

移动机器人长期自主环境适应研究进展和展望

曹风魁¹ 庄严¹ 闫飞¹ 杨奇峰^{2,3,4} 王伟¹

摘要 真实世界中存在光照、天气、季节及场景结构等复杂环境因素, 这些因素的改变对移动机器人基本行为和任务能力带来巨大挑战. 随着机器人与人工智能技术的不断发展, 如何使移动机器人在长期运行中与复杂多变的环境条件相适应是智能机器人领域的研究热点. 本文重点从地图构建与动态维护、重定位及场景理解等移动机器人基本行为能力的系统综述入手, 对移动机器人长期自主环境适应的前沿技术与研究方向进行了着重论述与分析. 最后对该领域的研究重点和技术发展趋势进行了探讨.

关键词 环境适应, 长期自主, 地图构建, 重定位, 场景理解, 移动机器人

引用格式 曹风魁, 庄严, 闫飞, 杨奇峰, 王伟. 移动机器人长期自主环境适应研究进展和展望. 自动化学报, 2020, 46(2): 205–221

DOI 10.16383/j.aas.c180493



开放科学(资源服务)标识码(OSID):

Long-term Autonomous Environment Adaptation of Mobile Robots: State-of-the-art Methods and Prospects

CAO Feng-Kui¹ ZHUANG Yan¹ YAN Fei¹ YANG Qi-Feng^{2,3,4} WANG Wei¹

Abstract In real-world applications, mobile robots will work in complex open environments, where there are various changing factors such as time-varying illumination, weather, seasons and scene layout, which are all challenging tasks for a mobile robot with a long-term autonomous environment adaptation ability. With the rapid development and improvement of the technology in mobile robotics and artificial intelligence, how to make mobile robots adapt to changing environments is a hot research topic. This paper starts with a systematic review of the basic capabilities of mobile robots such as mapping and map updating, relocalization and scene understanding in dynamic environments, and then focuses on the cutting-edge technologies of long-term autonomous environment adaptation of mobile robots. The research emphases and prospective technical development trends are also presented at the end of this paper.

Key words Environment adaptation, long-term autonomy, mapping, relocalization, scene understanding, mobile robots

Citation Cao Feng-Kui, Zhuang Yan, Yan Fei, Yang Qi-Feng, Wang Wei. Long-term autonomous environment adaptation of mobile robots: state-of-the-art methods and prospects. *Acta Automatica Sinica*, 2020, 46(2): 205–221

近年来, 随着机器人产业的快速发展, 多种类的移动机器人正逐步从实验室走进人类的生活^[1]. 这意味着移动机器人的工作环境发生了由单一的实验测试场景到复杂的真实人类世界的巨大转变, 同时其工作时间也由有限的周期性测试转变为连续的长

期运行. 相比相对稳定的测试场景, 真实世界会存在动态干扰、光照条件变化、天气季节转变、环境改造等诸多不确定性因素. 真实世界中这些渐进或剧烈的环境变化给机器人环境感知与行为决策带来巨大的挑战. 如何使移动机器人适应真实世界中复杂多样的变化, 保证其长期自主环境适应能力是长久以来国内外学者密切关注的挑战性研究问题^[2–3].

在移动机器人领域, 对一个机器人系统进行长期自主环境适应性的测试与验证, 其必要条件是测试周期必须要完整覆盖其工作环境的周期性变化^[4–7]. 对于工作在室内环境的移动机器人系统, 其工作周期通常以天为单位, 因此通常通过连续多周期的测试来验证系统的稳定性. 而对于工作在室外开放环境下的移动机器人系统, 因为光照、季节以及人为干扰等环境条件的改变, 环境的表象、场景结构乃至景物的分布均会发生不同程度的变化, 因此

收稿日期 2018-07-17 录用日期 2019-03-08
Manuscript received July 17, 2018; accepted March 8, 2019
国家自然科学基金(61503056, U1608253)资助
Supported by National Natural Science Foundation of China (61503056, U1608253)
本文责任编辑 吴立刚
Recommended by Associate Editor WU Li-Gang
1. 大连理工大学控制科学与工程学院 大连 116024 2. 中国科学院沈阳自动化研究所 沈阳 110016 3. 中国科学院大学 北京 100049
4. 沈阳新松机器人自动化股份有限公司 沈阳 110168
1. School of Control Science and Engineering, Dalian University of Technology, Dalian 116024 2. Shenyang Institute of Automation, China Academy of Sciences, Shenyang 110016 3. University of Chinese Academy of Sciences, Beijing 100049 4. Shenyang SIASUN Robot & Automation CO., LTD., Shenyang 110168

与之对应的测试周期要以年为单位,也就是要历经不同天气、季节等诸多环境条件的检验。

真实世界中复杂的环境条件变化对地图构建与维护、重定位及场景理解等移动机器人基础的自主行为能力带来巨大的挑战。例如,环境的表象及结构变化导致移动机器人对环境当前的观测与先前构建的地图发生矛盾,致使先验地图失效;移动机器人在不同时间重复访问同一场景时,光照、季节等环境条件的变化或城市设施改建等人为因素会导致其两次对场景的观测不一致,进而导致其重定位的失败;真实世界环境中存在大量的行人及车辆,且其行为存在偶然性与不确定性,这给移动机器人尤其是高速移动机器人的自主导航造成巨大的挑战;移动机器人的场景理解主要是对其所处环境进行目标检测、语义分割及语义关联等任务,光照、季节等环境条件及对环境观测视角等的改变都会降低场景理解相关任务的准确性。

要保证移动机器人在长期运行中具有良好的环境适应能力,则主要需要其在地图构建与动态维护^[8]、重定位^[9]和自主场景理解^[10]这三方面行为任务均能够长期有效应对环境条件变化。因此本文主要从这三方面分析论述移动机器人长期自主环境适应的研究进展与前沿动态。

1 移动机器人地图构建与动态维护

移动机器人在人工设计的实验室测试场景中的地图构建方法^[11-16]已被广泛研究,也已趋于成熟。但这些传统方法大都默认环境的外观和结构是基本不变的,因此移动机器人基于先验地图来完成定位及导航等相关任务。此外,早期针对动态环境的地图构建方法^[17-20]将移动物体看作环境中离群数据,进而从观测数据中过滤或跟踪移动的物体,但它们还是基于机器人所处环境的外观与结构是静态不变的假设,这种假设使得这类方法在环境表象或结构会发生巨大变化的真实世界中不再适用。

由于目前尚无统一的地图构建框架可以在多种类场景中均具有通用性,基于近年来相关文献查阅结果^[21-22],我们将移动机器人的工作场景大致分为功能单一的封闭环境、高度动态的室内环境和大范围室外开放环境。

1.1 功能单一的封闭环境

诸如仓库、停车场等小范围功能性场所是现阶段移动机器人有较多成功应用案例的环境。这类场景最大的特点就是环境模型发生着频繁的变化,如货物装卸及车辆流动所导致的环境改变。图1摘自文献^[23],其显示了不同时刻车辆流动给停车环境带来的明显变化。研究学者将该环境中的物体定义

为三类:静态物体、半静态物体和动态物体,其中半静态物体(Semi-static)是指货物、车等可能被移动的静态物体。功能单一的封闭环境中半静态物体位置的改变可能导致机器人先前构建的地图过时失效,进而致使机器人定位的失败甚至造成严重的后果。如何应对半静态物体位置变化对构建环境模型的冲击及动态障碍的干扰,是实现移动机器人在该类环境下长期稳定自主运行的关键。



(a) 12:00 (b) 18:00

图1 不同时间下的停车场^[23]

Fig.1 The same parking lot at different time^[23]

Meyer-Delius 等^[24]利用临时的局部地图(Temporary local maps)表述来跟踪因半静态物体位置改变而造成的异常观测。虽然该方法仍是基于环境静态不变性假设的,但其当环境改变时便建立临时的局部地图对环境变化进行表述并对原始先验地图进行临时扩展,然后依赖扩展后的地图利用粒子滤波算法估计机器人位姿。实验显示通过检测由半静态物体引起的异常观测,可有效提高机器人定位的鲁棒性。

针对此类功能相对单一的结构化封闭场景,基于概率估计模型的构图和定位方法体现出良好的优越性。Saarinen 等^[25]提出一种基于栅格模型的动态环境构图方法。该方法将每个栅格定义为拥有两个状态的独立马尔科夫链,并利用基于近因权重(Recency weighting)的策略在线学习状态转换模型参数。这种基于概率的栅格地图不仅可以表述栅格是否被占用,同时还可以对栅格动态进行建模。同样基于概率估计模型,Tipaldi 等^[23]利用二维栅格的隐式马尔科夫模型来表述被占有栅格的动态属性。通过利用观测数据对表述模型参数的训练学习,该方法可有效辨别动态、半静态和静态等不同物体的动态属性。作者利用该方法有效对停车场环境进行地图构建和动态维护。

Biswas 等^[26]提出基于插曲(Episodic)的非马尔科夫过程定位方法,该方法在整个机器人的运动轨迹中估计观测数据的置信度以及观测数据置信度与地图中未存储的静态物体、动态物体和存储的静态物体之间的关联。机器人在长期运行过程中从观测数据中提取长期特征来匹配静态地图,提取短期特征用于匹配历史观测中的短期特征。与文献^[15]和^[16]两种方法相比,该方法在停车场等数据集上

取得了更好的表现效果。

Saarinen 等^[7]提出了一种新颖的在线地图构建方法。作者结合了 NDT (Normal distributions transform) 和占用栅格 (Occupancy grid) 两种常用的地图表述模型的优势: NDT 地图模型表述简洁高效, 能实现在线的地图构建, 但是不具备占用栅格地图可被递归更新的特性; 占用栅格可被递归更新, 并根据重复的观测释放或占用栅格空间, 这种特性可有效表述环境的变化。该方法在某牛奶生产车间持续 17 个小时的测试中显示出其良好的表现。

1.2 高度动态的室内环境

近年来, 家庭服务、银行导引、餐厅送餐等多种服务机器人产品陆续推向市场。与仓库、停车场等功能单一的场所相比, 家庭、办公室及餐饮场所等生活场景的组成元素更加多样化, 同时还存在诸如门窗开闭、家具移动及行人无规则运动等复杂因素。为了应对室内环境的复杂变化, 许多地图构建与动态维护方法被提出以提升机器人的自主环境适应能力。

Biber 等^[27]提出一种基于采样的环境表述方法对环境变化进行建模。基本的思想是用不同时间段对同一场景的环境建模组合描述当前场景, 随着机器人的运行会不断加入最新的观测模型, 剔除时间最久远的环境表述。这种思想被称为“记忆消退” (Memory decay)。当机器人自主定位时, 将当前传感器数据与地图中不同时间段的环境模型比对, 选出与当前传感器数据最匹配的环境表述。这使得定位更加鲁棒, 地图不断更新永不过时。为了验证方法的有效性, 在包含若干实验室、办公室、走廊和大厅的整个楼层中, 作者每天进行三次 (早上、午后、傍晚) 测试, 并坚持了长达 5 周共计 75 次测试的周期。

受人类记忆的多重存储模型启发, Dayoub 等^[28]利用移动机器人对环境的“短时记忆” (Short-term memory) 和“长期记忆” (Long-term memory) 来构建地图。作者利用注意力机制将感知信息中感兴趣的特征转移到短时记忆地图中, 同时利用训练机制去选择稳定的特征更新到长期记忆地图表述中。为了更好地适应变化环境及节省存储空间, 他们提出罢免机制从拓扑节点中移除过时无效的信息。类似地, Morris 等^[29]同时构建短时记忆和长期记忆两个三维八叉树地图 (OctoMap) 来表述移动机器人周围环境。短时记忆地图存储移动机器人当前状态下的观测信息, 长期记忆地图是对短时记忆地图的累积更新, 但只有短时记忆地图中永久的环境信息才被更新转移到长期记忆地图中。实验证实, 利用多种地图表述能更好地实现移动机器人在复杂多变的环境中自主定位及导航。

Krajník 等^[30]认为人类的许多日常活动都是周期性明显, 有规律可循的。因此作者将未知的环境变化过程定义为周期性函数, 进而以栅格地图为基础, 利用频谱对环境的时空变化进行建模表述。将环境变化过程定义为周期性函数, 可利用傅里叶变换有效识别。同时将时域转换为频域, 可高效地辨认分析并保存规律性的环境变化过程。对环境变化过程的周期性假设, 能减轻机器人地图构建及维护的复杂度, 同时可使长航时机器人具有对环境变化的预测能力, 有助于更好地长时间自主运行。在实验过程中, 作者利用移动机器人在室内环境中以巡逻的模式采集一周的数据验证方法的准确性和效率。结果显示, 该方法在任意时间长度数据中测试, 只需要恒定且少量的存储空间, 这一特性正是长航时机器人长期运行所需要的。

1.3 大范围室外开放环境

随着室外移动机器人技术的不断发展, 如何对室外准结构化或完全非结构化开放环境进行建模及动态维护成为当前研究的热点。室外开放环境中移动机器人地图构建所面临的主要挑战是如何有效地检测行人、车辆等高危动态目标, 以及如何有效地应对光照、季节等环境条件的转变。激光雷达传感器因为利用主动光源感知环境, 不受光照条件影响, 可很好地解决光照条件变化造成的环境外貌变化, 且测距精度高, 有利于构建长期有效的高精度地图模型。谷歌、百度等无人车均采用三维激光雷达感知环境, 取得了非常显著的研究成果。值得一提的是, Zhang 等^[31]提出的基于三维激光的实时定位与构图方法, 取得了非常有竞争力的构图效果。

在室外开放环境中, 在地图中精确定位、预测和跟踪人车等高动态障碍是室外移动机器人安全的基础和必要保障。传统的动态障碍物检测方法大都是基于人为设计特征的分类方法。但机器人长期的运行过程中, 因为观测视角的变化, 环境中互相遮挡等问题, 使得利用基于人为设计特征分类的方法稳定性降低。Wang 等^[32]提出一种端到端的三维激光点云分割方法: 输入是原始三维点云, 输出是可直接被分类的点云簇。作者利用欧氏最小生成树 (Euclidean minimum spanning tree) 实现基于图的点云聚类算法对点云进行分割, 并设计了基于 RANSAC 的边界选择规则。他们指出并证实, 针对车辆、行人等文中提及的特定物体类别, 解决二分类问题要比直接解决多分类问题高效的多。这主要是因为鲁棒的多目标物体分类需要足够的形状信息, 而通过预分割获得的目标点云簇中, 部分点云簇缺乏足够的满足多分类要求的形状信息。正因为如此, 在观测视角受限或其他限制条件下, 局部被观测的

动态物体也能被很好地检测与定位。

Tanzmeister 等^[33]提出一种新颖的基于栅格的环境描述模型,可有效检测静态物体,同时评估其速度及不确定性。该方法利用粒子滤波算法直接对观测数据的速度概率分布进行评估,并不依赖现有的目标跟踪算法来过滤与创建和更新地图无关的数据点,也不依赖动态物体分类算法对观测数据预分类。由于没有特定的特征提取和形状假设,任意的移动物体都能被有效定位及跟踪。

机器人在长期自主运行过程中,应充分利用先前积累的经验教训,不断提升其开放环境中的构图能力。Ott 等^[34]提出一种实时的无监督增量式的学习方法,利用 AP (Affinity propagation) 聚类算法对三维激光点云构建基于物体类别层面的环境模型,并通过聚类物体与机器人的交互来判断该物体是否为障碍物,并预测发生冲突的可能性。这种机器人与环境增量式交互学习的构图方法,满足长航时机器人长期运行的需求,同时通过长期交互学习,机器人会更加稳定地适应复杂多变的开放环境。

为了构建长期有效的地图,Pomerleau 等^[35]在苏黎世联邦理工学院校园中坚持了长达七个月的测试。他们认为终生地图应该根据环境变化维护更新,同时应包含语义层信息指导机器人的运动规划。他们的长期定位与构图系统包含位姿估计、全局地图维护和速度估计三个模块,通过对环境的重复观测推断每个三维点的动静态属性,并评估动态点的速度。该方法的优点是不依赖离线训练的目标检测模型及点云聚类过程,直接对每个三维点进行估计,节省计算成本。

移动机器人长期的自主运行不仅要考虑环境的条件变化,还需要考虑计算效率与存储空间。当机器人不断重复地访问场景时,传统的方法会受到计算效率和存储空间的制约而降低算法的性能。为了利用有限的存储空间对地图进行最有效的表述,Dymczyk 等^[36]将场景识别与地图维护独立,实现在线场景识别和离线地图维护的交替作业。作者提出一种简化的地图表述模型 (Summary map),该地图只保存对场景识别有用的路标特征,并裁剪机器人重叠的位姿轨迹。他们认为经常被观测到的路标特征是需要被保存的,因为他们更容易被重复检测与匹配,且随着时间的推移更加稳定。

研究表明,对环境进行目标检测^[37]和场景理解^[38-39],构建带有更高层次语义信息的地图模型^[40-41],有助于在物体类别层面对机器人定位与导航进行指导。语义地图的构建使机器人感知利用环境中更为稳定的语义信息,同时有助于实现移动机器人的类人决策,更好地融入人类生活。Drouilly 等^[42]利用场景理解算法对全景相机图像进行物体

分类,构建局部语义子地图;进而通过位姿估计算法将各子地图连接构建全局拓扑语义地图。他们提出一种基于语义内容的定位算法,在用语义信息描述路径的同时,也可利用语义内容检索完成定位与导航。此外,构建语义地图可实现移动机器人与人更方便直接地人机交互,进而更高效地完成指定任务。

1.4 小结

真实世界中复杂的环境条件变化给移动机器人地图构建及动态维护带来巨大挑战。为了应对这一挑战,国内外专家学者提出的移动机器人地图构建及动态维护方法主要分为三类。第一类是基于概率的地图构建与维护方法^[7, 23, 25-26],此类方法通常依赖栅格地图模型,利用移动机器人对环境重复的观测不断更新维护地图模型;但此类方法通常要求移动机器人具有较准确的位姿跟踪或全局定位能力。在不依赖全局定位系统的前提下,移动机器人基于轮式/视觉/激光里程计^[43]进行位姿跟踪时,其累积误差会随着机器人的长期运行而不断增长。因此,此类基于概率的方法通常只适用于较小范围的应用场景。第二类是基于采样或“记忆”的方法^[27-29],此类方法通常同时利用对环境的多次观测来构建环境地图。同时,将多次观测数据中相对稳定的环境信息更新到地图中。但此类方法需要尽可能多地访问同一场景以覆盖环境的条件变化,且大范围场景下易产生信息爆炸。第三类是基于学习的方法^[32, 34-35],此类方法利用移动机器人与环境长期的交互与学习,从而提升移动机器人自主环境适应能力。随着深度学习和强化学习的不断发展,基于学习的方法具有广阔的研究前景。

表 1 对文献 [7-34] 中提出的方法进行了归纳和总结:针对不同类型的应用场景和其相应的环境条件变化,国内外学者提出了相应的解决方案。尽管不同的方法存在其不足之处,但均在相应的应用场景中取得不错的实验效果。

2 移动机器人长期运行中的重定位

移动机器人在长期运行中不可避免地要重复访问同一场景,因此准确地识别历史访问场景并估计机器人在该场景中的位姿(即重定位),是移动机器人必不可少的基本能力^[44]。在不依赖全局定位系统的前提下,场景识别^[45-46](Place recognition)技术可检索出与机器人当前所处场景相同的历史访问场景,然后利用场景配准技术可直接计算出机器人在该场景中的精确位姿,这是实现机器人重定位的前提。但移动机器人长期运行中的鲁棒重定位对场景识别的精度提出了很高要求。

现阶段,在条件稳定环境中的重定位研究已

表 1 动态环境下长航时机器人地图构建与长期维护方法

Table 1 Methods of mapping and long-term maintenance for mobile robots in dynamic environment

测试场景	关键词	主要思想	局限性	文献
停车场	临时地图	用临时地图跟踪由环境中半静态物体引起的矛盾观测, 并临时地扩展环境的静态地图, 进而利用粒子滤波算法实现机器人的定位	该算法默认动态物体被成功地检测并滤除	[24]
	动态栅格地图	将每个栅格定义为隐式马尔科夫模型, 利用初始状态分布和观测模型预测状态转换概率	模型假设与真实环境中的动态变化的一致性问题; 栅格与栅格之间的独立假设不合理	[23]
	非马尔科夫过程; 插曲片段	释放了马尔科夫过程的独立性假设, 分析观测数据与地图中未标记的物体之间的关联; 利用关联分析将对环境的观测分为长期、短期和动态特征	依赖提前构建的地图先验信息	[26]
生产车间	独立马尔科夫链; 栅格地图	将每个栅格定义为有两个状态的独立马尔科夫链; 栅格状态转换被定义为两个泊松过程, 且转换模型参数通过在线学习获得	算法模型假设环境中的动态变化属于齐次过程, 而事实上动态变化是有时间依赖性的	[25]
	正态分布变换栅格地图	结合正态分布变换地图的简洁和栅格地图的稳定来表述动态环境, 并定义了精确的栅格状态递归更新模型; 该方法可用于多分辨率地图	该方法依赖一个外部准确的定位系统	[7]
	记忆消退	利用多重时间尺度下的环境模型表述环境; 随时间推移, 新的环境模型不断更新, 旧的不断消退	需要不断地对环境进行访问和建模; 只适应渐进的环境变化	[27]
室内环境	多重记忆存储模型	模仿人类多重记忆模型, 用选择机制将环境观测分为传感器记忆、临时记忆和永久记忆; 并利用永久记忆对地图进行更新维护	地图中缺少尺度度量信息, 且只考虑了环境中有限的环境变化	[28]
	多重记忆存储模型/3D 构图	该方法利用短期记忆和永久记忆机制, 保证只利用永久的环境信息构建地图	环境中稳定的元素需要经常被观测和识别才能加入到永久记忆中	[29]
	频谱分析	利用频谱对环境中的时空动态变化进行建模; 较小的存储需求适用于大范围环境	假设人类的行为是有规律可循的, 只适用于部分情形	[30]
	摘要地图	摘要地图中只保存被认为有用的路标信息	地图中有限的环境信息, 只适用于解决特定的任务	[36]
城市环境	端到端分割聚类	搭建了输入是原始点云数据、输出是分割聚类结果的端到端架构; 利用二分类解决多分类问题	只考虑了车辆、行人和自行车三类动态目标	[32]
	统一的栅格环境模型	构建了新颖的基于栅格的环境模型, 其中对动静物体及其不确定性、速度等特征进行统一建模	利用栅格地图表述城市环境, 存在栅格分辨率选择和边缘混淆问题	[33]
校园环境	无监督增量学习	利用 AP 聚类算法对三维点云进行聚类; 并通过机器人与环境的交互得到聚类目标是障碍的概率	该方法只判断聚类目标是否为障碍, 没有进一步估计其速度等特性	[34]
	地图长期维护	构建了包含位姿估计、全局地图维护和速度估计三个模块长期定位与构图系统, 通过对环境的重复观测, 直接对三维点进行状态预测与更新	该方法假设动态目标运动平滑	[35]

取得巨大的成功. 令人印象最深刻的当属基于环境表象 (Appearance-based) 的重定位方法, 如 FAB-MAP^[47] 和 DBOW^[48], 这类方法都是通过度量从环境的表象中提取的人工设计特征来实现场景匹配, 进而实现移动机器人的重定位. 但开放环境中光照、天气和季节等环境条件变化会导致场景表象差异显著. 图 2 摘自文献 [49], 展现了不同天气条件下同一场景的表象差异; 图 3 所展现的是 Nordland 数据集^[50] 中不同季节中同一场景的表象差异. 这些表象

差异会导致从同一场景提取的特征差异显著, 进而导致移动机器人重定位的失败. 因此如何从复杂多变的场景中提取对环境条件变化鲁棒的特征^[51] 是实现鲁棒重定位的关键.

描述场景的特征类型多样, 按获取方式可分为人工设计特征^[52] (Handcrafted features) 和基于学习获得特征^[53] (Learnt features); 按特征维度可分为二维图像特征和三维点云特征. 考虑到点云特征种类远少于图像特征, 为了便于讨论, 本章分别从人

工设计的图像特征、基于学习得到的图像特征和从三维点云中提取的三维特征这三类来综述机器人重定位自主环境适应的发展现状。

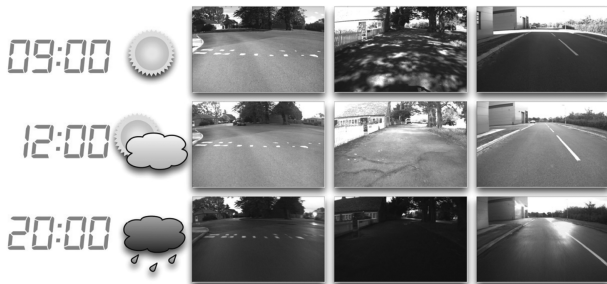


图 2 不同天气条件下的同一场景^[49]

Fig. 2 The same scene under different weather conditions^[49]

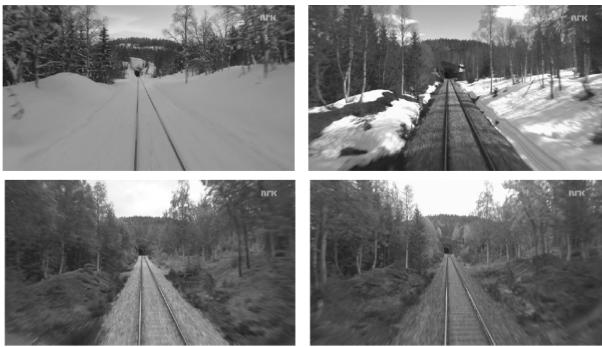


图 3 Nordland 数据集中不同季节条件下的同一场景^[50]

Fig. 3 The same scene under different season conditions in Nordland dataset^[50]

2.1 人工图像特征

尽管人工图像特征在变化的环境中不够稳定,但考虑到其模型简洁且计算效率高,许多学者对基于人工设计特征的方法进行改进以适应真实世界中复杂的环境条件变化。

为了应对光照条件变化,McManus 等^[54]将 RGB 图像转换为光照不变性图像 (Illumination-invariant images),并利用 FAST 特征检测器和 BRIEF 特征描述符实现场景匹配。如图 4 所示^[55],该图像模型对不同的光照条件具有良好的适应性。同样,Arroyo 等^[56]利用光照不变性图像模型和多种人工设计的图像特征实现不同版本的基于视觉的拓扑定位方法。该方法同时利用图像序列取代单幅图像并提取全局二进制描述符来实现高效匹配。SeqSLAM^[50]的作者坚信,利用足够长的图像序列(文中在不同的数据集中图像序列长度为 20 或 30 张)来代替单幅图片可有效提升场景匹配的准确度,提升场景识别算法对条件变化环境的适应能力,进而提高机器人场景识别的鲁棒性。

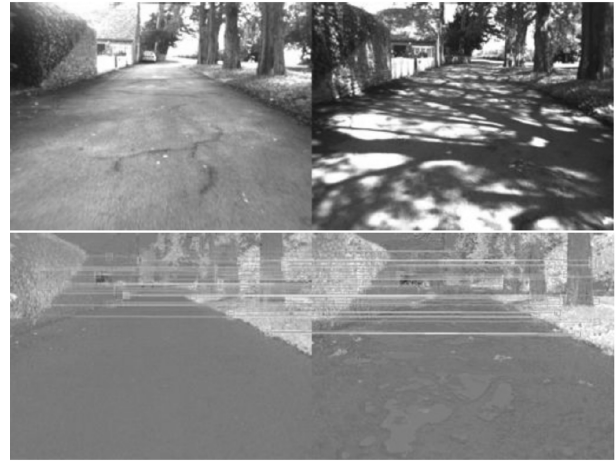


图 4 上:不同光照条件下同一场景的图像;下:相应的具有光照不变性的图像间的匹配^[55]

Fig. 4 Top: the images of the same scene with different lighting condition; Bottom: the matching result using illumination-invariant images^[55]

为了应对周期性季节转换,Neubert 等^[57]提出一种基于学习的环境外观预测方法。自然条件下环境外观的变化呈现一定的周期性,因此是可预测的。该方法通过学习环境外观随着时间的变化,来预测不同条件下的环境外观。作者利用该预测方法对 FAB-MAP^[58]、BRIEF-Gist^[59]、SeqSLAM 等传统的基于人工设计图像特征的重定位算法进行了改进,取得了显著的效果提升。实验结果表明基于环境外观预测的方法很好地适应自然条件的转变。以基于人工设计图像特征 (FAST detector and BRIEF descriptor) 的场景匹配算法为基础,Churchill 等^[49]提出了一种基于经验的定位方法。该方法将访问过的场景看作一条经验,并利用视觉里程计将观测经验串联在一起。如果同一场景中新的外观表述没有与历史外观表述匹配成功,则该新外观表述将被视为新的场景。通过机器人足够长时间的运行和积累,能对同一场景不同条件下的环境模型进行有效的表述,同时实现长航时机器人精确的重定位。但这种贪婪的方法需要尽可能多地访问同一个场景,才能完全适应同一场景不同条件下的变化,且在大范围开放环境中可能出现信息爆炸。

Johns 等^[60]对场景中的动态元素对局部特征的产生与消失的影响和特征所对应的潜在的外观进行了丰富的理论分析。作者将场景中的动态元素分为三类:行人、车辆等短期动态元素;建筑物重建或季节转换对树木的影响等长期动态元素;光照条件等重复出现和消失的周期性动态元素。分析各种动态元素对场景中局部特征的影响,能更好地学习动态场景中的静态特性,提升动态环境中的场景识别算法。

表 2 对基于人工设计图像特征的重定位方法进行了深入的总结和分析. 因为摄像头被动感知的特性, 光照条件变化和季节转换是基于视觉的重定位方法所面临的主要挑战. 基于光照不变性图像的方法只能解决光照变化引起的场景差异, 而基于经验的重定位方法和基于图像序列^[61]的重定位方法对光照条件和季节转换都有良好的适应性.

2.2 基于学习获取图像特征

传统的场景识别方法利用基于关键点的特征盲目地对所有的存储图片进行匹配. McManus 等^[62]反其道而行之, 对每一个场景无监督地学习特定的特征检测器, 这使得场景识别任务变成组织一大批附有空间索引的特征检测器. 传统的特征表述的是边、角点等低层面的结构, 而场景独立的 (Place-dependent) 的特征检测表述的是场景中像窗户、建筑物、轮廓等中等层面易识别的视觉元素, 对条件变化更加鲁棒. 实验证明, 这种方法可有效提高场景识别定位算法的鲁棒性, 虽然在定位尺度度量精度上有所降低, 但是可以接受的. 类似地, Linegar 等^[63]训练针对特定场景的线性 SVM 分类器来识别环境中非常有辨识度的部分. 他们在牛津数据集 (Oxford robotcar) 近 10 km 的路线中做了测试, 利用定制的地标 (Bespoke landmarks) 来实现机器人的自主定位可在变化的环境中大大减少定位失败的几率. 这种方法利用提取场景中长期稳定的特征来实现条件变化场景中的场景识别, 在缓慢变化的场景取得了很好的表现效果.

近年来, 深度学习^[64-66]在图像分类和目标检

测等领域取得了突破性进展, 成为目前最有效的自动特征学习方法. 深度学习模型具有强大的表征和建模能力, 逐层自动学习目标特征表示, 将原始数据经过一系列非线性变换, 生成高层次的抽象表示, 避免了手工设计特征的繁琐低效. 鉴于深度学习在计算机视觉领域不可思议的成功, 将卷积网络特征应用在视觉导航或场景识别等机器人领域引起了极大的兴趣.

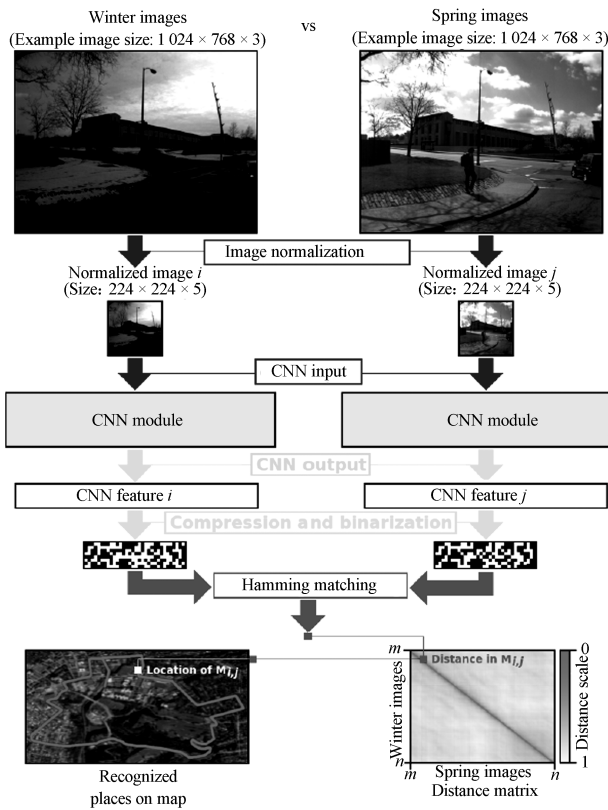
如图 5 所示, Arroyo 等^[67]利用卷积神经网络 (CNN) 获得的卷积特征代替人工设计图像特征, 该特征融合了不同卷积层的图像信息. 为了更快速实现场景识别, 卷积特征被压缩为二进制描述符并利用汉明距离实现场景匹配. 尽管卷积特征被用来提高在表象变化场景中的场景识别精度, 但是如何利用高维的卷积网络特征实现高效的场景匹配仍然是个挑战. Hou 等^[68]利用视觉词典算法将卷积网络特征转换为数值词典向量来实现场景的高效匹配; 同时为了解决利用词典向量实现场景匹配的感知混淆现象, 作者首先选取 k 个候选对象, 然后利用基于哈希的投票策略来确认最优的匹配场景.

文献 [69] 综合评估比较了三种先进的卷积网络在场景识别中的应用. 得出以下结论: 1) 从卷积网络较高的卷积层获取的特征暗含场景中的语义信息, 可以通过分割搜索空间来提升场景识别的效率; 2) 利用局部敏感哈希算法将卷积网络特征简化为二进制特征, 并利用二进制描述符的汉明距离来近似特征间的余弦距离, 这样可在很小的精度损失的情况下显著地提升匹配效率; 3) 在解决场景识别问题时, 针对场景理解训练的卷积网络比针对目标识别

表 2 移动机器人基于人工设计图像特征的重定位方法

Table 2 Visual methods of relocalization based on hand-crafted features for mobile robots in term of long-term autonomy

目标环境	关键词	主要思想	发表年份	文献
光照/季节/环境结构变化	基于经验描述的重定位	将环境的模型表述定义为一条“经验”, 利用视觉相对定位将场景的多种经验串联起来	2013	[49]
光照/季节变化	基于图像序列的重定位	利用较长的图像序列代替单幅图像实现场景匹配, 完成重定位	2012	[50]
光照条件变化	光照不变性图像	将 RGB 图像转换为具有光照不变性的图像, 进而利用 FAST 特征检测器和 BRIEF 特征描述符实现场景匹配	2014	[54]
光照/季节变化	光照不变性图像; 图像序列	利用光照不变性图像和图像序列等技术手段实现重定位; 同时提取全局二进制描述符来提高效率	2018	[56]
光照/季节变化	外观变化预测	鉴于自然条件下环境外观呈现周期性变化, 该方法通过预测不同条件下的环境外观来实现长期重定位	2013	[57]
光照/季节/动态因素	场景动态模型	通过分析场景中各种动态元素对局部特征的影响, 学习并利用场景中稳定的静态特征, 实现鲁棒的重定位	2013	[60]

图5 基于卷积特征的场景匹配系统框架^[67]Fig. 5 Framework of place recognition based on ConvNet features^[67]

训练的表现优异; 4) 从卷积网络中间层提取的特征可有效应对因光照、天气或季节引起的场景外观变化, 而从最高层提取的特征对观测视角的变化具有更高的鲁棒性.

2.3 三维点云特征

虽然摄像机成本低廉且具有丰富的色彩纹理等感知信息, 但其易受到光照条件影响, 甚至在黑暗环境中失去感知能力. 鉴于摄像机对光照条件敏感的特性, 许多学者利用激光雷达替代摄像机, 从数据源头解决光照条件变化对重定位算法的影响.

直接从三维点云尤其是稠密点云中提取特征并用于场景匹配, 需要很大的时间开销, 这严重影响机器人重定位的实时性. 为了降低搜索时间, 文献 [70] 中对原始三维点云进行降采样, 利用关键点投票的方法实现亚线性时间增长的场景识别算法. Magnusson 等^[71] 从三维点云中提取线性特征、面性特征以及球性特征, 并利用对不同类型特征的分布统计来实现场景间的高效匹配. 但直接在三维空间中对点云进行处理, 时间复杂度远远超过图像的二维空间. 受基于视觉的场景识别算法的启发, 文献 [72–74] 中将三维点云转换为二维的图像模型, 并利用基于视觉的方法来提升场景识别效率.

基于局部特征的场景识别算法易遭受环境变化的影响, 而基于全局特征的场景识别算法又易遭受观测视点的干扰. Dubé 等^[75] 提出一种基于点云分割的场景识别算法. 基于点云分割的方法在基于局部特征和基于全局特征的方法之间取得了很好的平衡, 能够做到扬二者之长避二者之短. 但基于点云分割的方法依赖在大范围场景中鲁棒的分割方法, 且在停车场等环境结构频繁变化的环境中不再适用. He 等^[76] 提出一种新颖的三维点云全局描述符用于场景识别. 作者通过将点云向球体坐标内多个平面进行投影, 利用点云在各平面中的分布构造描述符. 因为采用 PCA 技术寻找点云的主方向, 该描述符具有一定的观测位置和视角不变性. 此外, 该方法只关注点云的整体分布, 并不关注细节特征, 因此可有效适应因季节变换引起的环境变化.

对移动机器人来讲, 应对环境中的人车甚至建筑物改建等结构变化是至关重要的能力. 受文献 [49] 的启发, Maddern 等^[77] 将基于历史经验思想扩展到基于激光的构图和定位方法. 该方法通过重复访问, 收集环境中不同条件下的模型表述. 同时作者提出一种基于概率的定位方法, 利用多种不同的环境模型表述实现精确定位. 本文利用廉价的二维激光在牛津大学校园中长达一年的测试中实现了实时的持续改进的高精度地图构建和定位.

尽管国内外研究学者针对复杂多变的环境提出很多表现优异的场景识别算法, 但在机器人长期运行的过程中仍很难完全避免出现错误的重定位, 尤其是假阳性的误定位. Latif 等^[78] 提出一种增量式基于一致性分析的闭环验证方法 RRR (Realizing, reversing, and recovering), 利用视觉里程计的连续位姿约束与检测闭环约束之间的一致性分析来验证闭环准确性. 通过移除误匹配的假阳性闭环检测结果, 可保证机器人位姿地图的准确性.

表 3 对移动机器人基于三维点云的重定位方法进行了总结与对比. 尽管基于三维点云的方法普遍具有光照条件不变性, 但其针对环境结构、季节转换等条件变化的方法并不成熟. 复杂多变的环境中鲁棒高效的重定位方法亟待研究完善.

2.4 小结

为提高移动机器人对环境的长期自主适应能力, 基于视觉的移动机器人重定位方法主要利用以下技术手段: 1) 将视觉图像转换为具有光照不变性图像^[54–56], 降低光照、天气等条件变化对场景感知的影响; 2) 利用图像序列^[50–56] 替代单一图像来描述场景, 通过对连续图像的建模描述, 可充分利用各场景中稳定的环境特征来提高移动机器人场景描述的鲁棒性; 3) 利用深度学习等机器学习方

表 3 移动机器人基于三维点云的重定位方法

Table 3 Methods of relocalization based on point clouds for mobile robots in term of long-term autonomy

算法类型	主要思想	局限性	发表年份	文献
直接法	对三维点云进行降采样, 利用关键点投票的方法实现高效的场景配准	不能解决环境的结构变化	2013	[70]
特征法	从三维点云中提取线性特征、面性特征以及球形特征, 并利用对不同 类型特征的分布统计来实现场景间的高效匹配	众多的参数调节, 时间成本高	2009	[71]
	将三维点云转换为二维图像, 利用视觉的方法实现机器人的重定位	生成二维图像不具有视点不变性	2018	[73]
	将三维点云向若干个平面投影, 统计每个平面上点的投影分布生成全局特征, 进而实现场景匹配	对季节变化敏感	2016	[76]
分割聚类	对三维点云进行分割聚类, 利用聚类目标替代特征点实现场景匹配	对车辆等可能被移动的半静态物体敏感	2017	[75]
基于经验的方法	通过对同一场景的重复观测, 捕获其在结构或外观上的变化, 进而利用基于经验的方法实现机器人的长期稳定重定位	为了捕获某一场景的所有变化, 需要不断对该场景重复访问, 代价昂贵, 在大范围场景中易造成信息爆炸	2015	[77]

法^[62-63, 67-69], 并依赖大量的人工标注数据进行训练, 从而提取对环境条件变化稳定的特征来描述场景, 进而实现移动机器人重定位对环境条件变化的自主适应能力. 近年来, 随着大量图像数据集的涌现和深度学习、迁移学习等机器学习方法的不断发展, 移动机器人基于机器学习的重定位方法展现了良好的前景.

基于激光雷达的移动机器人长期运行中重定位方法主要是基于特征的方法^[70-71, 75], 此类方法在三维空间对点云提取局部特征或全局特征. 但特征提取过程通常需要大量的时间成本, 且不能有效应对视点和季节的转换. 受深度学习在视觉领域的启发, 基于激光的端到端的移动机器人重定位方法^[79] 具有很大的发掘潜力. 此外, 基于多元感知数据融合的移动机器人重定位方法也是今后发展的趋势.

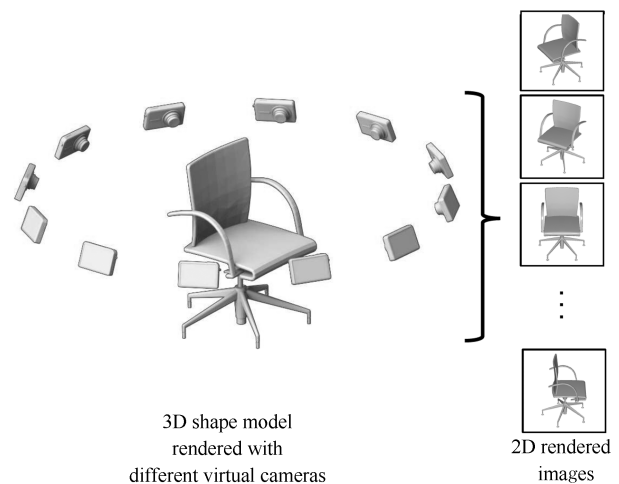
3 移动机器人长期自主场景理解

如何使移动机器人场景理解具有对开放环境的长期自主适应能力, 是移动机器人领域另一个极具挑战性的研究课题^[80]. 移动机器人在长期运行中, 对环境观测的时间、视点、距离和角度的不同, 会导致同一场景或同一物体的观测表象差异显著, 这对传统场景理解方法^[81-83] 造成巨大的困扰. 此外, 室外开放环境中存在未知性和不确定性, 环境类型、条件的多样性和不均衡还会给移动机器人场景理解带来小样本、样本类别不均衡等严峻挑战.

3.1 室内环境自主场景理解

当室内服务机器人重复访问同一个房间、办公室或其他场所时, 机器人对同一环境的观测视点差异性、室内多目标间的相互遮挡等问题都对机器人

进行目标检测和语义分割造成了较大影响. 从图 6 可以看出^[84], 对同一物体不同视角的观测结果存在显著差异. Zhuang 等^[85] 提出一种基于二维方位角 (Bearing angle) 图模型的室内场景三维目标检测方法, 该方法利用 Gentle-AdaBoost 算法训练多元分类器并提出了一种基于三维点云空间分割的尺度协调 (Scale coordinate) 技术进行多尺度的目标检测. 为了保证检测结果的准确性, 利用语义信息剔除错误的检测结果, 进一步提高了目标检测的准确率. 基于深度学习的方法^[86-87] 利用海量数据预训练神经网络模型, 也可有效解决视角、尺度的多样性问题, 在多目标识别和语义分割领域取得了很大成功.

图 6 同一物体多视角观测图像^[84]Fig. 6 Multiple view images of the same object^[84]

在复杂的室内环境中, 移动机器人需要在不同类型的室内场景间不断切换, 因此其自主场景理解还要应对由多样化场景切换所带来的跨领域场景理

解问题. 为了应对跨领域特征分布差异对目标识别带来的挑战, 文献 [88] 提出两种在线的特征变换算法通过在线方式来学习跨领域相似性度量函数, 并结合 k -NN 分类器实现了跨领域的多类物体识别.

移动机器人长期运行中不可避免地会引入少量新类别, 因此如何提升机器人场景理解对类不平衡问题和小样本问题的适应性成为长航时机器人自主场景理解领域的一个极具挑战性的研究课题. Zhang 等^[89] 提出了一种跨领域的人工合成加权实例的 WSMOTE-TrAdaboost 算法, 通过跨领域动态合成实例和“有选择性”地进行实例迁移来再平衡目标领域中训练数据的类分布. 由于基于小规模训练样本训练的分类器易于过拟合, 文献 [90] 通过共享方位角图碎片特征的方式联合训练多个目标检测器, 并且提出了一个基于数据采样技术的类不平衡算法, 对低置信率检测输出矩形框的再分类进一步提高了多类目标检测的准确率.

3.2 室外开放环境自主场景理解

在室外开放环境中, 移动机器人自主场景理解不仅要应对观测视角和尺度变化, 还要应对光照条件的变化甚至季节的转换. 在基于深度学习的场景理解算法^[64-66] 成熟之前, 学者一般利用条件随机场 (Conditional random field)^[91] 和随机森林 (Random forest) 等分类方法进行场景理解. 传统的分类方法需要人为设计特征训练分类器, 而人为设计的特征在条件变化的环境中稳定性差, 因而导致机器人场景理解在复杂多变的真实世界中可靠性降低.

前向摄像机采集图像存在显著的尺度多样性, 如图 7 所示, 近处的汽车相较远处的汽车占用图像中更多像素, 这对目标识别和语义场景理解造成巨大的挑战. 为了应对图像中显著的尺度多样性, 文献 [92] 结合深度学习和立体视觉重构算法实现了具有尺度不变性的语义分割方法. 该方法的核心在于在网络结构中添加了尺度选择层 (Scale selection layer), 该层网络在与相应的立体重构深度信息相匹配的尺度上提取卷积特征. 具有尺度不变性的特征表述能有效应对图像中尺度变化, 使得像素级 (Pixel-level) 的分类器不需要学习透视的相关规则. 针对同样问题, Ansari 等^[93] 利用量化的深度信息学习网络估计充足的深度信息, 进而利用深度信息调整分割网络中池化域的尺寸, 进而实现了具有尺度不变性的语义场景分割网络.

现阶段主流的语义场景分割方法大都利用摄像机作为主要的传感器. 但因为光照、天气甚至季节的转变可能导致训练数据集和测试数据集存在巨大差异, 进而影响语义分割的精度. 而作为主动感知传感器 (Active sensors), 激光雷达可有效应对此类变

化. Kim 等^[94] 利用图像和三维点云联合训练语义分割网络结构, 最终输出图像的语义分割结果. 作者设计了投影模块 (Projection module) 使得多线程 (Multi-modal) 网络学习 2D 和 3D 特征, 并在不同的领域将特征有效结合, 实现具有季节不变性的语义场景分割方法.



图 7 室外环境中不同尺度的车辆

Fig. 7 Cars in different scales in outdoor environments

3.3 小结

移动机器人长期自主场景理解方法主要基于机器学习理论和技术. 基于传统机器学习的目标分类和场景认知方法, 需要根据训练样本人为设计特征和训练分类器; 为了适应环境变化, 只有增加不同环境条件下的训练样本来训练多视角、多尺度分类器等^[95]. 但人工设计的特征存在局限性和片面性, 并不能全面地描述目标物体, 因此基于深度学习的目标分类和场景认知方法得到了广泛的关注并取得了优越的成果. 此类方法通过利用大量的标注数据, 训练端到端的深度神经网络, 跳过了特征设计环节, 充分利用了大数据的优越性.

虽然研究学者公开了许多视觉数据集, 但相对于不同的科学问题, 数据集的数量和种类仍是有限的. 基于迁移学习的方法, 将现有利用大规模数据集训练的网络架构迁移应用到小规模目标数据集中, 实现对目标问题的有效解决. 在未来一段时间内, 基于深度学习和迁移学习的移动机器人长期自主场景理解方法仍会占据主导地位.

表 4 对文献 [85-94] 所列举的研究工作进行了总结与对比. 随着深度学习的不断发展, 利用丰富的训练数据训练深度网络解决环境多样性的问题已日趋成熟; 同时, 迁移学习利用已有的模型映射到小样本甚至无法标注的领域, 为解决跨领域、类不平衡和小样本等问题提供了便利.

表 4 动态环境下长航时机器人自主场景理解方法

Table 4 Methods of scene understanding for mobile robots in term of long-term autonomy in dynamic environment

关键词	主要思想	目标问题	发表年份	文献
经典学习方法	将三维点云转换为二维图模型, 利用图像纹理和尺度协调技术解决多尺度分类问题	尺度不变性	2015	[85]
	利用大量训练数据训练目标识别模型, 利用训练数据的多样性解决多视角/多尺度目标识别	视角/尺度不变性	2015	[87]
深度学习方法	结合深度学习和立体视觉重构算法实现了具有尺度不变性的语义分割方法	尺度不变性	2016	[92]
	利用深度信息学习网络估计充足的深度, 进而利用深度信息调整分割网络中池化域的尺寸, 进而实现了具有尺度不变性的语义场景分割网络	尺度不变性	2018	[93]
	利用图像和三维点云联合训练语义分割网络结构, 学习并结合 2D 和 3D 特征, 实现鲁棒的语义场景分割方法	光照/季节不变性	2018	[94]
迁移学习方法	提出基于在线学习的跨领域特征变换算法, 并结合 k -NN 分类器实现了跨领域的多类物体识别	跨领域多分类	2018	[88]
	通过跨领域动态合成实例和“有选择性”地进行实例迁移来再平衡目标领域中训练数据的类分布	类不平衡问题	2017	[89]
	通过共享碎片特征的方式联合训练多个目标检测器, 并且提出了一个基于数据采样技术的类不平衡算法, 对低置信率检测输出矩形框的再分类进一步提高了多类目标检测的准确率	类不平衡/小样本问题	2018	[90]

4 研究重点和技术发展趋势

近年来, 得益于传感器技术、模式识别及机器学习等领域的突飞猛进, 国内外学者对移动机器人长期自主环境适应的研究展开了探索性研究工作, 并取得一些阶段性成果. 尽管如此, 但仍有相当数量的基础理论和关键技术问题亟待解决. 如何保证移动机器人地图构建与动态维护、重定位及场景理解等基础任务能力对复杂多变环境的长期适应性, 仍然是一个开放的问题. 基于对国内外移动机器人长期自主环境适应研究现状及发展趋势的分析, 要提高移动机器人对开放环境的长期自主适应性, 需要特别考虑以下几方面问题.

4.1 传感器数据融合与信息增强

对于移动机器人来说, 单一类型的传感器难以满足其长期自主环境适应的需求. 数据融合与信息增强可发挥多传感器协作环境感知的优势. 现阶段激光与视觉原始数据级的融合方法已比较成熟, 但其特征级和语义级等高层次的融合方法仍处于探索阶段. 近年来, 深度学习被广泛用于对图像或激光点云数据进行逐层目标特征的自动学习. 但如何利用图像和激光点云数据联合训练深度学习网络, 发挥多传感器的优越性仍有待研究. 通过多层次传感器数据融合得到的环境信息拥有更强大的环境表征能

力, 能为机器人提供更可靠的决策依据.

4.2 迁移学习与环境适应

移动机器人需要应对外部工作环境中光照、天气、季节和观测视角等因素的干扰和影响. 不同条件下, 同一场景或同一物体的特征分布存在明显差异. 近年来, 用于移动机器人长期验证的数据集较少, 目前只有 Nordland^[50]、CMU^[96]、Oxford^[97] 和 NCLT^[98] 等几个数据集能提供不同季节的环境感知数据. 此外, 这些数据集涵盖的环境类型很有限, 如 Nordland 只记录了火车轨道附近的野外环境, 而 CMU、Oxford 和 NCLT 数据集记录的只是校园结构化环境数据. 基于这些有限的数据集进行深度网络训练, 不足以保证移动机器人具有对开放环境适应能力的鲁棒性. 因此, 利用迁移学习将从有限数据集训练获得的知识迁移到开放环境的不同场景和不同任务中, 能有效增强移动机器人的长期自主环境适应能力.

激光雷达测距和视觉感知是长航时移动机器人常用的两种外部环境感知手段. 相比视觉传感器被动的环境感知方式, 激光雷达可实现主动的高精度测距, 且不受光照变化影响. 此外, 激光雷达还可以降低因季节改变导致的植被颜色变化对环境特征的影响. 但受硬件成本及采集难度的影响, 激光点云数

据集的规模远低于图像数据集,可用于移动机器人长期自主环境适应测试的数据集更是寥寥无几.虽然现有的深度图、强度图、方位角图等^[99]模型可将三维点云转换为二维图像化的表述方式,但由于上述模型与传统视觉图像的生成方式不同,从而造成各自的特征空间分布存在很大差异.若能将基于传统视觉图像训练所得知识迁移到基于三维点云的物体检测、场景语义分割或场景识别等任务中,则可解决因训练样本不足而导致的模型性能退化问题,将极大提高机器人对开放环境的适应性.

4.3 深层次语义信息挖掘与关联

现实世界中的物体都拥有独特的属性,且不同物体间存在着一定的联系.随着深度学习技术的发展,移动机器人的场景认知能力得到了很大的提升,但现阶段大部分开放环境中的场景认知研究工作仍停留在物体检测与场景语义分割的层面.场景中物体间基于属性的关联(如桌椅间的功能性关联、电脑主机与显示器间的依赖性关联),以及人与人之间社交性的关联,均有待进一步挖掘.此外,场景类别与场景中物体种类及布局间的关联存在一定的规律性,如卧室中通常有床和衣柜,厨房中通常有烟机灶具和厨具,花园环境中通常有大量的树木和花草,停车场环境中通常停有大量的汽车等.

在对场景中的物体进行精确识别的基础上,通过对物体与物体间、物体与场景间的关联进行挖掘与分析,可使机器人获得场景类别层面的认知,进而有针对性地选取与场景认知结果相对应的模型与参数,用于重定位与场景理解等任务.以厂区、社区中的长航时安防巡逻机器人为例,通过分析环境中车辆、建筑和植被等不同种类景物的数量和分布,可有效识别停车场、花园、居民区、厂房等具有特殊意义的场景类别,进而有针对性地完成指定任务.

4.4 高仿真测试环境与长航时训练

要使移动机器人具备稳定的长期自主环境适应能力,必须进行长时间的测试. Google、Uber 等公司的无人车进行不间断的路测,就是要提高其对环境的长期自主适应能力. 但移动机器人在真实环境中进行长时间的实测需要大量的人力物力,同时可能会承担昂贵的实测风险与代价. 现阶段,许多功能强大的实景仿真软件,如 Carla^[100]、Deepdrive^[101]、AirSim 和 AutoSim,为移动机器人自主行为模型的训练与测试提供了便利.

以 Carla 为例,该仿真软件主要用于城市环境中移动机器人系统的构建、训练以及验证.该平台可模仿真实城市环境中的建筑、植被、行驶车辆和行人等主要组成元素,并支持环境中不同光照和天气条件的设置.图 8 展示了 Carla 平台中不同天气条

件下的城市仿真场景,可以看出其具有很高的仿真效果.此外,平台中搭建的机器人系统支持视觉、激光传感器的灵活配置,并能实现传感器数据的实时获取,同时可获取场景中真实的深度信息及像素级语义标注.如图 9 所示,Carla 平台中虚拟机器人实时获取的视觉图像、深度图像及像素级语义标注等数据具有非常高的质量.基于 Carla 平台,移动机器人可充分利用其城市高仿环境进行长期的自主环境适应训练与测试. Carla 的更详细的信息参见 <http://carla.org/>.



图 8 Carla 中不同天气条件下的同一城市场景^[100]

Fig. 8 The same urban scene under different weather conditions in Carla^[100]

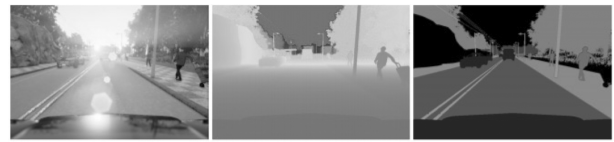


图 9 Carla 中机器人实时获取图像、深度图和语义标签^[100]

Fig. 9 The image, depth image and semantic labels gained by a mobile robot in real-time in Carla^[100]

4.5 增量式终生学习

仿真软件虽具有很高的仿真效果,但仿真环境并不能完全反映真实开放环境的复杂变化.移动机器人长期自主环境适应能力要接受真实环境的终生考验.在验证过程中,机器人应充分利用在线获取的真实环境数据及其与环境的交互学习,不断优化其自主行为模型以提高对开放环境的适应能力.此外,学习过程应该是终生的(Life-long),要覆盖移动机器人所有的工作时间.机器人在长期运行中会积累大量在线获取数据,这些数据具有非常好的时效性,但机器人只能获取如可行区域、不可行区域等简单粗略的标注信息,如何充分利用机器人弱标注的在线获取数据进行有效的学习是机器人终生学习的难点,也是使机器人具有长期自主环境适应能力的关键.

5 结束语

本文重点从地图构建与动态维护、重定位及场景理解等移动机器人基本的任务能力综述入手,对国内外移动机器人长期自主环境适应的研究现状进行了详细的梳理和评述,并对其未来的发展趋势进行了展望。

References

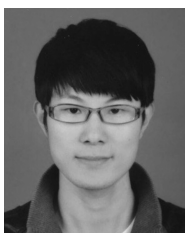
- Chai Tian-You. Development directions of automation science and technology. *Acta Automatica Sinica*, 2018, **44**(11): 1923–1930
(柴天佑. 自动化科学与技术发展方向. 自动化学报, 2018, **44**(11): 1923–1930)
- Barfoot T, Kelly J, Sibley G. Special issue on long-term autonomy. *The International Journal of Robotics Research*, 2013, **32**(14): 1609–1610
- Kunze L, Hawes N, Duckett T, Hanheide M, Krajník T. Artificial intelligence for long-term robot autonomy: a survey. *IEEE Robotics and Automation Letters*, 2018, **3**(4): 4023–4030
- Garg S, Jacobson A, Kumar S, Milford M. Improving condition- and environment-invariant place recognition with semantic place categorization. In: Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Vancouver, Canada: IEEE, 2017. 6863–6870
- Garg S, Suenderhauf N, Milford M. Don't look back: robustifying place categorization for viewpoint- and condition-invariant place recognition. In: Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA). Brisbane, Australia: IEEE, 2018. 3645–3652
- Naseer T, Burgard W, Stachniss C. Robust visual localization across seasons. *IEEE Transactions on Robotics*, 2018, **34**(2): 289–302
- Saarinen J P, Andreasson H, Stoyanov T, Lilienthal A J. 3D normal distributions transform occupancy maps: an efficient representation for mapping in dynamic environments. *The International Journal of Robotics Research*, 2013, **32**(14): 1627–1644
- Lázaro M T, Capobianco R, Grisetti G. Efficient long-term mapping in dynamic environments. In: Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Madrid, Spain: IEEE, 2018. 153–160
- Siva S, Zhang H. Omnidirectional multisensory perception fusion for long-term place recognition. In: Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA). Brisbane, Australia: IEEE, 2018. 5175–5181
- Zhu J L, Ai Y F, Tian B, Cao D P, Scherer S. Visual place recognition in long-term and large-scale environment based on CNN feature. In: Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV). Changshu, China: IEEE, 2018. 1679–1685
- Se S, Lowe D G, Little J J. Vision-based global localization and mapping for mobile robots. *IEEE Transactions on Robotics*, 2005, **21**(3): 364–375
- Pitzer B, Stiller C. Probabilistic mapping for mobile robots using spatial correlation models. In: Proceedings of the 2010 IEEE International Conference on Robotics and Automation. Anchorage, USA: IEEE, 2010. 5402–5409
- de la Puente P, Rodriguez-Losada D, Valero A, Matia F. 3D feature based mapping towards mobile robots' enhanced performance in rescue missions. In: Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems. St. Louis, USA: IEEE, 2009. 1138–1143
- Xin Jing, Gou Jiao-Long, Ma Xiao-Min, Huang Kai, Liu Ding, Zhang You-Min. A large viewing angle 3-dimensional V-SLAM algorithm with a Kinect-based mobile robot system. *Robot*, 2014, **36**(5): 560–568
(辛菁, 苟蛟龙, 马晓敏, 黄凯, 刘丁, 张友民. 基于 Kinect 的移动机器人视角 3 维 V-SLAM. 机器人, 2014, **36**(5): 560–568)
- Yang Hong, Qian Kun, Dai Xian-Zhong, Ma Xu-Dong, Fang Fang. Kinect-based 3D indoor environment map building for mobile robot. *Journal of Southeast University (Natural Science Edition)*, 2013, **43**(S1): 183–187
(杨鸿, 钱堃, 戴先中, 马旭东, 房芳. 基于 Kinect 传感器的移动机器人室内环境三维地图创建. 东南大学学报(自然科学版), 2013, **43**(S1): 183–187)
- Ding Wen-Dong, Xu De, Liu Xi-Long, Zhang Da-Peng, Chen Tian. Review on visual odometry for mobile robots. *Acta Automatica Sinica*, 2018, **44**(3): 385–400
(丁文东, 徐德, 刘希龙, 张大鹏, 陈天. 移动机器人视觉里程计综述. 自动化学报, 2018, **44**(3): 385–400)
- Montemerlo M, Thrun S, Whittaker W. Conditional particle filters for simultaneous mobile robot localization and people-tracking. In: Proceedings of the 2002 International conference on Robotics and Automation. Washington, USA: IEEE, 2002. 695–701
- Wolf D F, Sukhatme G S. Mobile robot simultaneous localization and mapping in dynamic environments. *Autonomous Robots*, 2005, **19**(1): 53–65
- Montesano L, Minguez J, Montano L. Modeling dynamic scenarios for local sensor-based motion planning. *Autonomous Robots*, 2008, **25**(3): 231–251
- Wang C C, Thorpe C, Thrun S, Hebert M, Durrant-Whyte H. Simultaneous localization, mapping and moving object tracking. *The International Journal of Robotics Research*, 2007, **26**(9): 889–916
- Henriques J F, Vedaldi A. Mapnet: an allocentric spatial memory for mapping environments. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 8476–8484
- Bürki M, Dymczyk M, Gilitschenski I, Cadena C, Siegwart R, Nieto J. Map management for efficient long-term visual localization in outdoor environments. In: Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV). Changshu, China: IEEE, 2018. 682–688

- 23 Tipaldi G D, Meyer-Delius D, Burgard W. Lifelong localization in changing environments. *The International Journal of Robotics Research*, 2013, **32**(14): 1662–1678
- 24 Meyer-Delius D, Hess J, Grisetti G, Burgard W. Temporary maps for robust localization in semi-static environments. In: Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems. Taipei, China: IEEE, 2010. 5750–5755
- 25 Saarinen J, Andreasson H, Lilienthal A J. Independent Markov chain occupancy grid maps for representation of dynamic environment. In: Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. Vilamoura, Portugal: IEEE, 2012. 3489–3495
- 26 Biswas J, Veloso M. Episodic non-Markov localization: reasoning about short-term and long-term features. In: Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA). Hong Kong, China: IEEE, 2014. 3969–3974
- 27 Biber P, Duckett T. Dynamic maps for long-term operation of mobile service robots. In: Proceedings of the 2005 Robotics: Science and Systems Conference. Cambridge, USA: Massachusetts Institute of Technology, 2005. 17–24
- 28 Dayoub F, Cielniak G, Duckett T. Long-term experiments with an adaptive spherical view representation for navigation in changing environments. *Robotics and Autonomous Systems*, 2011, **59**(5): 285–295
- 29 Morris T, Dayoub F, Corke P, Wyeth G, Upcroft B. Multiple map hypotheses for planning and navigating in non-stationary environments. In: Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA). Hong Kong, China: IEEE, 2014. 2765–2770
- 30 Krajník T, Fentanes J P, Cielniak G, Dondrup C, Duckett T. Spectral analysis for long-term robotic mapping. In: Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA). Hong Kong, China: IEEE, 2014. 3706–3711
- 31 Zhang J, Singh S. Low-drift and real-time lidar odometry and mapping. *Autonomous Robots*, 2017, **41**(2): 401–416
- 32 Wang D Z, Posner I, Newman P. What could move? Finding cars, pedestrians and bicyclists in 3D laser data. In: Proceedings of the 2012 IEEE International Conference on Robotics and Automation. Saint Paul, USA: IEEE, 2012. 4038–4044
- 33 Tanzmeister G, Thomas J, Wollherr D, Buss M. Grid-based mapping and tracking in dynamic environments using a uniform evidential environment representation. In: Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA). Hong Kong, China: IEEE, 2014. 6090–6095
- 34 Ott L, Ramos F. Unsupervised online learning for long-term autonomy. *The International Journal of Robotics Research*, 2013, **32**(14): 1724–1741
- 35 Pomerleau F, Krüsi P, Colas F, Furgale P, Siegwart R. Long-term 3D map maintenance in dynamic environments. In: Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA). Hong Kong, China: IEEE, 2014. 3712–3719
- 36 Dymczyk M, Lynen S, Cieslewski T, Bosse M, Siegwart R, Furgale P. The gist of maps – summarizing experience for lifelong localization. In: Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA). Seattle, USA: IEEE, 2015. 2767–2773
- 37 Zhuang Yan, Lu Xi-Bin, Li Yun-Hui, Wang Wei. Mobile robot indoor scene cognition using 3D laser scanning. *Acta Automatica Sinica*, 2011, **37**(10): 1232–1240 (庄严, 卢希彬, 李云辉, 王伟. 移动机器人基于三维激光测距的室内场景认知. *自动化学报*, 2011, **37**(10): 1232–1240)
- 38 Yan Fei, Zhuang Yan, Wang Wei. Outdoor scene comprehension of mobile robot based on multi-sensor information fusion. *Control Theory & Applications*, 2011, **28**(8): 1093–1098 (闫飞, 庄严, 王伟. 移动机器人基于多传感器信息融合的室外场景理解. *控制理论与应用*, 2011, **28**(8): 1093–1098)
- 39 Yu Miao, Hu Zhan-Yi. Higher-order markov random fields and their applications in scene understanding. *Acta Automatica Sinica*, 2015, **41**(7): 1213–1234 (余淼, 胡占义. 高阶马尔科夫随机场及其在场景理解中的应用. *自动化学报*, 2015, **41**(7): 1213–1234)
- 40 Zhu Bo, Gao Xiang, Zhao Yan-Nan. Place perception for robot indoor semantic mapping: a survey. *Acta Automatica Sinica*, 2017, **43**(4): 493–508 (朱博, 高翔, 赵燕喃. 机器人室内语义建图中的场所感知方法综述. *自动化学报*, 2017, **43**(4): 493–508)
- 41 Sun L, Yan Z, Zaganidis A, Zhao C, Duckett T. Recurrent-OctoMap: learning state-based map refinement for long-term semantic mapping with 3-D-lidar data. *IEEE Robotics and Automation Letters*, 2018, **3**(4): 3749–3756
- 42 Drouilly R, Rives P, Morisset B. Semantic representation for navigation in large-scale environments. In: Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA). Seattle, USA: IEEE, 2015. 1106–1111
- 43 Wang S, Clark R., Wen H K, Trigoni N. End-to-end, sequence-to-sequence probabilistic visual odometry through deep neural networks. *The International Journal of Robotics Research*, 2018, **37**(4–5): 513–542
- 44 Kendall A, Grimes M, Cipolla R. PoseNet: a convolutional network for real-time 6-dof camera relocalization. In: Proceedings of the 2015 IEEE international conference on computer vision (ICCV). Santiago, Chile: IEEE, 2015. 2938–2946
- 45 Lowry S, Sünderhauf N, Newman P, Leonard J J, Cox D, Corke P, et al. Visual place recognition: a survey. *IEEE Transactions on Robotics*, 2016, **32**(1): 1–19
- 46 Kim G, Kim A. Scan context: egocentric spatial descriptor for place recognition within 3D point cloud map. In: Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Madrid, Spain: IEEE, 2018. 4802–4809
- 47 Cummins M J, Newman P M. FAB-MAP: appearance-based place recognition and mapping using a learned visual vocabulary model. In: Proceedings of the 27th International Conference on Machine Learning. Haifa, Israel: Omnipress, 2010. 3–10

- 48 Galvez-López D, Tardos J D. Bags of binary words for fast place recognition in image sequences. *IEEE Transactions on Robotics*, 2012, **28**(5): 1188–1197
- 49 Churchill W, Newman P. Experience-based navigation for long-term localisation. *The International Journal of Robotics Research*, 2013, **32**(14): 1645–1661
- 50 Milford M J, Wyeth G F. SeqSLAM: visual route-based navigation for sunny summer days and stormy winter nights. In: Proceedings of the 2012 IEEE International Conference on Robotics and Automation. Saint Paul, USA: IEEE, 2012. 1643–1649
- 51 Tang L, Wang Y, Ding X Q, Yin H, Xiong R, Huang S D. Topological local-metric framework for mobile robots navigation: a long term perspective. *Autonomous Robots*, 2019, **43**(1): 197–211
- 52 Rublee E, Rabaud V, Konolige K, Bradski G. ORB: an efficient alternative to SIFT or SURF. In: Proceedings of the 2011 IEEE International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011. 2564–2571
- 53 Milford M, Lowry S, Sunderhauf N, Shirazi S, Pepperell E, Upcroft B, et al. Sequence searching with deep-learned depth for condition- and viewpoint-invariant route-based place recognition. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Boston, USA: IEEE, 2015. 18–25
- 54 McManus C, Churchill W, Maddern W, Stewart A D, Newman P. Shady dealings: robust, long-term visual localisation using illumination invariance. In: Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA). Hong Kong, China: IEEE, 2014. 901–906
- 55 Ratnasingam S, McGinnity T M. Chromaticity space for illuminant invariant recognition. *IEEE Transactions on Image Processing*, 2012, **21**(8): 3612–3623
- 56 Arroyo R, Alcantarilla P F, Bergasa L M, Romera E. Are you ABLE to perform a life-long visual topological localization. *Autonomous Robots*, 2018, **42**(3): 665–685
- 57 Neubert P, Sünderhauf N, Protzel P. Appearance change prediction for long-term navigation across seasons. In: Proceedings of the 2013 European Conference on Mobile Robots. Barcelona, Spain: IEEE, 2013. 198–203
- 58 Cummins M, Newman P M. Appearance-only SLAM at large scale with FAB-MAP 2.0. *The International Journal of Robotics Research*, 2011, **30**(9): 1100–1123
- 59 Sünderhauf N, Protzel P. BRIEF-Gist - closing the loop by simple means. In: Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems. San Francisco, USA: IEEE, 2011. 1234–1241
- 60 Johns E, Yang G Z. Dynamic scene models for incremental, long-term, appearance-based localisation. In: Proceedings of the 2013 IEEE International Conference on Robotics and Automation. Karlsruhe, Germany: IEEE, 2013. 2731–2736
- 61 Han F, Wang H, Huang G Q, Zhang H. Sequence-based sparse optimization methods for long-term loop closure detection in visual SLAM. *Autonomous Robots*, 2018, **42**(7): 1323–1335
- 62 McManus C, Upcroft B, Newman P. Learning place-dependant features for long-term vision-based localisation. *Autonomous Robots*, 2015, **39**(3): 363–387
- 63 Linegar C, Churchill W, Newman P. Made to measure: bespoke landmarks for 24-hour, all-weather localisation with a camera. In: Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA). Stockholm, Sweden: IEEE, 2016. 787–794
- 64 Zhang Hui, Wang Kun-Feng, Wang Fei-Yue. Advances and perspectives on applications of deep learning in visual object detection. *Acta Automatica Sinica*, 2017, **43**(8): 1289–1305 (张慧, 王坤峰, 王飞跃. 深度学习在目标视觉检测中的应用进展与展望. *自动化学报*, 2017, **43**(8): 1289–1305)
- 65 Chen L C, Papandreou G, Kokkinos I, Murphy K, Yuille A L. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, **40**(4): 834–848
- 66 Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(4): 640–651
- 67 Arroyo R, Alcantarilla P F, Bergasa L M, Romera E. Fusion and binarization of CNN features for robust topological localization across seasons. In: Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Daejeon, South Korea: IEEE, 2016. 4656–4663
- 68 Hou Y, Zhang H, Zhou S L. BoCNF: efficient image matching with Bag of ConvNet features for scalable and robust visual place recognition. *Autonomous Robots*, 2018, **42**(6): 1169–1185
- 69 Sünderhauf N, Shirazi S, Dayoub F, Upcroft B, Milford M. On the performance of ConvNet features for place recognition. In: Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Hamburg, Germany: IEEE, 2015. 4297–4304
- 70 Bosse M, Zlot R. Place recognition using keypoint voting in large 3D lidar datasets. In: Proceedings of the 2013 IEEE International Conference on Robotics and Automation. Karlsruhe, Germany: IEEE, 2013. 2677–2684
- 71 Magnusson M, Andreasson H, Nuchter A, Lilienthal A J. Appearance-based loop detection from 3D laser data using the normal distributions transform. In: Proceedings of the 2009 IEEE International Conference on Robotics and Automation. Kobe, Japan: IEEE, 2009. 23–28
- 72 Zhuang Y, Jiang N, Hu H S, Yan F. 3-D-laser-based scene measurement and place recognition for mobile robots in dynamic indoor environments. *IEEE Transactions on Instrumentation and Measurement*, 2013, **62**(2): 438–450
- 73 Cao F K, Zhuang Y, Zhang H, Wang W. Robust place recognition and loop closing in laser-based SLAM for UGVs in urban environments. *IEEE Sensors Journal*, 2018, **18**(10): 4242–4252

- 74 Kim G, Park B, Kim A. 1-day learning, 1-year localization: long-term LiDAR localization using scan context image. *IEEE Robotics and Automation Letters*, 2019, **4**(2): 1948–1955
- 75 Dubé R, Dugas D, Stumm E, Nieto J, Siegwart R, Cadena C. SegMatch: segment based place recognition in 3D point clouds. In: Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA). Singapore, Singapore: IEEE, 2017. 5266–5272
- 76 He L, Wang X L, Zhang H. M2DP: a novel 3D point cloud descriptor and its application in loop closure detection. In: Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems. Daejeon, South Korea: IEEE, 2016. 231–237
- 77 Maddern W, Pascoe G, Newman P. Leveraging experience for large-scale LIDAR localisation in changing cities. In: Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA). Seattle, USA: IEEE, 2015. 1684–1691
- 78 Latif Y, Cadena C, Neira J. Robust loop closing over time for pose graph SLAM. *The International Journal of Robotics Research*, 2013, **32**(14): 1611–1626
- 79 Uy M A, Lee G H. PointNetVLAD: deep point cloud based retrieval for large-scale place recognition. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: 2018. 4470–4479
- 80 Zhuang Yan, Chen Dong, Wang Wei, Han Jian-Da, Wang Yue-Chao. Status and development of natural scene understanding for vision-based outdoor mobile robot. *Acta Automatica Sinica*, 2010, **36**(1): 1–11
(庄严, 陈东, 王伟, 韩建达, 王越超. 移动机器人基于视觉室外自然场景理解的研究与进展. *自动化学报*, 2010, **36**(1): 1–11)
- 81 Li L J, Socher R, Li F F. Towards total scene understanding: classification, annotation and segmentation in an automatic framework. In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009. 2036–2043
- 82 Kumar M P, Koller D. Efficiently selecting regions for scene understanding. In: Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE, 2010. 3217–3224
- 83 Kim B S, Kohli P, Savarese S. 3D scene understanding by voxel-CRF. In: Proceedings of the 2013 IEEE International Conference on Computer Vision. Sydney, Australia: IEEE, 2013. 1425–1432
- 84 Su H, Maji S, Kalogerakis E, Learned-Miller E. Multi-view convolutional neural networks for 3D shape recognition. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 945–953
- 85 Zhuang Y, Lin X Q, Hu H S, Guo G. Using scale coordination and semantic information for robust 3-D object recognition by a service robot. *IEEE Sensors Journal*, 2015, **15**(1): 37–47
- 86 Eitel A, Springenberg J T, Spinello L, Riedmiller M, Burgard W. Multimodal deep learning for robust RGB-D object recognition. In: Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems. Hamburg, Germany: IEEE, 2015. 681–687
- 87 Wang A R, Lu J W, Cai J F, Cham T J, Wang G. Large-margin multi-modal deep learning for RGB-D object recognition. *IEEE Transactions on Multimedia*, 2015, **17**(11): 1887–1898
- 88 Zhang X S, Zhuang Y, Wang W, Pedrycz W. Transfer boosting with synthetic instances for class imbalanced object recognition. *IEEE Transactions on Cybernetics*, 2018, **48**(1): 357–370
- 89 Zhang X S, Zhuang Y, Hu H S, Wang W. 3-D laser-based multiclass and multiview object detection in cluttered indoor scenes. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, **28**(1): 177–190
- 90 Zhang X S, Zhuang Y, Wei W, Pedrycz W. Online feature transformation learning for cross-domain object category recognition. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, **29**(7): 2857–2871
- 91 Zhuang Y, Liu Y S, He G J, Wang W. Contextual classification of 3D laser points with conditional random fields in urban environments. In: Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Hamburg, Germany: IEEE, 2015. 3908–3913
- 92 Krešo I, Čaušević D, Krapac J, Šegvić S. Convolutional scale invariance for semantic segmentation. In: Proceedings of the 38th German Conference on Pattern Recognition. Hannover, Germany: Springer, 2016. 64–75
- 93 Ansari M D, Krauß S, Wasenmüller O, Stricker D. ScaleNet: scale invariant network for semantic segmentation in urban driving scenes. In: Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. Funchal, Madeira, Portugal: Scitepress, 2018. 399–404
- 94 Kim D K, Maturana D, Uenoyama M, Scherer S. Season-invariant semantic segmentation with a deep multimodal network. *Field and Service Robotics*. Cham, Germany: Springer, 2018. 255–270
- 95 Xiong Dan, Lu Hui-Min, Xiao Jun-Hao, Zheng Zhi-Qiang. Robust long-term object tracking with adaptive scale and rotation estimation. *Acta Automatica Sinica*, 2019, **45**(2): 289–304
(熊丹, 卢惠民, 肖军浩, 郑志强. 具有尺度和旋转适应性的长时间目标跟踪. *自动化学报*, 2019, **45**(2): 289–304)
- 96 Bansal A, Badino H, Huber D. Understanding how camera configuration and environmental conditions affect appearance-based localization. In: Proceedings of the 2014 IEEE Intelligent Vehicles Symposium. Dearborn, USA: IEEE, 2014. 800–807
- 97 Maddern W, Pascoe G, Linegar C, Newman P. 1 year, 1000 km: The Oxford RobotCar dataset. *The International Journal of Robotics Research*, 2017, **36**(1): 3–15

- 98 Carlevaris-Bianco N, Ushani A K, Eustice R M. University of Michigan North Campus long-term vision and lidar dataset. *The International Journal of Robotics Research*, 2016, **35**(9): 1023–1035
- 99 Liu Y S, Wang F, Dobaie A M, He G J, Zhuang Y. Comparison of 2D image models in segmentation performance for 3D laser point clouds. *Neurocomputing*, 2017, **251**: 136–144
- 100 Dosovitskiy A, Ros G, Codevilla F, Lopez A, Koltun V. CARLA: an open urban driving simulator. In: Proceedings of the 1st Annual Conference on Robot Learning. Mountain View, United States: PMLR, 2017.
- 101 Quiter C, Ernst M. deepdrive/deepdrive: 2.0 [Online]. available: <https://doi.org/10.5281/zenodo.1248998>, March 26, 2018.



曹风魁 大连理工大学博士研究生. 主要研究方向为 SLAM 和移动机器人长期自主定位技术.
E-mail: cfkybfq@mail.dlut.edu.cn
(**CAO Feng-Kui** Ph.D. candidate at the School of Control Science and Engineering, Dalian University of Technology. His research interest covers

SLAM and long-term autonomous localization for mobile robots.)

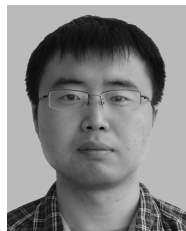


庄严 大连理工大学控制科学与工程学院教授. 主要研究方向为移动机器人面向复杂三维环境的自主感知、建模、地图构建、场景理解及机器学习技术在机器人领域的应用. 本文通信作者.
E-mail: zhuang@dlut.edu.cn

(**ZHUANG Yan** Professor at the School of Control Science and Engineering,

Dalian University of Technology. His research interest covers mobile robot 3D environment perception, mod-

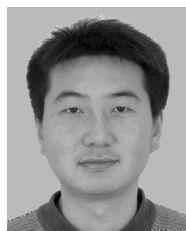
elling, mapping, scene understanding and machine learning in robotics. Corresponding author of this paper.)



闫飞 大连理工大学控制科学与工程学院副教授. 主要研究方向为移动机器人地图构建、路径规划、自主导航和场景理解. E-mail: fyan@dlut.edu.cn

(**YAN Fei** Associate professor at the School of Control Science and Engineering, Dalian University of Technology. His research interest covers mo-

bile robot map building, path planning, autonomous navigation, and scene understanding.)



杨奇峰 中国科学院沈阳自动化研究所, 中国科学院大学博士生. 主要研究方向为机器人学及智能控制.

E-mail: yangqifeng@siasun.com

(**YANG Qi-Feng** Ph.D. candidate at Shenyang Institute of Automation, Chinese Academy of Sciences, University of Chinese Academy of Sciences.

His research interest covers robotics and intelligent control.)



王伟 大连理工大学控制科学与工程学院教授. 主要研究方向为自适应控制、预测控制, 机器人学及机器人控制.

E-mail: wangwei@dlut.edu.cn

(**WANG Wei** Professor at the School of Control Science and Engineering, Dalian University of Technology. His research interest covers adap-

tive control, predictive control, robotics, and intelligent control.)