


基于视觉的多机器人协作SLAM问题

吕强, 林辉灿, 张洋, 马建业

装甲兵工程学院控制工程系, 北京 100072

 视觉SLAM仅采用图像作为外部信息,用于估计机器人位置的同时构建环境地图。SLAM是机器人自主性的基本前提,如今在小动态环境采用激光或者声呐传感器构建2D地图得到较好地解决。然而动态、复杂和大范围下的SLAM仍存在问题,使用视觉作为基本的外部传感器是解决问题的一个新颖热门的研究方法。在视觉SLAM中使用计算机视觉技术,如特征检测、特征描述和特征匹配,图像识别和恢复,还存在很多改善的空间。本文在视觉SLAM领域的最新技术的基础上,对基于视觉的多机器人协作SLAM领域的前沿技术进行综述。

移动机器人的自主导航问题可以分为定位、地图构造和路径规划3个部分^[1]。定位的任务是确定机器人在当前环境中的位姿,地图构建的任务是将对环境的零碎分布观测整合为一个联系的模型,路径规划的任务是确定最佳路径并导航机器人通过环境。

起初定位和地图构建是分开研究的,后来学者们意识到二者是相互关联的。同时定位与地图构建(simultaneous localization and mapping, SLAM)是指机器人在未知环境中运动时增量地构建周围环境地图,同时运用此地图对机器人自身位姿进行估计^[2]。SLAM是移动机器人系统最基本、最重要的一项能力,随着未知环境探测的日趋复杂和非结构化,SLAM已成为自主移动机器人亟待解决的重要基础性难题^[3,4],并被视为移动机器人领域的“圣杯”^[5]。当摄像头作为唯一的感知外部环境的传感器用于SLAM,就称为视觉SLAM。准确的环境地图有助于实现高精度的定位,而精确定位则有利于建立更加真实可信的环境地图^[6],两者相辅相成。

许多视觉SLAM系统在如下的环境中容易失败:外部环境、动态环境、显著特征太多或太少的环境、大尺度环境,以及当摄像头运动不稳和部分或全

部的闭锁现象。视觉SLAM系统成功的关键是具备处理上述困难的能力。近10年来,从发表的论文可以看出一个清晰的趋势,就是使用视觉作为唯一的外部传感器的感知系统用于解决SLAM问题^[7-11]。之所以会有这样的趋势,是因为基于摄像头的系统既能够获取尺度信息又能够恢复环境的外观、颜色和纹理,使得机器人能够完成更高级的功能,如在环境中识别人员。进一步说,摄像头更便宜、更轻且功耗更低。不幸的是相机数据存在误差,由于相机分辨率不足、光线变化、表面缺乏纹理和快速运动引起图像模糊等因素。许多学者对上述问题进行研究并取得不少成果,比较典型的有Klein等基于BA(bundle adjustment)的实时SLAM算法,即并行跟踪和构图(parallel tracking and mapping, PTAM)^[9], Engel等采用直接的方法构建大范围内的半稠密地图,即LSD-SLAM^[12],以及RGB-D SLAM^[13]和ORB SLAM^[14]等。RGB-D SLAM采用一种能够同时获取环境彩色信息(RGB)和深度信息(depth)的RGB-D传感器。

对于单个机器人的SLAM系统已经有不少的解决方案,然而多机器人的SLAM相较而言还是一个新领域,多机

器人SLAM系统可以提升构图的速度和精度,有利于路径规划等探索研究。该领域的核心问题是多机器人之间的通信拓扑、任务规划和地图融合。任务规划:当多个机器人用于共同探索同一个环境进行研究时,需要处理好各个机器人的任务规划问题。做好任务规划可以节省探索整个研究环境的时间。通信拓扑是指机器人之间的通信结构,有完全连接和部分连接等多种通信拓扑结构。对于完全连接需要使用外部设备以克服通信带宽和距离的限制,外部设备用于将一个机器人的数据发送给所有的机器人^[15]。部分连接采用两个机器人之间的一对一通信解决通信带宽和距离的限制问题,机器人仅能使用通信范围内的其他机器人的信息。地图融合是解决多机器人SLAM的关键问题,运用共享的信息改进全局地图的精度和构图范围,全局地图通过联合所有机器人的局部地图获得。当机器人之间的相对位置关系已知,只需要根据每个机器人的定位和构图进行拼接地图;而当机器人之间的相对位置未知时,需要对每个机器人的地图进行转换才能进行拼接地图。

总之,虽然需要处理上述难题,但是多机器人SLAM能够以更快的速度、



更高的精度探索未知环境。而且,多机器人SLAM具有更强的容错能力,环境条件的恶劣或者变化、由于错误的关联或假设模型引起的算法失败等问题,可以通过多机器人系统克服。因此,近几年相对单个机器人的SLAM应用研究,多机器人的SLAM应用研究更加吸引人和热门。

1 视觉SLAM

1.1 同时定位和地图构建

在1985—1990期间,Chatila和Smith提出地图构建和定位同时进行研究的问题。不久之后,该问题有了专业名称SLAM(simultaneous localization and mapping)。有关SLAM的历史可以参阅Durrant和Bailey撰写的有关SLAM发展历程的文章^[16,17]。SLAM指移动平台(机器人、运输载体等)能够构建所探索环境的全局地图,同时利用地图进一步确定自身位置。

为了对环境构建地图,机器人必须拥有用于感知和获取外部环境观测的传感器,可分为外部传感器和内部传感器两种。外部传感器包括声呐、激光扫描仪、相机以及GPS等,这些传感器都存在噪声和距离受限的问题。此外,前3种传感器只能获得环境的局部视图。激光传感器和声呐传感器可以获取精确而稠密的环境结构信息,然而,存在的问题是:在复杂性高的环境以及物体识别方面均不可用,质量大、体积大使得难以在飞行机器人或者人型机器人上应用。另一方面,GPS在狭窄街道、城市峡谷、水下和室内等环境的定位精度不能满足要求。内部传感器能够测量机器人的速度、位置变化和加速度,例如编码器、加速度计和磁罗盘等,但由于存在累积误差,并不能保证足够的估计精度。

在文献[18]~[20]中得到证实,为了获得机器人位置的高精度且鲁棒性好的估计,需要使用信息融合技术处理多个传感器的数据。然而,传感器增加引起的系统的造价过高、重量增加和功耗过大等问题,因此,探索如何仅仅依靠相机作为传感器实现同时定位和地图

构建就显得尤为重要。

1.2 相机在SLAM中的应用

第一个视觉导航的作品采用双目立体结构相机^[21,22],然而,多数情况下由于价格昂贵,很难获取双目或者三目立体相机。另一种解决方法是使用一对单目相机,导致需要考虑不懂方面的困难:采用硬件或软件的方式实现相机同步;每个CCD传感器对颜色或光线的响应略有不同;采取何种几何结构进行校准等。

不管采用什么类型的相机,在使用之前都必须校准(手动或者自动),校准需要估计相机的内部参数和外部参数,前者依赖于相机的几何结构(焦距和光点),后者由相机在空间中的位置决定(某个坐标系中的旋转和位移)。根据多视角棋盘校准图像,关联图像坐标与真实世界坐标可获取参数^[23],也有许多工具能够实现相机的校准:OpenCV的校准函数、Matlab的相机校准工具箱等。

使用多相机带有或者不带有视角重叠的研究^[24,25],使用特殊镜头的相机如广角^[26]或者全向^[27]相机增加视觉范围,进而一定程度上减小累积误差。近几年出现的RGB-D传感器应用于室内环境的地图构建,已被验证是一种非常有前途的方法^[28,29]。

德国弗莱堡大学的Engelhard和慕尼黑技术大学的Sturm等^[30]提出了一种基于RGB-D传感器的实时3D视觉SLAM方法,主要包括特征提取和匹配、位姿估计、位姿修正、位姿图优化4个步骤。该方法对Henry等提出的方法进行了3点改进:首先在特征提取时将SIFT特征替换为SURF特征,其次在位姿图优化时使用g2o方案代替TORO,最后在环境地图的表示时将点云地图转换成3D占据栅格地图。文献[31]中进一步对SIFT、SURF和ORB^[32]3种不同的特征描述方法进行了研究和实验分析对比,而且针对着色点云地图存储量大和计算量大的问题,使用基于Octree的地图构建框架OctoMap将点云地图体素化转换成3D占据栅格地图。此外,他们还首次构建了一个包含

RGB-D图像序列和真实摄像机轨迹的数据集用于评估SLAM系统的性能^[33]。基于该数据集进行了大量的实验,结果表明其位姿估计平均精度为9.7 cm和3.95°,平均每帧的处理时间为0.35 s,基本满足实时运行的要求。

自从单目SLAM的出现^[34],仅用一个相机实现同时定位和地图构建的方法就变得越来越流行。这是因为通过手机、数码相机或笔记本等可以轻松获得相机,比立体相机更加容易获取。单目方法的硬件简单灵活经济,数据处理时间也较短。仅依靠单目并不能直接获得足够的信息计算航标的深度,因此需要处理航标的初始化问题,通常有两种解决方案^[35,36]。为了从一个相机获取环境的三维信息,需要处理多个观测图像的特征跟踪问题。

2 图像数据关联

最先进的RGB-D SLAM系统大多数都采用迭代最近点(iterative closest point, ICP)算法去匹配点特征,这些点特征的空间位置是根据匹配的传感器深度数据计算得到的。然而,特征的深度测量数据经常被噪声所干扰,因为视觉特征倾向于在实物的边缘。为了减小估计误差,高翔等^[37]采用了可靠的深度数值提取和选择特征的方法,如平面点特征。平面点特征有利于提高传统的ICP的精度和鲁棒性,同时保持合理的实时应用的计算消耗。通过公开数据集的轨迹跟踪和构图结果以及实物机器人的实际试验,验证了一个基于平面的RGB-D SLAM的系统的有效性。

在单目视觉SLAM中,绝大多数采用基于特征的地图表示,因此此处主要对基于特征的运动估计方法进行分析。对该问题的研究,是计算机视觉领域中经典的从运动恢复结构(SFM)问题,也有学者将其等同于视觉里程计研究。基于2D图像特征的运动估计一般包括特征检测、特征匹配和运动估计。

2.1 特征提取

特征提取对于视觉SLAM是最为关键也最为基础的一个环节,因为后续的所有处理都以此为基础进行。SIFT^[38]

方法在 10 多年的应用中验证其良好的性能,成功应用于视觉特征、目标识别、图像匹配和视觉构图等方面^[32]。但由于采用 128 维向量,相较而言在计算和匹配过程中运算速度较慢,不利于在实时性要求很高的 SLAM 系统中的应用。SURF^[39]能够在保持与 SIFT 相同性能的辨别率的情况下,提高运算和匹配的效率,然而该描述符使用了 64 维向量的浮点数据即 256 位,当需要存储数以百万计的描述符时也是不能满足要求的。BRIEF^[40]使用二进制流作为关键点特征的描述符,该方法使用较少的位数就能够获得高性能且计算简单,文献[40]验证该方法在获得与 SURF 相当或者更好性能的情况下,需要的时间更少。ORB^[32]是一种基于 BRIEF 的快速二进制描述符,该方法具有旋转不变性和抗干扰性能,更加适用于视觉 SLAM 系统对实时性和精度的要求。ORB-SLAM^[41]是一个满足大、小范围,室内外环境的实时单目 SLAM 系统,且可以完全的自动初始化,采用适者生存的方法选择重构点和关键帧,生成稠密地图,该地图仅当场景内容改变时才更新,加强了终生地图构建的能力。

2.2 特征匹配

特征匹配的性能对 SLAM 效果具有决定性的影响,为快速获得高精度的匹配,主要有两类算法^[41]:一是灰度匹配,使用统计模型寻找点匹配;二是基于特征的方法,该方法基于从图像检测得到的特征(点、线或面)。两种方法各有千秋,灰度匹配的方法具有更高的精度但同时消耗更多的计算资源,目前来看并不满足视觉 SLAM 的实时性要求,另一方面,基于特征的方法需要的计算资源较少且抗干扰性能好。基于特征的方法需要先从图像中提取特征,在上一节中已经介绍过方法。基于特征的方法计算两个帧间的特征匹配,有两种思路:一是先提取前一帧图像的特征,然后在局部范围内的图像检索寻找对应点^[42],该方法具有的优点是算法速度快,局限性是仅在两帧图像变化较小的时候适用;二是对两幅图像都进行特征提取,在根据提取的特征建立联系,该

方法的优点是能够适用于帧间的变化较大的情况,缺点是计算效率较低及存在匹配不一致。对于匹配不一致的问题可以采用文献[38]和[43]的方法提高匹配一致性。

2.3 运动估计

Structure from Motion(SFM)技术能够从一系列图像中计算场景中的 3D 结构和相机位置^[44]。标准的流程是提取输入图像的特征,进而特征匹配,采用一种称为捆绑调整(bundle adjustment, BA)的非线性优化方法,最小化重投影误差^[45]。解决 SLM 递增问题的一种方法是视觉里程计,该方法仅使用序列图像就可以实时计算相机的位姿。Klein 等发表的称为 PTAM 的单目方法,基于关键帧,采用两个平行的处理线程,第一个线程执行的任务是鲁棒地跟踪很多特征,另一线程的任务是生成 3D 地图点。该方法在动态环境中存在相似纹理环境中出现跟踪失败^[9]。

Strasdat 等^[46]证实为了提高单目 SLAM 系统的位置精度,增加特征的数量比增加帧的方式更为有效,同时,BA 优化的方法比滤波器的方法效果更好,然而,该文章指出滤波器的方法在较高不确定的情况下更有利。理想的 SLAM 系统将同时利用 SFM 技术和概率滤波器的优点^[47]。

3 多机器人协作 SLAM

多机器人方案相对单个机器人方案有许多优点,例如,随着机器人数量的增加,整个系统执行任务需要的时间缩短,对错误的容忍度增加且地图的精度得到提高。Marjovi 等^[48]对多机器人在时间上的优势做了阐述,在不同环境中使用不同数量的机器人,以获得满意的构图结果。两个机器人的系统用于构建地图,需要的时间是单个机器人的 1/2,尽管存在一些优势,但是多机器人 SLAM 系统需要探索由多个机器人带来的问题,例如在线路径规划、同时定位和地图构建、特征提取和数据关联。然而,这些问题在单个机器人 SLAM 系统已经得到研究并取得一定的成果,此外,任务分配、通信拓扑和地图融合是

多机器人系统需要解决的重要难题。有关通信拓扑方面的综述,可以参阅文献[49],本文主要关注多机器人系统的地图融合问题。

多机器人系统中,每个机器人在进行 SLAM 的过程中都有自己的坐标系,因此由不同机器人创建的地图不能够直接用于融合。学者们已经研究出一些实现地图融合的方法,通常情况下,在获取两类必备条件的情况下就能够实现地图融合。一是机器人之间互相清楚对方的相对位置,如当机器人相遇时使用全向相机获得机器人之间的距离,进而计算地图的转换矩阵^[50]。二是机器人之间不清楚互相之间的位置,但是各自构建的地图之间有公共的区域,文献[51]匹配地图的点特征。

Gil 等^[52]则使用 RBPF 滤波对地图进行融合。Vidal-Calleja 等^[53]使用一个飞行器与一个地面机器人进行合作,在半结构化的地形中构建地图。Benedettelli 等^[54]让各个机器人独立进行 SLAM,生成独立的地图。在机器人相遇时,计算地图的相似处,进行合并生成全局地图,然后,每个机器人通过开发融合的地图开始单个机器人的 SLAM 算法。Michael 等^[55]对地震中毁坏的建筑里采用地面和空中机器人进行协作建图进行研究,通过采集多层结构的环境构建 3D 地图。他们提供了协作建图方法的细节,并且报告了试验中分别由单个机器人和机器人团队构建的地图。

Forster 等^[56]通过微型飞行器的单目相机对准由地面机器人的深度传感器构建的地图,得到稠密的 3D 重构地图,如图 1 所示。他们提出一种用于微型飞行器联合地面机器人的定位方法,解决了通过多个机器人上的不同传感器计算得到的 3D 地图的配准问题,由微型飞行器上的单目相机得到的稠密的 3D 重构与地面机器人上的深度传感器得到的地图配准。一旦配准,微型飞行器上的稠密重构将用于扩充由地面机器人计算得到的地图,使用航空视角所传递的信息进行扩充。整体的方法是新颖的,由移动相机得到的现场稠密



重构,解决了空地联合定位问题。

Saeedi 等^[57]使用神经网络拼合地图。他们先从网格地图中提取特征,根据特征计算两个地图间的旋转与平移。对于增强的自主性的机器人来说,在没有 GPS 信号的环境中导航是一项非常重要的要求。虽然在一些单个机器人的场景中已经得到解决,但是下一步的挑战是为多个机器人设计有效的构图和导航算法。实现编队自主性的关键需求是提供给机器人协作去精确的构建环境地图的能力,这个问题和协作的同时定位和地图构建密切相关,他们研究的构图过程扩展到多机器人新的占有栅格地图融合算法。地图融合通过转换若干单个地图到 Hough 空间得到, Hough 空间里地图用一种抽象的形式描述。Hough 变换的特性用于找到各个地图之间的公共区域,进而用于计算地图之间的未知变换。他们亦提出使用 Hough 峰来拼合地图的算法^[58],融合的结果如图 2 所示,其中图 2(a), (b)和(c)分别为子地图, (d)为融合得到的全局地图。

4 发展趋势

机器人在进入未知环境时,需要对所处环境进行有效地探测和感知,获取其在环境中的姿态信息,同时还要构建出环境地图。只有在了解环境真实信息的基础上才能进行避障、路径规划并

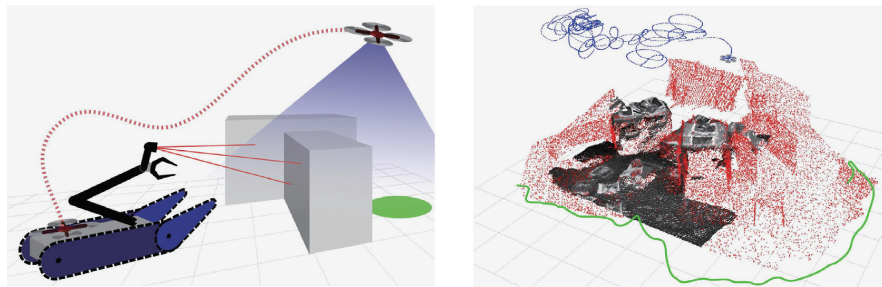


图1 空中和地面机器人的定位和构图

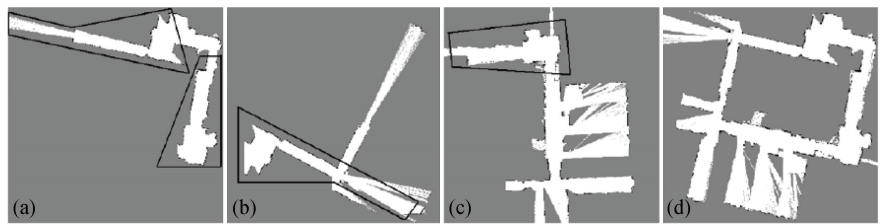


图2 3个子地图融合成1个全局地图

完成相应的目标任务。

1) 联合空中多旋翼飞行器和地面机器人,搭载 RGB-D、单目或多目等不同的视觉传感器,充分发挥各自的优点获取更多的环境信息,构建适用于室内外复杂动态环境的多机器人视觉 SLAM 系统。

2) 在多机器人系统的地图融合工作中,在子地图的重叠部分可在优化地图精度方面进一步研究,可用于闭环检测提高整体系统性能。另外,对部分机器人构图失败的情况,如何提高系统的

容错能力也值得进一步研究。

3) 深度学习技术能够让机器自主理解图像中存在的特征,并以此为基础去辨别图像的相似性,完成闭环检测的任务。然而深度学习技术在 SLAM 研究领域,远没有得到应用,仅仅在一些物体识别中得到应用^[59,60]。下一步可以基于视觉的 SLAM 系统中的大量图像信息,应用深度学习知识在闭环检测以及目标识别等高级别任务中。

参考文献 (References)

- [1] Stachniss C. Robotic mapping and exploration[M]. Springer Tracts in Advanced Robotics, 2009.
- [2] Durrant-Whyte H, Bailey T. Simultaneous localization and mapping: part I[J]. Robotics & Automation Magazine, IEEE, 2006, 13(2): 99-110.
- [3] 陈卫东, 张飞. 移动机器人的同步自定位与地图创建研究进展[J]. 控制理论与应用, 2005(3): 455-460.
- [4] 张亮. 移动机器人同步定位与地图重建算法研究[D]. 杭州: 浙江大学, 2009.
- [5] Bailey T, Durrant-Whyte H. Simultaneous localization and mapping (SLAM): Part II[J]. Robotics & Automation Magazine, IEEE, 2006, 13(3): 108-117.
- [6] 于金霞, 王璐, 蔡自兴. 未知环境中移动机器人自定位技术[M]. 北京: 电子工业出版社, 2011.
- [7] Paz L M, Piniés P, Tardós J D, et al. Large-scale 6-DOF SLAM with stereo-in-hand[J]. Robotics, IEEE Transactions on, 2008, 24(5): 946-957.
- [8] Davison A J, Reid I D, Molton N D, et al. MonoSLAM: Real-time single camera SLAM[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2007, 29(6): 1052-1067.
- [9] Klein G, Murray D. Parallel tracking and mapping for small AR workspaces[C]//Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on. IEEE, 2007: 225-234.
- [10] Sáez J M, Escolano F. 6dof entropy minimization slam[C]//Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on. IEEE, 2006: 1548-1555.
- [11] Piniés P, Tardós J D. Large-scale slam building conditionally independent local maps: Application to monocular vision[J]. Robotics, IEEE Transactions on, 2008, 24(5): 1094-1106.

- [12] Engel J, Schöps T, Cremers D. LSD-SLAM: Large-scale direct monocular SLAM[M]//Computer Vision-ECCV 2014. Springer International Publishing, 2014: 834-849.
- [13] Endres F, Hess J, Sturm J, et al. 3-d mapping with an rgb-d camera[J]. Robotics, IEEE Transactions on, 2014, 30(1): 177-187.
- [14] Mur-Artal R, Montiel J M M, Tardos J D. ORB-SLAM: A versatile and accurate monocular SLAM system[J]. arXiv preprint arXiv:1502.00956, 2015.
- [15] Ferranti E, Trigoni N, Levene M. Brick& Mortar: an on-line multi-agent exploration algorithm[C]//Robotics and Automation, 2007 IEEE International Conference on. IEEE, 2007: 761-767.
- [16] Whyte H, Bailey T. Simultaneous Localization and Mapping (SLAM) Part I The Essential Algorithms[J]. IEEE Robotics & Automation Magazine, 2006, 13(2): 99-110.
- [17] Bailey T, Durrant-Whyte H. Simultaneous localization and mapping (SLAM): Part II[J]. IEEE Robotics & Automation Magazine, 2006, 13(3): 108-117.
- [18] Castellanos J A, Neira J, Tardós J D. Multisensor fusion for simultaneous localization and map building[J]. Robotics and Automation, IEEE Transactions on, 2001, 17(6): 908-914.
- [19] Majumder S, Scheding S, Durrant-Whyte H F. Sensor fusion and map building for underwater navigation[C]// Proceedings of Australian Conference on Robotics and Automation. 2000: 25-30.
- [20] Nützi G, Weiss S, Scaramuzza D, et al. Fusion of IMU and vision for absolute scale estimation in monocular SLAM[J]. Journal of intelligent & robotic systems, 2011, 61(1-4): 287-299.
- [21] Se S, Lowe D, Little J. Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks[J]. The International Journal of Robotics Research, 2002, 21(8): 735-758.
- [22] Olson C F, Matthies L H, Schoppers M, et al. Rover navigation using stereo ego-motion[J]. Robotics and Autonomous Systems, 2003, 43(4): 215-229.
- [23] Hartley R, Zisserman A. Multiple view geometry in computer vision[M]. Cambridge: Cambridge University Press, 2003.
- [24] Kaess M, Dellaert F. Probabilistic structure matching for visual SLAM with a multi-camera rig[J]. Computer Vision and Image Understanding, 2010, 114(2): 286-296.
- [25] Carrera G, Angeli A, Davison A J. SLAM-based automatic extrinsic calibration of a multi-camera rig[C]//Robotics and Automation (ICRA), 2011 IEEE International Conference on. IEEE, 2011: 2652-2659.
- [26] Davison A J, Cid Y G, Kita N. Real-time 3D SLAM with wide-angle vision[C]//Proc. IFAC/EURON Symp. Intelligent Autonomous Vehicles. 2004: 31-33.
- [27] Scaramuzza D, Siegwart R. Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles[J]. Robotics, IEEE Transactions on, 2008, 24(5): 1015-1026.
- [28] Huang A S, Bachrach A, Henry P, et al. Visual odometry and mapping for autonomous flight using an RGB-D camera[C]//International Symposium on Robotics Research (ISRR). 2011: 1-16.
- [29] Hu G, Huang S, Zhao L, et al. A robust rgb-d slam algorithm[C]//Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on. IEEE, 2012: 1714-1719.
- [30] Engelhard N, Endres F, Hess J, et al. Real-time 3D visual SLAM with a hand-held RGB-D camera[C/OL]. [2015-09-31]. http://vision.informatik.tu-muenchen.de/_media/spezial/bib/engelhard11euron.pdf.
- [31] Endres F, Hess J, Engelhard N, et al. An evaluation of the RGB-D SLAM system[C]//Robotics and Automation (ICRA), 2012 IEEE International Conference on. IEEE, 2012: 1691-1696.
- [32] Rublee E, Rabaud V, Konolige K, et al. ORB: An efficient alternative to SIFT or SURF[C]//Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011: 2564-2571.
- [33] Sturm J, Magnenat S, Engelhard N, et al. Towards a benchmark for RGB-D SLAM evaluation[C]. Proc. of the RGB-D Workshop on Advanced Reasoning with Depth Cameras at Robotics: Science and Systems Conf.(RSS), Los Angeles, USA. 2011, 2: 3.
- [34] Davison A J. Real-time simultaneous localisation and mapping with a single camera[C]//Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on. IEEE, 2003: 1403-1410.
- [35] Lemaire T, Berger C, Jung I K, et al. Vision-based slam: Stereo and monocular approaches[J]. International Journal of Computer Vision, 2007, 74(3): 343-364.
- [36] Vidal-Calleja T, Bryson M, Sukkarieh S, et al. On the observability of bearing-only SLAM[C]//Robotics and Automation, 2007 IEEE International Conference on. IEEE, 2007: 4114-4119.
- [37] Gao X, Zhang T. Robust RGB-D simultaneous localization and mapping using planar point features[J]. Robotics & Autonomous Systems, 2015, 72:1-14.
- [38] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [39] Bay H, Tuytelaars T, Gool L V. SURF: Speeded up robust features[J]. Computer Vision & Image Understanding, 2006, 110(3): 404-417.
- [40] Lepetit V, Ozuysal M, Trzcinski T, et al. BRIEF: Computing a local binary descriptor very fast[J]. IEEE Transactions on Software Engineering, 2011, 34(7): 1281-1298.
- [41] Huijuan Z, Qiong H. Fast image matching based-on improved SURF algorithm[C]//Electronics, Communications and Control (ICECC), 2011 International Conference on. IEEE, 2011: 1460-1463.
- [42] Chli M, Davison A J. Active matching for visual tracking[J]. Robotics and Autonomous Systems, 2009, 57(12): 1173-1187.
- [43] Fischler M A, Bolles R C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography[J]. Communications of the ACM, 1981, 24(6): 381-395.
- [44] Pollefeys M, Van Gool L, Vergauwen M, et al. Visual modeling with a hand-held camera[J]. International Journal of Computer Vision, 2004, 59(3): 207-232.



- [45] Engels C, Stewénus H, Nistér D. Bundle adjustment rules[J]. *Photogrammetric computer vision*, 2006, 2: 124–131.
- [46] Strasdat H, Montiel J M M, Davison A J. Real-time monocular slam: Why filter?[C]//*Robotics and Automation (ICRA)*, 2010 IEEE International Conference on. IEEE, 2010: 2657–2664.
- [47] Fuentes–Pacheco J, Ruiz–Ascencio J, Rendón–Mancha J M. Visual simultaneous localization and mapping: a survey[J]. *Artificial Intelligence Review*, 2015, 43(1): 55–81.
- [48] Marjovi A, Nunes J G, Marques L, et al. Multi-Robot Exploration and Fire Searching[C]// *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2009:1929–1934.
- [49] 张国良, 汤文俊, 曾静, 等. 考虑通信状况的多机器人CSLAM问题综述[J]. *自动化学报*, 2014, 40(10): 2073–2088.
- [50] Michael N, Shen S, Mohta K, et al. Collaborative mapping of an earthquake-damaged building via ground and aerial robots[J]. *Journal of Field Robotics*, 2012, 29(5): 832–841.
- [51] Lee H C, Lee S H, Choi M H, et al. Probabilistic map merging for multi-robot RBPF-SLAM with unknown initial poses[J]. *Robotica*, 2012, 30(2): 205–220.
- [52] Gil A, Reinoso O, Ballesta M, et al. Multi-robot visual slam using a rao-blackwellized particle filter. *Robotics and Autonomous Systems*, 2010, 58(1): 68–80.
- [53] Vidal–Calleja T A, Berger C, Sola J, et al. Large scale multiple robot visual mapping with heterogeneous landmarks in semi-structured terrain. *Robotics and Autonomous Systems*, 2011, 59(9): 654–674.
- [54] Benedettelli D, Garulli A, Giannitrapani A. Cooperative SLAM using M-Space representation of linear features[J]. *Robotics and Autonomous Systems*, 2012, 60(10): 1267–1278.
- [55] Zhou X S, Roumeliotis S. Multi-robot SLAM with unknown initial correspondence: The robot rendezvous case[C]//*Intelligent Robots and Systems*, 2006 IEEE/RSJ International Conference on. IEEE, 2006: 1785–1792.
- [56] Forster C, Pizzoli M, Scaramuzza D. Air-ground localization and map augmentation using monocular dense reconstruction[C]//*Intelligent Robots and Systems (IROS)*, 2013 IEEE/RSJ International Conference on. IEEE, 2013: 3971–3978.
- [57] Saeedi S, Paull L, Trentini M, et al. Neural network-based multiple robot simultaneous localization and mapping[J]. *IEEE Transactions on Neural Networks*, 2011, 22(12): 2376–2387.
- [58] Saeedi S, Paull L, Trentini M, et al. Map merging for multiple robots using Hough peak matching[J]. *Robotics and Autonomous Systems*, 2014, 62(10): 1408–1424.
- [59] Kostavelis I, Gasteratos A. Learning spatially semantic representations for cognitive robot navigation[J]. *Robotics and Autonomous Systems*, 2013, 61(12): 1460–1475.
- [60] Bo L, Ren X, Fox D. Learning hierarchical sparse features for rgb-d object recognition[J]. *International Journal of Robotics Research*, 2014, 33(4): 581–599.

A survey on the cooperative SLAM problem of multi-robots systems based on visual

LÜ Qiang, LIN Huican, ZHANG Yang, MA Jianye

Department of Control Engineering, Academy of Armored Forces Engineering, Beijing 100072, China

Abstract Visual SLAM using only images as external information estimates the robot position while building the environment map. SLAM is a basic prerequisite for autonomous robots. Now it has been solved by using a laser or sonar sensor to build 2D map in a small dynamic environment. However, in a dynamic, wide range and complex environment there are still problems to be solved, and the use of vision as the basic external sensor is a new area of research. The use of computer vision techniques in visual SLAM, such as feature detection, characterization, feature matching, image recognition and recovery, has still much room for improvement. The paper offers a brief overview on visual SLAM about the latest and easy to understand technologies in the field. Multi-robot systems have many advantages over a single robot, which can improve the precision of SLAM system, and better adapt to the dynamic and complex environment. This paper expounds the methods of multi-robot SLAM, with emphasis on the map fusion methods.

Keywords multi-robots; SLAM ; computer vision; map merging

作者简介:吕强,研究方向为自主移动机器人,电子信箱:rokyou@live.cn

(责任编辑 刘志远)