



引文格式:徐连瑞,游雄.任务驱动视角下机器地图现状与发展[J].武汉大学学报(信息科学版),2024,49(4):609-623.DOI:10.13203/j.whugis20220578

Citation: XU Lianrui, YOU Xiong. A Task-Driven Perspective on Status and Development of Machine Map[J]. Geomatics and Information Science of Wuhan University, 2024, 49(4): 609-623. DOI: 10.13203/j.whugis20220578

任务驱动视角下机器地图现状与发展

徐连瑞¹ 游雄¹

¹ 信息工程大学地理空间信息学院,河南 郑州,450001

摘要:随着国家对机器人发展战略的不断推进,保障机器人认知和学习的机器地图近年来成为地图科学研究的新方向。针对目前机器地图理论研究相对匮乏的问题,基于测制用一体运行机制,阐述总结任务驱动下的机器地图关键技术和研究现状。数据获取与处理是地图模型构建和任务应用的基础支撑,围绕机器人平台、特征提取、语义分割、多传感器融合等技术现状进行归纳总结;地图模型构建承上启下,分析参考常用的地图模型架构及其特点,描述不同任务情境中如何搭建使用弹性、稳健、可靠的建图系统;任务应用是机器地图功能性的集中体现,介绍总结路径规划、目标检测、知识表达与推理等典型应用研究现状。论述展望了机器地图存在的问题及未来发展方向。

关键词:任务驱动;机器地图;测制用一体运行机制;认知学习

中图分类号:P208

文献标识码:A

收稿日期:2022-12-06

DOI:10.13203/j.whugis20220578

文章编号:1671-8860(2024)04-0609-15

A Task-Driven Perspective on Status and Development of Machine Map

XU Lianrui¹ YOU Xiong¹

¹ School of Geospatial Information, Information Engineering University, Zhengzhou 450001, China

Abstract: With the continuous promotion of national robot development strategy, machine map, which guarantees robot cognition and learning, has become a new direction of map science research in recent years. To address the lack of theoretical research on machine map, we summarize the key technologies and research status of task-driven machine maps based on the integrated sensing, mapping and decision-making mechanism. Data acquisition and processing are the basic support for map model construction and task application. And the state of robot platform, feature extraction, semantic segmentation and multi-sensor fusion are summarized. Map model construction is the top and bottom. The commonly used map model architecture and its characteristics are analyzed to describe how to build a resilient, robust, and reliable map building system in different task contexts. Task application is the central expression of the functionality of machine maps. The task applications are the concentrated manifestation of the functionalities of machine maps. And the current status of research on typical applications such as path planning, target detection, knowledge representation and reasoning is introduced and summarized. The problems and future development directions of machine map are discussed.

Key words: task-driven; machine map; integrated sensing, mapping and decision-making; cognitive learning

近十年来,随着科学研究对大脑认知机理的不断发掘,很多工作试图将人类对环境的认知思维和学习能力“赋予”机器人,使其具备自主感知、思考、决策的能力^[1-7]。机器地图作为一种全

新地图范式,以机器人为使用主体,按照一定的数学法则,综合多维传感器环境感知特征,构建以环境要素及其相互关系描述为重点的客观世界模型,服务于机器人认知、学习、决策等一系列

基金项目:国家自然科学基金(42130112);中原学者游雄科学家工作室科研项目(2020);信息工程大学科研项目(1064201)。

第一作者:徐连瑞,博士生,主要从事环境智能感知和机器地图建模工作。1793674773@qq.com

通讯作者:游雄,博士,教授。youarexiong@163.com

自主行为的规划与执行。国内针对机器人环境建模与功能性探索的研究起步较晚,20世纪90年代出现的以数字形式记录、存储的电子地图,让人们意识到位置、路径等信息的快速检索和应用给出行带来了极大的便利^[8],随后导航电子地图的发展和相关标准的制定为无人汽车基于任务的应用打下了坚实基础^[9]。高精地图脱胎于导航

电子地图,作为国家推动智能汽车创新发展战略的重要组成部分,是现阶段L3级及以上无人驾驶必备的数据基础^[10]。作为客观世界部分空间信息表达的载体,高精地图能够实现无人系统数据采集、数据一体运行表达、环境认知、轨迹规划、地图融合、矢量化表达、地图更新等功能,如图1所示。

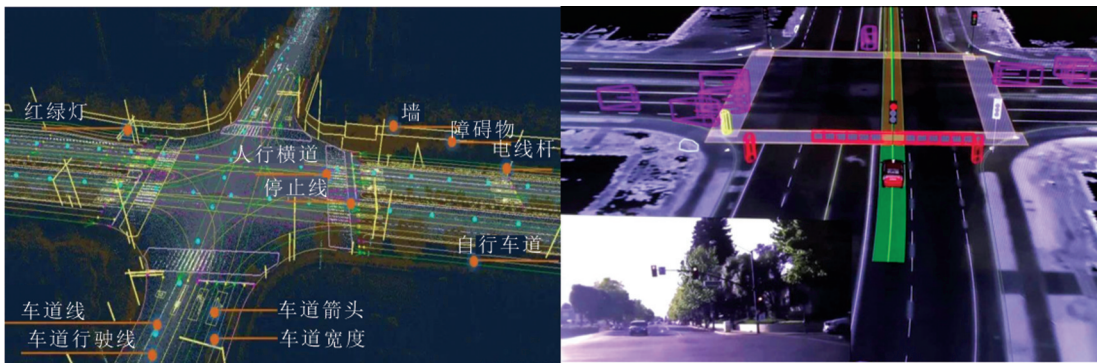


图1 高精地图示例

Fig. 1 Example of High Definition Map

目前,机器地图相关研究多侧重于提高地图模型的构建精度与效率,忽略了机器人任务执行的实时性与成功率。为有效叠加任务空间与地图空间,实现信息共通、共融、共享,机器地图测制用一体运行机制^[11]深度耦合测图、制图、用图环节,以用图任务为驱动,将几何建模、场景解译、制图建模、任务评估和空间决策等功能集成在高度一体的框架中,通过各个模块的密切联动、相互合作,共同完成智能体对环境的感知、认知和决策过程。同时,任务驱动的机器地图能够基于“按需、动态”的智能化建图理念,让机器通过学习的方法重点关注实例要素的可度量性、空间关系的可描述性、环境模型的可解释性^[12-14],将测制用一体运行机制置于循环过程中,通过学习几何、属性、拓扑、空间关系等要素扩展地图信息表达维度,形成有效信息回路。如图2所示,在前端完成感知信息从数字化到信息化再到知识化的转变,在后端嵌入任务模块,加强地理信息的获取更新、知识表达、空间推理、控制决策等集成应用。Apollo^[15]作为百度公司推出的自动驾驶开放平台,已在自动驾驶、智能汽车、智能交通三大领域提出解决方案,包含定位、感知、规划、导航四大核心模块,具备海量数据的仿真引擎,可以在使用过程中实现高效的信息共享与反馈,目前已为超过百万的用户提供在线或离线服务。由美国国防高级研究计划局发起的地下挑战赛(subterranean challenge, SubT)^[1],旨在促进机器人系统于没有先验信息的恶劣地下环境中自主

可靠地探索,分享反馈多模态数据,构建地图模型,并对指定物体提供“救援”,涉及自主探索、目标探测、定位建图、规划导航、多机器人协同、任务界面设计等子功能模块。机器人协同技术联盟^[2,16](robotics collaborative technology alliances, RCTA)项目由美国陆军研究实验室主持,以度量地图和语义地图双向链接为基础,针对无人系统认知度量世界模型、语义感知、自适应行为生成、元认知和深度学习等基本任务,构建“想”“看”“说”“动”“工作”拟人化功能模块,实现环境信息的学习、整合、推理,形成自适应、多尺度、可推理的数字地图,高效主动地与作战人员实现交互^[17-18]。

上述研究都为机器地图在世界模型、态势感知、信息交互、任务推理等方面提供了研究基础,通过人工智能、人机交互等关键技术的创新,有效促进了机器地图理论与应用体系的形成。本文基于任务驱动理念,结合机器地图相关概念和前景需求,主要围绕数据获取与处理、地图模型构建、任务应用3个方向总结阐述机器地图发展现状,为机器地图研究提供理论参考和技术支撑。

1 数据获取与处理

数据获取与处理在机器地图体系结构中发挥“基石”作用,是改善数据质量、增强数据适配性、提高数据和应用耦合度的有效手段。机器人

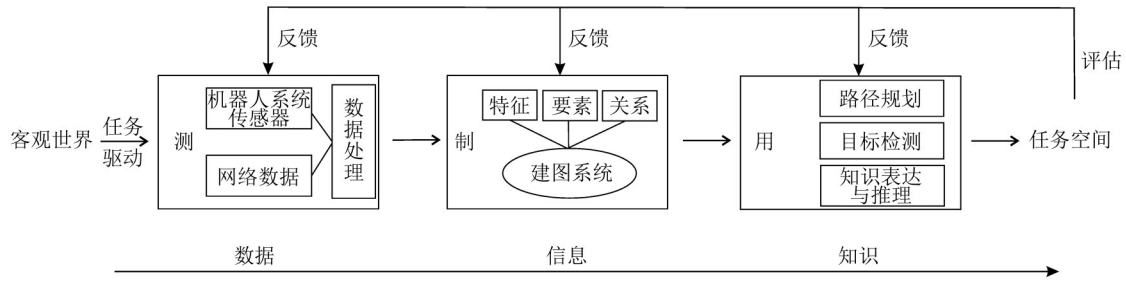


图2 测制用一体运行机制

Fig. 2 Integrated Sensing, Mapping and Decision-Making Mechanism

系统与客观环境实现交互获取数据,数据输入包括传感器数据和网络先验信息。传感器数据实时性强,不同数据类型之间形成有效互补;网络先验信息体量大、发散性强、语义丰富,能够与传感器数据交叉验证,缩小传感器盲区。两类数据通过前端数据处理共同形成机器地图高质量数据基础。

1.1 机器人平台

任务需求引领与设计理念更新是机器人硬件系统快速发展的重要成因。目前对于客观世界的探索,通常根据环境特点与任务需求,组合不同类型的机器人形成可靠的机器人团队^[19-20],如图3所示。

SubT挑战赛中,针对陡坡、泥泞、碎石、楼梯等地形障碍,CTU-CRAS团队充分考虑越障能力、工作时长、探测范围等因素,组建了一个包含履带式机器人Absolem、轮式机器人Husky、六足

机器人PhantomX Mark和多旋翼飞行器DJIF450的机器人团队,高效完成了机器人团队于恶劣地形条件下对未知环境的探索任务^[1]。CoSTAR团队在使用轮式机器人Husky、履带式机器人Telex Pro和四足机器人Spot的基础上,将四旋翼飞行器替换为可高速移动的遥控车,解决了多旋翼飞行器工作时长受限问题^[21]。CERBERUS团队将四足机器人作为探索主体,配合小型耐碰撞飞行平台和大型多旋翼系统,为探索复杂环境提供必要的障碍翻越能力和续航能力^[22]。相较于地下空间复杂的地形环境,地面道路体系与环境平整度、完备性较高,高精地图、RCTA项目的动态数据采集多依赖装备光探测和测距、相机、惯导等传感器的轮式车辆和无人机,配合卫星数据链的信息传输,形成对无人系统的有效控制^[2,23],完成导航、自适应战术推理等任务^[24]。



图3 SubT挑战赛中的机器人团队

Fig. 3 Robot Team in SubT Challenge

1.2 特征提取

当数据传入处理模块后,特征的有效提取尤为重要。围绕点云数据的特征提取,人工设计的方法有局部特征描述子、全局特征描述子和混合特征描述子,描述特征包括法向量、曲率、特征熵、发散性等^[25-32]。虽然人工设计的特征描述子在特定环境中具有较强的针对性,但泛化性不强,而深度学习和强化学习在点云数据特征提取

方面表现出巨大潜力,并具备一定的自主学习能力,以目标为导向最大化收益信号,能够高效地从浅层次数据信息中学习到深层次数据特征^[33-36]。考虑到点云的无序性和非结构化,基于深度学习方法提取点云特征可以分为直接法与间接法。间接法主要通过投影、区域划分将点云表示为具有有序结构的图像、体素、拓扑图等数据类型,弱化其无序性,再利用基于有序数据的

特征提取方法提取特征。文献[37]使用多视图卷积方法,将点云投影到不同视角下,对其视觉特征进行分组,通过特征之间的关联扩展特征信息的描述维度;文献[38]采用端到端框架,将点云数据体素化,使用三线性插值和全连接条件随机场将体素与点云的细粒度表示结合,增强特征提取的全局一致性;文献[39]根据点云数据密度,基于非平衡二叉树结构划分空间,在子节点内存存储特征,根据分辨率变化调整节点数量,在减少计算内存的前提下,实现更深层次的网络特征提取。但间接法在点云选取、划分、重组过程中,存在信息丢失、数据结构保存不完整的缺陷,因此直接法提取点云特征,跳过数据类型变换和划分重组步骤,保留数据结构完整性,通过局部点之间的空间依赖关系表征点的高层次空间特征与几何变换;文献[40]结合 PointSIFT 提出一种离散单元采样模块分割算法,抑制周围噪声点冗余影响,提高了网络对于复杂结构点云的分割效果;文献[41]围绕大规模点云数据特征提取问题,使用均匀分割场景中获取的超点图组织特征关系,并对邻域点提供充分的目标上下文描述。文献[42]为模拟点与点之间的相对关系,提出 PointWeb 描述点特征,使用自适应特征调整模块实现局部特征学习,并构建局部完全链接网络,形成特征聚合与提取。虽然直接法相对间接法在效率和精度上有了较大进步,但也在一定程度上忽略了点云邻域之间的相互作用,邻域之间的细节保留较少。

不同于点云数据,图像像素排列的有序性和规则化易于深度学习直接提取特征。卷积神经网络以卷积核滑动方式加权组合相邻感受野信息,配合池化层与全连接层输出最终特征向量^[43]。考虑到图像特征提取任务的差异性,不同的卷积方式能够有效提高图像特征的提取质量。空洞卷积^[44]通过在卷积核内注入空洞,扩大感受野,使卷积输出包含更大范围的特征信息,解决因图像分辨率低或特征丢失导致的问题;可变形卷积^[45]相较于空洞卷积具有更强的调整能力,通过在卷积核每个元素上添加方向参数获得变形偏移量,实现尺度和感受野的自适应,提高特征提取与应用需求的适配性;非局部卷积^[46]通过嵌入式高斯函数、点积等形式构造卷积结构,将位置的响应表示作为位置的特征进行加权平均,主要用于处理远距离像素间的相互关系,增强提取特征在空间上的关联性。这些更加精细化的卷

积方式能够在最大限度保证局部连接性和权值共享性的基础上处理图像,以适配不同的特征提取需求。同时,网络结构轻量化也是特征提取模块在机器人平台部署的关键因素,AlexNet^[47]、GoogleNet^[48]、ResNet^[49]、UNet^[50]、EspNet^[51]等都在减少网络参数、提高计算效率方面作出贡献,一定程度上解决了由于网络复杂导致的部署难、训练难、过拟合等问题。

1.3 语义分割

点云作为机器地图的主要表现形式,具备可度量、易表征、具象化的优点。相较于图像,点云采集不受光照、位姿影响,空间信息更丰富;相较于体素,点云空间利用率更高,不会因过度表示物体内部空间产生冗余^[52]。随着几何特征局限性的不断凸显,机器地图深层次语义特征往往更能代表局部点云的整体特性,因此语义分割在机器地图数据处理过程中占据重要地位^[53-54]。语义分割本质上是给每一个点分类并打上标签,根据特征相似性将点聚类,如图4所示。现阶段语义分割主要分为基于投影的方法和基于点的方法,基于投影的方法包括多视图表示、球面表示、体素表示、栅格表示、柱面表示、混合表示,基于点的方法包括邻域特征池化、不规则点云卷积、循环神经网络建模、图网络^[36]。

PointNet^[55]最先使用深度学习对点云进行语义分割,该方法基于点云置换不变性和旋转不变性,使用多层感知机提取每个点特征,通过最大池化聚合全局特征,实现点的类别预测。后续 PointNet++^[56]借鉴多层感受野思想,使用内积不断扩充感受野,叠加不同分辨率下提取的局部特征,解决了特征丢失的问题。点云近密远疏是由于 LiDAR 工作机制无法避免的问题,针对点云在空间中的不均匀分布,PolarNet^[57]从点云的网格划分入手,以自上而下的角度用球形网格划分点云,设计环卷积组成特征学习模块,输出点的量化预测。Cylinder3D^[58]改进了网格划分思想,设计圆柱坐标划分和非对称卷积网络,提高了网格划分的非空网格比例,有效提升了语义分割精度。围绕大规模场景,RandLA-Net^[59]基于海量点云数据,采用随机下采样方法提高逐点计算效率,配合局部特征聚合模块,通过编码解码结构组建网络。BLNet^[60]在 RandLA-Net 基础上,利用特征差和相对位置差,基于注意力机制学习聚合特征,增强了语义分割在大规模场景应用中的稳健性。还有方法创新性地引入视觉注意力机



图 4 语义分割示例

Fig. 4 Example of Semantic Segmentation

制、形状上下文等方法对点云进行语义分割,形成对周边环境的定性描述,也大大增强了机器地图对空间的认知解释能力^[61]。

1.4 多传感器融合

基于同时定位与地图构建(simultaneous localization and mapping, SLAM)框架实现的多传感器融合是处理多模态数据的常用技术手段。多传感器信息的有效融合不仅能够扩展环境感知的数据维度,还能够增强数据的准确性和可靠性。统一的时空基准是多传感器信息描述环境的必要前提,文献[62]在统一坐标系和观测物体的基础上,使用不同传感器局部坐标系换算保证了数据在空间上的一致性;文献[63]使用时间戳标记处理异构数据,基于样条函数拟合不同时刻下的数据,实现时间基准的统一。此外,针对传感器工作频率不一致问题,文献[64]采用预积分方法共同优化机器人状态量和图像关键帧位姿,保持传感器数据采集频率的一致性。但无论是主动还是被动传感器,都无法避免因外界环境因素干扰产生的数据噪声。早期对于噪声的处理多以卡尔曼滤波^[65]为主,逐渐发展为扩展卡尔曼滤

波^[66]、多状态约束卡尔曼滤波^[67]等方法。文献[68]对传统卡尔曼滤波和因子图优化方法进行改进,采取流水线多层次的处理策略融合数据,屏蔽失效传感器的数据传入,以此减少数据产生的噪声。文献[69]将基于特征法的关键帧与惯性传感器数据紧耦合,通过构建重投影误差函数优化视觉-惯性里程计数据,提高数据的可靠性。同时也有一些研究以弱监督或无监督的方式实现端到端的多传感器融合,该方式可以在未知传感器参数的情况下执行SLAM后端操作,实现信息融合,增强系统鲁棒性^[70-73]。目前,随着机器学习、概率模型、视觉处理等技术以及传感器硬件的快速发展,嵌入式处理器功耗不断降低,并行计算能力不断加强,基于SLAM的多传感器融合技术将在机器人平台遂行任务过程中发挥越来越重要的作用。

2 地图模型构建

地图模型构建是机器地图测制用一体运行机制的核心环节,在机器人运行回路中起到承上启下的作用。对于机器地图,高精度的几何观测和多维特征感知是建图的必要前提,复杂任务的适配性和多特征联合要素解译需求也对地图模型架构提出更高要求,因此,层次化、弹性、可靠的地图模型构建需要综合考量以上因素。

2.1 地图模型

机器地图需要满足准确性、完整性、可验证性、可扩展性的需求,高效的地图模型应能准确地表征客观环境,结合多尺度、多层次结构,实现数据标定、存储、索引、可视化等功能。目前国内研究对于机器地图模型的标准尚未统一,本文通过借鉴研究具有类似功能的地图模型形成相关总结认识。

Lanelets^[74]作为一种应用于自动驾驶的轻量级地图模型,包含地图最小单元 lanelet、规则元素、驾驶通道 3 个模块,最初奔驰公司将其作为路径规划系统为无人车辆提供服务。局部动态地图^[75]引入数据库集成管理传感器信息,根据数据特点形成分层地图模型:第一层底层数据主要由道路数据构成,第二层相对静态数据包含一些地图信息中不包含的信号,第三层相对动态数据包含交通拥堵状况等信息,第四层顶层数据由实时传感器数据组成。Lanelet2^[76]具备独立车道概念,更加注重实际应用,并在 Lanelets 基础上将地图模型划分为物理层、关联层、拓扑层,形成要素

的体系化表达,能够实现复杂环境下车辆的控制规划。文献[77]认为不同的导航任务对于地图精度的需求是不同的,提出了一个7层地图模型,包含道路层、交通信息层、路巷连接层、Lane图层、映射特征层、动态对象容积层、智能决策支持层,每一图层都包含不同的数据类型,不同图层之间可以相互独立调用或适配任务需求联合索引,共同支持各种导航任务类型。文献[78]认为智能高精地图的核心在于自适应场景变换和任务需求,通过评估反馈不断优化地图模型,因此地图模型数据逻辑结构中应包含丰富准确的语义信息及驾驶经验信息,并依此将模型结构划分为4层:静态地图层、实时数据层、动态数据层、用户模型层。

2.2 建图系统

SLAM框架经过古典时期、算法分析时期、鲁棒感知时期的不断发展^[79],智能化程度越来越高,学习能力越来越强,目前已经成为机器地图

建图系统的有力支撑。但在当前SLAM技术的发展中,对于地图的研究仅限于作为实现SLAM算法的中间过程,没有向“用图”方向延伸,地图内容简单、表达形式简略、质量有待提高,并且缺少对于机器地图形式与建图系统的深入研究,更多的是作为一种技术手段服务于机器地图的数据处理与模型构建,在建图系统中发挥类似“测图”的作用。

早期的建图系统多采用马尔可夫随机法、条件随机场、随机决策森林等方法,方便数据在地图上标记使用,但这些方法因为离线或精度问题并不能部署在机器人平台上实时处理数据^[80-82]。任务驱动在建图系统对于客观环境的表征有多种形式,如图5所示,大多地图模型采用分层次多图层结构,以度量地图和拓扑地图为底层,叠加高精度可解释的语义层,配合规则库以及实体与规则库交互的关联层,共同组成机器地图的整体架构。

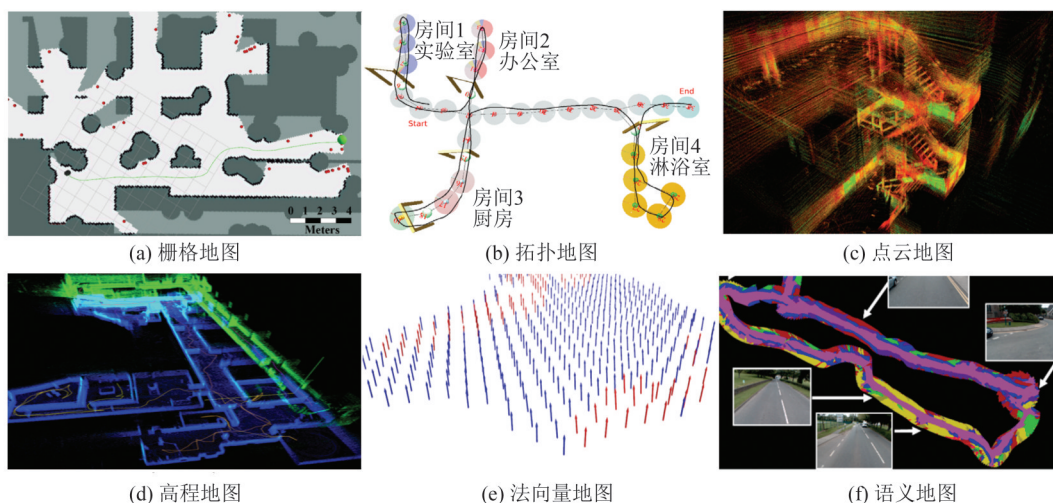
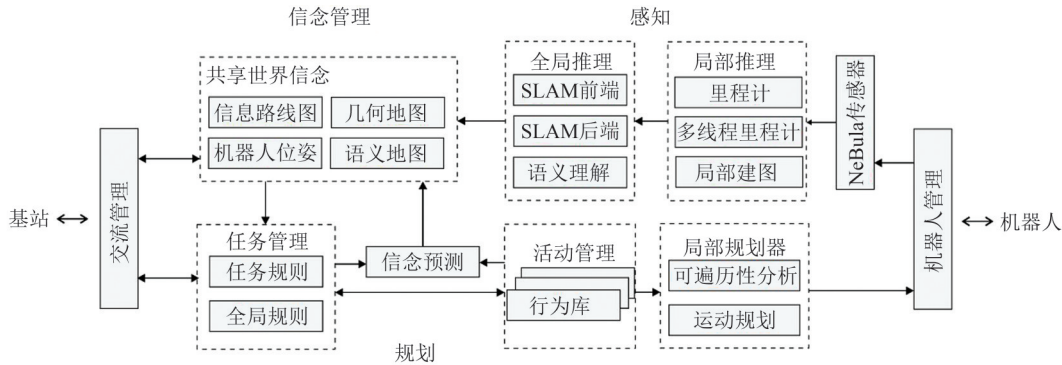


图5 常见地图表征类型

Fig. 5 Common Map Representations

围绕建图系统部署和计算资源问题,轻量化的 SEMANTIC-RTAB-MAP 算法^[83]采用 RTABMAP^[84]和 YOLO^[85]作为基础架构,通过目标检测估算实例位置,以非深度学习的方法完成建图。叠加多种图层能够充分表征环境特点,特别是在未知或有遮挡的环境中,为获得更高的环境表征效率,可以通过语义信息预测环境,根据环境特点判断拓扑地图的连通性,再对度量地图中的物体位置进行推理^[86-87]。针对环境中的不确定性,仅凭预测和推理不具备较高的置信度,需要结合相应的概率计算和表示。多保真度弹性建图系统 NeBula 原型由美国宇航局喷气推进实验室为探索月球、火星和其他行星体上的极端

表面和地下地形而设计,能够在不确定的环境中预测和评估各类结果和风险,协调建图模块可靠地探索未知地形^[88]。为应对 SubT 地下恶劣环境设计的“网络化自主信念感知系统”基于 NeBula 作出改进,将建图的前端模块组合成端到端系统,结合感知定位与映射、弹性状态估计、全局感知规划,创建机器地图的概率表示,计算风险,以最小代价完成地图构建^[21](见图6)。RCTA^[2]项目中,机器人以观测对象之间的空间关系作为约束,对于无遮挡的环境直接栅格化处理,得到高置信度的代价地图;对于有遮挡的环境通过概率预测的方式完成几何空间推理,根据上下文关系实现满足机动约束的代价地图构建。

图6 NeBula网络化自主信念感知系统^[88]Fig. 6 NeBula Autonomous Belief Awareness Network System^[88]

传统的地图模型构建更新维护过程无法适应不断变化的环境信息,并且需要消耗大量的资源和人力,导致其扩展性和实时性无法满足任务需求。文献[89]基于传感器信息的多样性和实时性,提出一种在线地图构建系统,包含透视图像编码器、神经视图变换器、点云编码器和地图元素解码器4个模块,通过地图元素的矢量化和语义化,完成语义分割、实例分割、道路预测等功能。为实现端到端的地图构建,减少人工参与,文献[90]提出了一种层次化建图网络,直接编码矢量化场景信息,以图网络表示对象之间交互关系,基于节点特征预测缺失信息,以此构建高精度的环境地图。综上,任务驱动的建图系统针对不确定性需要较多的计算资源,并且要对环境反馈数据进行实时处理,多形式表征的地图模型也是任务过程中机器人与环境交互的需要,减少人工参与的端到端建图系统是当前的重要发展方向。

3 任务应用

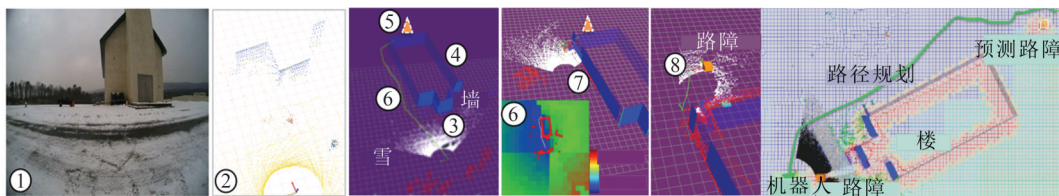
任务应用是机器地图功能性的集中体现,也是机器地图系统保障机器人动作执行的最终端。本文在梳理有关机器人任务应用相关文献的基础上,结合对任务应用关键词的查询,同时参考RCTA项目对无人平台保障系统的服务愿景,发现:在复杂的客观世界中,使用机器地图保障任务执行时,安全、可靠和适应性的移动是机器人首先需要解决的问题,也是规划开展任务应用的基础;在机器人与物理世界的交互过程中,目标检测为交互过程提供潜在的对象实体,减少目标搜索中的计算资源消耗,提高终端任务执行效率;实时融合和更新的地图模型记录了丰富的空间知识,包含空间关系、语义关系、空间格局等,

基于机器地图模型的知识表达与推理是机器人自主学习、积累、智能化的有效方法,也是一个渐进、迭代的交互过程,更是机器人形成拟人化认知经验与规则的底层驱动。因此,本文对路径规划、目标检测、知识表达与推理等3个机器地图代表性任务应用现状进行归纳总结。

3.1 路径规划

基于任务驱动的路径规划,需要正确、全面、高效地理解识别目标点,选择合适、安全、代价最短的路径,引导控制机器人执行运动指令,导航至目标点。传统方法通过A*算法^[91]、蚁群算法^[92]、Dijkstra算法^[93]等,寻找起始点与目标点之间的最短路径。任务驱动的路径规划需要考虑更多因素,高精地图^[78]重视路径规划中地图与人的交互,将道路语义和用户语义考虑其中,引入用户体验数据集平滑优化轨迹,直至生成满足用户需求的无碰撞可执行轨迹。为应对人机交互的扩展,机器人不再局限于已知目标点的路径规划,文献[2]尝试将以确定坐标表示的目标点替换为以环境中地物空间关系表示的目标点,例如“去到这栋楼的东南角”,以机器人感知到的世界模型变化作为基于任务的上下文推理器的触发条件,通过语义理解的方式完成空间几何推理,实现路径规划,如图7所示。围绕任务过程中目标点被遮挡的问题,文献[18]基于思维-理性自适应认知结构,提出了一种能够假设和预测环境不可见部分的方法,通过对先验知识进行启发性编码,再与声明性内存中的几何模式匹配,将检索匹配完成的部分投影回世界模型,作为可靠的目标点提供给路径规划系统。

尽管基于绝对或相对描述的路径规划已经有了长足进步,但针对缺乏先验知识的机器人自主路径规划还有许多待解决的难题。文献[21]构建了一种异构弹性位姿估计模块,为缺乏先验

图7 RCTA项目路径规划任务^[2]Fig. 7 Path Planning in RCTA^[2]

知识的机器人提供高置信度的定位服务功能。基于准确的定位和建图,高层任务规划器会在中层规划目标中寻找规划目标点,根据高分辨率代价地图进行优先级排序,获得底层目标的导向路径,最终通过测量生成规划路径。文献[94]对未知环境路径规划中的可通过性问题进行分析研究,将三维点云图转换为高程图,同时辅以相机的彩色与深度信息,通过A*算法与引导随机行进算法在高程图上进行路径规划,帮助机器人系统有效地避开规划路径中的障碍物,该方法对于多机器人团队具有一定的普适性。综上,路径规划是机器人的基础应用,也是各种高层次任务的应用前提,当前工作的发展方向主要集中在如何将语义知识和空间关系引入路径规划,强调基于非实时感知信息实现定位定姿,引导路径规划走向更高层次的应用研究。

3.2 目标检测

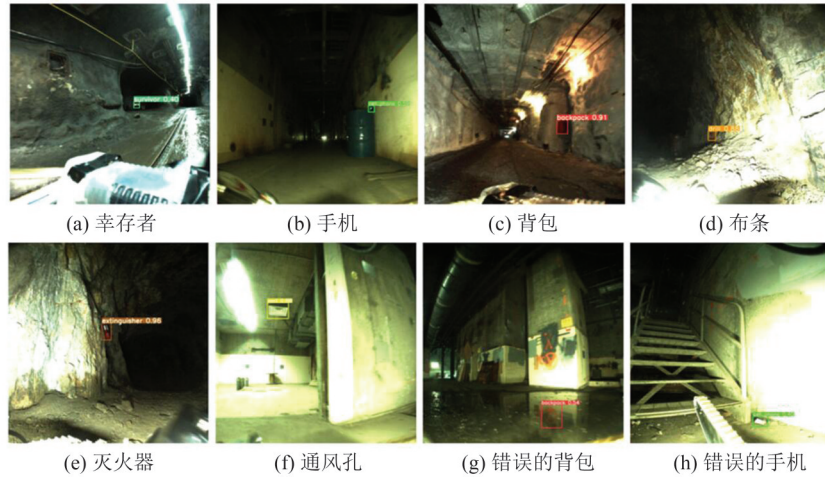
目标检测从图像中发现感兴趣的物体,并对物体类别进行判定,为机器地图与客观环境交互提供了实体对象,其成功率和精确率是机器地图实例应用的重要基础。随着深度学习网络发展的轻量化和并行计算能力的提升,基于深度学习的目标检测方法在机器地图应用中扮演着越来越重要的角色。有无生成目标候选框是分类目标检测算法的重要依据,有目标候选框生成的双阶段目标检测算法开始使用区域卷积神经网络(region-convolutional neural networks, R-CNN)^[95]作为基础网络,通过支持向量机分类器实现分类,后期的尺度金字塔池化网络^[96]和更快R-CNN网络^[97]分别对描述尺度和检测速度进行优化,计算效率大幅提升;无目标候选框的单阶段目标检测算法省略了候选框提取过程,使用批量归一化、多尺度训练、融合预测等方法,在保持计算效率的同时提高物体检测的准确率和定位精度。但是大多数目标检测算法是在有先验信息的情况下事先训练模型,利用环境验证,没有考虑到任务场景缺乏先验信息的情况,因此基于复杂场景的无监督学习方法可以采用生成式对抗

网络^[98]生成模型、判别模型,在博弈过程中产生最佳输出。

此外,目标检测也需要考虑地下、隧道、战场等复杂恶劣环境,其导致的不确定性会引发多目标检测问题,这种情况下神经网络对于数据良好的泛化能力将发挥重要作用。一些方法尝试从训练方式和网络角度优化目标检测算法,文献[99]提出的基于迁移学习的自适应多类识别方法能够实时微调检测数据,保证检测类别的精确度,同时借鉴混合模型训练策略,在残差网络的帮助下实现特征的跳跃连接,解码预测结果,并对概率框进行得分排序,减少目标检测算法的误判。在很多任务中,目标检测并不是最终目的,精确识别目标并对其定位对于灾害救援等任务更具有实际意义,如图8所示。文献[100]在检测到目标边界框后,将边界框投影到实际点云图中,通过时间一致性检测,应用卡尔曼滤波确定被检测对象的最终位置。文献[101]提出了一种由显著性注意力模型支撑的目标识别算法,结合了四周-中心凹视觉方法,能够在地图上标记目标对象的位置。文献[102]提出了基于对象的空间分层概率表示法,采用基于尺度不变特征变换的目标识别系统,使用立体摄像机捕捉目标并在度量地图上表示其坐标。目标检测与目标定位相辅相成,深度学习对于目标定位具有更好的鲁棒性,减少人为设定参数,因此具有更强的泛化性能。

3.3 知识表达与推理

知识表达是机器人用解决任务的形式表达世界模型的信息,机器人如何对任务中学习到的知识差异性和不确定性进行表达是一项很重要的研究。本体模型作为集成知识的框架,可以将低层次信息与高层次信息有效结合,形成知识描述和知识关联^[103],通过Jess引擎实现知识推理,大幅提高知识检索、存储、共享的效率^[104]。在构建本体模型的过程中,需要将本体模型分解成“类”“对象属性”“数据属性”“实例”等定义,建立关联关系,实现环境的语义化描述,服务于后续

图 8 SubT 挑战赛目标检测任务^[21]Fig. 8 Object Detection in SubT^[21]

的知识推理^[105]。

高效的信息查询与检索是知识表达和推理的前提,机器地图中的大部分信息都包含在度量地图、拓扑地图、语义地图等不同的地图模型中,机器人在任务过程中可以通过信息查询与检索实现知识挖掘和任务推理,突破传感器的感知范围限制。实际研究中发现,知识表达与推理往往与机器地图中的语义部分交织紧密。KnowRob^[106]是一个基于网络本体语言和逻辑编程语言 Prolog 的推理系统,可以部署在机器人的任务框架中,通过从传感器中获取信息形成语义进行推理,用以指导机器人的行动。文献[107]将公共数据库中检索到的物品使用方法组织在一个本体模型中,为机器人提供物品使用的语义信息,增强地图的功能性。文献[108]将机器人通过自然语言命令解析到的语义信息转化为二元特征,形成一种逻辑化的表达形式。文献[109]提出了一个基于时态流的知识推理框架,通过时态逻辑实现增量推理,并对其泛化性进行研究,可在多个机器人系统上实现推理应用。文献[110]结合图像局部和全局特征,从语义知识的角度推理所处环境的类型。文献[111]有效利用视觉信息,结合层次时间网络与注意力模型,可以从语义上推理机器人的空间位置。此外,还有方法从概率几何模型、语义分割、训练分类器等角度对从地图中学习到的知识进行表达和推理,对环境信息进行抽象化的提取与应用,更好地保障基于机器地图的任务规划^[112-114]。

4 结 语

本文围绕数据获取与处理、地图模型构建、

任务应用 3 个主要方向,阐述总结相关文献的研究现状、关键技术和存在问题。数据获取与处理是地图模型构建和任务应用的基础支撑,针对机器人平台、特征提取、语义分割、多传感器融合 4 个方面,基于现阶段研究工作和项目成果,结合深度学习等关键技术进行整体性归纳。地图模型构建整合尺度地图、拓扑地图和语义地图优势,围绕机器人平台作为用图主体的特殊性与局限性,借鉴参考高精地图相关数据模型,分析思考机器地图表达特点和整体架构,并对现有建图系统发展现状和存在问题进行归纳。任务应用从“按需”角度,阐述了路径规划、目标检测、知识表达与推理三类基础性应用的现状和关键技术,分析其中典型案例,为具体技术在任务实践中落地提供理论研究基础。

综上所述,机器地图是顺应技术发展与应用需求的前瞻性探索,但现如今仍属于发展的初始阶段。关于机器地图基础理论、运行机制、数据模型、表征形式至今还没有形成业界统一认识,本文描述的一些技术方法在机器地图相关项目中已得到应用,但在机器地图大框架下的实际效果还有待验证。此外,目前针对无人平台的数据获取和处理方法较为成熟,并形成了相应体系,地图模型构建和任务应用方向上欠缺较多,特别是针对多模态数据的组织形式、语义特征的关联使用、地图模型和任务应用的耦合、机器地图的自学习能力等方面还有很大的发展空间,已有的研究成果与机器地图需求之间存在不少差距。结合现有的研究基础,针对机器地图未来的发展方向和前景,还有以下几点认识:

1) 语义信息的优化。现阶段针对大规模点

云的语义分割算法获取的语义信息较为孤立,关联性和泛化性不强,一定程度上限制了同一场景语义信息在不同应用中的交互。未来工作应注重语义信息的数据关联、联合估计、空间推理,聚焦语义信息的空间范围和时序性,合理地充分利用语义信息为任务应用提供服务保障。

2) 机器地图评估方法。机器地图是一个复杂的系统,如何评估地图信息准确性、地图构建效果、功能模块实用性、任务模块关联度都是机器地图评估方法中应该着重考虑的问题。现有评估指标主要集中在绝对轨迹误差、相对位姿误差、均交并比、总体精度等,缺乏与任务结合的系统性评估体系。

3) 弱监督或无监督条件下的任务应用。一些任务场景机器人缺乏先验信息,或在通信受限条件下信息不能有效传输。基于此,如何摆脱模型训练对环境数据的依赖,在无过拟合的前提下,正确地选择嵌入函数,使用自适应机制调解网络参数,提升建图系统的实际性能和计算模型的泛化性是现阶段亟待解决的问题。

4) 数据组织与更新。机器地图对于获取到的环境信息,大多采用编码、属性描述等方式,以各种实例组织数据,表达知识。未来,复杂多维的环境信息和动态多变的组织规则需要机器地图加强对对象、位置和属性关系的识别推理,基于规范化组织的空间结构,保持实例空间属性随时间变化的有效更新,这种时序性的变化学习目前为止还没有一个完整的技术体系支撑,因此也是机器地图未来的主要研究方向。

可以预见,未来机器地图将在机器人面对复杂环境的空间描述与认知方学习方面发挥重要作用,为无人作战、服务保障、智慧城市等领域的关键技术研究提供积极支撑,具有广阔前景。

参 考 文 献

- [1] Rouček T, Pecka M, Čížek P, et al. Darpa Subterranean Challenge: Multi-robotic Exploration of Underground Environments[C]//The 6th International Conference on Modelling and Simulation for Autonomous Systems, Palermo, Italy, 2019.
- [2] Young S H, Patel D. Robotics Collaborative Technology Alliance Program Overview[C]//Unmanned Systems Technology XX, Orlando, USA, 2018.
- [3] Wong C C, Wang W W, Lee Y L. Soccer Robot Design for FIRA MiroSot League[C]//IEEE International Conference on Mechatronics, Taipei, China, 2005.
- [4] Obdrzalek D. Eurobot Junior and Starter: A Comparison of Two Approaches for Robotic Contest Organization [C]//The 1st Slovak-Austrian International Conference on Robotics in Education, Bratislava, Slovakia, 2010.
- [5] Iša J, Dlouhý M. Robotour-Robotika. cz Outdoor Delivery Challenge [C]//The 1st Slovak-Austrian International Conference on Robotics in Education, Bratislava, Slovakia, 2010.
- [6] Schneider S, Hegger F, Ahmad A, et al. The Rockin@Home Challenge [C]//The 41st International Symposium on Robotics, Munich, Germany, 2014.
- [7] Spurný V, Báča T, Saska M, et al. Cooperative Autonomous Search, Grasping, and Delivering in a Treasure Hunt Scenario by a Team of Unmanned Aerial Vehicles [J]. *Journal of Field Robotics*, 2019, 36(1): 125-148.
- [8] Wang Hong, Ai Tinghua, Zhu Guorui. Adaptive Strategy on the Visualization of Electronic Map [J]. *Geomatics and Information Science of Wuhan University*, 2004, 29(6): 525-528. (王洪, 艾廷华, 祝国瑞. 电子地图可视化中的自适应策略 [J]. 武汉大学学报(信息科学版), 2004, 29(6): 525-528.)
- [9] Song Ying, Li Qingquan. Fusion Expression of Real-Time Traffic Information and Navigation Electronic Map [J]. *Geomatics and Information Science of Wuhan University*, 2010, 35(9): 1108-1111. (宋莺, 李清泉. 实时交通信息与移动导航电子地图融合表达 [J]. 武汉大学学报(信息科学版), 2010, 35(9): 1108-1111.)
- [10] Liu Jingnan, Dong Yang, Zhan Jiao, et al. Thoughts and Suggestions on Autonomous Driving Map Policy [J]. *Strategic Study of CAE*, 2019, 21(3): 92-97. (刘经南, 董杨, 詹骄, 等. 自动驾驶地图有关政策的思考和建议 [J]. 中国工程科学, 2019, 21(3): 92-97.)
- [11] You Xiong. From VGE to Machine Maps: New Developments in Spatial Cognitive Tools [C]// Seminar on Maps and Spatial Awareness, Zhengzhou, China, 2018. (游雄. 从 VGE 到机器地图: 空间认知手段的新发展 [C]//地图与空间认知研讨会, 中国郑州, 2018.)
- [12] Nüchter A, Hertzberg J. Towards Semantic Maps for Mobile Robots [J]. *Robotics and Autonomous Systems*, 2008, 56(11): 915-926.
- [13] Lang D, Paulus D. Semantic Maps for Robotics [C]//The Workshop Workshop on AI Robotics at ICRA, Chicago, IL, USA, 2014.
- [14] Capobianco R, Gemignani G, Bloisi D D, et al. Au-

- tomatic Extraction of Structural Representations of Environments[C]//The 13th International Conference on Intelligent Autonomous Systems, Padua, Italy, 2016.
- [15] Xu J X, Luo Q, Xu K C, et al. An Automated Learning-Based Procedure for Large-Scale Vehicle Dynamics Modeling on Baidu Apollo Platform[C]//International Conference on Intelligent Robots and Systems, Macau, China, 2019.
- [16] Childers M A, Lennon C T, Bodt B, et al. US Army Research Laboratory (ARL) Robotics Collaborative Technology Alliance 2014 Capstone Experiment [R]. US Army Research Laboratory Aberdeen Proving Ground United States, Baltimore, USA, 2014.
- [17] Oh J, Suppé A, Duvallet F, et al. Toward Mobile Robots Reasoning Like Humans[C]//The Twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, Texas, USA, 2015.
- [18] Oh J, Howard T M, Walter M R, et al. Integrated Intelligence for Human-Robot Teams[C]//International Symposium on Experimental Robotics, Tokyo, Japan, 2016.
- [19] Williams J, Jiang S, O'Brien M, et al. Online 3D Frontier-Based UGV and UAV Exploration Using Direct Point Cloud Visibility[C]//IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, Karlsruhe, Germany, 2020.
- [20] Oleynikova H, Taylor Z, Fehr M, et al. Voxblox: Incremental 3D Euclidean Signed Distance Fields for On-Board MAVPlanning[C]//International Conference on Intelligent Robots and Systems, Vancouver, Canada, 2017.
- [21] Thakker R, Alatur N, Fan D D, et al. Autonomous Off-Road Navigation over Extreme Terrains with Perceptually-Challenging Conditions[C]//The 17th International Symposium on Experimental Robotics, Valletta, Malta, 2020.
- [22] Tranzatto M, Miki T, Dharmadhikari M, et al. CERBERUS in the DARPA Subterranean Challenge [J]. *Science Robotics*, 2022, 7(66): 9742.
- [23] Lennon C, Bodt B, Childers M, et al. RCTA Capstone Assessment [C]//Unmanned Systems Technology XVII, Baltimore, MD, USA, 2015.
- [24] He Yong, Lu Hao, Wang Chunxiang, et al. Generation of Precise Lane-Level Maps Based on Multi-sensors[J]. *Journal of Chang'an University(Natural Science Edition)*, 2015, 35(S1): 274-278. (贺勇, 路昊, 王春香, 等. 基于多传感器的车道级高精度地图制作方法[J]. 长安大学学报(自然科学版), 2015, 35(S1):274-278.)
- [25] Xu Y S, Ye Z, Yao W, et al. Classification of LiDAR Point Clouds Using Supervoxel-Based Detrended Feature and Perception-Weighted Graphical Model[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, 13: 72-88.
- [26] Hackel T, Wegner J D, Schindler K. Contour Detection in Unstructured 3D Point Clouds[C]//IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016.
- [27] Rusu R B, Blodow N, Beetz M. Fast Point Feature Histograms (FPFH) for 3D Registration [C]//IEEE International Conference on Robotics and Automation, Kobe, Japan, 2009.
- [28] Drost B, Ulrich M, Navab N, et al. Model Globally, Match Locally: Efficient and Robust 3D Object Recognition[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 2010.
- [29] Kasaei S H, Tomé A M, Seabra Lopes L, et al. GOOD: A Global Orthographic Object Descriptor for 3D Object Recognition and Manipulation [J]. *Pattern Recognition Letters*, 2016, 83: 312-320.
- [30] Rusu R B, Bradski G, Thibaux R, et al. Fast 3D Recognition and Pose Using the Viewpoint Feature Histogram [C]//International Conference on Intelligent Robots and Systems, Taipei, China, 2010.
- [31] Lehtomäki M, Jaakkola A, Hyyppä J, et al. Object Classification and Recognition from Mobile Laser Scanning Point Clouds in a Road Environment[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(2): 1226-1239.
- [32] Alhamzi K, Elmogy M, Barakat S. 3D Object Recognition Based on Local and Global Features Using Point Cloud Library [J]. *International Journal of Advancements in Computing Technology*, 2015, 7(3): 43-54.
- [33] Chen Y S, Jiang H L, Li C Y, et al. Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(10): 6232-6251.
- [34] Dara S, Tumma P. Feature Extraction by Using Deep Learning: A Survey[C]//The Second International Conference on Electronics, Communication and Aerospace Technology, Coimbatore, India, 2018.
- [35] Shi Yile, Yang Wenzhong, Du Huixiang, et al. Overview of Image Captions Based on Deep Learning

- [J]. *Acta Electronica Sinica*, 2021, 49(10): 2048-2060. (石义乐, 杨文忠, 杜慧祥, 等. 基于深度学习的图像描述综述[J]. 电子学报, 2021, 49(10): 2048-2060.)
- [36] Long Xiaoxiao, Cheng Xinjing, Zhu Hao, et al. Recent Progress in 3D Vision [J]. *Journal of Image and Graphics*, 2021, 26(6): 1389-1428. (龙霄潇, 程新景, 朱昊, 等. 三维视觉前沿进展[J]. 中国图象图形学报, 2021, 26(6): 1389-1428.)
- [37] Feng Y F, Zhang Z Z, Zhao X B, et al. GVCNN: Group-View Convolutional Neural Networks for 3DShape Recognition [C]//IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018.
- [38] Tchapmi L, Choy C, Armeni I, et al. SEGCloud: Semantic Segmentation of 3DPoint Clouds[C]//International Conference on 3D Vision, Qingdao, China, 2017.
- [39] Riegler G, Ulusoy A O, Geiger A. OctNet: Learning Deep 3D Representations at High Resolutions[C]//IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017.
- [40] Li Z, Wang J, Qu X, et al. 3D Point Cloud Segmentation for Complex Structure Based on PointSIFT [C]//The Third Chinese Conference on Pattern Recognition and Computer Vision, Nanjing, China, 2020.
- [41] Xie S N, Liu S N, Chen Z Y, et al. Attentional Shape ContextNet for Point Cloud Recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018.
- [42] Zhao H S, Jiang L, Fu C W, et al. PointWeb: Enhancing Local Neighborhood Features for Point Cloud Processing [C]//IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 2019.
- [43] Zhou Feiyan, Jin Linpeng, Dong Jun. Review of Convolutional Neural Network[J]. *Chinese Journal of Computers*, 2017, 40(6): 1229-1251. (周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.)
- [44] Zhang Z, Wang X, Jung C. DCSR: Dilated Convolutions for Single Image Super-resolution[J]. *IEEE Transactions on Image Processing*, 2018, 28(4): 1625-1635.
- [45] Dai J F, Qi H Z, Xiong Y W, et al. Deformable Convolutional Networks [C]//IEEE International Conference on Computer Vision, Venice, Italy, 2017.
- [46] Wang X L, Girshick R, Gupta A, et al. Non-local Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018.
- [47] Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks [C]//The 25th International Conference on Neural Information Processing Systems, Lake Tahoe, Nevada, 2012.
- [48] Szegedy C, Liu W, Jia Y Q, et al. Going Deeper with Convolutions [C]//IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015.
- [49] He K M, Zhang X Y, Ren S Q, et al. Deep Residual Learning for Image Recognition [C]//IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016.
- [50] Ronneberger O, Fischer P, Brox T. UNet: Convolutional Networks for Biomedical Image Segmentation [C]//The 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 2015.
- [51] Mehta S, Rastegari M, Caspi A, et al. EspNet: Efficient Spatial Pyramid of Dilated Convolutions for Semantic Segmentation [C]//European Conference on Computer Vision, Munich, Germany, 2018.
- [52] Grilli E, Menna F, Remondino F. A Review of Point Clouds Segmentation and Classification Algorithms [J]. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2017, 42: 339-344.
- [53] Wang L, Huang Y C, Hou Y L, et al. Graph Attention Convolution for Point Cloud Semantic Segmentation [C]//IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 2019.
- [54] Xie Y X, Tian J J, Zhu X X. Linking Points with Labels in 3D: A Review of Point Cloud Semantic Segmentation [J]. *IEEE Geoscience and Remote Sensing Magazine*, 2020, 8(4): 38-59.
- [55] Charles R Q, Hao S, Mo K C, et al. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation [C]//IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017.
- [56] Qi C R, Yi L, Su H, et al. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space [C]//The 31st International Conference on Neural Information Processing Systems, Long Beach, California, USA, 2017.
- [57] Zhang Y, Zhou Z X, David P, et al. PolarNet: An Improved Grid Representation for Online LiDAR Point Clouds Semantic Segmentation [C]//IEEE

- Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2020.
- [58] Zhu X G, Zhou H, Wang T, et al. Cylindrical and Asymmetrical 3D Convolution Networks for LiDAR Segmentation [C]//IEEE Conference on Computer Vision and Pattern Recognition, Nashville, USA, 2021.
- [59] Hu Q Y, Yang B, Xie L H, et al. RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds [C]//IEEE Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2020.
- [60] Han W, Wu H, Wen C, et al. BLNet: Bidirectional Learning Network for Point Clouds [J]. *Computational Visual Media*, 2022, 8(4): 585-596.
- [61] Wu Jun, Cui Yue, Zhao Xuemei, et al. SSA-point Net++: A Space Self-attention CNN for the Semantic Segmentation of 3D Point Cloud [J]. *Journal of Computer-Aided Design and Computer Graphics*, 2022, 34(3): 437-448. (吴军, 崔玥, 赵雪梅, 等. SSA-PointNet++: 空间自注意力机制下的 3D 点云语义分割网络 [J]. 计算机辅助设计与图形学学报, 2022, 34(3): 437-448.)
- [62] Zhang Bo, Wei Zhenzhong, Zhang Guangjun. Rapid Coordinate Transformation Between a Robot and a Laser Tracker [J]. *Chinese Journal of Scientific Instrument*, 2010, 31(9): 1986-1990. (张博, 魏振忠, 张广军. 机器人坐标系与激光跟踪仪坐标系的快速转换方法 [J]. 仪器仪表学报, 2010, 31(9): 1986-1990.)
- [63] Alam N, Dempster A G. Cooperative Positioning for Vehicular Networks: Facts and Future [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2013, 14(4): 1708-1717.
- [64] Forster C, Carlone L, Dellaert F, et al. On-Manifold Preintegration for Real-Time Visual Inertial Odometry [J]. *IEEE Transactions on Robotics*, 2017, 33(1): 1-21.
- [65] Jahanchahi C, Mandic D P. A Class of Quaternion Kalman Filters [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, 25(3): 533-544.
- [66] Bailey T, Nieto J, Guivant J, et al. Consistency of the EKF-SLAM Algorithm [C]//IEEE International Conference on Intelligent Robots and Systems, Beijing, China, 2006.
- [67] Mourikis A I, Roumeliotis S I. A Multi-state Constraint Kalman Filter for Vision-Aided Inertial Navigation [C]//IEEE International Conference on Robotics and Automation, Rome, Italy, 2007.
- [68] Zhang J, Singh S. Visual-LiDAR Odometry and Mapping: Low-Drift, Robust, and Fast [C]//IEEE International Conference on Robotics and Automation, Seattle, USA, 2015.
- [69] Leutenegger S, Furgale P, Rabaud V, et al. Keyframe-Based Visual-Inertial Slam Using Nonlinear Optimization [C]//Robotics Science and Systems, Berlin, Germany, 2013.
- [70] De Tone D, Malisiewicz T, Rabinovich A. SuperPoint: Self-supervised Interest Point Detection and Description [C]//IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, USA, 2018.
- [71] Maye J, Furgale P, Siegwart R. Self-supervised Calibration for Robotic Systems [C]//IEEE Intelligent Vehicles Symposium, Gold Coast, Australia, 2013.
- [72] Kahn G, Abbeel P, Levine S. BADGR: An Autonomous Self-supervised Learning-Based Navigation System [J]. *IEEE Robotics and Automation Letters*, 2021, 6(2): 1312-1319.
- [73] Wang X L, Zhang R F, Shen C H, et al. Dense Contrastive Learning for Self-supervised Visual Pre-training [C]//IEEE Conference on Computer Vision and Pattern Recognition, Nashville, USA, 2021.
- [74] Bender P, Ziegler J, Stiller C. Lanelets: Efficient Map Representation for Autonomous Driving [C]//IEEE Intelligent Vehicles Symposium Proceedings, Dearborn, USA, 2014.
- [75] Shimada H, Yamaguchi A, Takada H, et al. Implementation and Evaluation of Local Dynamic Map in Safety Driving Systems [J]. *Journal of Transportation Technologies*, 2015, 5(2): 102-112.
- [76] Poggenhans F, Pauls J H, Janosovits J, et al. Lanelet2: A High Definition Map Framework for the Future of Automated Driving [C]//The 21st International Conference on Intelligent Transportation Systems, Maui, USA, 2018.
- [77] Jiang K, Yang D G, Liu C R, et al. A Flexible Multi-layer Map Model Designed for Lane-Level Route Planning in Autonomous Vehicles [J]. *Engineering*, 2019, 5(2): 305-318.
- [78] Liu Jingnan, Zhan Jiao, Guo Chi, et al. Data Logic Structure and Key Technologies on Intelligent High Precision Map [J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(8): 939-953. (刘经南, 詹骄, 郭迟, 等. 智能高精地图数据逻辑结构与关键技术 [J]. 测绘学报, 2019, 48(8): 939-953.)
- [79] Cadena C, Carlone L, Carrillo H, et al. Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age [J].

- IEEE Transactions on Robotics*, 2016, 32(6): 1309–1332.
- [80] Lai K, Bo L F, Fox D. Unsupervised Feature Learning for 3D Scene Labeling[C]//IEEE International Conference on Robotics and Automation, Hong Kong, China, 2014.
- [81] Kundu A, Li Y, Dellaert F, et al. Joint Semantic Segmentation and 3D Reconstruction from Monocular Video[C]//The 13th European Conference, Zurich, Switzerland, 2014.
- [82] Hermans A, Floros G, Leibe B. Dense 3D Semantic Mapping of Indoor Scenes from RGB-D Images [C]//IEEE International Conference on Robotics and Automation, Hong Kong, China, 2014.
- [83] Mao M Y, China B U B, Zhang H W, et al. SEMANTIC-RTAB-MAP (SRM) : A Semantic SLAM System with CNNs on Depth Images [J]. *Mathematical Foundations of Computing*, 2019, 2(1): 29–41.
- [84] Labbé M, Michaud F. Appearance-Based Loop Closure Detection for Online Large-Scale and Long-Term Operation [J]. *IEEE Transactions on Robotics*, 2013, 29(3): 734–745.
- [85] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger [C]//IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017.
- [86] Oberlander J, Uhl K, Zollner J M, et al. A Region-Based SLAM Algorithm Capturing Metric, Topological, and Semantic Properties [C]//IEEE International Conference on Robotics and Automation, Pasadena, USA, 2008.
- [87] Xie L H, Markham A, Trigoni N. SnapNav: Learning Mapless Visual Navigation with Sparse Directional Guidance and Visual Reference [C]//IEEE International Conference on Robotics and Automation, Paris, France, 2020.
- [88] Bouman A, Ginting M F, Alatur N, et al. Autonomous Spot: Long-Range Autonomous Exploration of Extreme Environments with Legged Locomotion [C]//IEEE International Conference on Intelligent Robots and Systems, Las Vegas, USA, 2020.
- [89] Li Q, Wang Y, Wang Y L, et al. HDMaPNet: An Online HDMaP Construction and Evaluation Framework [C]//International Conference on Robotics and Automation, Philadelphia, USA, 2022.
- [90] Liu Y, Yuan T, Wang Y, et al. Vectormapnet: End-to-End Vectorized HD Map Learning [C]//International Conference on Machine Learning. PMLR, Honolulu, Hawaii, USA, 2023.
- [91] Hart P E, Nilsson N J, Raphael B. A Formal Basis for the Heuristic Determination of Minimum Cost Paths [J]. *IEEE Transactions on Systems Science and Cybernetics*, 1968, 4(2): 100–107.
- [92] Dorigo M, Birattari M, Stutzle T. Ant Colony Optimization [J]. *IEEE Computational Intelligence Magazine*, 2006, 1(4): 28–39.
- [93] Latombe J C. Robot Motion Planning [M]. New York: Springer Science and Business Media, 2012.
- [94] Chung T H, Orekhov V, Maio A. Into the Robotic Depths: Analysis and Insights from the DARPA Subterranean Challenge [J]. *Annual Review of Control, Robotics, and Autonomous Systems*, 2023, 6: 477–502.
- [95] Girshick R, Donahue J, Darrell T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [C]//IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, 2014.
- [96] He K M, Zhang X Y, Ren S Q, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904–1916.
- [97] Girshick R, Iandola F, Darrell T, et al. Deformable Part Models Are Convolutional Neural Networks [C]//IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015.
- [98] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Nets [C]//The 27th International Conference on Neural Information Processing Systems, Montreal, Canada, 2014.
- [99] Hoffman J, Guadarrama S, Tzeng E, et al. LSDA: Large Scale Detection Through Adaptation [C]//The 27th International Conference on Neural Information Processing Systems: Volume 2, Montreal, Canada, 2014.
- [100] Tan L, Lv X, Lian X, et al. YOLOv4_Drone: UAV Image Target Detection Based on an Improved YOLOv4 Algorithm [J]. *Computers and Electrical Engineering*, 2021, 93: 107261.
- [101] Meger D, Forssén P E, Lai K, et al. Curious George: An Attentive Semantic Robot [J]. *Robotics and Autonomous Systems*, 2008, 56(6): 503–511.
- [102] Case C, Suresh B, Coates A, et al. Autonomous Sign Reading for Semantic Mapping [C]//IEEE International Conference on Robotics and Automation, Shanghai, China, 2011.
- [103] Lim G H, Suh I H, Suh H. Ontology-Based Unified Robot Knowledge for Service Robots in Indoor

- Environments [J]. *IEEE Transactions on Systems, Man, and Cybernetics: Part A: Systems and Humans*, 2011, 41(3): 492-509.
- [104]Manzoor S, Rocha Y G, Joo S H, et al. Ontology-Based Knowledge Representation in Robotic Systems: A Survey Oriented Toward Applications [J]. *Applied Sciences*, 2021, 11(10): 4324.
- [105]Zhu Xiaomin. Research on Key Techniques in Positioning and Path Planning of Indoor Mobile Robots Based on UWB and Semantic Map [D]. Shanghai: East China University of Science and Technology, 2020. (朱晓民. 基于UWB及语义地图的室内移动机器人定位与路径规划关键技术研究[D]. 上海: 华东理工大学, 2020.)
- [106]Tenorth M, Beetz M. KnowRob: A Knowledge Processing Infrastructure for Cognition-Enabled Robots [J]. *International Journal of Robotics Research*, 2013, 32(5): 566-590.
- [107]Tenorth M, Kunze L, Jain D, et al. KNOWROB-MAP-Knowledge-Linked Semantic Object Maps [C]//The 10th IEEE-RAS International Conference on Humanoid Robots, Nashville, USA, 2010.
- [108]Matuszek C, Herbst E, Zettlemoyer L, et al. Learning to Parse Natural Language Commands to a Robot Control System [C]//Experimental Robotics: The 13th International Symposium on Experimental Robotics, Québec City, Canada, 2013.
- [109]Heintz F. Semantically Grounded Stream Reasoning Integrated with ROS [C]//IEEE International Conference on Intelligent Robots and Systems, Tokyo, Japan, 2013.
- [110]Viswanathan P, Southey T, Little J, et al. Place Classification Using Visual Object Categorization and Global Information [C]//Canadian Conference on Computer and Robot Vision, St. John's, Canada, 2011.
- [111]Kostavelis I, Amanatiadis A, Gasteratos A. How Do You Help a Robot to Find a Place? A Supervised Learning Paradigm to Semantically Infer About Places [C]//The 8th International Conference on Hybrid Artificial Intelligent Systems, Salamanca, Spain, 2013.
- [112]Levesque H J. Knowledge Representation and Reasoning [J]. *Annual Review of Computer Science*, 1986, 1: 255-287.
- [113]Sridharan M, Gelfond M, Zhang S Q, et al. REBA: A Refinement-Based Architecture for Knowledge Representation and Reasoning in Robotics [J]. *Journal of Artificial Intelligence Research*, 2019, 65: 87-180.
- [114]Sadik A, Urban B. An Ontology-Based Approach to Enable Knowledge Representation and Reasoning in Worker-Cobot Agile Manufacturing [J]. *Future Internet*, 2017, 9(4): 90.