

室内移动机器人的 SLAM 算法综述



田野 陈宏巍 王法胜 陈兴文

大连民族大学信息与通信工程学院 辽宁 大连 116000

(17713291470@163.com)

摘要 SLAM(Simultaneous Localization and Mapping),即同时定位与地图构建,目前被广泛应用于机器人领域。SLAM 算法使得机器人处于陌生环境时,能够通过自身搭载的传感器来感知环境信息并建立环境地图,并完成对自身位姿的计算,从而能够在未知环境中进行移动。随着研究者们对 SLAM 问题的深入研究,SLAM 领域相关成果已非常丰富,但是有关室内场景 SLAM 的论述还不够系统。通过对现有的关于 SLAM 算法发展成果的总结和对比,对室内 SLAM 进行了综合性的阐述。首先介绍了 SLAM 的技术现状和室内场景 SLAM 在不同传感器下的分类问题;其次介绍了 SLAM 的经典框架;然后根据相关传感器种类的不同,简要介绍了不同传感器下常见的 SLAM 算法的原理,同时讨论了传统室内 SLAM 算法中存在的诸多局限性,引出了基于多传感器融合技术的 SLAM 和基于深度学习技术的 SLAM 两个研究方向;最后介绍了 SLAM 的未来发展趋势和应用领域。

关键词: 室内;定位与建图;激光雷达;相机;多传感器;深度学习

中图分类号 TP242

Overview of SLAM Algorithms for Mobile Robots

TIAN Ye, CHEN Hong-wei, WANG Fa-sheng and CHEN Xing-wen

Department of Information and Communication, Dalian Minzu University, Dalian, Liaoning 116000, China

Abstract As a localization and map construction method, SLAM(Simultaneous Localization and Mapping) is widely used in the field of robots. SLAM algorithm enables the robot to perceive environmental information and establish environmental map through sensors carried by the robot itself in an unfamiliar environment, and calculate its own posture. In this way, the robot can move in an unknown environment. With the in-depth study of SLAM, the research results in the field of SLAM have been very rich. However, the discussion on indoor SLAM is not comprehensive enough. Through the summary and comparison of the existing development results of the SLAM method, a comprehensive statement is shown. In this paper, the technical status of SLAM and the classification problem of SLAM under different sensors in indoor scenes are firstly introduced. Secondly, the classic framework of SLAM is revealed. Thirdly, the principles of SLAM algorithms with different sensors are described according to the different types of related sensors. Fourthly, the limitations of the traditional indoor SLAM algorithms are discussed and two research directions-SLAM based on multi-sensor fusion technology and SLAM based on deep learning technology are led out. Finally, the future development trend and application field of SLAM are suggested.

Keywords Indoor, Positioning and mapping, Lidar, Camera, Multi-sensor, Deep learning

1 引言

我国制造行业的不断发展,使基于高新技术的自动化产业逐步取代了传统的低端制造业。由于现代社会对提高生产效率的需要和人们对智能生活方式的向往,移动机器人技术得到了发展,并广泛应用于生产和生活。控制移动机器人需要计算机控制系统、传感器、定位系统和路径跟踪规划^[1],从 20 世纪 80 年代概率理论的广泛应用开始,SLAM 的发展也进入了“快车道”,概率方法和 SLAM 的结合取得了很大突

破,出现了以卡尔曼滤波、粒子滤波及最大似然估计为基础的研究方法^[2]。在 SLAM 算法后续的发展过程中,出现了许多开源的算法,有的基于激光雷达,有的基于相机。目前,人们对于移动机器人功能的要求越来越高,希望移动机器人能完成室内生活场景下的特定任务,为人们服务。移动机器人在室内的应用有扫地机器人、交互机器人等,室内场景下的定位与建图技术因此得到了发展。

随着 SLAM 算法的发展,各类 SLAM 算法均表现出了优秀的性能:基于激光的 SLAM 算法在静态小场景下,更能

收稿日期:2020-07-24 返修日期:2020-09-23 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:辽宁省自然科学基金指导计划项目(2019-ZD-0171)

This work was supported by the Natural Science Foundation Guidance Plan Project of Liaoning Province, China(2019-ZD-0171).

通信作者:陈宏巍(porschen@qq.com)

发挥其稳定性,有着稳定的测距性能,受光照强度的影响较小,操作简单;基于视觉的 SLAM 算法在纹理细节丰富的环境中,可以快速地识别环境中的纹理信息,能够应对复杂多变的动态环境。但是,由于机器人所处的室内环境越来越复杂,在某些情况下,基于单一传感器的 SLAM 算法显现出局限性。首先,当长时间在动态室内环境下建图时,移动机器人会面临物体移动、光照强度多变等影响因素,单一的激光或视觉传感器都无法在此类环境中稳定工作。其次,在硬件的稳定性方面,激光雷达的长时间旋转扫描会对内部机械结构产生损坏,移动机器人在环境中的测距工作会因震动而产生误差。最后,在算法实现与计算量方面,小场景下激光类算法的计算量低于视觉类,但随着区域面积的增加,计算量也会随之加大,进而影响移动机器人系统的实时建图性能。针对上述单一传感器中存在的不足,可以从几种可能的方向出发,如考虑改进传感器硬件系统的稳定性、环境适应性等,也可以考虑多传感器融合的研究方向,使各传感器协同工作,以改善单一传感器的局限性。另外,目前深度学习技术在图像处理方面显现出优势,通过神经网络对数据的训练可以省略提取特征环节,简化计算流程,使定位与建图过程变得简捷高效。视觉 SLAM 与深度学习的结合会提高移动机器人在有光照变化或动态大场景下实时建图的鲁棒性。

本文综合分析了研究者的研究成果,阐述了 SLAM 的框架和分类,以及在室内环境下每种类别所包含的算法和优缺点。然后通过对具体算法示例的论述到整体分类,从宏观上更好地把握每种类别的传感器功能和各种 SLAM 算法在室内场景下的优劣。随后,进一步论述了为了改善单一传感器局限性而采取的研究方案,最后讨论了室内 SLAM 的发展趋势。

2 SLAM 技术现状

Smith 等^[3]最早提出 SLAM 的概念,定位与建图是 SLAM 算法要解决的问题。当移动机器人处于未知环境中时,要依靠自身搭载的传感器采集得到距离、图像等环境信息,对自身位姿进行估计,同时建立环境地图。对自身位姿和周围环境的感知是移动机器人定位导航系统必备的能力。通过 30 多年的稳定发展,SLAM 被应用于人们生产,生活的多个领域,如机器人、虚拟现实、无人驾驶等。

自 2015 年起,SLAM 进入了鲁棒感知阶段^[4],与发展之初相比,其更加注重智能化、自动化,为移动机器人在室内环境下的运行指明了技术方向。移动机器人通过添加自适应模块来增强 SLAM 系统的自适应性,提高系统对故障的感知能力,同时,几何建图也在朝着语义建图方向发展,着重增强移动机器人系统的智能感知能力^[5]。目前,激光 SLAM 和视觉 SLAM 都在不断地发展和优化,近年来在多个领域都有所应用。在激光 SLAM 中,Lenac 等提出了一种基于 3D 姿态的快速 SLAM 系统^[6],汽车通过搭载 360°的激光雷达可以更好地估算车辆的轨迹;Pierzchala 等针对森林技术,提出使用基于 graph-SLAM 的同时定位和映射算法来生成森林的局部图^[7],通过激光扫描来启用自动 3D 映射;Drosechel 等提出了一种激光 SLAM 方法^[8],将 3D 激光扫描仪的结果汇总到本

地多分辨率地图中,可以解决局部地图中存在的漂移问题。在视觉 SLAM 中,Belbachir 等提出了一种基于单目相机的视觉 SLAM 系统^[9],同时添加了辅助传感器并应用于车载系统,提高了定位系统在驾驶中的精度;Schneider 等提出了一个开放的视觉惯性映射框架 Maplab^[10],其免费的源代码和高效的平台使视觉 SLAM 的研究受益。

早期,国外研究者提出了多种可用于室内的 SLAM 算法,室内激光 SLAM 的经典算法有 Gmapping, Cartographer 等,视觉 SLAM 算法有 ORB-SLAM, SVO-SLAM 等。国内研究的成果也在增加,并对经典算法进行了优化和改进。工业方面,随着近年来智能制造业的进步,室内 SLAM 已在多个领域有所应用,如医院、商场、仓储等。国内广州港南沙港区四期工程集装箱码头的无人驾驶智能引导车采用了激光雷达 SLAM、视觉 SLAM 和北斗导航系统的相互融合技术,创新性地使用了多种传感器结合的信息采集方式^[11];阿里菜鸟无人物流车,针对室内复杂的仓储环境搭载了固态激光雷达,拥有 3D 环境感知能力,提高了物流仓储分类的运作效率。国外 Lenac 等应用改进的 Hector 算法完成了其高校建筑图的绘制;Facebook 计算机团队的 AR 技术可精确地追踪多个点的深度,提高了实时性。

现阶段,由于深度学习在图像处理方面的优势,视觉 SLAM 和深度学习的结合有着很好的表现,但此方法的实际应用还不太广泛,大多处于研究阶段。人工智能的发展更加强调工作方式的智能化,面对不理想的场景,如光照强度变化大、检测目标移动速度快等,研究者们在 SLAM 算法中引入了深度学习技术。卷积神经网络可以直接输入图像的数据信息,避免了人工对图像进行处理和提取特征,较传统的几何方法提了效率,但目前多处于实验研究阶段。深度学习与视觉 SLAM 的结合主要体现在以下 3 个方面:帧间估计、闭环检测和语义地图的建立^[12]。视觉 SLAM 与深度学习结合,有效改善了在室内光照变化大的环境中相机的劣势,同时省略了特征提取的过程,深度学习的模型训练还在物体移动速度快的环境中提高了系统的鲁棒性。整体上看,此方法简捷高效,避免了复杂的计算量。

总之,早期的室内 SLAM 都有了研究和改进,各类算法都有各自擅长的领域,目前,人们正在探索如何采用其他技术与 SLAM 相融合,研究出对环境适应能力更高的定位导航技术。

3 SLAM 分类

根据移动机器人搭载的传感器不同,可以将 SLAM 分为激光 SLAM 和视觉 SLAM,为了更好地适应环境,目前还出现了多传感器融合的 SLAM。

3.1 激光 SLAM

激光 SLAM 搭载的是 2D 或 3D 激光雷达,根据所处的不同环境特点来选择适合的激光雷达。考虑到部分室内环境相对简单和对移动机器人的功能要求较低,移动机器人可选择 2D 激光雷达;而当室内环境因素多变且更为复杂时,移动机器人多采用 3D 激光雷达来动态扫描三维空间。

激光雷达在工作过程中,通过扫描环境信息获得分散的点,这些点中包含各种信息,如距离和角度,此类分散的点聚集时被称为点云。每两个点中的信息存在差异,机器人系统要将每个点的信息进行对比来获得距离和角度的变化程度,通过计算距离和角度的差异可以得出移动机器人的位置变化情况,从而知晓机器人的位姿。

目前激光雷达技术已趋于成熟,也应用于非常多的领域。激光雷达的使用有很多优势,如相比于视觉SLAM来说,激光点云计算量较小,计算简便,测距准确,而且在光照变化较大时不易受影响。但激光雷达仍有待改进,其简单的旋转式机械装置的长时间工作不利于内部结构的稳定。固态激光雷达的研究已有很多成果,如机电系统(MEMS)、面阵闪光(flash)技术和光学相控阵(OPA)技术^[13],可以在一定程度上改善传统激光雷达硬件结构的不足,使视角和信息采集能力有所提升,满足其基本功能,但是移动机器人搭载的固态激光雷达成本昂贵且不易维护。

3.2 视觉SLAM

视觉SLAM依靠相机来进行移动机器人的自身定位,其计算量相对于激光SLAM较大。按照相机工作方式的不同,分为单目相机、双目相机和深度相机3类^[14],它们有各自的特点。

单目相机拥有一个摄像头,记录二维的空间信息,通过采集环境的图像信息,再运用视觉几何原理获得机器人的位姿变换^[15]。单目相机由于结构问题,在三维空间或静态空间中无法获取到距离信息,容易产生误差,但其成本低且结构简单的优势备受初学者关注。双目相机有两个摄像头,类似人的双眼来判断距离,解决了单目相机易产生误差的问题,但计算量较单目相机有所增加。由于双目相机测量环境深度信息的距离有限,双目视觉SLAM往往不适合在大场景下工作。深度相机则运用了红外传感器技术,类似激光雷达通过发射并返回光来判断距离,相比于单目和双目相机,深度相机更容易获得环境的深度信息。其操作流程简单,但成本高、搭载困难,使得应用场景有限,更适合室内的定位和导航,故深度相机目前还存在诸多局限。

总体上来说,视觉SLAM的环境描述更加精细,可比对环境的纹理信息,提高了精确度,可以辨识两处环境的差异度。与激光SLAM相比,视觉SLAM搭载的相机虽然在内部结构方面更加稳定,但仍然存在局限性,移动机器人在室内环境只搭载单一的相机作为传感器,会受到环境条件的制约,鲁棒性较差,无法胜任在复杂环境中建图的任务。

3.3 多传感器融合的SLAM

针对室内环境下单一传感器的局限性问题,多传感器融合的SLAM技术方案有了一定的研究进展。除了激光雷达、相机外,另一种硬件设备——惯性测量单元(Inertial Measurement Unit, IMU)作为SLAM的辅助传感器,可以对角速度和加速度进行测量,在移动机器人运动过快时解决漂移问题。其与视觉SLAM系统的融合将会与相机优势互补,既能应对慢速移动时丰富的信息采集,又能解决快速移动时的轨迹漂移问题。

Wang等提出了一种基于激光雷达与双目视觉的移动机器人SLAM算法,该算法通过改进RBPF-SLAM算法,采用双目视觉与激光雷达,并结合里程计数据进行信息融合^[16],进而进行机器人的定位与导航。在该算法中,首先对激光采集得到的数据进行分析处理,得到距离信息;其次采用ORB算法提取双目相机测得的特征点,通过比对特征点得到角度信息;然后将距离信息与角度信息融合,得到机器人的实时位姿;最后根据激光雷达和双目相机的标记信息建立地图。该算法的提出改善了单一传感器的局限性问题,既改善了激光雷达机械结构不稳定的问题,又充分利用了视觉SLAM中图像采集精度较高的优势,提高了机器人建图的鲁棒性。Pan等提出一种单目相机与IMU融合的SLAM算法^[17],既利用了IMU对角速度和加速度的准确测量,又利用了相机在图像方面丰富的信息采集。Fan等在单目视觉和IMU融合的基础上加入了磁力计,将磁力计的数据融合到单目视觉SLAM算法中,不但可以解决纯旋转情况下的轨迹漂移问题,还可以提高精度^[18]。该方法在室内SLAM中可以改进单一传感器存在的局限性问题,并且可以使精度有所提高。

更多的多传感器融合方案仍在研究发展中,目前关于多传感器融合的问题还有待解决,如多种传感器工作时如何协同工作并实时反馈信息、如何判断某传感器是否故障等。

3.4 与深度学习技术结合的SLAM

深度学习于2006年被提出,其通过模拟人的大脑结构以构造复杂的神经网络模型,利用大量数据进行训练来模拟人的学习过程。计算机视觉领域发展至今,深度学习已在图像分类和目标检测方面有较好的表现,其实时能力强、适应能力好的特点引起了研究者关注,可与室内的视觉SLAM相结合,利于移动机器人的定位与导航。

传统的视觉SLAM由于相机的某些局限性,在建图过程中存在许多问题。例如,相机不适宜在光照强度变化大的环境中工作;在缺乏照明或环境纹理信息不明显时,相机的工作效果不佳;移动机器人系统运动过快或位于动态环境中时,相机容易跟丢目标。这些都会影响定位与建图的准确度。深度学习在图像处理方面的优势为SLAM带来了很大的提升空间,弥补了传统视觉SLAM的不足,提高了鲁棒性。

近年来,卷积神经网络在图像处理方面得到了广泛应用,与SLAM算法也有了很好的结合。总体来看,深度学习与SLAM算法的结合较基于传统几何方法的SLAM算法的效率有所提高,但现在深度学习仍然无法替代SLAM的全部工作,只能替代其中的一个或几个环节。目前常见的深度学习与SLAM结合的方法有帧间估计、闭环检测和语义地图构建。

3.5 几种室内SLAM算法的性能分析和比较

室内SLAM算法按照传感器的不同可以分为激光SLAM、视觉SLAM、多传感器融合的SLAM和基于深度学习的SLAM,在算法实现过程中,每个类别的算法性能存在差异,对环境的适应性、结果的准确度以及计算量都有所不同。

由于激光雷达在测距方面的优势,激光SLAM的测距相对准确,而对纹理信息则不能准确采集;视觉SLAM与激光

SLAM 在某些方面相反,相机能够采集环境的纹理信息,但是受自身结构的限制,易受到光照的影响,从而对相机的成像造成干扰;多传感器融合的 SLAM 则可以兼具激光雷达和相机的优点,补齐两者的短板,在工作过程中表现出稳定的性能,但是传感器数量的增加也会加大其计算量和故障率;基于深度学习的 SLAM 是一个新颖的方向,目前还处于研究之中,深度学习方法还不能完全替代整个 SLAM 系统,但是用深度学习来替代或帮助 SLAM 中的某个环节,会使工作效率有所提高,能减小工作量且提高鲁棒性。

表 1 列出了几种室内 SLAM 算法的性能比较。

表 1 室内 SLAM 算法实现过程中的性能比较

Table 1 Performance comparison of indoor SLAM algorithms in implementation process

SLAM 算法类别	优点	缺点
激光 SLAM	计算量小,测距准确,受光照影响较小	在纹理信息丰富的环境中不能很好地采集特征信息
视觉 SLAM	在纹理信息丰富的环境中能很好地采集特征信息,精确度高	计算量较大,受光照影响较大
多传感器融合 SLAM	降低了传感器的局限性,可适应多种环境	数据融合复杂度较高,计算量较大,不易维护
基于深度学习的 SLAM	卷积神经网络对图像处理更加智能,鲁棒性较好	不能完全替代整个 SLAM 框架,目前有待进一步研究

4 SLAM 的经典框架

SLAM 的经典框架主要分为以下部分:传感器、视觉里程计、后端、闭环检测和建图。

在视觉 SLAM 中,传感器和视觉里程计被纳入前端。传感器负责采集环境信息获得数据,数据会传到视觉里程计,视觉里程计对相邻图像或点的数据进行估算,根据部分图像或点的对比形成局部地图,估计位姿。移动机器人在行进过程中会采集多个位置的信息,后端通过将视觉里程计测得的不同时刻的位姿进行优化,得到全局地图。闭环检测的目的是通过前后信息对比来判断机器人走过的位置是否重合,避免漂移。如果位置重合,则将信息反馈至后端。最后,根据前面 4 步的信息分析得出机器人的运动轨迹,并进行建图。SLAM 算法的框架如图 1 所示。

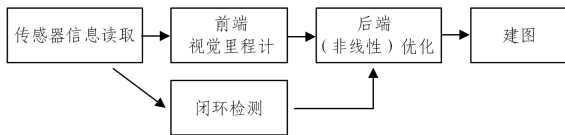


图 1 SLAM 算法的经典框架

Fig. 1 Classic framework of SLAM algorithm

SLAM 的经典框架经过几十年的研究和运用已经趋于成熟,在理想环境下工作效率可观。当 SLAM 算法与深度学习相结合时,深度学习技术会替代传统 SLAM 框架的一个或几个^[19]。基于深度学习的视觉里程计进行帧间估计时,较传统的帧间估计省去了复杂的运算过程,使效率提高。传统的后端闭环检测多采用词袋模型的方式,需要人为提取特征并进行图像信息的匹配来检测闭环。而基于深度学习的闭环检测利用了深度学习在图像比对与处理方面的优势,通过训练

的数据信息来自发地进行特征提取,较传统检测更加快速和准确。基于深度学习的建图过程较传统几何建图更加明确环境信息,带有环境标签信息的语义地图能够更好地进行机器人与人的交互。

5 基于激光雷达的 SLAM 算法

在激光雷达由军用转为民用的过程中,人们研究出了几种有效的室内激光 SLAM 算法。其中基于滤波的激光 SLAM 主要采用 Gmapping 算法,基于图优化的激光 SLAM 算法主要采用 Cartographer 算法。

5.1 Gmapping 算法

2007 年,Grisetti 等以 Fast-SLAM 方案为基本原理提出了 Gmapping 算法。Gmapping 是一种基于粒子滤波的二维 SLAM 算法,搭载了激光雷达传感器,目前常用于对室内环境进行二维的定位与建图。

粒子滤波的核心思想是随机采样,其主要分为以下几个阶段:初始化阶段、搜索阶段、决策阶段及重采样阶段。初始化阶段对移动机器人的位姿进行初始化;搜索阶段随机分布粒子,随后获得反馈的目标相似度信息;决策阶段利用一系列随机样本的加权和近似后验概率密度函数,通过求和来获得近似积分,然后进行粒子的权值计算,为选择性地重采样作准备;重采样阶段按照粒子权值在整体粒子权值中的占比复制粒子,有目的地重新分布粒子。粒子滤波算法将重复以上过程,最后进行地图估计^[20]。粒子滤波算法的主要工作流程如图 2 所示。

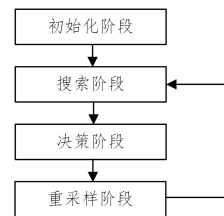


图 2 粒子滤波算法流程

Fig. 2 Particle filter algorithm process

相比于粒子滤波算法,Gmapping 算法更多地使用了统计方法。粒子滤波算法主要根据建立环境地图所需的粒子数量来衡量算法的复杂度,因此如何使用较少的粒子来构建较为精确的地图是该算法需解决的主要问题^[21]。为了改进粒子滤波算法由于粒子数过多而计算量大的问题,Gmapping 算法做了两个主要的改进:改进提议分布和选择性重采样。其中改进提议分布的作用是降低粒子数量,选择性重采样则解决了之前重复采样引发的粒子耗散问题。经过多年的研究与改进,Gmapping 算法已具备简单、易实现和低成本的特点,Mao 等曾利用基于 Raspberry Pi 板的室内移动机器人平台进行关于 Gmapping 算法的室内地图构建^[22],其结果显示了导航的鲁棒性。目前,关于 Gmapping 算法还有多种完善方法,如 Wongswan 等提出的矫正梯度细化算法,是一种新的改进基于粒子过滤器定位的方法^[23],其扩展了传统的粒子滤波算法,更具普适性。

在实现过程中,Gmapping 算法适用于室内小场景和低特

征环境下的定位与建图,精度较高且计算量小。Gmapping 不需要太多的粒子并且没有闭环检测,因此计算量小于 Cartographer,而且精度相差不大。当需要构建大场景或复杂环境地图时,Gmapping 算法则不再适用,随着场景的扩大,更多的粒子需求会加大计算量,在其本身没有闭环检测的情况下容易出现累计误差,使建图出现精度偏差。

5.2 Cartographer 算法

基于图优化的 SLAM 的核心思想是在移动机器人建图过程中实时构建位姿图,机器人的位姿作为节点,各个节点之间的转换关系作为边。与基于滤波方法的 SLAM 不同的是,基于图优化的 SLAM 分为前端和后端两个模块^[24],后端引入了闭环检测环节,因此相比于粒子滤波类算法,图优化类算法可以适应面积较大的场景,利用闭环检测可以消除误差,避免误差累积造成的建图不准确。

Cartographer 算法是谷歌公司开发的基于图优化的开源 SLAM 算法,在背包上面搭载激光雷达传感器和 IMU,可以实现二维和三维建图,并将测得的环境数据生成分辨率为 5 cm 的栅格地图。基于图优化的 SLAM 的出现解决了大尺度场景建图的问题。

Cartographer 框架主要分为两个部分:前端(Local SLAM)和后端(Global SLAM)。较其他激光雷达 SLAM 算法不同的是,Cartographer 算法的前端引入了子图(submap)的概念,前端进行数据提取和数据关联时,激光雷达每扫描一次会形成一个子图,每次扫描而得的数据帧会与上一次得到的子图进行比对并且插入上一次得到的子图中,子图的更新优化依赖于数据帧的不断插入,当没有数据帧插入时则形成完整优化的子图,此处主要应用非线性最小二乘法来进行求解。如此反复,获得若干个子图,即局部地图。后端首先进行闭环检测^[25],再对前端获得的若干个子图进行优化。通过全局计算得到优化后的位姿,可用来消除累计误差^[26],得到最优的全局地图,其工作流程如图 3 所示。

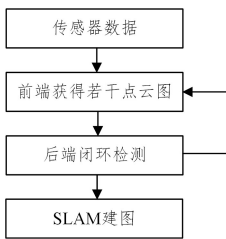


图 3 Cartographer 的工作流程

Fig. 3 Cartographer work process

Hess 等^[27]在此算法背景下介绍了一种基于二维激光雷达的实时闭环检测方法,采用分支定界法进行搜索,使搜索范围不断变小。当获得新的扫描图时,在其周围搜索可匹配的帧,如果匹配成功,则产生一个闭环。其中的优化环节采用了非线性最小二乘方法,将闭环检测的位姿优化问题转化为求最优解问题。

在实际应用中,Cartographer 算法对激光雷达的硬件要求不高,体现了易操作的特点。硬件水平的提高还会改善 Cartographer 算法的建图性能。Jian 等在移动机器人系统中

采用了工业级计算机作为核心控制器^[28],结合了激光雷达进行 ROS 平台下的室内建图,高效地完成了建图导航任务。Cartographer 算法在室内场景下还可以实现手持激光雷达建图,具有一定的优越性。一方面,Cartographer 算法前端可通过激光扫描获得大量的子图,因此可胜任较大室内场景的建图任务。另一方面,该算法采用基于图匹配的闭环检测方法,使计算量减小,同时提高了工作效率。

5.3 Hector 算法

Hector 算法也是基于图优化的 SLAM 算法,采用了高斯-牛顿法,通过搭载激光雷达传感器进行地图构建。

与 Cartographer 算法流程类似,Hector 算法也分为前端和后端,前端负责对机器人的运动进行估计,后端对位姿进行优化,但后端缺少了闭环检测环节。Liu 等在此基础上提供了一种可用的、具有完整的 6 自由度运动的平台,可以应用到空中飞行状态下的建图工作中。在建图过程中,首先前端进行激光扫描,获得栅格地图。每当激光雷达获得新的数据时,将其与上一时刻的地图进行匹配。为使激光雷达数据映射到栅格地图中,采用双线性插值的方法来获得连续的栅格地图。后端采用高斯-牛顿法对邻近帧进行匹配,使地图数据误差最小,得到优化的地图。最后选用了 3 种分辨率的地图,其分辨率逐个增加,目的是为了匹配值局部最小^[29]。对于室内走廊等退化环境,激光雷达无法获得准确的环境特征,会使精确度降低,因此 Yu 等于 2019 年提出了一种基于信息融合的改进型 Hector 算法,其结合里程表和惯性测量单元辅助机器人建图,采用卡尔曼滤波法进行信息融合^[30],提高了输入信息的准确度,最终提升了移动机器人在室内退化环境中的定位及建图性能。

在算法实现过程中,Hector 算法的基本流程与 Cartographer 算法大致相同,但与 Gmapping, Cartographer 算法不同的是,Hector 算法不需要里程计,但对激光雷达的精度要求较高,精度至少要达到 40 Hz 的帧率。Hector 算法对硬件要求较高的特点使其在室内场景应用时存在局限性,但该算法能估计 6 个自由度的位姿,可以胜任崎岖路面环境及空中的定位与建图工作,也可应用于无人机室内导航。

5.4 几种室内激光 SLAM 算法的比较

常用于室内的激光 SLAM 算法(如 Gmapping, Cartographer, Hector 等)由于各自的原理和应用情况不同,因此在实际应用中会表现出不同的性能。

Gmapping 算法作为一种经典的、可应用于室内的激光 SLAM 算法,具有轻便易用的优点,可用于小型室内移动机器人,但由于缺少闭环检测,其仅能够在简单的、低特征的室内环境下保持可靠的精度;Cartographer 算法与 Gmapping 算法相比,增加了闭环检测环节,在建图精度方面有所改善,同时能适应大场景下的建图导航工作;Hector 算法的大致流程与 Cartographer 算法相似,但是缺少闭环检测,使精度下降,同时由于其对激光雷达精度要求较高,使得 Hector 算法的室内应用不够广泛。

表 2 对几种常见的室内激光 SLAM 算法在实现过程中的性能进行了比较。

表2 室内激光 SLAM 算法在实现过程中的性能比较

Table 2 Performance comparison of indoor laser SLAM algorithms in implementation process

激光 SLAM	优点	缺点
Gmapping	易实现;在室内小场景且特征较少时表现优良;改善了粒子耗散问题	缺少闭环检测环节,精确度较低
Cartographer	可进行闭环检测,利于在室内较大场景下建图,精度较高	优化时计算量较大
Hector	不需要里程计	对激光雷达要求较高;缺少闭环检测环节

6 基于视觉的 SLAM 算法

基于视觉的定位与建图(V-SLAM)同样是移动机器人导航的热门研究方向。视觉 SLAM 算法依托相机获得环境信息,搭载不同的相机进行特征提取、选择关键帧^[31],获得不同的建图效果。单目相机、双目相机和 RGB-D 相机在实际运用过程中应结合各自的特点。

视觉 SLAM 算法拥有经典的框架,包含前端视觉里程计、后端优化、闭环检测和建图 4 个主要部分^[32]。早期的 Mono-SLAM 是世界上第一个单目视觉 SLAM 算法^[33],后端采用卡尔曼滤波,对前端稀疏特征点进行跟踪,导致信息容易丢失。后期的 PTAM 算法^[34]使用非线性优化取代了卡尔曼滤波,出现了前端和后端的区分,使得定位与建图能够同时进行。下面就近些年几种可用于室内环境的 V-SLAM 算法进行简述。

6.1 ORB-SLAM

ORB-SLAM 算法(Oriented FAST and Rotated BRIEF)是 Mur-Arta 提出的一种 V-SLAM 系统^[35],是一种应用于单目相机的 SLAM 算法,也是一种基于特征的 SLAM 算法,它保持着和早期其他算法同样的效果:定位和建图同时进行。与早期的 SLAM 算法相比,其添加了闭环检测功能来检测闭合回路,以减小累计误差。

ORB-SLAM 算法的核心思想为基于 ORB 描述量进行特征匹配和重定位。该算法分 3 个线程运行,包括特征检测与跟踪、建立地图、闭环检测与校正。特征检测与跟踪线程,跟踪环节首先从环境图像中提取 ORB 特征,根据前一帧的图像对当前帧进行优化,估计当前的位姿;其次通过相邻的地图点匹配更多的特征来进行位姿优化;最后选取关键帧为建立局部地图作准备。建立地图线程,建图部分主要完成局部地图的构建,通过加入关键帧来更新局部地图,同时剔除多余或重复的关键帧,以此对构建的局部地图进行优化。闭环检测与校正线程,闭环检测进行闭环的检测和修正。闭环检测使用了词袋模型(Bow2)方法,相当于将特征集合构建为字典,首先从采集到的图像中提取特征点并对其进行描述,其次使用 K-means 算法或贝叶斯算法进行聚类操作,最后以树的方式表示字典,以便搜索闭环。基于词袋模型方法的闭环检测的优势是可以通过检测图像相似度快速地检测到闭环。闭环校正主要是闭环融合和 Essential Graph 的图优化^[36]。ORB-SLAM 的工作流程如图 4 所示。

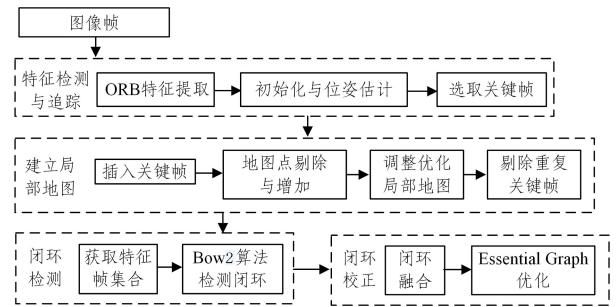


图4 ORB-SLAM 的工作流程

Fig. 4 Work process of ORB-SLAM

单目相机容易产生视差,因此对环境深度的判断会产生误差,而双目相机由于基线的存在,可以在一定程度上改善单目相机视差的缺陷。但双目相机基线的应用场景并不是任意的,双目相机基线的长度有限,当相机距离环境物体或边缘较远时,基线会失去作用,此时的双目相机类似于单目相机。RGB-D 相机能更好地改善这一缺陷。2017 年, Tardos 等发布了 ORB-SLAM2 算法^[37], ORB-SLAM2 是一套基于单目相机、双目相机以及 RGB-D 相机的系统。在 ORB-SLAM 的基础上, ORB-SLAM2 能够应用于搭载双目相机和 RGB-D 相机的移动机器人^[38],解决了测量环境深度的问题。与 ORB-SLAM 不同的是, ORB-SLAM2 在后端采用的是基于单目和双目的光束法平差优化(bundle adjustment)的方法,提高了精确度。ORB-SLAM2 作为较流行的 V-SLAM 系统,能创建相对简单的 3D 点云图^[39]。

ORB-SLAM 作为在特征丰富的室内环境中表现优秀的一种 SLAM 算法,除了可用于室内地面移动机器人外,近年来还常应用于室内场景的飞行器,结合其他方式协同工作。Haddadi 等在 GPS 受限的环境中进行室内导航时,将 ORB-SLAM 与 EKF-SLAM 相结合^[40],使用低成本四旋翼无人机进行室内导航,提高了估计位姿的鲁棒性和轨迹跟踪的准确性。Santos 等提出了一种将鲁棒控制器与鲁棒的视觉反馈相结合的方法,视觉反馈部分采用了 ORB-SLAM 算法,作者使用了车载摄像头来采集视觉传感器数据并实时检测是否存在干扰^[41]。实验表明此方法的结果令人满意,飞行器不仅可以完成室内导航任务,还可以应用于有风的室外场景。

对于室内场景下的定位与建图,基于特征的 ORB-SLAM 算法能够根据检测到的室内环境特征来构建局部地图,然后进行优化。但由于其依赖环境特征,具有一定的局限性,因此在纹理特征缺失的环境下无法得到所需的特征点,使建图陷入困境。

6.2 SVO-SLAM 算法

基于特征的方法是依靠提取图像中显著的特征来描述环境的特征信息,根据描述进行邻近帧间的特征匹配,该方法的优点是精度高,缺点是速度较慢;基于直接的方法是利用图像像素进行像素匹配,无需关联特征数据,与特征法相比,速度有所提升。因此, Forster 等于 2014 年将特征点法和直接法相结合^[42],提出了一种半直接法视觉里程计算法 SVO(Semi-Direct Monocular Visual Odometry)。

SVO-SLAM 算法是半直接法(稀疏直接法)中的常用算法,不同于利用图像特征点匹配的基于特征的方法,以及对整个图像直接匹配的直接法,半直接法的核心思想是对图像中特征点周围的小图像块进行直接匹配^[43],提高了效率和精度。其主要分为两个线程——前端运动估计和后端地图构建。当单目相机获得环境信息后,运动估计线程首先根据相邻帧之间的像素灰度差异,建立优化函数来优化光度误差,然后优化特征点位姿,最后分别优化三维的空间点和相机的位置,构建像素重投影误差的优化函数。后端通过深度估计来建立地图,当出现新的关键帧时,选取若干个特征点,将选取的点通过矩阵变换获得各自相应的极线,在极线上找到特征点的对应点,再使用三角测量方法计算深度;然后不断更新,直到深度估计收敛到一定的范围,在地图中融入该三维点坐标。SVO-SLAM 算法通过同时进行前端和后端,可以同时进行定位与建图,这使得 SVO-SLAM 算法系统能够快速和高效地对相机运动姿态进行估计,同时能够迅速、及时地对周围环境地图进行拓展和延伸^[44]。SVO-SLAM 算法的工作流程如图 5 所示。

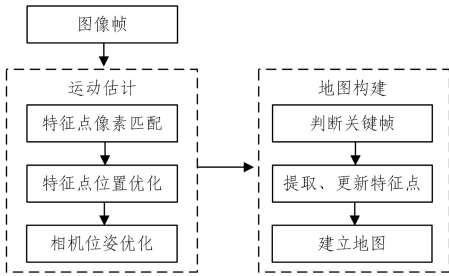


图 5 SVO-SLAM 算法工作流程

Fig. 5 Work process of SVO-SLAM algorithm

2017 年,SVO-SLAM 算法从 1.0 版本升级为 2.0 版本。Forster 等提出了 SVO 2.0,与早期的 SVOS-LAM 算法相比,SVO 2.0 能够支持多相机和宽视场相机,同时增加了运动先验,在跟踪边缘像素和弱角点方面具有鲁棒性^[45],实验表明,该算法的每帧跟踪仅需要 2.5 ms,其速度和精度都有所改善。

Huang^[46]介绍了 SVO-SLAM 算法相比直接法和特征点法的优势,即 SVO 算法避免了在提取、匹配特征点上耗费过多时间,也不需要整幅图像进行匹配,所以在计算资源有限的机载微型计算机上也能实现很快的运算速度。SVO-SLAM 算法还可实现对移动机器人的位置、姿态、速度、未知尺度、惯性测量单元零偏及重力方向的估计。其实验结果表明,SVO-SLAM 算法计算量小,实时性好,适用于无人机系统,也可用于室内机器人建图或手持建图。

6.3 几种常用的室内视觉 SLAM 算法的比较

在视觉 SLAM 算法中,可用于室内的算法有 ORB-SLAM,SVO-SLAM 等。其中,ORB-SLAM 算法发展至今,对各种类型相机的兼容度较高,如单目相机、双目相机、RGB-D 相机等,可针对室内环境的不同而选用不同类型的相机;SVO-SLAM 算法创新性地采用半直接法,计算量减小,准确度有所提高。但与 ORB-SLAM 算法相比,SVO-SLAM 算法不能很好地适应多种类型的相机。

表 3 对两种室内视觉 SLAM 算法进行了性能比较。

表 3 室内视觉 SLAM 算法在实现过程中的性能比较

Table 3 Performance comparison of indoor visual SLAM algorithms in implementation process

视觉 SLAM	优点	缺点
ORB-SLAM	精确度高;可兼容多种类型相机	特征不丰富,环境建图效果较差
SVO-SLAM	采用半直接法提高了效率和精度;计算量较小,实时性好	不能兼容多种类型相机

7 基于多传感器信息融合的 SLAM 算法

多传感器信息融合技术是一个新的研究领域,融合了多个学科门类,在 SLAM 领域中主要用于解决单一传感器在特殊环境下存在的局限性问题,目前研究中还存在诸多难点。近年来,出现了激光视觉融合、视觉 IMU 融合、激光视觉 IMU 融合等传感器融合方式,并且实验证明,基于多传感器融合的 SLAM 算法较单一传感器的 SLAM 算法在面对复杂特殊环境时更具灵活性和可靠性,系统更具鲁棒性,实验结果也更加精确可信。多传感器融合的常用方法包括加权平均法、卡尔曼滤波法、贝叶斯估计法等,下面就几种常见的方法进行简要介绍。

7.1 加权平均法

由于多种传感器输入的数据较单一传感器复杂度更高且具有冗余性,因此需要将不同来源的数据加以整合,同时剔除存在的冗余信息,提高信息融合后的系统性。加权平均法具有简便、直观的特点,该方法直接对数据源进行操作,可将一组传感器提供的冗余信息进行加权平均,其结果作为融合值。该方法适用于构建动态实时系统。

假设存在 k 个传感器,对同一目标进行检测,输出观测数据 x_1, x_2, \dots, x_k ,此时应确定各个数据对应的权重 w_k ,则最终结果可用下式表示:

$$\bar{x} = \frac{x_1 w_1 + x_2 w_2 + \dots + x_k w_k}{\sum_{i=1}^k w_i} \quad (1)$$

如何确定各传感器输入数据的权重是加权平均法的核心问题。早期 Li 等提出了一种使用评价函数确定加权因子从而获得融合数据的方法,同时证明了此方法的融合数据优于使用算术平均值方法得到的数据^[47]。后期在此基础上出现了二次加权的方法,引入了最优比例权重的概念,先对单个传感器进行加权,再对整体进行加权,目的是使算法性能达到最优,精度相应提高。但由于传感器的方差固定不变,因此在融合时要考虑传感器的方差对融合权重的影响。

加权平均法作为一种可对数据直接操作的方法。在实际应用过程中使数据融合更为简便,同时可适用于复杂的室内环境,实时观测数据变化。

7.2 卡尔曼滤波法

卡尔曼滤波主要用于融合低层次实时动态的多传感器冗余数据。该方法用测量模型的统计特性得到统计意义下的数据最优融合方式和数据估计。假设有 n 个观测方程, w 为噪

声,则多传感器信息融合系统方程可定义为:

$$x(k+1) = Ax(k) + Bu(k) + \omega \quad (2)$$

$$y_1(k) = C_1 x(k) + v_1 \quad (3)$$

...

$$y_n(k) = C_n x(k) + v_n \quad (4)$$

卡尔曼滤波法的核心问题是如何建立滤波所需的状态方程和观测方程。在实际工作过程中,卡尔曼滤波法将为多传感器融合提供最优估计,其前提是移动机器人系统应具备线性动力学模型,使用卡尔曼滤波法可以简化计算量,并且节约数据存储。现阶段针对卡尔曼滤波算法,出现了一些改进方案。Ayabakan等在卡尔曼滤波器的基础上,研究了联邦卡尔曼滤波器(Federated Kalman Filter, FK)的性能,有集中式和分散式结构的卡尔曼滤波器被广泛应用于室外定位和导航应用^[48]。近年来,联邦卡尔曼滤波器由于其室外导航性能而被应用于室内定位。在室内环境中,由于环境的差异会产生各种数据融合方面的困难。如果存在障碍物,会涉及反射问题,Kaya等为了解决反射过程中接收信号强度指示器的精度较低的问题,对位置的估计采用了无线传感器网络中RSSI测量的方法,该传感器网络由3个信标节点、1个移动节点和卡尔曼滤波器中的加速度计测量组成^[49]。实验结果表明此方法在障碍物较多的环境中可以提高位置估计的精度。

在算法实际运行过程中,采用单一的卡尔曼滤波器对多传感器组合系统进行数据统计时还存在问题,如单个传感器的增加会使故障率随之增加;在单个系统出现故障时会影响整个系统,使系统的可靠性降低。

7.3 贝叶斯估计法

贝叶斯估计用来描述两个条件概率之间的关系,为数据融合提供了一种手段,是融合静态环境中的多传感器高层信息的常用方法。其一般公式如下:

$$P(A_i | B) = \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)} \quad (5)$$

在早期的研究中,Chen等提出了一种基于贝叶斯估计的多传感器数据融合方法,此方法主要利用多个同类型的传感器对同一待测参数进行测量^[50]。在该方法中进行数据融合主要分为3个步骤。第1步引入置信距离理论,由于在系统拥有多个传感器时,各传感器的输出数据服从正态分布,因此可通过置信距离理论对传感器输出数据进行判断。第2步选择最佳融合数,选择一个临界值来划分置信距离,判断两个传感器输出的数据是否互相支持。第3步采用贝叶斯估计进行融合计算,使传感器信息依据概率原则进行组合,用条件概率结果来表示传感器观测的不确定性。当多个传感器的观测坐标一致时,可以直接对传感器的数据进行融合,但大多数情况下,传感器测量数据要以间接方式采用贝叶斯估计来进行数据融合。

Zhang等考虑了在传感器数据融合中存在的各来源数据的不确定性和不一致性的问题,提出了一种优化的贝叶斯估计多传感器数据融合方法——基于卡尔曼滤波器的贝叶斯融合算法^[51]。根据滤波器在传感数据和融合数据方面的应用

方式,他们提出了3种不同的技术,即前向滤波法、后向滤波法和前后向滤波法。其结果证明,该方法可以解决传感器数据的不确定性和不一致性问题。经过不断的研究,人们发现在多传感器数据融合的实现过程中,贝叶斯估计法较其他统计方法如最大似然估计法、矩法估计法等存在一定的便捷性。

7.4 几种多传感器数据融合算法的对比

在加权平均法、卡尔曼滤波法、贝叶斯估计法这3种多传感器融合方法中,每种方法在实际运行时都有各自的特点。加权平均法作为一种较简便的多传感器融合方法,操作简单且易实现,但要根据实际情况准确设置每个传感器的权重以使得融合效果更优;卡尔曼滤波法在应用时要尽量避免数据出现冗余,否则会对数据融合造成影响,降低实时性和准确度;贝叶斯估计法作为数据融合的经典方法,引入了条件概率,与卡尔曼滤波法相比更加简便快捷。

表4对3种常用的多传感器数据融合算法进行了性能比较。

表4 常用的多传感器数据融合算法在实现过程中的性能比较

Table 4 Performance comparison of multi-sensor data fusion algorithm in implementation process

多传感器融合算法	特点
加权平均法	简单、直接、易操作
卡尔曼滤波法	避免了复杂的计算,对数据进行了最优估计。但存在冗余数据时,实时性较差
贝叶斯估计法	用条件概率表示信息的不确定性,更加便捷

8 基于深度学习的SLAM算法

目前,深度学习技术与SLAM的结合主要体现在3个方面:帧间估计、闭环检测和语义地图构建。

8.1 深度学习与帧间估计

帧间估计即视觉里程计,作为SLAM前端的主要任务,其根据两个相邻帧的图像信息来判断相机的位姿。在传统的视觉SLAM中,当相机采集到图像信息时,前后相邻的两帧图像作为第一和第二关键帧,将第二帧与第一帧图像进行匹配来确定下一个关键帧^[52]。此后每当获得一帧图像时,先进进行与前一帧相邻图像的特征匹配,并与设定的阈值相比较来确定下一个关键帧。移动机器人在运动过程中会不断得到图像帧,将特征提取与匹配算法处理相邻帧的特征用于相机的位姿估计。

基于深度学习的帧间估计较传统SLAM中的帧间估计显得更加直接,省去了特征提取的环节。根据训练方法和数据集标签化程度的不同,将基于深度学习的帧间估计方法分为监督学习、无监督学习、半监督学习3类^[53]。监督学习的方法通过采集到的图像帧来估计位姿,然后通过卷积神经网络(CNN)学习图像帧的特征,再比对相邻帧间的信息,这种基于深度学习的数据训练方法避免了传统几何方法的大量计算;无监督学习则可以在获得单帧图像的同时估计出图像深度和相机位姿;半监督学习方法需要人为干预来标记信息,在一些特定环境中改善了无监督学习的不适应性。

近些年,基于深度学习的帧间估计方法得到了研究,

Konda 等提出了一种预测相机运动方向和速度的方法,使用了深度神经网络架构,该端到端的架构对视觉里程计信息进行了提取并学习^[54],其实验表明这种无监督的方式能够适应连续帧的位姿估计。Costante 等利用卷积神经网络在图像方面的优势进行帧间估计,他们对比了 3 种卷积神经网络(CNN-1b,CNN-4b,P-CNN)之后,提出了用 Brox 算法提取连续两帧的稠密光流特征的方法,作为卷积神经网络的输入^[55],其实验结果展示了该算法在应对图像运动模糊及光照变化方面的鲁棒性。

目前,在实际应用中,融合了深度学习的帧间估计与传统 SLAM 中的帧间估计相比,显得更加简单,避免了复杂的计算。但基于深度学习的帧间估计尚不完全成熟,如要实现更优的性能,还有待继续研究。

8.2 深度学习与闭环检测

闭环检测用来检测机器人在运动过程中是否到达过曾经到过的位姿,类似于图像相似度检测问题。如果机器人到达过曾经到过的位姿,则看作一个闭环。闭环检测在 SLAM 中是一个关键环节,因为其检测结果的精确度直接影响最终建图的准确度。

在 SLAM 中,传统的闭环检测受信息检索技术的影响,多采用基于词袋模型的方法^[56]。词袋模型方法的信息检索多依赖于人为干预和设定,而且词袋规模较大时效率较低,占用资源。在视觉 SLAM 中,人为提取的特征信息无法适应特殊环境,在光照强度变化剧烈或有移动物体的环境中无法胜任建图工作。而在深度学习中,卷积神经网络在特征提取和分类方面的优势可以改进上述问题,给基于深度学习的闭环检测方法带来了更多研究空间。Yi 等提出了一种基于卷积神经网络的词袋模型方法,这种用于可扩展鲁棒视觉场景识别的 CNN 特征词袋图像匹配方法以传统的词袋模型法为基础,利用卷积神经网络在图像识别方面的优势加速特征匹配^[57],在处理室内大场景地图时效率较高。目前,还涌现出其他深度学习方法,例如,Hou 等利用 Caffe 框架下的特征提取使用了 Alex Net 模型,其优点是即使在光照强度变化较大的环境下,特征提取也较为准确^[58];Sunderhauf 等在 Caffe 框架的基础上,使用以 ImageNet 数据库预训练好的 Alex Net 模型进行特征提取,提出了两种改进思路,一种是利用局部敏感散列搜索算法进行高效的特征搜索并提取,另一种是根据语义信息对搜索区域进行分割以减少搜索对象,以此解决了 SLAM 闭环检测也中的实时性问题^[59]。

近年来,随着人们对神经网络的深入研究,神经网络在闭环检测方面的应用越来越多,SLAM 闭环检测也越来越智能化。如 Merrill 等将手工制作特征和卷积神经网络提取特征两种方式相结合,以一种新颖的方式来进行闭环检测,并且从专门针对位置识别任务训练的定制深度 CNN 中提取所有信息模式,以此来解决只有改变用户选择的阈值才能真正闭合回路的偶然性问题^[60]。其实验表明,在此方法下闭环检测系统不需要设定阈值,并且闭环检测的准确度有所提高。

传统的闭环检测方法依靠人为设定阈值,将当前帧中的

信息与历史帧信息进行比较,检验是否超出阈值,由此判断是否形成闭环。由于此方法中阈值设定具有人为干预的因素,因此可能导致结果的不准确性。基于深度学习的闭环检测原理是利用深度学习自发学习先前从大量关键帧中提取的特征,然后将当前的信息与之前的学习结果进行比对,判断是否存在闭环^[61],基于神经网络的判断比人为处理特征信息更加精确,可以得到信息更丰富的图像^[62]。因此深度学习技术的引入提高了定位建图的导航系统的实时性,使移动机器人在室内光照变化大或动态大场景下建图的鲁棒性有所提高。

关于基于传统方法和深度学习方法的闭环检测效果如表 5 所列。

表 5 两种闭环检测方法的性能对比

Table 5 Performance comparison of two closed-loop detection

methods		
	基于传统方法的 闭环检测	基于深度学习的 闭环检测
智能化程度	低	高
准确度	低	高
局限性	高	低
特殊环境下的适应能力	弱	强

8.3 深度学习与语义地图构建

语义 SLAM 是近些年深度学习技术与 SLAM 相结合的产物,人们希望对地图进行标签化。传统的 SLAM 只考虑机器人的两个问题:“我在哪儿”“我该如何行走”,但对“周围环境中的物体究竟是什么”没有研究,只依靠传感器进行几何建图。随着现在智能化水平的提高,语义建图成为研究目标。机器人在建图过程中不仅能够获得环境的几何信息,同时还能识别环境中的单个物体,明确每个物体的属性,这样既能提高移动机器人的智能化水平,又能使机器人与人进行交互。

认知地图^[63]在 2007 年被提出,其认为图像中应包含语义信息,帮助机器人定位。后续研究者认为二维的地标和射频技术也可以生成语义地图。Yao 等在构建环境地图方面,通过对动静态区域的分析,筛选关键帧,得到绘图关键帧,并对其点云进行叠加,得到动态环境下的静态环境信息^[64],提高了位姿估计的鲁棒性。Sunderhauf 等利用了稀疏的基于特征的 RGB-D SLAM/基于图像的深度学习目标检测方法和三维无监督的分割方法^[65],辅助机器人理解场景信息并进行交互。Ma 等提出了一种在 RGB-D 图像序列中进行语义分割的方案,使用了深度学习方法,用一种自监督方式训练网络用于预测多视角一致的语义信息^[66],此方法在单视角和多视角方面都取得了不错的效果。

语义 SLAM 的出现在 SLAM 实际运行过程中丰富了传统 SLAM 的内容,使地图信息更加全面;室内环境建图时,语义 SLAM 可以为物体建立标签信息,有助于机器人与环境进行交互。

9 室内 SLAM 发展趋势

通过几十年的积累,基于不同传感器的 SLAM 具有各自的应用领域。现阶段,激光 SLAM 广泛应用于扫地机器人等

室内机器人, SVO-SLAM 的计算速度快, 适用于无人机系统, ORB-SLAM2 的定位导航方案完善, 适用于室内移动机器人来提取环境特征。SLAM 技术的不断完善使移动机器人更加智能, 可以更好地执行设定的任务。但是针对复杂的室内环境, 传统的单一传感器存在一定的局限性, 会受到光照、结构或非结构障碍物、环境纹理等的影响, 使最终的建图效果下降。

针对上述问题, 首先在实际应用中如果仅改善传感器硬件的机械系统, 会使得稳定性提升, 但同时成本也增加了, 导致其无法胜任复杂室内环境中的导航任务。其次, 如果将深度学习和 SLAM 相结合, 虽然可以提高特征提取和选择的精度, 但目前来说, 数据训练的环节仍然无法替代 SLAM 的前端, 传统的几何建图仍是主流, 因此基于深度学习的 SLAM 算法还有待提升。最后, 多传感器的方案较以上两种方案显现出了更多优势, 在复杂室内环境下表现更优, 目前常见的传感器组合有激光视觉融合、视觉惯性融合、激光视觉和 IMU 三者融合等。一方面, 由于 IMU 自身含有陀螺仪和加速度计, 可测得加速度和角速度, 因此常用于三维空间的导航, 可应用于较为复杂的室内环境中; 另一方面, 激光和视觉传感器各有优劣, 激光、视觉传感器二者相互结合, 不仅在传感器硬件结构的稳定性方面优势互补, 而且在采集环境信息时, 激光点云与相机图像信息的融合提高了定位的精确度。因此, 激光、视觉和 IMU 相互融合既可以通过 IMU 测得移动机器人的位姿, 又可以利用激光雷达和相机采集距离和环境特征, 是室内多传感器融合 SLAM 的一个发展趋势。

在未来, 为了解决室内环境下的传感器局限性问题, 基于多传感器融合的 SLAM 算法将会是室内 SLAM 的主要发展趋势, 激光、视觉和 IMU 融合的方式更具优势。多种传感器通过信息融合可以更好地适应移动机器人所处的复杂性环境和人们对其功能的多样性需求, 而且目前工业界的机器人 SLAM 算法正朝着多传感器融合的方向发展。同时在多种传感器融合的研究中仍然面临诸多问题, 如多传感器协作运行过程中数据的实时处理、降低计算复杂度等, 还需要进一步研究。

结束语 本文通过对 SLAM 算法框架、分类以及经典室内 SLAM 算法的论述, 讨论了现阶段 SLAM 算法在室内场景中存在的问题, 提出了多传感器数据融合的发展方向。综上所述, 室内简单环境中的单一传感器研究已逐渐成熟, 各种算法也得到了广泛应用。基于多传感器融合的 SLAM 算法将是未来的发展方向, 会为 SLAM 算法带来更多的研究空间, 能够提高移动机器人定位与建图的精确度。基于多传感器融合的 SLAM 算法作为未来有价值的研究方向, 也会为室内移动机器人的定位与建图带来更好的解决方案。

参考文献

- [1] LIU Z. Research status and future of mobile robot technology [J]. Digital Technology and Application, 2019, 37(10): 214-215.
- [2] MA Z G, ZHAO Y G, LIU C Y, et al. Review of Research on Laser and Vision Fusion SLAM Method[J]. Computer Measurement and Control, 2019, 27(3): 1-6.
- [3] SMITH R C, CHEESEMAN P. On the representation and estimation of spatial uncertainty[J]. The International Journal of Robotics Research, 1986, 5(4): 56-68.
- [4] SCARAMUZZA D, CADENA C, LEONARD J, 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.
- [5] LI Y T, MU R J, SHAN Y Z. A brief analysis of the development status of unmanned system vision SLAM technology [J/OL]. Control and Decision, 2020: 1-10. http://kzyjc.alljournals.cn/ch/reader/download_new_edit_content.aspx?file_no=201908110000001&edit_id=20200106145046001&flag=2&year_id=&quarter_id=.
- [6] LENAC K, KITANO A, CUPEC R, et al. Fast Planar Surface 3D SLAM Using LIDAR[J]. Robotics and Autonomous Systems, 2017, 92: 197-220.
- [7] PIERZCHALA M, GIGUERE P, ASTRUP R. Mapping forests using an unmanned ground vehicle with 3D Lidar and graph-SLAM[J]. Computer and Electronics in Agriculture, 2018, 145: 217-225.
- [8] DROSECHEL D, BEHNKE S. Efficient Continuous-Time SLAM for 3D Lidar-Based Online Mapping[C]//2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2018: 1-9.
- [9] BELBACHIR N, NOORI N, AKDEMIR B. Real-Time Vehicle Localization using on-Board Visual SLAM for Detection and Tracking[C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). IEEE, 2019: 4291-4295.
- [10] SCHNEIDER H, DYMZYK M, FEHR M, et al. Maplab: An Open Framework for Research in Visual-Inertial Mapping and Localization[J]. IEEE Robotics and Automation Letters, 2019, 8(29): 1418-1425.
- [11] DENG S Y, GUO C J. Research on real-time location and map construction method based on multi-sensor fusion[C]//Proceedings of the 11th China Satellite Navigation Annual Conference-S13 Autonomous Navigation, Academic Exchange Center of China Satellite Navigation System Management Office. 2020: 6.
- [12] ZHAO Y, LIU G L, TIAN G H, et al. A Summary of Visual SLAM Based on Deep Learning[J]. Robot, 2017, 39(6): 889-896.
- [13] CHEN J Y, SHI Y C. Research progress of solid-state lidar[J]. Opto-Electronic Engineering, 2019, 46(7): 1-11.
- [14] GAO X, ZHANG T. The fourteen lectures on visual SLAM [M]. Beijing: Publishing House of Electronics Industry, 2017: 97-107.
- [15] ZHAO L S, FENG Y, CAO Y. Optimization of SURF algorithm parameters in monocular visual positioning[J]. Computer Technology and Development, 2012, 22(6): 6-9.
- [16] WANG X W, HE L L, ZHAO T. Research on Mobile Robot SLAM Based on Lidar and Binocular Vision[J]. Journal of

- Transduction Technology, 2018, 31(3):394-399.
- [17] PAN L H, TIAN F Q, YING W J, et al. Monocular camera-IMU external parameter automatic calibration and online estimation of visual-inertial navigation SLAM[J]. Chinese Journal of Scientific Instrument, 2019, 40(6):56-67.
- [18] FAN J B, SUN J, FAN H H, et al. Autonomous positioning method based on magnetometer, IMU and monocular vision[J]. Overall Aerospace Technology, 2019, 3(6):39-45.
- [19] LIU R J, WANG X S, ZHANG C, et al. A Summary of Visual SLAM Based on Deep Learning[J]. Journal of System Simulation, 2020, 32(7):1244-1256.
- [20] WU Z Y, ZHANG C, LIN Y. Optimal design of laser SLAM algorithm based on RBPF [J]. Computer Engineering, 2020, 46(7):294-299.
- [21] WANG H, WANG P. Research and Implementation Based on RBPF-SLAM Algorithm [J]. Computer Application System, 2019, 28(7):169-173.
- [22] MAO S L, DI W, HAO Y D. Indoor 2D Laser SLAM on a Raspberry Pi-Based Mobile Robot[C]//2019 11th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC). 2019.
- [23] WONGSUWAN K, SUKVICHAI K. Generalizing corrective gradient refinement in RBPF for occupancy grid LIDAR SLAM [C]//2017 IEEE International Conference on Robotics and Biomimetics(ROBIO). IEEE, 2017.
- [24] ZHANG L, LIU Z Y, CAO J Y, et al. Cartographer algorithm and system realization based on enhanced pose fusion of sweeping robot[J]. Journal of Software, 2020, 31(9):1-13.
- [25] WU C D, YAO J M, HU H L. Cartographer 2D SLAM algorithm-research[J]. Cable TV Technology, 2018(4):20-22.
- [26] SANTOS J M, PORTUGAL D, RUI P R. An evaluation of 2D SLAM techniques available in Robot Operating System [C]// IEEE International Symposium on Safety, Security and Rescue Robotics. IEEE, 2014:1-6.
- [27] HESS W, KOHLER D, RAPP H, et al. Real-time loop closure in 2D lidar SLAM [C]//Proc of IEEE International Conference on Robotics and Automation. Piscataway, NJ: IEEE, 2016: 1271-1278.
- [28] JIAN J Z, LI X. Design and Implementation of ROS-Based Autonomous Mobile Robot Positioning and Navigation System [C]//2019 18th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES). 2019.
- [29] KOHLBRECHER S, VON S O, MEYER J, et al. A flexible and scalable slam system with full 3d motion estimation[C]//2011 IEEE International Symposium on Safety, Security and Rescue Robotics. IEEE, 2011: 155-160.
- [30] YU N, ZHANG B. An Improved Hector SLAM Algorithm based on Information Fusion for Mobile Robot[C]//2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems(CCCIS). IEEE, 2019:279-284.
- [31] QUAN M X, PU S H, LI G. Visual SLAM Review[J]. Journal of Intelligent Systems, 2016, 11(6):768-776.
- [32] LV L H. Summary of Vision-based Real-time Location and Map Reconstruction(V-SLAM)[J]. New Technology, 2019:67-70.
- [33] DAVISON A J, REID I D, MOLTON N D, et al. Mono-SLAM: real-time single camera SLAM[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(6):1052-1067.
- [34] KLEIN G, MURRAY D. Parallel Tracking and Mapping for Small AR Workspaces[C]//IEEE and ACM International Symposium on Mixed and Augmented Reality. 2007:225-234.
- [35] MUR-ARTAL R, MONTIEL J M M, TARDOS J D. ORB-SLAM: A Versatile and Accurate Monocular SLAM System[J]. IEEE Transactions on Robotics, 2015, 31(5):1147-1163.
- [36] LIU H M, ZHANG G F, BAO H J. Overview of simultaneous positioning and map construction methods based on monocular vision [J]. Journal of Computer Aided Design and Graphics, 2016, 28(6):855-868.
- [37] MUR-ARTAL R, TARDOS J D. ORB-SLAM2: An Open-Source Slam System for Monocular, Stereo and RGB-D Cameras [J]. IEEE Transactions on Robotics, 2017, 33(5):1255-1262.
- [38] ZHANG Y, LI H S, MA L, et al. Research on real-time positioning of underwater robot based on ORB-SLAM2 algorithm[J]. Bulletin of Surveying and Mapping, 2019(12):1-7.
- [39] LU J W, WANG Y L. Construction of real-time grid map based on ORB-SLAM2[J]. Application Research of Computers, 2019, 36(10):3124-3131.
- [40] HADDADI S J, CASTELAN E B. Visual-Inertial Fusion for Indoor Autonomous Navigation of a Quadrotor Using ORB-SLAM [C]//2018 Latin American Robotic Symposium, 2018 Brazilian Symposium on Robotics(SBR) and 2018 Workshop on Robotics in Education(WRE). 2018.
- [41] SANTOS M, SOTO O, BUSTAMANTE V P. ORB-SLAM Based Active Disturbance Rejection Control for Quadrotor Autonomous Flight[C]//2018 XX Congreso Mexicano de Robótica (COMRob). 2018.
- [42] FORSTER C, PIZZOLI M, SCARAMUZZA D. SVO: Fast semi-direct monocular visual odometry[C]//IEEE International Conference on Robotics and Automation. IEEE, 2014:15-22.
- [43] XIAO Y, RUAN X G, ZHU X Q, et al. A Micro-UAV Monocular Vision SVO/INS Integrated Navigation Method[J]. Journal of Chinese Inertial Technology, 2019, 27(2):211-219.
- [44] MENG Y, WANG W, HAN H, et al. A visual / inertial integrated landing guidance method for UAV landing on the ship[J]. Aerospace Science and Technology, 2019, 85:474-480.
- [45] FORSTER C, GASSNER M, WERLBERGER M, et al. SVO: Semi-Direct Visual Odometry for Monocular and Multi-Camera Systems[J]. IEEE Transactions on Robotics and Automation, 2017, 33(2):249-265.
- [46] HUANG T. Augmented reality system based on monocular visual odometer [D]. Guangzhou: South China University of Technology, 2018.
- [47] LI X N, XU Z Y, ZHANG Y H. Determination of weight factors

- for multi-sensor data fusion[J]. *Journal of Spacecraft TT&C Technology*, 2005, 24(1): 65-67.
- [48] AYABAKAN T, KERESTECIOGLU F. Indoor positioning using federated Kalman Filter[C]// 26th Signal Processing and Communications Applications Conference(SIU). 2018.
- [49] KAYA S B, ALKAR A Z. Indoor localization and tracking by multi sensor fusion in Kalman filter[C]// 26th Signal Processing and Communications Applications Conference(SIU). 2018.
- [50] CHEN X L, WANG J L, SUN L, et al. A Method of Data Fusion for Multi-sensor Based on Bayesian Estimation[J]. *Journal of Electronic Measurement and Instrument*, 2009.
- [51] ZHANG P, DONG W H, GAO D D. An Optimal Method of Data Fusion for Multi-Sensors Based on Bayesian Estimation[J]. *Chinese Journal of Sensors and Actuators*, 2014, 27(5): 643-648.
- [52] KE L H. Research and Implementation of SLAM Based on Deep Learning[D]. Xi'an: Xi'an Polytechnic University, 2019.
- [53] LI S P, ZHANG T. Overview of the application of deep learning in visual SLAM[J]. *Space Control Technology and Application*, 2019, 45(2): 1-10.
- [54] KONDA K, MEMISEVIC R. Learning visual odometry with a convolutional network[C]// Proceedings of the 10th International Conference on Computer Vision Theory and Applications. Lisbon, Portugal; SCITCC Press, 2015: 486-490.
- [55] COSTANTE G, MANCINI M, VALIGI P, et al. Exploring representation learning with CNNs for frame-to-frame ego-motion estimation[J]. *IEEE Robotics and Automation Letters*, 2016, 1(1): 18-25.
- [56] HE Y L, CHEN J T, ZENG B. Fast closed loop detection method based on simplification convolutional neural network[J]. *Computer Engineering*, 2018, 44(6): 182-187.
- [57] YI H, HONG Z, ZHOU S. BoCNF: Efficient image matching with bag of Conv-Net features for scalable and robust visual place recognition[J]. *Autonomous Robots (S0929-5593)*, 2017, 42(9): 1-17.
- [58] HOU Y, ZHANG H, ZHOU S L. Convolutional neural network based on image representation for visual loop closure detection [C]// IEEE International Conference on Information and Automation. Piscataway, USA; IEEE, 2015: 2238-2245.
- [59] SUNDERHAUF N, SHIRAZI S, DAYOUB F, et al. On the performance of Conv-Net features for place recognition[C]// IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, USA; IEEE, 2015: 4297-4304.
- [60] MERRILL N, HUANG G. CALC2.0: Combining Appearance, Semantic and Geometric Information for Robust and Efficient Visual Loop Closure[C]// 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS). 2019.
- [61] LUO S X, ZHANG S J. Research on Closed Loop Detection Algorithm Based on Deep Learning[J]. *Computer and Digital Engineering*, 2019, 47(3): 494-502.
- [62] GAO X, ZHANG T. Unsupervised learning to detect loops using deep neural networks for visual SLAM system[J]. *Autonomous Robots*, 2017, 41(1): 1-18.
- [63] VASUDEVAN S, GACHTER S, NGUYEN V, et al. Cognitive maps for mobile robots-an object based approach[J]. *Robotics and Automatic Systems*, 2007, 55(5): 359-371.
- [64] YAO E L, ZHANG H X, SONG H T, et al. Robust SLAM algorithm based on semantic information and edge consistency[J]. *Robot*, 2019, 41(6): 751-760.
- [65] SUNDERHAUF N, PHAM T T, LATIF Y, et al. Meaningful Maps with Object-Oriented Semantic Mapping [C] // 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS). 2017.
- [66] MA L, STUCKLER J, KARL C, et al. Multi-View Deep Learning for Consistent Semantic Mapping with RGB-D Cameras [C]// 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS). 2017.



TIAN Ye, born in 1996, postgraduate. His main research direction include robot and SLAM algorithm.



CHEN Hong-wei, born in 1981, Ph. D, lecturer. His main research direction include robot and SLAM algorithm, etc.