Fig. 5 Network structure of PointNet

间变换矩阵和特征空间变换矩阵),可将三维点云调整到一个更有利于分类和分割的角度。2)采用最大池化(max pool)对三维点云提取整体特征,解决了三维点云的无序性问题。

PointNet 网络大都是对三维点云中的单个采样点进行处理,而后再将所有的点连接起来,该网络只考虑了每个点的全局特征,对于局部信息考虑不足。针对此问题,众多学者在局部特征提取方面进行了很多研究,提出了以 PointNet++ 为代表的网络结构。该网络结构使用分层抽取特征思想,首先选取一些比较重要的点作为每个局部区域的中心点,再在这些中心点周围选取 k 个邻近点作为局部点云,最后采用 PointNet 网络提取这些局部点云的特征。PointNet++ 网络结构解决了三维点云采集密度不均匀的问题。

2017 年,苹果公司的 Zhou 等^[40]提出了一种直接处理点云的端到端检测网络(VoxelNet),其架构如图 6 所示(引用于 Zhou 的 VoxelNet 中)。该方法仅依靠三维点云就能实现对车辆、行人等障

碍物较好的检测效果,避免了因人工变换特征而导致的检测精度不佳等问题。该架构由特征学习网络、卷积中间层,以及候选区域网络(RPN)3 个功能模块组成。特征学习网络首先在原始点云信息(深度 D ,高度 H ,宽度 W)上按照体素划分(voxel partition),在体素化过程中,每个体素尺寸表示为 $v_D \times v_H \times v_w$,将点云按照在体素空间中的疏密分布进行分组(grouping),再利用随机采样(random sampling)减少计算量,接着引入体素特征编码(VFE)层,提取每个体素内的逐个点特征以及局部聚合特征,将该空间表示为稀疏 4D 张量($C \times D' \times H' \times W'$,其中 C 表示维度, $D' = D/v_D$, $H' = H/v_H$, $W' = W/v_w$);在卷积中间层中采用 3D 卷积层、BN 层和 ReLU 层,对特征学习网络输出的稀疏 4D 张量特征,通过逐渐扩大的感受域聚合体素特征;最后连接到 RPN 上生成障碍物分类检测和位置回归结果。随后,在 VoxelNet 的基础上,苹果公司对其进行改进,推出了障碍物检测效果更佳的 VoxelNet++。

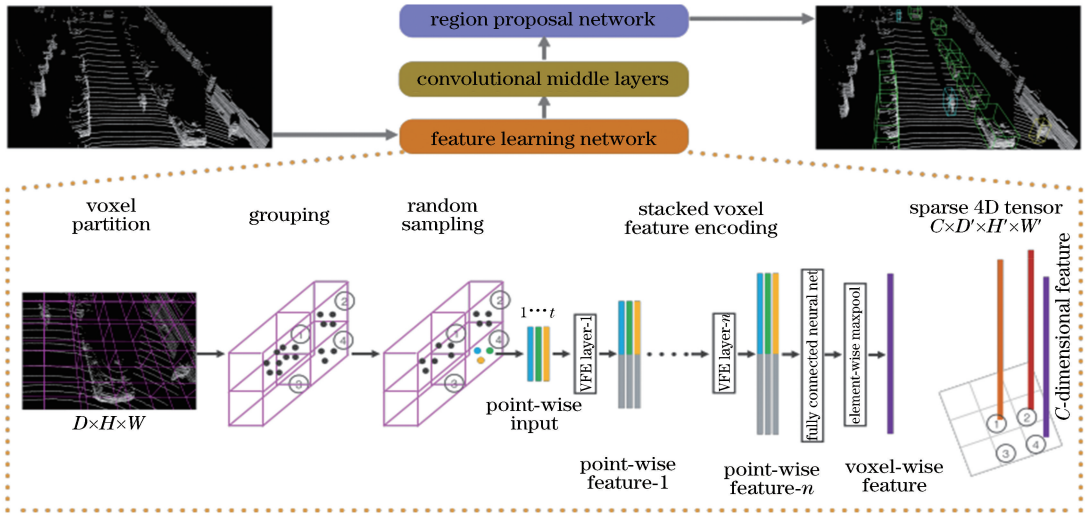


图 6 VoxelNet 网络架构

Fig. 6 Network structure of VoxelNet

4.3 点云深度学习数据集

数据是深度学习的 3 大核心要素之一。无人驾驶领域的三维物体检测常用的数据集包括 KITTI 数据集^[41]和 ApolloScape 数据集^[42]。

KITTI 数据集是由德国卡尔斯鲁厄理工学院(KIT)和丰田美国技术研究院(TTIC)于 2012 年联合创办的,也是目前国际上最大的自动驾驶场景下的计算机视觉算法评测数据集。该数据集包含市区、乡村和高速公路等真实场景下采集的数据(包含图像和激光雷达数据),可用于评测目标(车辆、自行车及行人

等)检测、目标跟踪(车辆跟踪、行人跟踪)和路面分割等计算机视觉技术在车载环境下的性能。针对三维物体检测,该数据集可提供三维激光雷达高精度的三维数据,并且能对路面上的车辆、自行车,以及行人进行相应的三维坐标位置标注。KITTI 数据集采集平台是由一辆大众帕萨特旅行车改装而成的,装备有 1 个 Velodyne64 线激光雷达、2 个灰度摄像机、2 个彩色摄像机、4 个光学摄像头,以及 1 个 GPS 导航系统,如图 7 所示(引用于 KITTI^[41]中)。

ApolloScape 是百度 Apollo 自动驾驶开放平台



图7 KITTI数据集采集平台

Fig. 7 Collection platform of KITTI dataset

的专题项目之一,也是目前行业内环境最复杂、标注最精准、数据量最大的三维自动驾驶开放数据集,该数据集包含有感知、仿真场景、路网数据等数十万帧逐像素语义分割标注的高分辨率图像数据,以及与其对应的像素语义标注、稠密点云、立体图像和立体全景图像。ApolloScape采用装备有RIEGL VMX-1HA移动测绘系统的多功能越野车进行数据采集,该系统包括2个LiDAR传感器(每秒500条扫描线,覆盖420 m内的360°视角)、1部INS/GNSS单元,以及2个前向相机(VMX-CS6, 3384×2710),采集平台如图8所示(引用于ApolloScape^[42]中)。ApolloScape中的三维点云相较于KITTI数据集中的点云要更加精确、稠密,标注的精细度也明显优于KITTI数据集。



图8 ApolloScape数据采集平台

Fig. 8 Collection platform of ApolloScape dataset

5 结束语

三维激光雷达是一种用于精确获取目标三维位置信息的传感器,近年来,已被广泛应用于自动驾驶技术中的障碍物检测之中。而研究更加高效的深度学习网络,以用于无人驾驶领域的环境感知,从而提高障碍物检测的精度和实时性,已经成为计算机视

觉领域的研究热点。但由于三维点云的数据量较大,应用深度学习进行障碍物检测时,对于计算的要求较高,因此,可以先确定待检测物体相关区域,再对相关区域进行障碍物检测,以减少计算时长,同时也可提高检测效果。另外,仅使用三维点云来进行环境感知还存在一定的缺陷,比如:交通信号灯及交通标志无法利用三维激光雷达进行检测,需要依靠相机来感知;同时,激光雷达受大雾等极端天气影响较大,需要使用毫米波雷达来辅助避障。综上所述,对于无人驾驶场景环境感知,以三维激光雷达为主的多传感器数据融合感知方式将是未来发展的重点,也是无人车实现完全自主行驶的必经之路。

参 考 文 献

- [1] Wang S F, Dai X, Xu N, *et al.* Overview on environment perception technology for unmanned ground vehicle[J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2017, 40(1): 1-6.
王世峰,戴祥,徐宁,等. 无人驾驶汽车环境感知技术综述[J]. 长春理工大学学报(自然科学版), 2017, 40(1): 1-6.
- [2] Liu J. Research on key technologies in unmanned vehicle driving environment modelling based on 3D LiDAR [D]. Hefei: University of Science and Technology of China, 2016.
刘健. 基于三维激光雷达的无人驾驶车辆环境建模关键技术研究[D]. 合肥: 中国科学技术大学, 2016.
- [3] Liu Y. Environment perception for intelligent vehicle based on combined millimeter-wave radar [D]. Yantai: Yantai University, 2016.
刘曰. 基于组合毫米波雷达的智能车环境感知方法[D]. 烟台: 烟台大学, 2016.
- [4] Wang N. Machine vision based front vehicle detection and tracking [D]. Chengdu: Southwest Jiaotong University, 2018.
王宁. 基于机器视觉的前方车辆检测与跟踪[D]. 成都: 西南交通大学, 2018.
- [5] Chen T T. Research on 3D LiDAR-based environment perception for ALV [D]. Changsha: National University of Defense Technology, 2011.
谌彤童. 三维激光雷达在自主车环境感知中的应用研究[D]. 长沙: 国防科学技术大学, 2011.
- [6] He Y, Jiang H, Fang H, *et al.* Research progress of intelligent obstacle detection methods of vehicles and their application on agriculture[J]. Transactions of the Chinese Society of Agricultural Engineering,

- 2018, 34(9): 21-32.
- 何勇, 蒋浩, 方慧, 等. 车辆智能障碍物检测方法及其农业应用研究进展[J]. 农业工程学报, 2018, 34(9): 21-32.
- [7] Thrun S, Montemerlo M, Dahlkamp H, *et al.* Stanley: the robot that won the DARPA grand challenge[J]. Journal of Field Robotics, 2006, 23(9): 661-692.
- [8] Montemerlo M, Becker J, Bhat S, *et al.* Junior: the Stanford entry in the Urban Challenge[J]. Journal of Field Robotics, 2008, 25(9): 569-597.
- [9] Douillard B, Underwood J, Melkumyan N, *et al.* Hybrid elevation maps: 3D surface models for segmentation [C] // 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 18-22, 2010, Taipei, Taiwan, China. New York: IEEE, 2010: 1532-1538.
- [10] Xiao L, Dai B, Liu D X, *et al.* CRF based road detection with multi-sensor fusion [C] // 2015 IEEE Intelligent Vehicles Symposium (IV), June 28-July 1, 2015, Seoul, Korea. New York: IEEE, 2015: 192-198.
- [11] Lu W L, Murphy K P, Little J J, *et al.* A hybrid conditional random field for estimating the underlying ground surface from airborne LiDAR data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2009, 47(8): 2913-2922.
- [12] Lei Z, Xi X H, Wang C, *et al.* Building point clouds extraction from airborne LiDAR data based on decision tree method [J]. Laser & Optoelectronics Progress, 2018, 55(8): 082803.
雷钊, 习晓环, 王成, 等. 决策树约束的建筑点云提取方法[J]. 激光与光电子学进展, 2018, 55(8): 082803.
- [13] Samples M, Michael R J. Learning a real-time 3D point cloud obstacle discriminator via bootstrapping [EB/OL]. [2018-11-05]. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.385.6290&rep=rep1&type=pdf>.
- [14] Ren G T, Wang X Q, Huang J, *et al.* Detection of accessible area by laser radar based on support vector machine [J]. Journal of Computer Applications, 2017, 37(S2): 82-84, 106.
任国亭, 王新晴, 黄杰, 等. 基于支持向量机的激光雷达可通行区域检测[J]. 计算机应用, 2017, 37(S2): 82-84, 106.
- [15] McDaniel M W, Nishihata T, Brooks C A, *et al.* Terrain classification and identification of tree stems using ground-based LiDAR [J]. Journal of Field Robotics, 2012, 29(6): 891-910.
- [16] Spinello L, Siegwart R. Human detection using multimodal and multidimensional features [C] // 2008 IEEE International Conference on Robotics and Automation, May 19-23, 2008, Pasadena, CA, USA. New York: IEEE, 2008: 3264-3269.
- [17] Azim A, Aycard O. Detection, classification and tracking of moving objects in a 3D environment [C] // 2012 IEEE Intelligent Vehicles Symposium, June 3-7, 2012, Alcalá de Henares, Spain. New York: IEEE, 2012: 802-807.
- [18] Zhang X M, Wan W G, Li X, *et al.* Mean shift clustering segmentation and RANSAC simplification of color point cloud [C] // 2014 International Conference on Audio, Language and Image Processing, July 7-9, 2014, Shanghai, China. New York: IEEE, 2014: 837-841.
- [19] Hao G T, Du X P, Song J J, *et al.* Reconstruction of scannerless 3D LiDAR range image based on the adaptive block grayscale-range Markov random field model with optimizing weights [J]. Acta Optica Sinica, 2015, 35(8): 0811001.
郝刚涛, 杜小平, 宋建军, 等. 基于权值优化分块自适应灰度-距离 Markov 随机场的无扫描 3D 激光雷达距离图像重构 [J]. 光学学报, 2015, 35(8): 0811001.
- [20] Cai R T, Zhu P. Face tracking with multi-feature based on Markov random field [J]. Laser & Optoelectronics Progress, 2017, 54(2): 021002.
蔡荣太, 朱鹏. 基于马尔科夫随机场的多特征人脸跟踪算法 [J]. 激光与光电子学进展, 2017, 54(2): 021002.
- [21] Byun J, Na K I, Seo B S, *et al.* Drivable road detection with 3D point clouds based on the MRF for intelligent vehicle [J]. Field and Service Robotics, 2015, 105: 49-60.
- [22] Wu Q Q, Cai Y. 3D reconstruction algorithm based on Gaussian process regression and Markov random field [J]. Applied Laser, 2017, 37(6): 881-887.
吴倩倩, 蔡艳. 基于高斯过程回归与马尔科夫随机场的三维深度图像重构算法 [J]. 应用激光, 2017, 37(6): 881-887.
- [23] Himmelsbach M, Hundelshausen F V, Wuensche H J. Fast segmentation of 3D point clouds for ground vehicles [C] // 2010 IEEE Intelligent Vehicles Symposium, June 21-24, 2010, San Diego, CA, USA. New York: IEEE, 2010: 560-565.

- [24] Wang X, Wang J Q, Li K Q, *et al.* Fast segmentation of 3-D point clouds for intelligent vehicles[J]. *Journal of Tsinghua University (Science and Technology)*, 2014, 54(11): 1440-1446.
王肖, 王建强, 李克强, 等. 智能车辆 3-D 点云快速分割方法[J]. *清华大学学报(自然科学版)*, 2014, 54(11): 1440-1446.
- [25] Asvadi A, Premevida C, Peixoto P, *et al.* 3D LiDAR-based static and moving obstacle detection in driving environments: an approach based on voxels and multi-region ground planes [J]. *Robotics and Autonomous Systems*, 2016, 83: 299-311.
- [26] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C] // *The 25th International Conference on Neural Information Processing Systems*, December 3-6, 2012, Lake Tahoe, Nevada, USA: Curran Associates Inc., 2012, 1: 1097-1105.
- [27] Girshick R, Donahue J, Darrell T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation [C] // *2014 IEEE Conference on Computer Vision and Pattern Recognition*, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 580-587.
- [28] Maturana D, Scherer S. 3D convolutional neural networks for landing zone detection from LiDAR [C] // *2015 IEEE International Conference on Robotics and Automation (ICRA)*, May 26-30, 2015, Seattle, WA, USA. New York: IEEE, 2015: 3471-3478.
- [29] Li B. 3D fully convolutional network for vehicle detection in point cloud [C] // *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, September 24-28, 2017, Vancouver, BC, Canada. New York: IEEE, 2017: 1513-1518.
- [30] Engelleke M, Rao D, Wang D Z, *et al.* Vote3Deep: fast object detection in 3D point clouds using efficient convolutional neural networks [C] // *2017 IEEE International Conference on Robotics and Automation (ICRA)*, May 29-June 3, 2017, Singapore. New York: IEEE, 2017: 1355-1361.
- [31] Maturana D, Scherer S. VoxNet: a 3D convolutional neural network for real-time object recognition [C] // *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, September 28-October 2, 2015, Hamburg, Germany. New York: IEEE, 2015: 922-928.
- [32] Yang B, Luo W J, Urtasun R. PIXOR: real-time 3D object detection from point clouds [C] // *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 7652-7660.
- [33] Zeng Y M, Hu Y, Liu S C, *et al.* RT3D: real-time 3-D vehicle detection in LiDAR point cloud for autonomous driving [J]. *IEEE Robotics and Automation Letters*, 2018, 3(4): 3434-3440.
- [34] Wu B C, Wan A, Yue X Y, *et al.* SqueezeSeg: convolutional neural nets with recurrent CRF for real-time road-object segmentation from 3D LiDAR point cloud [C] // *2018 IEEE International Conference on Robotics and Automation (ICRA)*, May 21-25, 2018, Brisbane, QLD, Australia. New York: IEEE, 2018: 1887-1893.
- [35] Caltagirone L, Scheidegger S, Svensson L, *et al.* Fast LiDAR-based road detection using fully convolutional neural networks [C] // *2017 IEEE Intelligent Vehicles Symposium (IV)*, June 11-14, 2017, Los Angeles, CA, USA. New York: IEEE, 2017: 1019-1024.
- [36] Beltrán J, Guindel C, Moreno F M, *et al.* BirdNet: a 3D object detection framework from LiDAR information [C] // *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, November 4-7, 2018, Maui, HI, USA. New York: IEEE, 2018: 3517-3523.
- [37] Minemura K, Liao H, Monroy A, *et al.* LMNet: real-time multiclass object detection on CPU using 3D LiDAR [C] // *2018 3rd Asia-Pacific Conference on Intelligent Robot Systems (ACIRS)*, July 21-23, 2018, Singapore. New York: IEEE, 2018: 28-34.
- [38] Charles R Q, Su H, Mo K C, *et al.* PointNet: deep learning on point sets for 3D classification and segmentation [C] // *2017 IEEE Conference on Computer Vision and Pattern Recognition*, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 77-85.
- [39] Charles R Q, Li Y, Hao S, Leonidas J G. Pointnet++: deep hierarchical feature learning on point sets in a metric space [EB/OL]. (2017-06-07)[2018-12-01]. <https://arxiv.org/abs/1706.02413>.
- [40] Zhou Y, Tuzel O. VoxelNet: end-to-end learning for point cloud based 3D object detection [C] // *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 4490-4499.

- [41] Geiger A, Lenz P, Stiller C, *et al.* Vision meets robotics: the KITTI dataset [J]. The International Journal of Robotics Research, 2013, 32(11): 1231-1237.
- [42] Huang X Y, Cheng X J, Geng Q C, *et al.* The ApolloScape dataset for autonomous driving [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 18-22, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 1067-10676.