

DOI: 10.19650/j.cnki.cjsi.J2210526

语义信息增强的 3D 激光 SLAM 技术进展*

周治国, 邱顺帆, 冯 新

(北京理工大学集成电路与电子学院 北京 100081)

摘要:由于激光雷达可直接获得测距信息且相较于视觉传感器对光照等环境变化更具鲁棒性等优点,激光同步定位与建图(SLAM)技术近年来得到广泛发展。传统激光 SLAM 已取得很多研究成果,但其仅利用几何特征,对场景的理解有限,难以应对复杂任务,除此之外,当前 SLAM 应用场景已由传统静态场景向复杂动态场景过渡,传统方法由于动态元素干扰大多难以获得较好的性能。因此,语义信息增强的三维(3D)激光 SLAM 技术愈发受到研究学者们的关注,通过赋予点云语义标签与纯几何特征进行融合,一方面借助语义信息滤除潜在运动对象以解决静态环境假设问题,另一方面以语义信息辅助激光里程计获得高精度的定位与建图。综述了语义信息增强的 3D 激光 SLAM 技术研究进展,提出了该技术通用框架,分模块对该领域的突出研究成果及应用进行重点介绍,最后对该领域发展方向进行了总结与展望。

关键词: 激光雷达点云;点云语义分割;激光里程计;同步定位与建图;语义信息增强

中图分类号: TH702 TP242 **文献标识码:** A **国家标准学科分类代码:** 510.70

Advances in SIE 3D Lidar SLAM technology

Zhou Zhiguo, Di Shunfan, Feng Xin

(School of Integrated Circuits and Electronics, Beijing Institute of Technology, Beijing 100081, China)

Abstract: Because Lidar can directly obtain ranging information and is more robust than visual sensors to environmental changes such as illumination, the technology of laser synchronous location and mapping (SLAM) has been widely developed in recent years. The traditional laser SLAM has made a lot of research achievements. But, it only uses geometric features, has limited understanding of the scene, and is difficult to deal with complex tasks. In addition, the current SLAM application scenarios have transited from traditional static scenes to complex dynamic scenes, and traditional methods are mostly difficult to achieve good performance due to interference of dynamic elements. Therefore, the 3D laser SLAM technology of semantic information enhancement has attracted more and more attention of researchers. The point cloud semantic tags are integrated with pure geometric features. On the one hand, the potential moving objects are filtered out with semantic information to solve the problem of static environmental assumptions. On the other hand, semantic information is used to assist the laser odometer to obtain high-precision positioning and mapping. This article summarizes the research progress of 3D laser SLAM technology for semantic information enhancement, puts forward a general framework for this technology, focuses on the outstanding research achievements and applications in this field in modules, and finally summarizes and prospects the development direction of this field.

Keywords: Lidar point cloud; point cloud semantic segmentation; laser odometer; simultaneous localization and mapping; semantic information enhancement

0 引言

作为一种在未知环境中确定自身位姿和构建环境一

致性地图的方法,同步定位与建图(simultaneous location and mapping, SLAM)技术在无人领域一直备受关注。目前,SLAM 技术被广泛应用于无人机、无人驾驶、增强现实(augmented reality, AR)等领域^[1],通过传感器对周围

环境的感知实现自主定位、建图、导航等功能。当前,移动机器人 SLAM 系统可以分为两类,一类是基于激光雷达的激光 SLAM,另一类则是基于视觉传感器的视觉 SLAM。相较于视觉 SLAM,激光 SLAM 技术构建的地图精度高,并且激光探测和测距(light detection and ranging, Lidar)传感器可以准确的记录深度信息和强度(从物体表面反射的量度),无论光照条件如何,其都具有很高的密度和帧速率,已经成为自动驾驶等关键任务的可靠信息来源,激光雷达传感器技术的最新进展使得从环境中产生高质量、低噪声和密集扫描的场景理解成为可能。同时,由于多线激光雷达具有多线束和高旋转速率,可以获得更丰富的环境信息,因此基于多线激光雷达的 3D 激光 SLAM 技术越来越受到研究者的关注。

3D 激光 SLAM 技术的基本原理是利用激光束的发射和反射获取周围环境内障碍物的距离信息,通过激光雷达传感器获得三维空间内的点云数据,再对相邻扫描帧进行配准获得位姿估计,并构建出完整的点云地图^[2]。基于多线激光雷达的 SLAM 算法有着巨大的发展潜能,其在帧间匹配过程中匹配方式丰富且鲁棒性更好,还能同图像等信息融合,以提升定位精度。但激光 SLAM 也有一定的局限性,纯几何激光雷达 SLAM 构建的位姿可能出现错误,同时在动态变化比较大的环境当中,容易出现定位丢失的情况。因此,研究者们考虑将语义信息引入 3D 激光 SLAM,借助语义信息实现 3D 激光 SLAM 技术的性能提升。

传统的 3D 激光 SLAM 方法在处理原始点云时是直接处理的,其构建的环境地图只能表达环境中的拓扑信息和几何信息,而无法获取和描述环境中的语义信息,机器人也无法真正理解环境^[3]。而语义信息增强的 3D 激光 SLAM 方法可以通过利用语义分割算法赋予点云数据

类标签语义,一方面通过语义滤除潜在运动对象类点云的方式解决静态环境假设问题,另一方面通过融合语义信息和几何特征的方式提出语义信息辅助的激光里程计方法来提升位姿估计的精确度。另外,基于激光雷达的 SLAM 系统的回环检测一直是一个有待解决的问题。回环检测,也称地点识别、位置识别,回环检测的关键就在于描述符的设置。传统的激光 SLAM 在回环检测中有两个重要问题需要克服:1) 无论视点如何变化都需要保证描述符来实现其旋转不变性;2) 由于点云的分辨率随距离变化,描述符对噪声的处理。而语义信息增强的激光 SLAM 可以将语义引入描述符,从而提高定位精度,有效检测回环。与传统 3D 激光 SLAM 相比,融入语义信息的 3D 激光 SLAM 方法在同步定位与建图方面的精度更高、鲁棒性更好。

本文由传统的 3D 激光 SLAM 的不足引出融入语义的 3D 激光 SLAM 技术,并总结近年该领域技术进展提了语义信息增强的 3D 激光 SLAM 技术 (semantic information enhancement 3D lidar SLAM, SIE 3D Lidar SLAM) 通用框架。SIE 3D Lidar SLAM 技术在前端过滤掉动态点云,只保留绝对静态点云传递给激光里程计,并将语义信息引入到点云配准过程中,采用语义信息辅助实现激光里程计的方法,结合语义特征和几何特征进行配准,从而得到精确且高效率的位姿估计和点云地图构建。

1 SIE 3D Lidar SLAM 技术通用框架

SIE 3D Lidar SLAM 技术的系统框架由输入部分、预处理模块、语义分割网络模块、语义辅助的激光里程计模块、回环检测模块、位姿融合/全局地图构建模块和输出部分组成,系统框架图如图 1 所示。

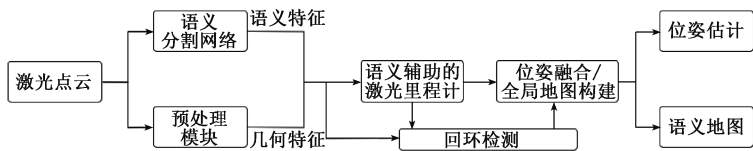


图 1 SIE 3D Lidar SLAM 系统通用框架

Fig. 1 General framework for the SIE 3D Lidar SLAM system

1.1 通用框架各模块功能

整个系统的输入部分可以是采集的原始激光点云数据,也可以是原始点云投影成得到的距离图像(range image)。输出部分则是构建完成的语义点云地图和位姿信息。

激光雷达作为传感器测量自身与周围环境边界的距离从而形成一系列空间点^[4],每个点都包含了丰富的信息,包括三维坐标、强度值、时间戳等,在某个坐标系下的

点集即为激光点云,一次扫描获取的点云称为 1 帧。激光 SLAM 通过对帧间点云的匹配进行位姿推算并构建完整的点云地图。激光雷达获取数据的方式主要分为 3 大类,星载、机载和地面。星载激光雷达主要依靠卫星平台,可测量到地球的每个角落,具有运行轨道高和观测范围广的优点;机载 Lidar 主要借助无人机大规模的采集点云数据,通过向地面发射激光信号收集地面反射的激光信号实现,该类方式常用于三维重建、农林普查、地质监

测等;地面 Lidar 包含地上固定、车载、手持等多种方式,此类方式目前应用最为广泛,车载 Lidar 日益成为自动驾驶领域的主流技术。

预处理模块可以对输入的原始激光点云进行运动补偿修正,并可以处理激光雷达生成的深度、法向量、语义等信息。在对上述信息进行预处理的过程当中,将各类类型的信息进行初始融合,主要是借助深度信息去除语义分割中的噪声点。

语义分割网络模块的作用是对无序、不规则的点云数据附加语义信息。为了使从激光雷达收集的3D数据在感知模块中得到充分利用,需要在点级别提供类别标记(语义)的点云,点云中的语义分割任务需要给输入模式中的点云数据中的每个数据点分配一个类标签,以记录对点云场景的理解。

在自主导航系统中激光里程计对于定位与建图功能起着重要作用,其主要是解决点云配准问题。语义辅助的激光里程计模块的作用是在进行点云特征点匹配前,利用语义信息过滤掉潜在运动对象点云,再经过统计滤波和直通滤波的方式过滤掉一些离群点和不稳定点,只保留健壮的鲁棒性高的静态点云,随后同时利用基于语义信息和几何特征点云配准流程实现帧间点云的位姿估计。

回环检测也称闭环检测,当运动距离增加、时间增长时,测量噪声的存在和系统误差的积累造成SLAM系统产生漂移,回环检测的存在就是为了消除漂移,因而在SLAM系统中占据重要地位。回环检测是指当机器人返回运动过程中去过的地方时,通过当前姿态与历史姿态的相对变换来判断历史位置,获得当前数据与历史数据的相关性,作为后端优化的附加约束,从而消除漂移。

在进行语义点云地图构建之前,需要通过激光里程计提供的估计位姿进行点云地图的合并更新,然后将该帧点云与已有点云地图进行匹配输出低频率但更高精度的位姿估计。位姿融合里程计模块和回环检测模块,利

用低频高精度的位姿来修正里程计高频的位姿输出,最终输出高频又高精度的估计位姿。语义点云地图构建主要是通过进行点云数据的二次匹配来完成点云地图的累加更新。在二次匹配过程中,一开始需要利用之前激光里程计模块提供的估计位姿将当前帧点云数据转至全局坐标系下,之后再使用之前构建的全局点云地图与当前帧进行点云匹配,配准方法与激光里程计的配准方式相同,二次配准的好处在于填补帧间配准没有考虑整体关系的情况,并且输出的更精确的估计位姿可以修正之前里程计的位姿。

1.2 激光雷达传感器参数标定及全局坐标归一化

对 Lidar 的参数标定可分为内参标定和外参标定两类。

内参标定由 Lidar 自身坐标系与内部的激光发射器坐标系之间的转换关系得到,这一步骤通常在出厂前已完成;外参标定由 Lidar 坐标系与所需坐标系之间转换关系获取,自动驾驶领域中通常为 Lidar 坐标系与车体坐标系的关系。Lidar 与车体为刚性连接,即其间的相对位姿固定不变,若想得到 Lidar 与车体间的坐标转化关系,就要对 Lidar 的位置进行标定,使得点云数据由 Lidar 坐标系一至车辆坐标系,坐标转换示意图如图2所示。以 Velodyne 激光雷达的 VLP-16 为例,其坐标系的 z 轴为地面垂直向上, y 轴负方向为电缆线接口方向, x 轴由右手法则确定;车体坐标系的坐标原点在车的后轴中心处,正上方为 z 轴,朝前为 x 轴,右手坐标系,确定坐标系 y 轴方向。两个三维空间直角坐标系之间的转换关系可以用旋转矩阵加平移矩阵表示,设图2中点 P 在 $Oxyz$ 坐标系下坐标为 (x, y, z) ,点 P' 在 $O'x'y'z'$ 坐标系下坐标为 (x', y', z') ,这两点的坐标转换关系如下:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{R} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} + \mathbf{T} \quad (1)$$

进一步推导出:

$$\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\beta \cos\gamma & \cos\alpha \cos\gamma - \cos\gamma \sin\alpha \sin\beta & \sin\alpha \sin\gamma + \cos\alpha \cos\gamma \cos\beta & \Delta x \\ -\cos\beta \sin\gamma & \cos\alpha \cos\gamma + \sin\alpha \sin\beta \sin\gamma & \cos\alpha \sin\beta \sin\gamma & \Delta y \\ -\sin\beta & -\cos\beta \sin\alpha & \cos\alpha \cos\beta & \Delta z \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} \quad (2)$$

通过实验采集同一点在两个坐标系下的真实坐标建立方程组,可求得未知数,从而得到变换矩阵,实现 Lidar 外参标定。

在激光 SLAM 中,点云数据的配准是将不同坐标系下的点云数据统一至同一坐标系下,即实现全局坐标归一化,其实质是对不同坐标系的平移和旋转,点云配准方法主要分为3类^[5]:1)基于点的点云配准方法;2)基于特征的点云配准方法;3)基于数学特性的点云配准方法。

基于点的点云配准方法直接操作扫描得到的原始数据点,其中迭代最近点(iterative closest /corresponding point, ICP)算法应用最广泛,由 Besl 等^[6]于1992年提出,基于最小二乘法的最有匹配原则,重复“确定对应关系点集”、“计算最优刚体变换”两个过程,直到满足收敛条件。该算法计算较为简单,但存在解容易陷入局部最优等缺陷,近年来研究学者提出了多种 ICP 改进算法,如 P2P1-ICP 算法^[7]、IDC 算法^[8]、ICL 算法^[9]、GICP 算法^[10]等。

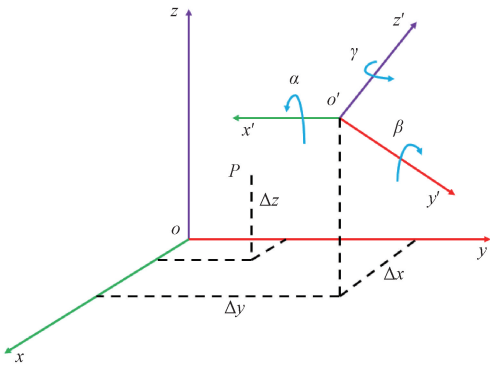


图2 Lidar坐标系与车体坐标系示意图
Fig. 2 Diagram of the Lidar coordinate system and the vehicle body coordinate system

基于特征的点云配准方法使用数据中的某些关键元素进行比较,例如几何特征点^[11]、线^[12]、面^[13]等及它们的组合形式,或法向量、曲率等特征描述符及自定义的各类特征描述符^[14-15]。根据检测到的特征,可以快速实现对应匹配,并且能够在不提供初始值的情况下解得位姿变换。

基于数学特性的点云配准方法采用各种数学性质求得帧间的位姿变化,其中正态分布变换(normal distributions transform, NDT)方法被人们广泛应用。Biber等^[16]于2003年第1次提出了NDT算法,用于二维平面点云数据之间的匹配,该方法将一次扫描中的离散点转换为由一组易于计算的正态分布组成的,定义于二维平面上的分段连续并可微的概率密度,最大化另一帧扫描帧中点的配准,并在此密度下评估其是否符合NDT,然后使用牛顿法获得优化求解。Magnusson等^[17]于2006年提出了3D-NDT算法,使得NDT算法突破了维度的限制,随后多位研究学者在此基础上提出了改进算法,如ML-NDT算法^[18]、NDT-OM算法^[19]等。

2 语义信息在前端里程计的应用

语义信息增强的3D激光SLAM方案可以在进行激光里程计模块前,先通过点云语义分割模块对激光雷达的原始点云进行语义分割,然后利用分类结果过滤掉潜在动态对象点云,避免动态点云干扰造成的匹配精确度的下降,从而提升SLAM位姿估计精确度并构建全局一致的点云地图。

2.1 原始点云的语义分割

传统的语义对象分割方法通常需要对图像样本进行学习,训练出分类器,利用分类器对图像中的语义对象进行检测,然后根据检测结果进行分割。由于分割过程依赖于检测,这种方法得到的分割结果将直接受到分类器

性能的影响^[20]。传统的语义分割种类比较多,机器学习领域中常见的分类器主要有归一化分割(normalized-cut)、随机森林(random forest, RF)、条件随机场(conditional random field, CRF)和支持向量机(support vector machine, SVM)等。由于在特征提取和分类过程中没有考虑大范围的上下文信息,导致语义分割缺乏部分信息,同时这种方法语义标注类别有限,准确性低。因此,传统的分割方法有很大的局限性,分割效果不够好,分割效率也比较低,完成一次分割耗时较长,难以应用于实时驾驶系统等。

深度学习在特征层具有非常强大的学习能力,逐渐成为点云语义分割的主流方法。2014年提出的使用全卷积网络(fully convolutional networks, FCN)进行语义分割^[21]成为许多研究的基础,通过增加卷积层的数量或调整网络结构,例如DeepLab^[22]、ICNet^[23]和SegNet^[24],进一步提高了语义分割的准确性。其中SegNet在解码器中采用去池化操作对特征图进行上采样,在分割中保证高频细节的完整度,由于编码器不包含全连接层,SegNet是一个轻量级网络,参数较少。SegNet的语义分割架构如图3所示。随着semantic KITTI^[25]、Semantic3D^[26]等公开数据集的出现,深度学习在点云场景理解和具体的语义分割方面的应用研究在过去几年中有所增加。Semantic KITTI认证语义分割任务有效类别如图4所示^[25],官方评估忽略了其中只有少数扫描点的不常见类,并合并了具有不同迁移状态的类,因此最终认证了19个有效的类。基于深度学习的点云语义分割方法与传统语义分割方法相比更加高效准确,这些方法依据其处理特点可以分为基于原始点云的、基于体素的、基于投影的以及基于混合手段的4种范式方法^[27]。基于点的分割方法直接处理原始不规则的3D点,无需进行任何额外的变换或预处理。依赖于体素表示的方法将3D空间离散为3D体积空间(即体素),并将每个3D点分配给相应的体素^[28-30]。基于投影的方法依赖于将3D数据投影到2D空间中,这种2D渲染图像表示更紧凑密集,是克服激光雷达数据稀疏性的一个有效方法。融合方法,通过使用基于原始点、基于投影或基于体素操作的融合网络来处理点云,过去的研究较少,但随着更多内存效率更高设计的可用,产生了更有竞争力的结果。

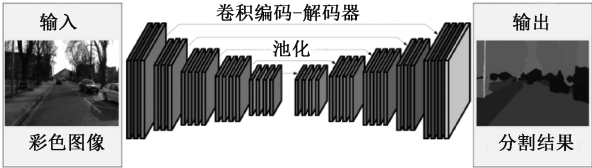


图3 SegNet 语义分割架构
Fig. 3 SegNet semantic segmentation architecture

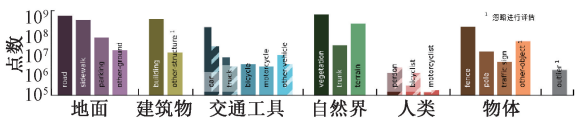


图4 SemanticKITTI 认证语义分割任务有效类别

Fig.4 SemanticKITTI authentication semantic segmentation task effective category

为了使从激光雷达收集的 3D 数据在感知模块中得到充分利用,需要在点级别提供类别标记(语义)的点云,点云中的语义分割任务需要给输入模态中的点云数据中的每个数据点分配一个类标签,以记录对点云场景的理解。

由于 3D 语义分割的性能低于 2D 语义分割,因此语义信息增强的 3D 激光 SLAM 方案在原始点云的语义分割方面,首先分别构建激光点云数据到投影图像像素以及体素的映射关系,再分别借助成熟的图像分割主干网络和稀疏体素卷积提取点云的二维特征和三维特征信息,通过创建多视图索引系统和多视图特征传播,实现高效的多视图特征交互学习,然后利用通道注意力模块提取各个特征通道间的依赖关系,在进行多尺度特征融合时,为不同细粒度的特征图计算和分配相应的注意力权

重,最终从语义图像中恢复全体原始点云的语义标签。range image 具有许多优点,可以将稀疏点云表示为密集图像,因此处理点云可采用许多经典的图像处理技术,实现与图像处理类似的效果,并且可以将三维 Lidar 点云投影为二维图像,从而减少处理点云占用的空间并提高处理速度。Chen 等^[31]利用 range image 将三维激光点云投影到二维图像中,之后以基于深度学习的图像分割算法进行语义分割,然后将其投影回三维点云中,得到每个测量点的语义标注,并将分割的结果与原始点云相结合,得到基于激光点云的分割结果。RangeNet++^[32]为用于对旋转雷达传感器记录的点云进行语义分割的框架,通过对输入点云的球面投影,利用距离图像和 2D 卷积和 GPU 加速的后处理方法实现了对激光雷达点云快速准确的语义分割。在激光雷达点云的语义分割方面,该方法在距离图像上操作的改进 2D 深度卷积神经网络(convolutional neural networks, CNN)优于当前的技术水平。在预处理过程中 OverlapNet^[33]框架利用激光雷达点云生成含有法向量、强度、深度和语义这些不同信息的距离图像,依靠基于二维图像的 CNN 来实现各种功能。

表 1 为 7 个典型的分割网络模型^[34-40]在 SemanticKITTI 数据集基准测试上的单次扫描结果。

表 1 各分割网络模型在 SemanticKITTI 数据集基准测试上的单次扫描结果

方法	PointNet ^[34]	SPGraph ^[35]	SPLATNet ^[36]	PointNet++ ^[37]	PointASNL ^[38]	PointConv ^[39]	FusionNet ^[40]
均交并比	14.6	17.4	18.4	20.1	46.8	51.2	61.3
平均准确率/%	-	-	-	-	-	87.1	91.2
道路	61.6	45.0	64.6	72.0	87.4	88.9	91.8
人行道	35.7	28.5	39.1	41.8	74.3	68.4	77.1
停车位	15.8	0.6	0.4	18.7	24.3	58.9	68.8
其他地面	1.4	0.6	0.0	5.6	1.8	19.7	30.8
建筑物	41.4	64.3	58.3	62.3	83.1	84.6	92.5
汽车	46.3	49.3	58.2	53.7	87.9	93.1	95.3
卡车	0.1	0.1	0	0.9	39.0	37.8	41.8
自行车	1.3	0.2	0	1.9	0	20.7	47.5
摩托车	0.3	0.2	0	0.2	25.1	22.9	37.7
其他车辆	0.8	0.8	0	0.2	29.2	38.1	34.5
植被	31.0	48.9	71.1	46.5	84.1	79.9	84.5
树干	4.6	27.2	9.9	13.8	52.2	62.3	69.8
地面	17.6	24.6	19.3	30.0	70.6	60.7	68.5
人	0.2	0.3	0	0.9	34.2	46.2	59.5
骑自行车的人	0.2	2.7	0	1.0	57.6	39.1	56.8
栅栏	12.9	20.8	23.1	16.9	43.9	52.0	69.4
杆子	2.4	15.9	5.6	6.0	57.8	48.1	60.4
交通标志	3.7	0.8	0	8.9	36.9	44.7	66.5

2.2 动态点云的滤除

当前的众多 SLAM 系统用几何信息表示环境,大多数方案都将移动平台所在的环境假设为静态的,从而忽略场景中运动对象所造成的点云畸变,这导致基于该假设的 SLAM 方案提取的场景特征往往具有一定误差且造成错误的数据关联,使得该类 SLAM 方法难以实现高精度的姿态估计和最终全局一致性的地图。

在语义信息增强的 3D 激光 SLAM 方案当中,点云语义分割算法网络可以将原始点云数据处理成附加语义类标签的点云数据,处理后的点云数据作为输入,从而直接锁定潜在动态对象类点云,过滤掉动态点云,只保留绝对静态点云传递给激光里程计。滤除环境中的运动对象可以最小化数据关联的误差,以提高姿态估计的精度。用来进行激光里程计的输入点云滤除了大部分动态物体点云,降低了动态点云的干扰,语义信息辅助的激光里程计的位姿估计精度和计算效率都有一定的提升。在大规模室外环境中,通常单独使用激光雷达或与 RGB 相机结合使用。Wang 等^[41]为了更准确地区分环境中的动态和静态元素,并为 SLAM 姿势估计提供具有更好鲁棒性的特征根据现实环境中人类空间认知的特点,建立了 6 大类 14 子类环境要素的动态量化指标,其对应值从 0~1,从动态到静态,如图 5 所示。SuMa++^[31]框架专注于生成具有大量语义类的语义地图,并使用这些语义过滤由动态对象(如移动车辆和人类)引起的异常值,以提高地图和里程计的准确性,即全局语义地图的构建利用了每帧点云的逐点语义标签,通过 RangeNet++^[32]有效滤除场景中的动态对象,构建出精度更高的全局地图。面元地图用同一个面元来表示点云信息中法向量信息相同的相邻点,在减少存储点数的同时包含有更丰富的几何信息,是点云地图的扩展。面元信息的一致性检测包括几何空间和语义信息一致性,其中几何空间的一致性理解为同一物体在不同观测下的空间信息的一致,包括有关位置和法向量方向信息;语义信息的一致性理解为同一静态对象在不同观测中语义标签的一致性,如运动对象即便在同一位置,对应的语义标签也可能会有变化。在过滤场景中的动态物体时,将潜在的移动对象全部删除有助于长期定位系统,但对于 SLAM 算法,停在路边的静止车辆于系统而言也是有意义的。在 SuMa++^[31]算法在移除处于运动状态的对象时保留一些潜在移动对象,是基于 SuMa^[42]的改进,因此在建图的环节中沿用了之前的面元地图。SuMa++^[31]算法在更新面元稳定性的同时进行面元的一致性检测,即同步检测几何空间和语义信息的一致性。不同过滤方法的效果对比如图 6 所示。

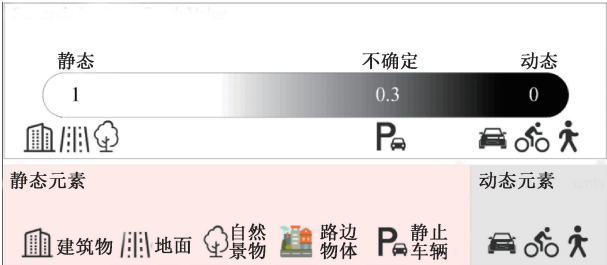


图 5 基于先验知识的环境要素分类
Fig. 5 Classification of environmental elements based on prior knowledge

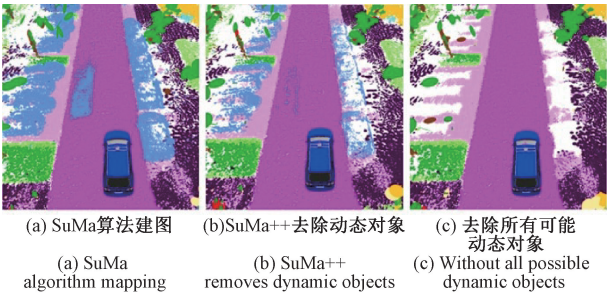


图 6 语义 SLAM 建图示意图
Fig. 6 Diagram of semantic SLAM mapping

2.3 结合语义特征的位姿估计

在 3D 激光 SLAM 中的点云配准算法通常需要执行两次,可分为粗配准和精配准。粗配准应用于相邻帧间的点云匹配,得到局部的位姿估计,但会由于误差累计造成整体精度相对不高;精配准过程是将粗配准的位姿与点云地图中的数据进行配准,得到全局地图下的可靠位姿估计。Parkison 等^[43]提出了一种点云配准算法,直接将基于图像的语义信息结合到两帧点云间的相对变换估计中。Zaganidis 等^[44]完成了将纯 Lidar 点云与结合图像的 Lidar 点云相结合的三维 Lidar 点云配准。但是这两种方法处理时间长,无法用于在线操作。在语义信息增强的 3D 激光 SLAM 方案中,帧间点云的位姿估计是通过基于语义信息和几何特点的点云配准流程实现的。配准流程主要实现 3 个步骤:1) 构建特征点匹配对应关系;2) 设定基于特征点几何误差和语义相似度误差的联合误差函数;3) 迭代求解匹配的位姿变换矩阵以使得误差最优。构建特征点匹配对应关系流程如图 7 所示。

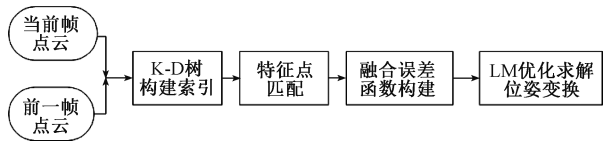


图 7 构建特征点匹配对应关系流程
Fig. 7 Process of building feature point matching correspondence

在构建特征点匹配对应关系的过程中,对于前一帧点云与当前帧点云中的边缘特征点集合和平面特征点集合,均需采用K-D树(k-dimensional tree)方式构建索引以提高查询效率。构建好相邻两帧点云的特征点匹配关系后,需要通过构建的误差函数来约束相邻帧的相对运动变换。在构建几何误差和语义相似度的融合误差函数时,如果直接通过约束语义类别一致来约束对应特征点对,这种方式虽然逻辑简单计算,但是对于语义分割中的不确定性结果和误差影响不能处理,会严重影响点云配准的精确度。因此,可以采用连乘的方式融合几何误差和语义相似度,当前特征点的语义相似度直接作为相应几何误差的权重,语义相似度权重高则相应特征点对的几何误差贡献度就高。最后,可以采用Levenberg-Marquardt^[45]方法进行优化,从而求解传感器的运动矩阵。

3 语义信息在后端回环的应用

回环对于在未知场景当中纠正SLAM中的累计误差是非常有必要的,传统的3D激光雷达SLAM方案已经是一种比较精确和稳定的系统了,但是这种方案中的闭环检测仍然是一个需要解决的问题。三维点云语义分割技术的不断发展使得研究者们能够更加方便的获取点云语义信息,通过语义信息辅助3D激光雷达SLAM进行回环检测,极大的提高了该方案的性能。

在SA-LOAM^[46]为一种新的基于LOAM的语义辅助Lidar-SLAM,该方案充分利用了里程计和回环检测中的语义信息,是一种语义辅助的ICP方案,包括语义匹配、下采样和平面约束,并在回环检测模块中集成了基于语义图的位置识别方法,充分利用了语义信息,可以提高定位的精度,从而进行有效的闭环检测,甚至在大规模场景中也可以构造一个全局一致的语义地图。回环检测模块主要包括回环候选生成、相似性评分、几何验证和位姿图优化,候选生成部分根据里程计提出潜在的环路候选,相似性评分是通过一个图相似性网络快速估计成对语义图的相似性,对遮挡和视点变化具有鲁棒性,然后利用ICP进行几何验证,排除可能的错误方案,避免灾难性的误报闭环,最后通过位姿图优化来更新更精确的位姿。整个回环检测模块的实现是将3D场景转换成语义图,然后通过图匹配深度学习网络获得场景相似度,再将其与回环候选生成、几何验证相结合,并维护一个轻量级的语义图,以实现高效、稳定的回环和检测,同时消除累积的错误。实验证明基于语义的环路闭合检测模块可以有效地减少累积误差,帮助建立一个全局一致的地图。SA-LOAM在

KITTI序列13和Ford序列01上构建的语义图如图8所示。

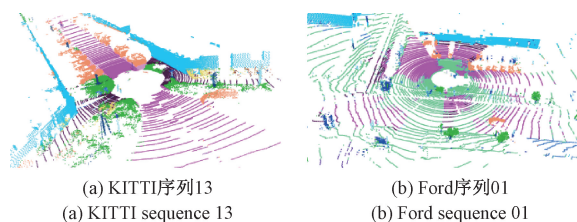


图8 SA-LOAM构建的语义图

Fig. 8 Semantic map constructed by SA-LOAM

SmSLAM + LCD (semantic SLAM + Loop closure detection)^[47]是一种很好的语义闭环方法,利用来自SLAM系统的高层语义和低层几何信息来进行环路检测和漂移校正。SmSLAM+LCD在映射的3D语义对象上构建对象共视图,并不断地用最新观测对其进行更新。在执行环路检测时,该方法通过比较与环路候选帧相关联的对象共视图来检查基于底层几何特征的环路候选,并通过一种从粗到精的方法计算环闭合帧之间的变化,从而纠正累积的漂移。SmSLAM+LCD能够在检测到环路闭合后实现更准确的漂移校正,并区分相似场景,避免闭环环路误报。

LIO-CSI^[48]中将点云帧编码为语义辅助扫描上下文图像,在构建语义辅助扫描上下文图像时,点云以扇区块的形式组织,径向被距离等分,称为环,也被一个称为扇形的角平均分割。扇形区域也以这种方式形成,如图9所示。两步循环闭包搜索方法如下:首先,从语义辅助扫描上下文图像中提取环密钥,并构建K-D树以快速查找候选项,最接近的多语义辅助扫描上下文图像在这些候选图像中进行过滤;然后,计算这些语义辅助扫描上下文图像的相似度得分,并将得分最高的图像对应的帧检测为循环闭合帧;最后,如果第*i*帧和第*j*帧之间有一个循环闭包对,则可以在因子图中添加一条观察边,通过语义辅助循环闭包检测方法消除累积错误,生成了全局一致的地图。

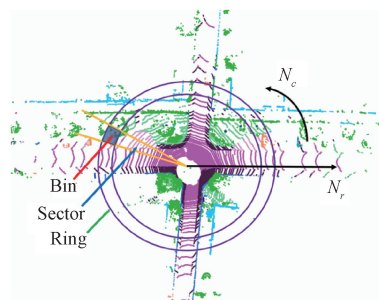


图9 语义辅助扫描上下文图像编码过程

Fig. 9 Semantic-assisted scan-context image encoding process

4 语义信息在地图构建的应用

传统激光雷达 SLAM 方案利用激光雷达探测获得的目标位置和速度等特征信息,结合机器人里程计数据和惯性测量单元(inertial measurement unit, IMU),利用粒子滤波算法估计其位置。基于点云特征进行目标匹配或基于深度学习的检测算法由于缺乏语义信息,目前发展并不成熟^[49]。

在语义信息增强的 3D 激光 SLAM 方案中,点云地图构建模块需要通过进行点云数据的二次匹配来完成点云地图的累加更新,二次匹配的示意图如图 10 所示。如果直接进行点云帧在点云地图上的累加,虽然可以得到连续且稠密的点云地图,但会增加存储空间的消耗和计算代价,因此可以采取以低频采样的方式来构建点云地图,在实际算法流程中,激光里程计的输出频率为 10 Hz,假设每 5 帧点云进行一次采样,选取出的点云帧再进行二次匹配,那么语义点云地图构建模块就能够以 2 Hz 的频率再输出高精确度的估计位姿。如果在姿态估计和地图构建的过程中将每一帧激光点云数据都并入到构建的点云地图中,不仅会导致维护的点云地图存储代价增大,而且在后端优化进行位姿修正时极大的拖累计算效率,因此许多研究学者们采用选取关键帧的方式来代替所有帧激光点云构建点云地图,而选取的关键帧必须能够保证所构建的场景地图信息不丢失。通常选取能够代表帧点云的关键帧条件如下:1)当前选取的关键帧点云与之前选取的关键帧点云的时间差有个阈值范围;2)当前选取的关键帧点云与之前选取的关键帧点云的位移差有个阈值范围;3)选取的关键帧点云鲁棒性强,能够反映局部点云场景的环境信息。

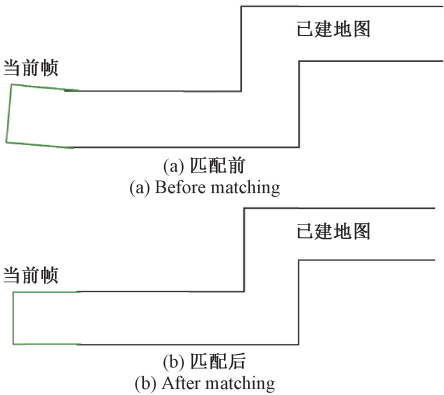


图 10 当前帧与已构建地图的匹配示意图
Fig. 10 Matching diagram between the current frame and the built map

语义地图可以分为两种,面向场景的语义地图和面向对象的语义地图。

4.1 面向场景的语义分割

基于场景的语义地图对场景中的每个区域进行语义标注。由于不同区域通常由门分隔开,因此 Vasudevan 等^[50-51]采用概率方法识别和注释门和特定对象,Rituerto 等^[52]进一步实现对门、楼梯和电梯的语义注释,Liu 等^[53]使用马尔科夫过程模型的非监督聚类方法来注释室内环境,并尝试使用不同的聚类方法进行语义注释^[54-56]。Pronobis 等^[57]在 2011 年提出了多模态地点分类系统,并继续完善这部分工作,为室内环境建立了语义注释模块^[58]。Goeddel 等^[59]应用 CNNST (convolutional neural networks spatial transformer) 实现占用网格上不同区域的语义标注,Hiller 等^[60]根据拓扑图推断并注释了占用网格。Chen 等^[31]提出了一种通过基于激光的点云语义分割实现、无需任何视觉数据构建语义地图的方法,通过这些信息提升在其他模糊和具有挑战性的环境下估计姿态的准确性,该方法利用扫描和映射之间的语义一致性滤除动态物体,在迭代最近点计算中提供更高级别的约束,在计算其信息矩阵时,对各个残差进行加权。同时使用基于几何信息和语义信息的权重来确定当前帧同模型帧之间的语义标签是否一致,若一致则直接使用语义分割估计出的概率进行加权计算,否则使用极小的权重值以降低噪点或权重。图 11 所示为面向场景的语义地图^[41]。

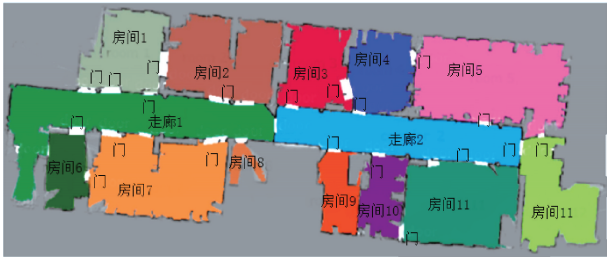


图 11 面向场景的语义地图
Fig. 11 Scenario-oriented semantic map

4.2 面向对象的语义地图

面向对象的语义地图使用场景识别和图像分割等技术方法在地图上精确标记场景中的对象。早期的研究主要使用机器学习的方法,Limketkai 等^[61]使用关系马尔科夫网络对场景中具有明显线条特征的对象进行语义标记。Andreas 在 2008 年提出语义地图的概念,以场景构建的 3D 点云为基础,利用室内建筑结构的特征分析,标记天花板、墙壁、地板。Sunderhauf 等^[62]提出了一种面向对象的语义地图构建方法,单个对象实例是映射中的关键实体,所生成的环境地图丰富了单独对象实体形式的

语义信息,这些对象以类标签和置信度的形式携带几何和语义信息,其中对象是完全独立于地图的非对象部分的独立实体,这实现了更高级的场景理解。Jeong等^[63]结合3D激光雷达和相机,创建了1个包括7个标签的语义地图,这些标签对应的对象占据了大部分环境。在语义地图构建的过程中,点云由3D激光雷达测量,并转换为全局一致,以生成3D地图,通过使用从摄像机获得的图像来执行基于CNN的2D语义分割,然后执行增量语义标记以整合这两个阶段的结果。完成语义建图后,执行地图细化以提高地图质量。在生成全局3D地图的过程中,需要将点云转换为相应的里程计并进行配准,因此要实时估计每个帧的里程。图12所示为面向对象的语义地图^[41]。

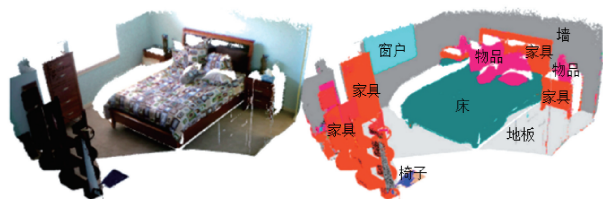


图12 面向对象的语义地图

Fig. 12 Object-oriented semantic map

5 结 论

3D激光SLAM作为一种借助激光雷达传感器实现自主定位的方案,可以帮助移动机器人在不借助外部定位源的情况下对未知环境进行地图构建和实现自我定位,提供给无人平台在移动过程中相对于当前场景的参考位姿。然而目前的主流SLAM研究文献和项目都是基于应用场景静态的假设,没有考虑到无人平台实际移动过程中采集到的动态对象点云的干扰,因此基于此种假设下的SLAM算法的精确度难以继续提升。此外,仅仅依靠激光点云数据的几何信息,虽然可以帮助机器人实现自身定位,但无法让机器人理解场景的周围环境信息,而理解点云场景环境的语义信息可以进一步实现更加精确的定位和地图构建一致性,甚至可以帮助机器人实现更加复杂的智能化任务。

本文对语义信息增强的3D激光SLAM技术的进展进行了系统的总结和分析,首先提出SIE 3D SLAM技术的通用框架并对其进行了详细的介绍,着重介绍了框架中的激光里程计模块、回环检测模块和语义地图构建模块,然后对语义信息在前端里程计、后端回环和地图构建中的应用进展进行了详细的阐述。通过对课题的深入研究和探索,认识到语义信息引入在3D激光SLAM技术中的应用有待更加深入的发掘。比如在激光点云场景下的

多目标动态跟踪任务中更加深入的划分动态点云,这种更深入的划分有助于我们不忽略细小对象的信息特征,尽可能的最大程度感知原始点云场景的环境信息,这不仅对于同步定位与建图任务本身有益,更有益于后续无人平台跟踪规划的复杂任务。另外,闭环检测作为当前完整SLAM流程重要的一个模块,可以在全局地图层面进行数据关联,帮助算法建立闭环约束,更进一步修正预测轨迹的累计误差和全局一致性的地图映射,结合语义特征和几何特征的闭环检测算法也是目前研究的重点和难点。语义信息在3D激光SLAM方面的应用前景非常广泛,意义重大。

参考文献

- [1] 焦传佳,江明,徐劲松,等. 基于激光信息的移动机器人定位研究[J]. 电子测量与仪器学报, 2021, 35(9): 1-9.
JIAO CH J, JIANG M, XU J S, et al. Research on positioning of mobile robot based on laser information[J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(9): 1-9.
- [2] 周治国, 曹江微, 邸顺帆. 3D激光雷达SLAM算法综述[J]. 仪器仪表学报, 2021, 42(9): 13-27.
ZHOU ZH G, CAO J W, DI SH F. Overview of SLAM algorithm for 3D lidar[J]. Chinese Journal of Scientific Instrument, 2021, 42(9): 13-27.
- [3] 蒋林, 刘奇, 雷斌, 等. 激光与视觉融合识别并构建语义地图改善定位性能[J]. 中国激光, 2022, 49(18): 140-154.
JIANG L, LIU Q, LEI B, et al. Laser and vision fusion identify and build semantic map to improve positioning performance[J]. China Laser, 2022, 49(18): 140-154.
- [4] 王常虹, 窦赫喧, 陈晓东, 等. 无人平台SLAM技术研究进展[J]. 导航定位与授时, 2019, 6(4): 12-19.
WANG CH H, DOU H X, CHEN X D, et al. Research progress in simultaneous location and mapping for unmanned vehicles [J]. Navigation Positioning and Timing, 2019, 6(4): 12-19.
- [5] GAO Y, LIU S, ATIA M M, et al. INS/GPS/LiDAR integrated navigation system for urban and indoor environments using hybrid SCAN matching algorithm[J]. Sensors, 2015, 15(9): 23286-23302.
- [6] BESL P J, MCKAY N D. A method for registration of 3D shapes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(2): 239-256.
- [7] CHEN Y, MEDIONI G. Object modelling by registration of multiple range images [J]. Image and Vision

- Computing, 1992, 10(3): 145-155.
- [8] LUF, MILIOS E E. Robot pose estimation in unknown environments by matching 2D range scans[J]. Journal of Intelligent and Robotic Systems, 1997, 18(3): 249-275.
 - [9] CENSI A. An ICP variant using a point-to-line metric[C]. IEEE International Conference on Robotics and Automation (ICRA), 2008: 19-25.
 - [10] SEGAL A, HAEHNEL D, THRUN S. Generalized-ICP[C]. Robotics: Science and Systems (RSS), 2009, 2(4): 435.
 - [11] JENSELT P, KRISTENSEN S. Active global localization for a mobile robot using multiple hypothesis tracking [J]. IEEE Transactions on Robotics and Automation, 2001, 17(5): 748-760.
 - [12] LIU S, ATIA M M, GAO Y, et al. Adaptive covariance estimation method for LiDAR-aided multi-sensor integrated navigation systems[J]. Micromachines, 2015, 6(2): 196-215.
 - [13] ZHANG J, SINGH S. Low-drift and real-time lidar odometry and mapping[J]. Autonomous Robots, 2017, 41(2): 401-416.
 - [14] RUSU R B, BLODOW N, BEETZ M. Fast point feature histograms (FPFH) for 3D registration [C]. IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2009: 3212-3217.
 - [15] SHU L, XU H, HUANG M. High-speed and accurate laser scan matching using classified features[C]. IEEE International Symposium on Robotic and Sensors Environments (ROSE), IEEE, 2013: 61-66.
 - [16] BIBER P, STRABER W. The normal distributions transform: A new approach to laser scan matching[C]. Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No. 03CH37453), IEEE, 2003, 3: 2743-2748.
 - [17] MAGNUSSON M, LILIENTHAL A, DUCKETT T. Scan registration for autonomous mining vehicles using 3D-NDT[J]. Journal of Field Robotics, 2007, 24(10): 803-827.
 - [18] CIHAN U, TEMELTA H. 3D multi-layered normal distribution transform for fast and long range scan matching[J]. Intelligent and Robotic Systems, 2013, 71(1): 85-108.
 - [19] SAARINEN J, ANDREASSON H, STOYANOV T, et al. Normal distributions transform occupancy maps: Application to large-scale online 3D mapping[C]. IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2013: 2233-2238.
 - [20] 陈天堂. 语义对象分割方法研究[D]. 成都: 电子科技大学, 2012.
 - CHEN T T. Research on semantic object segmentation method[D]. Chengdu: University of Electronic Science and Technology, 2012.
 - [21] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431-3440.
 - [22] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
 - [23] ZHAO H, QI X, SHEN X, et al. ICNet for real-time semantic segmentation on high-resolution images [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 405-420.
 - [24] BADRINARAYANAN V, KENDALL A, CIPOLLA R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
 - [25] BEHLEY J, GARBADE M, MILIOTO A, et al. Semantickitti: A dataset for semantic scene understanding of lidar sequences[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 9297-9307.
 - [26] HACKEL T, SAVINOV N, LADICKY L, et al. Semantic3D.net: A new large-scale point cloud classification benchmark[J]. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2017, DOI: 10.5194/isprs-annals-IV-1-W1-91-2017.
 - [27] GAO B, PAN Y, LI C, et al. Are we hungry for 3D LiDAR data for semantic segmentation? A survey of datasets and methods [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(7): 6063-6081.
 - [28] ZHANG C, LUO W, URTASUN R. Efficient convolutions for real-time semantic segmentation of 3D point clouds[C]. International Conference on 3D Vision, IEEE, 2018: 399-408.
 - [29] 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.
- [30] TCHAPMI L, CHOY C, ARMENI I, et al. Segcloud: Semantic segmentation of 3D point clouds [C]. International Conference on 3D Vision, IEEE, 2017: 537-547.
- [31] CHEN X, MILIOTO A, PALAZZOLO E, et al. Suma++: Efficient lidar-based semantic slam [C]. IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2019: 4530-4537.
- [32] MILIOTO A, VIZZO I, BEHLEY J, et al. Rangenet++: Fast and accurate lidar semantic segmentation [C]. IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2019: 4213-4220.
- [33] CHEN X, LÄBE T, MILIOTO A, et al. OverlapNet: Loop closing for LiDAR-based SLAM [J]. Autonomous Robots, 2021, 46: 61-81.
- [34] 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.
- [35] LANDRIEU L, SIMONOVSKY M. Large-scale point cloud semantic segmentation with superpoint graphs [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4558-4567.
- [36] SU H, JAMPANI V, SUN D, et al. Splatnet: Sparse lattice networks for point cloud processing [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 2530-2539.
- [37] QI C R, YI L, SU H, et al. Pointnet ++: Deep hierarchical feature learning on point sets in a metric space [C]. Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 5105-5114.
- [38] YAN X, ZHENG C, LI Z, et al. Pointasnl: Robust point clouds processing using nonlocal neural networks with adaptive sampling [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 5589-5598.
- [39] WU W, QI Z, FUXIN L. Pointconv: Deep convolutional networks on 3D point clouds [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 9621-9630.
- [40] VAN GANSBEKE W, NEVEN D, DE BRABANDERE B, et al. Sparse and noisy lidar completion with rgb guidance and uncertainty [C]. International Conference on Machine Vision Applications (MVA), IEEE, 2019: 1-6.
- [41] WANG W, YOU X, ZHANG X, et al. LiDAR-based SLAM under semantic constraints in dynamic environments [J]. Remote Sensing, 2021, 13 (18): 3651.
- [42] BEHLEY J, STACHNISS C. Efficient surfel-based SLAM using 3D laser range data in urban environments [C]. Robotics: Science and Systems, 2018: 59.
- [43] PARKISON S A, GAN L, JADIDI M G, et al. Semantic iterative closest point through expectation-maximization [C]. British Machine Vision Conference, 2018: 280.
- [44] ZAGANIDIS A, SUN L, DUCKETT T, et al. Integrating deep semantic segmentation into 3-D point cloud registration [J]. IEEE Robotics and Automation Letters, 2018, 3 (4): 2942-2949.
- [45] KELLEY C T. Iterative Methods for Optimization [M]. Philadelphia: Society for Industrial and Applied Mathematics, 1999.
- [46] LI L, KONG X, ZHAO X, et al. SA-LOAM: Semantic-aided LiDAR SLAM with loop closure [C]. IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2021: 7627-7634.
- [47] QIAN Z T, FU J, XIAO J. Towards accurate loop closure detection in semantic SLAM with 3D semantic covisibility graphs [J]. IEEE Robotics and Automation Letters, 2022, 7 (2): 2455-2462.
- [48] WANG G, GAO S, DING H, et al. LIO-CSI: LiDAR inertial odometry with loop closure combined with semantic information [J]. Plos One, 2021, 16 (12): e0261053.
- [49] 何松, 孙静, 郭乐江, 等. 基于激光SLAM和深度学习的语义地图构建 [J]. 计算机技术与发展, 2020, 30 (9): 88-94.
- HE S, SUN J, GUO L J, et al. Semantic map construction based on laser SLAM and deep learning [J]. Computer Technology and Development, 2020, 30 (9): 88-94.
- [50] VASUDEVAN S, GCHTER S, NGUYEN V, et al. Cognitive maps for mobile robots—An object based approach [J]. Robotics and Autonomous Systems, 2007, 55 (5): 359-371.
- [51] VASUDEVAN S, SIEGWART R. Bayesian space conceptualization and place classification for semantic maps in mobile robotics [J]. Robotics & Autonomous Systems, 2008, 56 (6): 522-537.
- [52] RITUERTO A, MURILLO A C, GUERRERO J J. Semantic labeling for indoor topological mapping using a

- wearable catadioptric system [J]. Robotics and Autonomous Systems, 2014, 62(5):685-695.
- [53] LIU M, COLAS F, POMERLEAU F, et al. A Markov semi-supervised clustering approach and its application in topological map extraction [C]. IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2012: 4743-4748.
- [54] LIU M, COLAS F, SIEGWART R. Regional topological segmentation based on mutual information graphs [C]. IEEE International Conference on Robotics and Automation, IEEE, 2011: 3269-3274.
- [55] BRUNSKILL E, KOLLAR T, ROY N. Topological mapping using spectral clustering and classification [C]. IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2007: 3491-3496.
- [56] LIU Z, CHEN D, VON WICHERT G. Online semantic exploration of indoor maps [C]. IEEE International Conference on Robotics and Automation, IEEE, 2012: 4361-4366.
- [57] PRNOBIS A, MARTINEZ MOZOS O, CAPUTO B, et al. Multi-modal semantic place classification [J]. The International Journal of Robotics Research, 2010, 29(2-3):298-320.
- [58] PRNOBIS A, JENSFELT P. Large-scale semantic mapping and reasoning with heterogeneous modalities [C]. IEEE International Conference on Robotics and Automation, IEEE, 2012: 3515-3522.
- [59] GOEDDEL R, OLSON E. Learning semantic place labels from occupancy grids using CNNs [C]. IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2016: 3999-4004.
- [60] HILLER M, QIU C, PARTICKE F, et al. Learning topometric semantic maps from occupancy grids [C]. IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2019: 4190-4197.
- [61] LIMKETKAI B, LIAO L, FOX D. Relational object maps for mobile robots [C]. International Joint Conference on Artificial Intelligence, 2005: 1471-1476.
- [62] SÜNDERHAUF N, PHAM T T, LATIF Y, et al. Meaningful maps with object-oriented semantic mapping [C]. IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2017: 5079-5085.
- [63] JEONG J, YOON T S, PARK J B. Towards a meaningful 3D map using a 3D lidar and a camera [J]. Sensors, 2018, 18(8):2571.

作者简介



周治国 (通信作者), 1998 年于华中科技大学获得学士学位, 2004 年于北京理工大学获得硕士学位, 2009 年于北京理工大学获得博士学位, 现为北京理工大学副教授, 主要研究方向为实时半实物仿真和智能无人飞行器信息感知与导航。

E-mail: zhiguozhou@ bit. edu. cn

Zhou Zhiguo (Corresponding author) received his B. Sc. degree from Huazhong University of Science and Technology in 1998, M. Sc. degree and Ph. D. degree both from Beijing Institute of Technology in 2004 and 2009, respectively. He is currently an associate professor at Beijing Institute of Technology. His main research interests include real-time hardware-in-the-loop simulation, information perception and navigation of intelligent unmanned vehicle.