硕士学位论文

野外环境下两栖机器人的自主视觉定位算法 研究

RESEARCH ON AUTONOMOUS VISUAL LOCALIZATION ALGORITHM OF AMPHIBIOUS ROBOT IN FIELD ENVIRONMENT

宋瑾

哈尔滨工业大学 2020年1月 国内图书分类号: TP242 国际图书分类号: 621

工学硕士学位论文

野外环境下两栖机器人的自主视觉定位算法 研究

- 硕士研究生:宋瑾
- **导** 师:陈浩耀 副教授
- 申请学位:工学硕士
- **学**科:控制科学与工程
- 所在单位:哈尔滨工业大学(深圳)
- 答辩日期: 2019年12月
- **授予学位单位**:哈尔滨工业大学

Classified Index: TP242 U.D.C: 621

A dissertation submitted in partial fulfillment of the requirements for the academic degree of Master of Engineering

RESEARCH ON AUTONOMOUS VISUAL LOCALIZATION ALGORITHM OF AMPHIBIOUS ROBOT IN FIELD ENVIRONMENT

Candidate:	Song Jin	
Supervisor:	Asso.Prof. Chen Haoyao	
Academic Degree Applied for:	Master of Engineering	
Specialty:	Control Science and Engineering	
Affiliation:	Harbin Institute of Technology, Shenzhen	
Date of Defence:	December, 2019	
Degree-Conferring-Institution: Harbin Institute of Technology		

摘要

两栖机器人可以直接为具有两栖作业需求的水产、勘探及救灾救险等产业服务,是提高我国两栖相关产业和国防能力的一个有效途径。两栖机器人要在野外环境下进行高效的自主作业极度依赖于稳定,准确的自主定位功能。在无 GPS 的环境中,机器人实现自主定位的一种主流方法是通过相机采集的图像序列来实现视觉里程计。但对于在野外环境中作业的两栖机器人而言,其自身运动特性、野外起伏不平的地形和复杂的光线条件会造成视觉信息缺失和相机视野丢失等问题,对传统的视觉里程计算法带来挑战。

为了实现两栖机器人在野外环境下准确、稳定的视觉定位,本文设计了一种 两栖主动视觉惯导里程计(AmphiActiveVIO)系统。该系统以基于滑动窗的紧融 合视觉惯导里程计为基础,通过相机采集的图像和 IMU 测得的加速度与角速度来 实时估计两栖机器人的位姿。同时,为了解决视野丢失问题,利用已探索的场景特 征点和机器人当前运动方向计算一个最优视角,该最优视角可以保证稳定视觉定 位的同时实现对场景最大限度的探索。然后通过一种主动曝光控制算法来解决野 外复杂光线所导致的图像视觉信息缺失问题,该算法通过最大化可导梯度百分比 指标来计算下一时刻最合适的曝光时间,从而保证下一帧图像不会出现过曝光或 欠曝光的情况。最后,为了解决由于两栖机器人运动所导致的特征追踪丢失问题, 通过将新检测的特征点和局部地图进行匹配来获得额外的视觉约束,能有效增加 视觉惯导里程计的稳定性。

为了验证两栖主动视觉惯导里程计系统的有效性,首先在仿真环境下对基于 最优视角的定位算法进行了验证,该算法在能保证稳定定位的同时实现对场景最 大限度的探索。然后通过在两栖机器人上加装三自由度云台,并在野外爬坡环境 下验证了主动视角调整能有效防止视野丢失。在野外 HDR 环境下对主动曝光控 制算法进行的实验表明,主动曝光控制相对于自动曝光可以保留更多的视觉信息, 并实现稳定的特征追踪,提高视觉定位算法的精度。在公开数据集 Euroc 和野外 的实验表明,相对于 Vins-Mono,增加额外特征匹配的 AmphiVIO 可以提高视觉定 位的精度和稳定性。

关键词:两栖机器人;视觉惯性里程计;主动定位;曝光控制

Abstract

Amphibious robots can directly serve many applications, e.g. the aquatic products, environmental exploration, and disaster rescue, etc. It is an effective way to improve China's amphibious related industries and national defense capabilities. However, the efficient and autonomous operation of amphibious robots in field environments relies heavily on the stable and accurate localization ability. In GPS-denied environments, traditional solutions to achieve autonomous localization is to utilize visual odometer technology through a sequence of consecutive images acquired by the camera. However, for amphibious robots operating in field environments, the motion characteristics, undulating terrain and complex lighting conditions cause many problems including lack of visual information and lost of field of view. The problems prevent the traditional localization algorithms from directly applying for amphibious robots.

To achieve an accurate and stable visual localization system for amphibious robots in the field environments, the thesis designs an amphibious active visual inertial odometer system (AmphiActiveVIO). The system is designed based on a tightly coupled visual inertial odometer with a sliding window. The image acquired by the camera and the data measured by the inertial measurement unit are fused to estimate the pose of the amphibious robot in real time. At the same time, to solve the problem of lost of field of view, an optimal viewing angle is calculated by using the perceived scene feature points and the current motion direction of the robot. The optimal viewing angle ensures stable visual localization while maximizing the exploration of the scene. Then, an active exposure control algorithm is developed to solve the missing problem of visual information caused by complex illumination conditions in field environments. The algorithm calculates the most suitable exposure time at the next frame by maximizing an predefined image gradient metric to ensure that the next frame is not overexposure or underexposure. Finally, to solve the problem of feature tracking lost caused by the robot's motion, several additional visual constraints are obtained by matching the newly detected feature points with the local map, and the method effectively increases the stability of the visual inertial odometer.

To verify the effectiveness of the proposed system, the localization algorithm based

on the optimal viewing angle was first verified in a simulated environment. The algorithm ensures the stable localization while achieving the maximum exploration of the scene. Then by equipping a three-degree-of-freedom gimbal on the amphibious robot, it is verified that the active view adjustment prevents the lost of field of view during climbing motions, and performs much better than that with the fixed view solution. Further, experiments on the active exposure control (AEC) algorithm in the fields under high dynamic lighting environments show that the active exposure control can preserve more image details than the automatic exposure of the camera. The AEC algorithm achieves stable feature tracking, and improves the accuracy of the visual localization algorithm. Experiments on the public data set Euroc also show that the developed AmphiVIO with additional feature matching can improve the accuracy of visual localization compared to Vins-Mono. Finally, the effectiveness of the whole AmphiActiveVIO localization system proposed in the thesis is verified with our amphibious robot platform in actual field environments.

Keywords: amphibious robot, visual inertial odometry, active localization, exposure control

目 录

摘 要	I
Abstract	II
第1章 绪论	1
1.1 课题研究的背景和意义	1
1.2 国内外研究现状	2
1.2.1 两栖机器人研究现状	
1.2.2 视觉里程计研究现状	5
1.2.3 主动视觉里程计研究现状	6
1.3 本文的主要研究内容	7
第2章 系统框架简介	9
2.1 整体系统框架	9
2.2 主动视角调整算法	11
2.2.1 基于优化的最大信息视角计算	11
2.2.2 基于最优视角的主动定位	16
2.3 本章小结	21
第3章 主动曝光控制算法	22
3.1 光度标定	22
3.1.1 相机光度成像过程	23
3.1.2 离线光度标定算法	24
3.2 主动曝光控制	
3.2.1 基于梯度的图像评价指标	27
3.2.2 基于优化的曝光控制算法	29
3.2.3 自适应光度矫正	
3.3 本章小结	

第 4	章	章 两栖视觉惯导里程计	
4.	1	视觉惯导预处理	
4.2	2	AmphiVIO 系统初始化	
4.	3	基于局部地图匹配的视觉约束	40
4.4	4	紧融合的 VIO 优化	46
4.:	5	本章小结	
第5	章	章 实验验证与结果分析	
5.	1	实验平台搭建	
5.2	2	主动视角调整实验	
5.	3	主动曝光控制实验	
5.4	4	两栖视觉惯导里程计实验	
5.:	5	本章小结	
结	论	€	60
参考	文	て献	
攻读	硕	贡士学位期间发表的论文	67
哈尔	滨	宾工业大学学位论文原创性声明和使用权限	
致	谢	射	

第1章 绪论

1.1 课题研究的背景和意义

中国拥有近 300 万平方公里的海域与 32000 公里长的海岸线,有大量的近海 作业需求,例如近海侦察以及排雷、农业生产、环境污染探测、资源探测、近海生 物研究、以及救灾等。两栖机器人的作业环境包含从浅水到拍岸浪区,以及滩涂 区域等多种野外环境。这些区域往往范围特别广泛,条件恶劣,导致依靠人力的 两栖作业十分困难。此外,为了应对未来多样的国际局势,保护好中国绵长的海 岸线安全,发展海军力量特别是两栖作战力量是我国的必然选择。近海及其他滩 涂两栖作业中,非常需要能在复杂水域、崎岖陆地以及水陆交界的泥、沙等过渡 环境松软介质中自主作业的两栖机器人。通过机器人携带高性能作业和勘探设备, 实现自主且快速有效的两栖作业,是提高我国两栖相关产业和国防能力的一个有 效途径。

为了适应两栖作业的需求,两栖机器人的野外自主导航、移动机构与推进机理 和集群协作等关键技术急需针对性的深入研究。其中野外自主导航技术负责确定 两栖机器人在野外环境中的位姿,规划一条到目标位置的可行路径,并控制移动 机构依照这条轨迹安全到达目标位置,是两栖机器人能实现自主作业的前提。高 效的自主导航功能需要自主定位、轨迹规划与轨迹跟踪等多个模块协同工作才能 实现。其中自主定位模块通过融合机器人本体传感器或外部传感器的数据确定机 器人在世界坐标系下的位姿。该模块是整个自主导航技术的基石,对最终自主导 航的性能有重要影响。轨迹规划和轨迹跟踪模块只有在自主定位模块提供的准确 位姿下才能正常工作。本课题需要解决的就是两栖机器人在野外环境下的自主定

自主定位可以使用如 GPS 这样的外部传感器实现,但当 GPS 失效时,所有依赖于该外部传感器的机器人将全部丧失定位功能,这在国防和其他重要场合是不被允许的。而仅依赖自身传感器的自主定位算法则不受此影响,极大地提高机器人自主作业的抗干扰能力,更加适用于两栖机器人这样需要稳定运行且外部环境复杂的应用。不依赖于外部传感器和提前制造好的地图,仅依靠自身传感器实现在未知环境下的自主定位,需要用到 SLAM (Simultaneous Localization and Mapping)技术,中文译作"同时定位与地图构建"。随着最近几年机器人,无人驾驶和 VR

的火热,SLAM 技术逐渐被人们所熟知。SLAM 技术致力于同时解决定位和环境 建图这两个相辅相成的研究问题,其中准确的定位依赖于高精度的地图,而高精 度地图的构建又依赖于准确的机器人定位结果。SLAM 是移动机器人实现自主化 中最重要的技术,一直是机器人领域的研究热点。目前,SLAM 中的主流的传感 器是激光测距仪和视觉系统,相比于激光测距仪,视觉系统更为廉价,而且可以 提供丰富的图像信息。近年来,基于图像光度的视觉 SLAM 算法可以构建半稠密 乃至稠密的地图,同时将视觉和惯导结合形成的视觉惯性导航系统可以获得很高 的定位精度。因此不论是在研究领域还是在实际应用中,视觉 SLAM 俨然成为了 主流趋势。

虽然现有视觉 SLAM 方法有很多,比如基于特征点法的 ORB-SLAM2^[1] 和基 于直接法的 DSO^[2],但是对于两栖机器人在复杂野外环境下的应用还有很多问题。 首先是两栖机器人自身运动带来的问题,现有的视觉 SLAM 往往是用在移动机器 人或者无人机上,而两栖机器人的一个特点是存在剧烈的运动,本文使用的两栖 机器人存在三足和六足两种运动模式。在三足运动模式下,运动较为平稳,但在 六足运动模式下,由于六足一起运动,两栖机器人的运动呈现周期性的上下晃动。 同时由于地形起伏不定,当两栖机器人步态出现紊乱时,也会导致视野出现剧烈 晃动,视角发生大幅度的变化,并伴有一定程度的模糊。这会严重降低依靠前后 帧图像块匹配来实现定位的视觉 SLAM 方法的性能。其次,野外环境不同于室内 环境的一个很重要的地方是野外环境的光线较为复杂,比如在正午时,整个环境 中由于太阳直射,亮度很高,而被植物遮挡住的阴暗位置则亮度很低,但是相机 由于感光范围有限,无法在一张图片中展现这么大的动态范围,会导致图像出现 过暗或者过亮的情况,使视觉特征丢失,从而导致视觉 SLAM 失效。最后,由于 相机安装在两栖机器人的前方,当两栖机器人进行爬坡运动时,相机会看向天空, 由于天空几乎没有特征,这会导致视觉 SLAM 直接失效。

针对两栖机器人在野外环境下,由于两栖机器人自身的运动特点、野外环境 崎岖的地形和复杂的光照条件导致已有视觉 SLAM 方法失效的问题。本文希望提 出一种以视觉为主的自主定位算法,该算法主要利用视觉信息,通过结合 IMU 信 息来完成基础定位功能。同时能够解决由于野外环境光线复杂,两栖机器人运动 幅度剧烈,地形崎岖不平对视觉惯导里程计的影响。

1.2 国内外研究现状

根据本文的研究方向,首先回顾两栖机器人在国内外的研究现状,主要关注

于导航方向的进展。然后回顾视觉里程计的发展历史,特别是近期视觉惯导里程 计的发展。最后回顾主动视觉里程计的发展,这里主动的含义为通过主动改变相 机的参数,如曝光时间、相机视角等,使捕获的图像中含有更丰富的视觉信息。

1.2.1 两栖机器人研究现状

近年来,受泥沙等两栖介质环境特种作业需求的驱动,许多研究机构和公司都 推出了针对性的两栖机器人。这些基于生物优良运动特性设计的仿生两栖机器人, 不仅可以实现高效、高机动的跨介质运动,而且不易被环境中原有生物发现,使 作业具有环保性或者突袭性。目前国际上和国内都有许多不错的成果,如图1-1所 示。国外做的较好的有瑞士洛桑联邦理工大学的仿蝾螈机器人 AmphiBot^[3],日本 东京工业大学的仿蛇形机器人 ACM-R5^[4] 等。国内虽然对两栖机器人的研究起步 较晚,但也有很多不错的成果,其中包括中科院自动化所与石河子大学合作研制 的两栖仿生机器鱼 "AmphiRobot"^[5],以及北航王田苗等人提出的仿鳄鱼水陆两栖 机器人^[6]。上述的仿生两栖机器人还未走向实际应用,主要原因在于:(1)多数 两栖机器人使用两套推进机构实现水中及陆地上的行走,但速度、效率、通过性 都比较低,距离实用差距较大;(2)对机器人在两栖介质中的行进性能,例如在 沙或泥中的行进性能以及两栖环境的功能转换,研究仍然较少,无法支撑两栖机 器人走向实用。针对这些问题,中科大开发的 AmphiHex-I^[7], 做了一些有益的尝 试,AmphiHex-I 通过主动变形实现水下和陆地两种有效推进方式的统一,使用一 套机构即可实现水下和陆地推进。在此基础上,中科大研发了 AmphiHex-II^[8],通 过六条变刚度的被动变形柔性腿可以避免介质切换时的主动腿型变化的复杂控制, 同时对其结构形态根据推进机构与松软介质互相作用机理进行了的优化设计。

现阶段国内外虽然有不少研究机构都在研究两栖机器人,但多数研究都只局限于机器人在两栖环境下的推进机理和机构研究,很少有人对两栖机器人导航系统进行深入研究。主要原因是两栖环境包括水下环境、滩涂环境和陆地环境等多种复杂的野外环境,在不同的介质中要使用对应的传感器的信息,最后实现多传感器融合的定位。对两栖机器人的定位而言,大多数研究还都是基于 GPS 的,如国内的尹咸阳^[9] 在 2006 年使用 GPS 对水陆两栖机器人进行定位,国外的 Kaznov等^[10] 在 2010 年使用 GPS 对球形两栖机器人进行定位。对于两栖机器人在水下环境的定位,由于水下环境的特殊性,基于视觉的方法无法检测到足够的视觉特征。主流方法是基于声呐的 EKF-SLAM,如 Ribas 等制作的 Ictineu 水下机器人通过机械扫描成像声呐提取线特征进行自主定位^[11], Aulinas 等在 2010 年使用侧扫声呐



图 1-1 国内外两栖机器人 (a) AmphiBot 机器人 (b) ACM-R5 机器人 (c) AmphiRobot 机器人 (d) AmphiHex-I 机器人 (e) AmphiHex-II 机器人

提取水下环境中的显著点特征实现了 REMUS-100 机器人在非结构环境中水下自 主定位^[12],但基于声呐的方法缺点在于水下声呐系统体积较大,价格非常昂贵。

两栖机器人在陆地环境下的自主视觉定位同足式机器人在野外环境下的视觉 定位较为类似。在四足机器人中,比较有代表性的是波士顿动力公司的BigDog^[13]。 该机器人能够稳定地行走于丛林、山地等复杂崎岖的地形上,通过融合GPS、LI-DAR、IMU 和双目视觉相机等传感器的信息进行定位,可以实现机器人在复杂野 外环境下长航时的自主定位。

有部分研究者尝试利用机器人的足式里程计 LO(Leg Odometry)来提高定位 系统的鲁棒性和精度。2013 年, Bloesch 在一个四足机器人上利用一个 EKF 框架 融合基于立足点运动方程的足部里程计和 IMU 来估计机器人的位姿^[14],在直线运 动 1m 的实验中,位置精度在 10% 左右。Reinstein 和 Hoffman 同样也只利用 LO 和 IMU 在 EKF 框架下进行位姿估计^[15],在大概 10m 的运动中最终的位置误差在 2.0% 左右。2015 年,波士顿动力将双目视觉传感器,惯性导航单元,腿足里程计 和可选的 GPS 融合到一个 EKF 框架中实现了一个鲁棒的、低延迟的、定位精度高 的自主定位系统^[16]。该系统能够在野外无 GPS 的环境下 24 小时工作,并能达到 1% 的定位精度。

总的来说,直接针对两栖机器人在水陆两栖环境下均可用的视觉定位算法几 乎没有,现有的针对足式机器人在野外环境下的自主定位方案,一般都需要融合 多传感器信息来实现自主定位。同时两栖机器人相对于足式机器人而言并没有完整的运动学,无法实现足式里程计来进行辅助定位。

1.2.2 视觉里程计研究现状

自主定位是机器人实现导航和自主作业的前提。SLAM 中主流的传感器是激 光测距仪和视觉传感器,相比于激光测距仪,视觉传感器价格低廉,同时能在无结 构的场景中使用。视觉 SLAM 最早的一个较完善的工作是 A.J.Davison 在 2007 年提 出来的 MonoSLAM^[17]。这是第一个实时的单目视觉 SLMA 系统, 但存在诸如应用 场景窄,容易丢失的缺点。2007 年,Klein 等人提出了 PTAM (Parallel Tracking and Mapping^[18], PTAM 首次实现了跟踪和建图部分的并行化, 同时也是第一个使用非 线性优化而不是使用传统的滤波器作为后端的 SLAM 系统。2015 年, Mur-Artal 等 提出了非常完善和易用的 ORB-SLAM^[19], 创新性的三线程结构取得了非常好的跟 踪和建图效果,能够保证轨迹和地图的全局一致性。2014年, J.Engle 等人提出了 LSD-SLAM^[20] (Large Scale Direct monocular SLAM) 标志着单目直接法在 SLAM 中 的成功应用。SVO (Semi-direct Visual Odometry)^[21] 是由 Forster 等人于 2014 年提出 的一种基于稀疏直接法的视觉里程计,创新性的提出了深度滤波器的概念。除了单 目 SLAM 外,还有一些 RGB-D 传感器上的 SLAM 方案,如 RTAB-MAP(Real Time Apperance-Based Mapping)^[22], 实现了一个比较完整的 SLAM 系统。国内 SLAM 中 做的较好的有浙江大学章国峰团队提出的 RKSLAM^[23] 和 RDSLAM^[24],这两个系 统在 AR 应用中都取得了不错的效果。

由于运动剧烈导致的图像模糊以及相邻两帧之间无共视导致的特征追踪丢失 会极大地降低视觉里程计的精度和稳定性。而惯性传感器可以测量传感器本体的 角速度和线加速度,与相机传感器具有明显的互补性。因此,在实际使用中,通 常结合惯性传感器(IMU)的信息组成视觉惯导里程计(Visual-Inertial odometry, VIO)进行联合估计,可以显著提高 SLAM 系统的稳定性和精度。2011年,Kelly 和 Sukhatme 提出一个融合 IMU 并可以在线估计外参数 VIO 系统^[25]。2015年, Leutenegger 描述了一个紧融合的 VIO 系统^[26],采用滑动窗口 Bundle Adjustment 可以实时估计重投影误差和机器人轨迹。2007年,Mourikis提出了MSCKF^[27],是基 于 EKF 滤波的紧融合方法,主要是引入了滑动窗口来管理特征点。2017年,Forster 阐述了一种精简的预积分方法,能有效的减少 VIO 的计算量^[28]。最近 VIO 方法逐 渐趋于成熟,基于滤波的方法中比较成熟的是 Michael 在 2017提出的 ROVIO^[29],基 于优化的方法中比较成熟的是国内香港科技大学 Tong Qin 在 2017 提出的 VINS^[30] 框架。

总体来说国内外在视觉 SLAM 上的研究已经较为成熟了,基本框架和算法已 经趋于完善,但是现有的算法在鲁棒性方面还有待提升,为了解决由于机器人运 动较快使图像模糊而导致的场景视觉特征丢失问题,国内外学者普遍将 IMU 测量 融合到视觉信息中进行联合状态估计。相对于单纯的视觉里程计,使用 VIO 在大 多数情况下可以获得更鲁棒,精度更高的位姿估计值。但是当机器人的运动存在 剧烈振动时会导致加速度计失效,同时在野外环境下由于光线原因或者视觉丢失 等问题会导致图像中根本无法提取有效的视觉信息,这也会直接导致 VIO 失效。 一部分学者尝试考虑机器人自身的运动学里程计来进行补偿,虽然精度不是很高, 但是也可以提供一种有效的估计。

1.2.3 主动视觉里程计研究现状

上述的视觉里程计方法,都只简单地考虑了如何利用已采集的相机图像来对 机器人的位姿进行估计,本文称这种方法为被动视觉里程计,与其相对的是主动 视觉里程计,即主动对相机的曝光时间或视角进行调整,使采集的图像中包含更 多的视觉信息,实现更鲁棒和准确的视觉里程计。对于在野外环境下进行作业的 两栖机器人而言,主要考虑两部分,第一部分是主动调节相机的曝光参数来最大 化相机采集图像中的视觉信息;第二部分是主动调整相机的视角,将相机视角从 无纹理的区域转换到富含纹理的区域,从而防止视觉里程计失效。

现阶段已有一些面向视觉里程计的曝光控制算法,2014年,Shim 等人^[31]提 出使用图像梯度作为最优曝光时间的指标,通过在原始图像上应用不同的伽马矫 正来寻找能最大化梯度信息的伽马矫正值,然后利用这个最优伽马矫正值计算最 优曝光时间。在2017年,Zhang 等人^[32]使用另一种更稳定的基于梯度的度量标 准,并结合相机的光度响应曲线通过梯度下降的方法来控制曝光时间。相机的光 度响应曲线可以通过对同一HDR 场景捕获的不同曝光时间下的图像计算得到^[33]。 使用主动曝光控制会破坏光流跟踪的光度不变假设,可以利用相机响应曲线对图 像进行光度矫正,提高光流跟踪的准确性,从而提高视觉惯导里程计的准确性。

当机器人利用自主导航功能对未知场景进行探索时,定位和探索这两个任务 是互相矛盾的,因为良好的定位结果需要机器人看向已探索的区域,而机器人要 对场景进行完整的探索需要让相机不断看向未探索的区域。为了解决这个矛盾,需 要在机器人的路径规划过程中考虑机器人的定位需求,这称为主动 SLAM。主动 SLAM 早期的工作包括最大化机器人规划路径中特征点观测的费雪信息^[34],或者 对 SLAM 前端特征匹配过程中的特征点进行主动搜索^[35]。在主动 SLAM 的路径 规划过程中,需要评价每一个位姿下机器人定位效果的好坏,在基于采样或基于 优化的路径规划方法中,同一位姿能观测的信息量大小需要进行多次计算,带来 了额外的计算量。因此一个较好的方法是将空间划分为固定大小的体素,提前计 算好每个体素位置上的信息量。在计算每个体素的信息量时,Roy 等人^[36]使用全 向相机来计算信息量,避免了在不同视角下信息量不同的问题,降低了计算的维 度。但是实际使用中的相机一般不是全向相机,为了解决这个问题,Ichter 等人^[37] 使用神经网络来构建一个感知费用地图用于路径规划。但是这会使得地图中的每 个体素只包含在不同朝向下费用的平均值,并不能用来估计在任一位姿下的费用。 最近,Zhang 等人^[38]提出了一种新的信息量表示方式,在这种表示方式下,信息 量可分解为位置相关和朝向相关的两部分的,基于位置部分可以高效地构建一个 信息场,但是由于在该过程中,建模相机可视模型使用了余弦函数,带来了相机 可视模型误差。

主动 SLAM 往往需要针对具体的机器人和场景形式来进行研究。现有的研究 主要面向于室内环境下的移动机器人,尚无有关两栖机器人在野外环境下进行主 动视觉定位的相关研究。野外环境与室内环境的主要区别在于两点,第一是阳光 和树木阴影形成了更大的亮度动态范围,第二是野外环境的复杂地形和两栖机器 人的运动形式对主动视觉定位提出了特别的要求。

1.3 本文的主要研究内容

针对两栖机器人在野外环境下进行自主视觉定位存在的一些问题,例如,由 于自身运动特性,野外崎岖地形和光线变化所造成的视觉信息缺失和相机视野丢 失等问题。本文主要从以下几个方面展开研究,并设计了相应的实验对所提出的 算法进行验证:

(1) 尝试设计一种两栖主动视觉惯导里程计(AmphiActiveVIO)系统。该系统 以紧融合视觉惯导里程计为基础,通过相机采集的图像和惯性测量单元测得的加 速度和角速度数据来实时估计两栖机器人的位姿。

(2)当两栖机器人在野外崎岖地形中爬坡时,相机视角会看向无纹理的天空, 导致视野丢失。为了解决这个问题,本文尝试提出了一种主动视角调整算法,该 算法利用已探索的场景特征点和当前两栖机器人的位置,计算一个最大信息视角。 为了平衡两栖机器人的定位和探索需求,结合机器人当前的运动方向计算一个最 优视角,使机器人在这个最优视角下能保证稳定视觉定位的同时实现对场景最大限度的探索。

(3)野外复杂的光线会导致图像过暗或过亮,从而无法提取足够的视觉特征 点。因此尝试通过主动控制相机的曝光时间来使图像中包含足够的视觉特征,该算 法通过最大化可导图梯度百分比指标来迭代计算下一时刻最合适的曝光时间,从 而保证下一帧图像不会出现过曝光或欠曝光的情况。

(4)当两栖机器人进行俯仰运动时会导致相机视野发生大幅变化,这会使得视觉惯导里程计前端的光流跟踪无法实现稳定的特征匹配。因此尝试将每个新检测的特征点和局部地图中的特征点进行匹配,如果匹配成功,则在滑动窗优化中新增加一个视觉约束项。

(5)设计相关的实验对本文提出的算法进行验证,在仿真数据上对基于最优视角的定位算法进行测试,同时让两栖机器人在野外环境中进行爬坡运动,验证使用三自由度云台的主动视角调整算法的有效性。在野外存在阳光和树木阴影的HDR环境中测试主动曝光控制算法对视觉惯导里程计精度的影响。最后在公开数据集 Euroc 上对基于匹配的视觉惯导里程计的精度进行评价,在野外实际环境中对 AmphiActiveVIO 系统的有效性进行实际测试。

第2章 系统框架简介

为了实现两栖机器人在野外环境下的自主视觉定位,本文设计了一种两栖主动视觉惯导里程计(AmphiActiveVIO)系统。该系统以紧融合视觉惯导里程计为基础,在其上综合利用信息理论和机器人的运动方向计算一个最优视角,通过三自由度云台来主动控制相机视角到最优视角从而防止视野丢失。然后利用图像梯度信息和相机响应曲线来计算合适的曝光时间,通过主动设置相机的曝光时间,使相机捕获的图像中能包含丰富的纹理信息。最后通过将当前图像新检测的特征点和局部地图进行匹配来增加额外的视觉约束,提高视觉惯导里程计后端优化的稳定性。本章接下来首先介绍 AmphiActiveVIO 系统中各个模块的作用和它们互相之间的联系,然后介绍用于防止视野丢失的最优视角定位算法。



2.1 整体系统框架



整个 AmphiActiveVIO 系统框架如图2-1所示。主要分为两栖视觉惯导里程计 (AmphiVIO), 主动曝光控制和主动视角调整三个模块。相关的传感器包括能提供

图像信息的单目相机和能提供高频加速度和角速度信息的惯性测量单元(IMU)。 执行机构为一个能控制相机视角的三自由度云台。

整个 AmphiActiveVIO 系统启动之后需要先进行初始化。初始化是为了给视觉 惯导里程计滑动窗中的所有状态量提供一个合理的初始值,因为视觉惯导里程计 是一个高度非线性的系统,没有合理的初始值,该系统很难优化收敛。初始值是 通过将运动恢复结构(Structure From Motion)计算得到的不带尺度的视觉轨迹与 IMU 估计的带尺度的轨迹进行对齐得到的。初始化过程中需要确定的状态量包括 陀螺仪和加速度计的偏置、重力矢量以及初始特征点地图中特征点的位置和相机 的位姿。

系统初始化完成后,三个主要模块,主动视角调整、主动曝光控制和AmphiVIO 将开始执行各自的功能。接下来从图像信息流的角度对整个AmphiActiveVIO的定 位过程进行梳理。首先是主动视角调整模块,该模块的目的是将相机调整到一个 合适的视角,防止视野丢失,即使视野中含有足够的视觉特征,不会看向天空等 无纹理的区域。该模块主要利用视觉惯导里程计估计的局部特征点地图来逐步构 建一个全局栅格地图,然后利用信息理论从全局栅格地图中计算出一个最大信息 视角,该最大信息视角朝向已探索场景中特征最丰富的区域。但如果一直控制相 机朝向最大信息视角,则机器人为了定位功能会放弃对场景的探索,即机器人会 一直看向场景中特征丰富的区域,而不会看向其他方向实现对场景的探索。因此 考虑根据当前定位的状态和机器人的运动方向计算一个最优视角。当定位状态良 好时,最优视角重靠近于最大信息视角。计算出来的最优视角则会被发送给一个 三自由度云台来主动控制相机视角。

当将相机视角调整到合适的角度后,需要利用主动曝光控制模块将相机的曝 光值设置到一个合适值,使视野中呈现足够的视觉信息,不会出现欠曝光或过曝 光的情况。该主动曝光控制模块首先对相机的响应曲线进行离线标定,获得图像 亮度和曝光时间之间的函数关系。然后对当前帧图像提取梯度,计算一个可导百 分比梯度指标,用基于优化的方法来最大化可导百分比梯度指标,计算得到下一 帧合适的曝光时间给相机。通过不断地调整相机曝光时间,可以使相机采集的图 像序列中含有丰富的视觉信息。

含有丰富视觉信息的图像和高频 IMU 数据接下来会输入到 AmphiVIO 模块中。输入图像需要先进行光度修正来解决由于曝光变化导致的前后帧图像亮度不

一致问题, 光度矫正后的图像通过光流跟踪来获得前后帧之间的特征匹配关系。由 于两栖机器人的俯仰运动特点, 会导致相机视角变化太大而使前端光流追踪丢失。 因此通过将特征点与局部特征点地图进行匹配来获得额外的视觉测量,并将作为 一个约束项加入到后端优化中。对 IMU 数据,则将其进行预积分后加入到后端优 化中。最后对整个滑动窗口中的状态量和观测量进行紧融合的 VIO 优化来获得低 频率的机器人位姿。为了之后主动视角调整模块计算最优视角,利用 IMU 的数据 对低频率的机器人位姿进行航迹递推可以得到高频率的机器人的位姿,将机器人 位姿的正方向(X 轴方向)当做机器人的运动方向。

2.2 主动视角调整算法

两栖机器人在野外环境中作业时,由于地形崎岖不平,当两栖机器人进行如 图2-2 (a)所示的爬坡运动时,位于两栖机器人前方的相机会看向几乎没有视觉特 征的天空,如图2-2 (b)所示,这会严重影响视觉惯导里程计的稳定性,导致定位 失败。为了解决这个问题,通过给两栖机器人装上一个三自由度云台来主动调整 相机视角至纹理丰富的区域,避免两栖机器人看向纹理缺失的区域。为了主动调 整相机视角,需要设计一种算法能够依据当前两栖机器人的运动状态和周围环境 中特征点的位置,实时计算出应朝何处调整视角,记该视角为最优视角。



a)两栖机器人在野外爬坡 b)两栖机器人爬坡时视角看向天空 图 2-2 两栖机器人野外爬坡视野丢失说明

2.2.1 基于优化的最大信息视角计算

为了计算最优视角,首先计算较为简单的最大信息视角,最大信息视角计算问题可以表述为:假设周围环境中所有N个特征点在世界坐标系下的位置 $\{p_i^w\}_{i=1}^N$

和机器人的当前位置 t_{wc} 已知,求机器人在当前位置 t_{wc} 上可以获得最大信息的视角 R_{wc}^{max} 。

对于该问题的一种直观解法是,沿着不同视角进行采样,分别计算在不同视 角下可以获得的信息量大小,能获得最大信息量的视角即为最大信息视角。为了 计算最大信息视角,需要先用数学的形式对信息大小进行度量。视觉里程计通过 相机图像中的信息来估计相机的运动,这本质上是一个参数估计的问题。对于参 数估计问题,费雪信息矩阵 (Fisher Information Matrix)^[39] 被用来表示观测量中含 有多少待估计状态量的信息。对于待估计的状态量 *x* 和对应的观测量 *z*,观测模 型可以用一个条件概率密度函数 *p*(*z*|*x*) 来表示,其中观测量对状态量提供的信息 大小由如式 (2-1) 所示的费雪信息矩阵来计算。

$$I_{x}(z) = \left(\frac{\partial}{\partial x} \log p(z|x)\right)^{\top} \left(\frac{\partial}{\partial x} \log p(z|x)\right)$$
(2-1)

如果测量过程中只引入了独立的零均值高斯白噪声 $N(0,\sigma^2)$,则公式(2-1)可以 写成:

$$I_{\mathbf{x}}(z) = \frac{1}{\sigma^2} \left(J_{\mathbf{x}} \right)^{\mathsf{T}} J_{\mathbf{x}}, \quad \nexists \oplus J_{\mathbf{x}} = \frac{\partial z}{\partial \mathbf{x}}$$
(2-2)

由公式(2-2)可以看出,如果测量噪声的标准差 σ 越大,则费雪信息矩阵越 小,这是因为测量噪声越大,则测量所包含的关于状态量的有用信息就越少。另一 方面,如果测量关于状态量的偏导数 J_x 越大,则费雪信息矩阵越大,这是因为如 果偏导数很大,那么当状态量变化较小时,测量量变化也会很大,也就是测量量 对状态量的变化很敏感。一般计算费雪信息矩阵时,x的取值为当前的估计值 \hat{x} 。 特别的,由费雪信息矩阵的逆,可以定义克拉美-罗(Cramfer-Rao)下界,也就是 该状态估计器估计的状态量的协方差可以达到的最小值。

利用费雪信息矩阵,可以最大信息视角计算问题建模为如式(2-3)所示的优化问题。

$$\boldsymbol{R_{wc}^{max}} = \underset{\boldsymbol{R_{wc}}}{\operatorname{arg\,max}} \sum_{i=1}^{N} v\left(\boldsymbol{T_{wc}}, \boldsymbol{p}_{i}^{w}\right) \boldsymbol{I}_{i}\left(\boldsymbol{T_{wc}}, \boldsymbol{p}_{i}^{w}\right)$$
(2-3)

其中 $T_{wc} = \{R_{wc}, t_{wc}\}$ 表示相机在世界坐标系下的位姿; $\{p_i^w\}_{i=1}^N$ 表示世界坐标系下所有地图点的位置; $I_i(T_{wc}, p_i^w)$ 表示在 T_{wc} 处观测到第 i 个特征点所对应的费雪信息矩阵。 $v(T_{wc}, p_i^w)$ 是一个二值函数,表示第 i 个特征点在位姿 T_{wc} 下是否可见,称为相机的可视模型。式(2-3)的优化目标为找到一个最大信息视角 R_{wc}^{max} 来使整个目标函数值最大。

仔细分析式(2-3)中目标函数的形式可以发现,该目标函数由许多小项求和 组成,每个小项和一个特征点相关,每个小项由两部分组成,第一部分表示该特 征点是否在当前位姿 *T_{wc}*下可见,第二部分则是该特征点对当前位姿 *T_{wc}*的费雪 信息矩阵。接下来首先分析费雪信息矩阵 *I_i*(·)的具体形式,并说明其和优化变量 是无关的,在优化过程中可以作为常量处理。然后介绍如何利用连续函数来逼近 二值函数 v(·)使最后的目标函数可导,能够使用梯度下降的方法进行优化求解。

相机对一个 3D 特征点的观测可以表示为不同的形式,如图像平面上的像素,或者是通过相机光心的方向向量 (bearing vector),其中方向向量相对于像素表示的优点是可以表示任意大小的视场角。使用方向向量的表示方法,相机对特征点 *p_i^w* 的测量模型可以写成:

$$f_{i} = \frac{p_{i}^{c}}{\|p_{i}^{c}\|_{2}} = \frac{1}{n_{i}}p_{i}^{c}, \quad p_{i}^{c} = T_{cw}p_{i}^{w}$$
(2-4)

由链式法则可得观测 fi 相对于相机位姿 Tcw 的雅克比:

$$\boldsymbol{J}_{i} = \frac{\partial \boldsymbol{f}_{i}}{\partial \boldsymbol{p}_{i}^{c}} \frac{\partial \boldsymbol{p}_{i}^{c}}{\partial \boldsymbol{T}_{wc}}$$
(2-5)

其中第一部分可以写成:

$$\frac{\partial f_i}{\partial \boldsymbol{p}_i^c} = \frac{1}{n_i} I_3 - \frac{1}{n_i^3} \boldsymbol{p}_i^c \left(\boldsymbol{p}_i^c \right)^\top$$
(2-6)

对于第二部分,使用李代数 se(3) 来表示位姿,则可写成:

$$\frac{\partial \boldsymbol{p}_{i}^{\boldsymbol{c}}}{\partial \boldsymbol{T}_{\boldsymbol{w}\boldsymbol{c}}} \rightarrow \frac{\partial \left(\exp\left(\boldsymbol{\xi}^{\wedge}\right)\boldsymbol{T}_{\boldsymbol{w}\boldsymbol{c}}\right)^{-1}\boldsymbol{p}_{i}^{\boldsymbol{w}}}{\partial \boldsymbol{\xi}} \stackrel{\text{gl}}{=} \frac{\partial \left(\boldsymbol{T}_{\boldsymbol{w}\boldsymbol{c}}\exp\left(\boldsymbol{\xi}^{\wedge}\right)\right)^{-1}\boldsymbol{p}_{i}^{\boldsymbol{w}}}{\partial \boldsymbol{\xi}}$$
(2-7)

其中 exp(·) 表示到特殊欧式群 SE(3) 上的指数映射。上述的两种形式实际上对应着 扰动项 δξ 是加在世界坐标系下的全局扰动,还是加在相机坐标系下的局部扰动。 使用加在世界坐标系下的全局扰动,则 J_i 可以写成如下形式:

$$\boldsymbol{J}_{i} = \frac{\partial \boldsymbol{f}_{i}}{\partial \boldsymbol{p}_{i}^{\boldsymbol{C}}} \boldsymbol{R}_{\boldsymbol{C}\boldsymbol{W}} \left[-\boldsymbol{I}_{3}, \quad \left[\boldsymbol{p}_{i}^{\boldsymbol{w}} \right]_{\times} \right]$$
(2-8)

在这种全局扰动形式下,对于只差一个旋转 $R_{cc'}$ 的两个位姿 T_{wc} 和 $T_{wc'}$ = $\{R_{wc}R_{cc'},t_{wc}\}$,二者雅克比矩阵之间的关系可以简单的写成 $J'_i = R_{C'}J_i$,那么二者拥有相同的费雪信息矩阵

$$\boldsymbol{I}_{i}^{\prime} = \frac{1}{\sigma^{2}} \boldsymbol{J}_{i}^{\top} \boldsymbol{R}_{\boldsymbol{C}\boldsymbol{C}^{\prime}}, \boldsymbol{R}_{\boldsymbol{C}\boldsymbol{C}^{\prime}} \boldsymbol{J}_{i} = \boldsymbol{I}_{i}$$
(2-9)

二者拥有相同费雪信息矩阵的原因在于公式(2-3)中, *I_i* 为不考虑相机的可 视模型时,第*i* 个特征点相对于相机位姿*T_{wc}* 的费雪信息矩阵。由于没有考虑相机 的可视模型,那么这个值和相机的视角 *R_{wc}* 是无关的,而只与相机的位置 *t_{wc}* 有

关。总而言之,通过选取方向向量作为观测量的表示形式,同时使用在世界坐标系下表示的全局扰动,费雪信息矩阵 $I_i(T_{wc}, p_i^w)$ 的值和相机视角 R_{wc} 是无关的,在优化过程中可以作为常量处理。

那么求解式(2-3)的关键在于如何使用连续函数来逼近相机的可视模型 $v(T_{wc}, p_i^w)$ 。假设相机水平,垂直和对角线方向的视场角相同,且均为 2α ,则理想 相机可视模型如图2-3所示,其中f表示特征点, $e_3 = [0,0,1]^T$ 表示相机的光轴, θ 表示特征点和光心的连线与光轴的夹角。



图 2-3 理想相机可视模型^[38]

可以发现函数 ν(·) 在正负 α 处是不可导的,为了解决这个问题, Zhang 等^[38] 提出将 ν(·) 建模为如式 (2-10) 所示的关于夹角 θ 的函数。该可视化模型的函数曲 线如图2-4 (a) 所示。建模成这样的主要原因是该函数能将信息场分解只和位置相 关与只和朝向相关的两部分。但是该模型存在一个主要问题是不在相机视野中的 特征点也占据了大量的比重,即使假设相机是视角达 120 度的广角相机,在视野 边缘外的特征点也有接近 0.5 的权重,而不是理想模型中的 0。

$$v_{\text{zhang}}\left(\boldsymbol{T}_{\boldsymbol{w}\boldsymbol{C}},\boldsymbol{p}_{i}^{\boldsymbol{w}}\right) \to v\left(\theta_{i}\right) = a_{2}\cos^{2}\theta_{i} + a_{1}\cos\theta_{i} + a_{0}$$
(2-10)

由于只需要计算当前相机位姿的最大信息视角,并不需要对整个场进行更新, 所以不需要将信息场分解为只和位置相关和只和朝向相关的两部分。因此,本文 设计如公式(2-11)所示的相机可视模型。

$$v_{\text{our}}\left(\boldsymbol{T}_{\boldsymbol{w}\boldsymbol{c}},\boldsymbol{p}_{i}^{\boldsymbol{w}}\right) \to v\left(\boldsymbol{\theta}_{i}\right) = e^{k\cos\theta_{i}}$$

$$(2-11)$$

其中 k 是系数因子,用来控制相机可视曲线的形状,相机视场角越小则 k 值越大。 其函数曲线图如图2-4(b)所示,使用 e^x 有两方面的考虑,首先 $\cos\theta \in [-1,1]$,当

哈尔滨工业大学工学硕士学位论文



特征点从相机正前方逐渐移向相机后方,则夹角 θ 逐渐增大, cos θ 的值则从 1 逐 渐变到-1,对应的 e^{cos θ} 的值也逐渐减小,同时由于 e^x 函数能保证, x < 0 时, e^x 的值不会大于 1,这使得位于相机视野后方的特征点不会贡献太多权重;其次,这 个函数的导数很容易求得,方便之后使用梯度上升的方法求解最优视角。

通过上述公式确定了式(2-3)中各项的具体形式,在进一步求解该式时,首 先计算出所有特征点相对于当前机器人位置的信息矩阵,在计算信息矩阵时,由 于信息矩阵和姿态是无关的,可以将姿态直接设置为单位矩阵。然后设定当前机 器人运动方向作为最大信息视角的优化初始值,使用如高斯牛顿之类的梯度下降 方法即可迭代求解出最大信息视角。

在如图2-5所示的仿真环境中对本文提出的基于优化的最大信息视角定位算法 进行验证。首先在该仿真环境中生成两道特征虚拟墙,该特征墙通过在 x = 2.5m, y = 2.5m,高度范围 $z \in [-1,1]m$ 上随机采样 1000 个特征点形成。然后在 { $x \in [-2,2], y \in [-2,2], z \in [-1,1]m$ }的空间中随机采样一些点表示机器人的可能位置。 对这些位置点计算其所对应的最大信息视角,结果如图2-5所示,其中蓝色的点表 示空间中的特征点,绿颜色的箭头表示最大信息视角的方向,箭头颜色越亮,则 信息总量越大。图2-5 (a)所示的是 Zhang 等^[38]的方法所计算的结果,该方法先 对离散化的空间中的每一个位置计算一个和相机视角无关的核,然后对可能的视 角进行采样,选取能使信息量最大的视角作为最大信息视角。本文基于优化的方 法结果如图2-5 (b)所示。对比图2-5 (a)的结果,可以看到,在接近特征墙附近, 基于优化的方法计算的最大信息视角要更加平滑,而 Zhang 等^[38] 基于采样的方法则无法计算出正确的最大信息视角,原因是因为其相机可视化模型式(2-10) 赋予 了位于相机可视区域后方的特征点太大的权重。



a) 基于采样的最大信息视角计算结果^[38] b) 本文基于优化的最大信息视角计算结果 图 2-5 两种最大信息视角计算结果对比

在仿真实验中,两种方法的计算时间对比如表2-1所示,Zhang等^[38]的方法需要通过预处理对空间进行离散化,然后对每个离散化的体素计算信息核,这一步需要花费 18.28s,之后对每个离散的位置以 10 度为间隔对视角进行采样,计算最大信息视角,这一步需要花费 2.85s。本文基于优化的方法则不需要预处理过程,直接通过优化而不是采样的方法计算所有位置的最大信息视角,总共需要花费 40.43s,平均下来计算单个位置的最大信息视角需要 0.114s,时间主要花费在每次迭代后需要对 1000 个特征点求取新的残差和雅克比,在实际野外环境中,机器人四周不会有这么密集的特征点,通过将周围的特征点进行稀疏化处理可以显著降低计算量,提高计算速度。

2.2.2 基于最优视角的主动定位

在上一小节中,计算最大信息视角的一个假设是机器人已知周围环境中所有 特征点的位置。但是机器人刚开始对周围环境是一无所知的,机器人需要不断对 环境进行探索来感知环境中的特征点。在这种情况下,如果每次都让机器人看向 最大信息视角处,会导致机器人为了定位而放弃对环境的探索。

通过一个仿真实验来说明这个过程。如图2-6所示,机器人沿着黑线所示的圆

表 2-1 最大信息视角计算方法时间花费对比				寸比
	方法	预处理时间	计算最大信息视角时间	单个位置计算时间
	基于采样的方法	18.28s	2.85s	0.0598
	基于优化的方法	Os	40.43s	0.114

哈尔滨工业大学工学硕士学位论文

弧均匀走7步,记位姿为 $\{T_k\}_{k=1}^7$,每两步之间夹角为30度,机器人视场角设定为90度。机器人感知到的环境中的特征点 $p_{i=1}^N$ 用红色实心圆圈表示,蓝色直线表示机器人的运动方向 R_k^{motion} ,红色直线表示最大信息视角方向 R_k^{maxInfo} ,图2-6最后一幅小图中蓝色点表示场景中实际存在的所有特征点。

在位置 T_1 处,机器人将其正前方感知到的特征点加入到特征点地图 $p_{i=1}^N$ 中,机器人运动到了位置 T_2 处后,通过公式 (2-3)可以计算出最大信息视角 R_2^{maxInfo} ,设置相机视角 R_2^{camera} 为 R_2^{maxInfo} 并将新检测到的特征点加入到特征点地图 $p_{i=1}^N$ 中,然后机器人继续向下一个位置运动。整个过程如算法2-1所示。

对图2-6所示的定位结果进行分析可得,由于最大信息视角会迫使机器人一直 看向已经探索过的特征点丰富的地方,导致机器人无法探索新的场景,在位置 *T*。 处出现了机器人运动方向和视角方向完全相反的情况,导致机器人无法有效的对 前方环境进行观察。使用基于最大信息视角的定位方法,当机器人走过一个半圆



图 2-6 基于最大信息视角定位的仿真结果

哈尔滨工业大学工学硕士学位论文

时,对场景也只探索了四分之一多一点,由于机器人的定位精度与地图完整度和 准确度是密切相关的,这就会导致之后机器人的定位精度的下降直至定位失效。

算法 2-1 最大信息视角定位算法

Input: 机器人的运动轨迹:机器人在世界坐标系的 M 个位置 $\{T_k\}_{k=1}^{M}$ 和运 动方向 $\{R_k^{\text{motion}}\}_{k=1}^M$ **Output:** 建立的特征点的地图: $p_{i=1}^N$; 机器人 M 个时刻的最大信息视角 $\left\{ R_{k}^{\max Info} \right\}_{k=1}^{M}$ 和发给云台的控制视角 $\left\{ R_{k}^{\operatorname{camera}} \right\}_{k=1}^{M}$ 1 Initialization: $\{p_i^N\}_{i=1}^N \leftarrow empty$; 当前最大信息视角对应的信息矩阵值 $I_{\text{curr}}^{\text{maxInfo}} \leftarrow 0;$ **2** for k = 1 : M do update T_k and R_k^{motion} ; 3 if k = 1 then 4 $R_k^{\text{camera}} \leftarrow R_k^{\text{motion}}$; 5 $p_{i=1}^{N} \leftarrow \text{DetectFeatures}();$ 6 $R_k^{\text{maxInfo}} \leftarrow \text{CalculateMaxInfoView}(T_k, p_{i=1}^N);$ 7 $I_{\text{curr}}^{\max \text{Info}} \leftarrow \text{CalculateInfoMatrix}(T_k, R_k^{\max \text{Info}}, p_{i=1}^N);$ 8 else 9 $R_k^{\text{maxInfo}} \leftarrow \text{CalculateMaxInfoView}(T_k, p_{i=1}^N);$ 10 $I_{\text{curr}}^{\max \text{Info}} \leftarrow \text{CalculateInfoMatrix}(T_k, R_k^{\max \text{Info}}, p_{i=1}^N);$ 11 $R_k^{ ext{camera}} \leftarrow R_k^{ ext{maxInfo}};$ 12 $p_{i=1}^{N} \leftarrow p_{i=1}^{N} + \text{DetectFeatures}();$ 13 end 14 15 end

为了解决基于最大信息视角进行定位所带来的问题,平衡机器人的环境探索 (exploration)和定位(localization)的需求,本文设计了一种简单有效的最优视角 计算算法。该算法通过对机器人运动方向 *R*^{motion} 和最大信息方向 *R*^{maxInfo} 按照信息 量大小进行插值得到最优视角 *R*^{best},如公式(2-12)所示,

$$R_k^{\text{best}} = \alpha \cdot R_k^{\text{motion}} + (1 - \alpha) \cdot R_k^{\text{maxInfo}}$$
(2-12)

其中插值因子α由公式(2-13)计算得到。

算法 2-2 最优视角定位算法			
Input: 机器人的运动轨迹:机器人在世界坐标系的 M 个位置 $\{T_k\}_{k=1}^M$ 和运			
动方向 $\{R_k^{\text{motion}}\}_{k=1}^M$			
Output: 建立的特征点的地图: $p_{i=1}^N$; 机器人 M 个时刻的最优视角			
$\{R^{\text{best}}\}^{M}$,最大信息视角 $\{R^{\text{maxInfo}}\}^{M}$,和发给云台的控制视角			
$\{R_{k}^{\text{camera}}\}_{k=1}^{M}$			
1 Initialization: $\{n^N\}^N \leftarrow empty$: 当前最大信自视角对应的信自矩阵值			
$I^{\text{maxInfo}} = 0$: L-HJAEHJAHJAHAJCHALA			
$I_{curr} \leftarrow 0$, \bot 的刻敢忧忧用对应的同意地许但 $I_{last} \leftarrow 0$,			
2 IOF $k = 1$: <i>M</i> do			
3 update I_k and R_k^{indust} ;			
4 if $k = 1$ then			
5 $R_k^{\text{camera}} \leftarrow R_k^{\text{motion}};$			
6 $p_{i=1}^{N} \leftarrow \text{DetectFeatures}();$			
7 $R_k^{\max Info} \leftarrow \text{CalculateMaxInfoView}(T_k, p_{i=1}^N);$			
$I_{\text{curr}}^{\text{maxInfo}} \leftarrow \text{CalculateInfoMatrix}(T_k, R_k^{\text{maxInfo}}, p_{i=1}^N);$			
9 $I_{\text{last}}^{\text{best}} \leftarrow I_{\text{curr}}^{\text{maxInfo}};$			
10 $R_k^{\text{best}} \leftarrow R_k^{\text{maxInfo}}$			
11 else			
12 $R_k^{\max Info} \leftarrow \text{CalculateMaxInfoView}(T_k, p_{i=1}^N);$			
13 $I_{\text{curr}}^{\text{maxInfo}} \leftarrow \text{CalculateInfoMatrix}(T_k, R_k^{\text{maxInfo}}, p_{i=1}^N);$			
14 $\alpha \leftarrow I_{\text{last}}^{\text{best}} / (I_{\text{last}}^{\text{best}} + I_{\text{curr}}^{\text{maxInfo}});$			
15 $R_k^{\text{best}} \leftarrow \alpha \cdot R_k^{\text{motion}} + (1 - \alpha) \cdot R_k^{\text{maxInfo } \oplus};$			
16 $R_k^{\text{camera}} \leftarrow R_k^{\text{best}};$			
17 $p_{i=1}^{N} \leftarrow p_{i=1}^{N} + \text{DetectFeatures}();$			
18 $I_{\text{last}}^{\text{best}} \leftarrow \text{CalculateInfoMatrix}(T_k, R_k^{\text{best}}, p_{i=1}^N);$			
19 end			
20 end			

$$\alpha = I_{\text{last}}^{\text{best}} / (I_{\text{last}}^{\text{best}} + I_{\text{curr}}^{\text{maxInfo}})$$
(2-13)

如果上一位置按照最优视角可以获得较大的信息量 *I*^{best},则表明机器人当前定位 状况良好,那么应当优先考虑对环境进行探索。否则,如果 *I*^{best} 较小,那么表明 机器人当前定位状况不好,定位快要丢失了,这时应该让机器人看向最大信息方 向,防止定位丢失。最后用云台控制相机视角 *R*^{camera} 为 *R*^{best} 即可实现基于最优视 角控制的定位算法。该算法可以在保证定位效果的同时有效地完成对场景的探索。 基于最优视角的定位算法原理如算法2-2所示。

图2-7所示的仿真实验对该算法流程进行了说明。同样的,让机器人沿着黑线所示的圆弧均匀走7步,记为 $\{T_k\}_{k=1}^7$,每两步之间夹角为30度,机器人视场角设定为90度。机器人感知到的环境中的特征点 $p_{i=1}^N$ 用红色实心圆圈表示,其中蓝色直线指代机器人的运动方向 R_k^{motion} ,红色直线指代的是最大信息视角方向 R_k^{maxInfo} ,绿色直线代表的是最优视角方向 R_k^{best} ,图2-7最后一幅小图中蓝色点表示场景中实际存在的所有特征点。

对仿真结果图2-7进行分析可以发现,在位置 T_2 处最优视角 R_2^{best} 更加靠近最 大信息视角 R_2^{maxInfo} 而不是机器人的运动方向 R_2^{motion} ,这是因为在位置 T_1 处由于机 器人距离特征点墙较远,定位效果不好,所以在位置 T_2 更倾向于进行定位。在位 置 T_5 处,最优视角方向 R_5^{best} 更加靠近机器人的运动方向 R_5^{motion} ,这时因为在位置



图 2-7 基于最优视角定位的仿真结果

T₄处机器人距离特征点墙很近,定位效果好,因此在T₅处机器人更加偏向于对环 境进行探索。使用基于最优视角的定位方法,可以在保证稳定定位的同时实现对 环境更大范围的探索。从最终位置T₇处可以看到,当机器人围绕环境转半圈之后, 完成了对环境中一半以上特征点的检测,其检测的特征点数是基于最大信息视角 定位方法的两倍以上。

2.3 本章小结

本章先从整体上对本文所设计的用于两栖机器人在野外环境下进行自主视觉 定位的 AmphiActiveVIO 系统进行了介绍。主要介绍了该系统的两栖视觉惯导里程 计 (AmphiVIO)、主动曝光控制和主动视角调整模块的作用和互相之间的联系,其 中 AmphiVIO 模块用于实时估计机器人的位姿,实现自主定位;主动曝光控制模块 用于控制相机曝光时间,使图像不会出现过曝光或欠曝光的现象;主动视角调整 模块则用来保证机器人上坡时不会出现视角看向天空而导致的视野丢失问题。然 后详细介绍了主动视角调整模块,为了解决两栖机器人的视野丢失问题。然 后详细介绍了主动视角调整模块,为了解决两栖机器人的视野丢失问题。。 时,为了有效的平衡机器人的定位需求和环境探索需求,本文设计了一种最优视 角定位算法。该算法通过综合考虑机器人的运动方向和最大信息方向得到最优视 角。如果当前定位状况良好,则最优视角偏向于机器人的运动方向,加强机器人 对环境的探索。如果定位状况较差,视觉定位快要丢失了,则使最优视角偏向于 最大信息视角方向,保证视野中有足够的视觉信息,防止定位丢失。

第3章 主动曝光控制算法

野外环境中存在亮度很高的太阳光和亮度很低的树木阴影,整个场景亮度的 动态范围会非常大,这种亮度动态范围大的环境叫做 HDR (High Dynamic Range) 环境。由于普通相机能捕捉的亮度范围是有限制的,会导致如图3-1所示的图像中 部分区域过亮,部分区域过暗的问题,几乎无法从图像中提取有效的视觉信息。为 了解决相机动态范围小的问题,许多相机会使用自动曝光算法来改变曝光参数,但 相机自带的曝光控制算法通常是保持图像平均亮度或部分区域的平均亮度为给定 值,并没有考虑图像是否可以提取足够的视觉特征点。图3-2就是对自动曝光的图 片进行特征提取的结果,可以发现该图中拥有丰富纹理特征的树林部分处于欠曝 光状态,无法提取到任何有效的视觉特征点。因此,需要设计一种适应于野外复 杂光线环境的主动曝光控制算法,使图像中可以提取足够的视觉特征点。

本章首先介绍如何通过光度标定来获取相机的响应函数,然后介绍如何利用 图像梯度作为评价指标并使用优化的方法来主动控制相机的曝光时间,最后介绍 如何对图像进行光度矫正来保证前后帧图像的亮度一致性。



图 3-1 野外 HDR 环境

图 3-2 自动曝光下的特征检测

3.1 光度标定

为了主动控制相机的曝光时间,首先需要弄清楚曝光时间和最终图像亮度值 之间的函数关系。本节首先介绍相机的光度成像过程,建立光度成像模型,然后 介绍如何利用固定场景采集的多张不同曝光下的图像来离线标定光度成像模型的 参数。

3.1.1 相机光度成像过程

对于场景中的一个物体,其既有可能自己是光源,也有可能被其他光源照射 而反射光线到空间中。使用**辐射率**来描述从物体射出的所有光线,辐射率的物理 含义为每单位立体角每单位投射表面的辐射通量,单位为瓦特每球面度每平方米, 记为 *L*。假设这个物体射出的光线被像素坐标 *u* 对应的感光元件接收到,则该感 光元件每单位面积每单位时间接受到的能量大小叫做**辐照度**,单位为瓦特每平方 米,记为 *E*(*u*)。

如果场景中产生的反射都是漫反射,那么对于场景中拥有相同的辐射率的不同物体点,理想情况下它们在图像上应该拥有相同的辐照度。但是由于相机成像过程中存在由于镜头导致的光学晕影现象^[41],即光的强度由图像中心向周围逐渐减弱,导致即使相机拍摄一个亮度非常均匀的物体,图像中心和边缘的亮度值并不相等。当考虑光学晕影时,辐照度 E(u) 可以通过将场景点的辐射率 L 乘以一个与像素坐标 u 相关的晕影因子 $V: \Omega \rightarrow [0,1]$ 得到,其中 Ω 表示图像中所有像素可能的坐标:

$$E(u) = V(u)L \qquad u \in \Omega \tag{3-1}$$



图 3-3 光度成像过程

当以曝光时间 Δt 拍摄一副图像时, 传感器的辐照度 *E* 会在 Δt 的时间内累 积为辐射曝光量, $H(u) = \Delta t E(u)$ 。辐射曝光量 H(u) 通过相机响应函数 (CRF) $f : \mathbb{R} \rightarrow [0,255]$ 映射为图像亮度 I(u)。由于实际使用的相机感光范围是有限制的, 当辐射曝光量 H(u) 超过感光动态范围时, 对应图像点的亮度会被截断为 0 或 255。将场景中一个辐射率为 *L* 的物体点映射到图像亮度 *I* 的过程可以写成公式 (3-2), 整体成像过程如图3-3所示。

$$I = f(\Delta t V(u)L) \tag{3-2}$$

3.1.2 离线光度标定算法

光度标定的目标是确定相机响应函数 *f* 的具体形式。一种常用的离线标定方法^[33] 是通过不同曝光时间 Δ*t_j* 下静止且辐射率恒定场景捕获的多张图片计算得到。 图3-4为在不同曝光时间下对同一 HDR 场景采集的多张图片。



图 3-4 在不同曝光时间下对同一 HDR 场景采集的图片

假定场景是静止的,且辐射率恒定,则图像上每一个像素*i*所对应的物体点辐 照度 *E_i* 在不同曝光时间下都是相同的。定义 *I_{ij}* 为曝光时间 Δ*t_j* 的图像上索引为 *i* 处像素的亮度值。则公式 (3-2) 可以写成:

$$I_{ij} = f(E_i \Delta t_j) \tag{3-3}$$

如果假设相机响应曲线是单调的,则将式(3-3)取反并求其自然对数可得:

$$\ln f^{-1}\left(I_{ij}\right) = \ln E_i + \ln \Delta t_j \tag{3-4}$$

其中 *i* 遍历图像上的所有像素点, *j* 遍历所有的曝光时间。定义函数 $g = \ln f^{-1}$ 为 逆相机响应函数,则公式 (3-4) 可以写成:

$$g(I_{ij}) = \ln E_i + \ln \Delta t_j \tag{3-5}$$

由于 g(z) 的定义域为可能的像素亮度值,其取值范围为已知的离散值序列 {*I_{min}*, *I_{min}* + 1,..., *I_{max}* - 1, *I_{max}*},其中 *I_{min}* 为最小像素亮度值(一般为 0),*I_{max}* 为最大像素亮度值(一般为 255)。假设 N 为每副图像中选取的像素个数,P 为总 共拍摄的图片个数,则整个问题变为已知每一幅图像所选取像素的亮度 *I_{ij}* 和每一

幅图像的曝光时间 Δt_j , 估计 ($I_{max} - I_{min} + 1$) 个 g(I) 的值和 $N \uparrow E_i$ 的值来最小化式 (3-6) 所示的最小二乘问题。

$$O = \sum_{i=1}^{N} \sum_{j=1}^{P} \left[g\left(I_{ij} \right) - \ln E_i - \ln \Delta t_j \right]^2 + \lambda \sum_{I=I_{\min}+1}^{I_{\max}-1} g''(I)^2$$
(3-6)

该式的第一项为误差项,表示逆响应函数的估计值和实际测量值之间的误差; 第二项为平滑项,用于保证函数 g(I) 平滑。在这种离散形式中,使用 g''(I) = g(I - 1) - 2g(I) + g(I + 1) 来计算其导数。公式(3-6)是标准的线性最小二乘问题,可以 直接使用奇异值分解进行求解。但有一点需要注意的是,该式的解存在一个尺度 因子 α 上的不确定性,即如果 ln E_i 的所有值被 ln $E_i + \alpha$ 替换,同时函数 g 被 $g + \alpha$ 替换,则最终的目标函数不会发生变化。因此,为了使上式的解唯一,可以增加一 个额外的约束项 $g(I_{mid}) = 0$,其中 $I_{mid} = \frac{1}{2}(I_{min} + I_{max})$ 。另一方面,如图3-5所示, 由于 g(I) 通常在 I_{min} 和 I_{max} 处有很大的斜率,且不够平滑。为了解决这个情况, 可以使用如式(3-7)所示的权重函数 w(z) 来让目标函数更加关注整个曲线的中间 部分。

$$w(I) = \begin{cases} I - I_{\min} & \text{for } I \le \frac{1}{2} (I_{\min} + I_{\max}) \\ I_{\max} - I & \text{for } I > \frac{1}{2} (I_{\min} + I_{\max}) \end{cases}$$
(3-7)

加上权重函数后,离线光度标定的目标函数变为:

$$O = \sum_{i=1}^{N} \sum_{j=1}^{P} \left\{ w\left(I_{ij}\right) \left[g\left(I_{ij}\right) - \ln E_{i} - \ln \Delta t_{j}\right] \right\}^{2} + \lambda \sum_{I=I_{\min}+1}^{I_{\max}-1} \left[w(I)g''(I)\right]^{2}$$
(3-8)

为了求解 $N \wedge \ln E_i$ 的值以及函数 g 的 $(I_{max} - I_{min} + 1)$ 个采样值,只需要保证 $N \cdot P > (I_{max} - I_{min}) + N$ 即可。假定 $(I_{max} - I_{min}) = 255$, P = 11,则选取 N = 50 对于求解整个最小二乘问题是足够的,没有必要选取所有的像素点,反而增加了 求解的计算量。另外对于像素点的选取最好要保证其在图像上分布均匀同时能够 覆盖到足够范围的亮度值。

图3-5是利用上述离线光度标定方法对图3-4所示的曝光序列进行标定的结果,标定分别在 RGB 三个通道进行,从图中可以看到三个颜色通道的曲线在中间部分完全重合,但在两端部分略有差异。同时整个逆相机响应函数在像素亮度值接近 *I_{min}*和 *I_{max}*的地方斜率很大,这和感光芯片动态范围具有截断特性是相一致。



图 3-5 离线光度标定的逆相机响应函数 g

3.2 主动曝光控制

普通相机的自动曝光算法可以在光线变化时,通过改变曝光参数来自适应调整图像的亮度。但该方法往往是通过调节曝光参数来保证图像中选定区域的亮度 维持到恒定值,并没有考虑到图像信息的缺失。如图3-6(a)所示,虽然相机的自动曝光算法让图像看起来更加自然,但是会导致图像有纹理的区域过暗,整副图像只能检测到 230 个 FAST 角点。相反,主动曝光控制^[32]以最大化图像中的视觉 信息量为目的来调整曝光时间,如图3-6(b)所示,使用主动曝光的方法虽然太阳 光部分看起来不是很自然,但其有效地保留了场景中的视觉信息,总共能检测到 598 个 FAST 角点。

在上一节中,通过离线光度标定确定了式(3-2)的具体形式,即图像亮度和 曝光时间之间的一一对应关系。为了能够对曝光时间进行动态调整,需要设计一 种评价指标来判断图像中视觉信息的丰富程度,然后设计一种曝光控制算法来依 照该评价指标计算最合适的曝光时间。



 a) 自动曝光下检测 FAST 角点
 b) 主动曝光下检测 FAST 角点

 图 3-6 自动曝光和主动曝光下 FAST 角点检测对比

3.2.1 基于梯度的图像评价指标

在计算机视觉领域,人们往往使用图像的梯度来表示图像的细节,因为梯度 较大的地方能更有效的进行特征点的提取和追踪。给定一副曝光时间为 Δt 的图像 I,则在像素 u 处的图像梯度大小由式 (3-9) 给定,其中 $\nabla I(\cdot) = \left[\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}\right]^{\mathsf{T}}$ 。

$$G(\mathbf{I}, \mathbf{u}, \Delta t) = \|\nabla \mathbf{I}(\mathbf{u}, \Delta t)\|^2$$
(3-9)

一种简单衡量图像细节丰富程度的评价指标是直接计算图像上所有像素点的 梯度大小之和:

$$M_{sum} = \sum_{\mathbf{u}_i \in I} G\left(\mathbf{u}_i\right) \tag{3-10}$$

但是, Shim 等^[31] 指出,由于真实图像的大部分图像区域梯度都很小,同时还会存在噪声带来的不正确的梯度。为了平衡弱梯度值和强梯度值的不平衡,同时滤除噪声影响, Shim 等提出使用梯度信息而不使用梯度值,在图像位置 *u*_i 处梯度信息 *m*_{ui} 的计算公式如下式所示:

其中 *Ğ* 表示将梯度幅值 *G* 归一化到范围 [0,1] 的结果; *σ* 是用来滤除噪声影响的 阈值; λ 参数用来调节强梯度值和弱梯度值之间的关系。则评价指标可以设为图像 所有像素的梯度信息求和,实际上就是所有梯度值的加权和。

$$M_{shim} = \sum_{\mathbf{u}_i \in I} m_{\mathbf{u}_i} \tag{3-12}$$

Zhang 等^[32] 通过实验发现,使用所有梯度值在一个固定百分比处的值是一个

更加鲁棒的估计,因此提出了如公式(3-13)所示的梯度百分比指标。当p = 0.5时, M_{perc} 就指代的是所有图像像素梯度值的中位数。

$$M_{perc}(p) = \text{ percentile } \left(\{ G(\mathbf{u}_i) \}_{\mathbf{u}_i \in I}, p \right)$$
(3-13)

但是上述的百分比评价指标是不可导的,为了之后的优化求解,需要定义一种 可导的百分比指标。首先将所有的梯度值按照升序进行排序,排完序的梯度值 记为 {*G*_{ith}}_{i∈[0,S]},其中,*S* 表示图像中所有的像素点,则可计算可导百分比指标 *M*_{derivedPerc}(*p*) 为排好序的梯度值的权重和:

$$M_{derivedPerc}(p) = \sum_{i \in [0,S]} W_{ith}(p) \cdot G_{ith}$$
(3-14)

权重 {W_{ith}(p)}_{i∈[0,S]} 定义为

$$W_{i\text{th}} = \begin{cases} \frac{1}{N} \sin\left(\frac{\pi}{2p \cdot S}i\right)^k, & i \le p \cdot S\\ \frac{1}{N} \sin\left(\frac{\pi}{2} - \frac{\pi}{2}\frac{i - p \cdot S}{S - p \cdot S}\right)^k, & i > p \cdot S \end{cases}$$
(3-15)

其中 N 为归一化参数,用于将权重和归一化到 1。直观上可导百分比指标是百分 比指标的一个近似, k 值越大,近似效果越好。



将梯度映射到梯度信息的函数 *m*_{ui} 在不同 λ 参数下的曲线如图3-7(a) 所示, 可导百分比指标 *M*_{derived}*Perc* 中的权重函数 *W*_{ith} 在不同百分比 *p* 下的曲线如图3-7 (b) 所示。为了直观的对四种评价指标进行比较,我们在一个 HDR 环境下拍摄多 种不同曝光时间的图片,然后计算在每个曝光时间下,四种不同评价指标的值。结 果如图3-8所示,其中相机曝光时间从 0ms 变化到 10ms,四种评价指标几乎都在曝
光时间 3.1ms 处达到最大值,记使评价指标最大的曝光时间为最优曝光时间。从 四种评价指标的曲线中可以发现,使用百分比的评价指标比使用梯度和指标更加 平滑。同时, *p* = 0.8 的百分比指标相对于 *p* = 0.5 的百分比指标,低曝光和最优曝 光时间的评价指标差距更大,对比更明显。对最优曝光时间 3.1ms 处的图像以及 曝光时间为 0.6ms 和 8ms 的图像进行 FAST 角点提取,如图3-9所示,从图中可以 看到 FAST 角点特征数量与评价指标是相一致,曝光时间为 3.1ms 的图像其评价 指标最大,其也保留了最多的图像细节,能检测到最多的 FAST 角点。



3.2.2 基于优化的曝光控制算法

在上两小节中,首先对相机的响应函数进行了离线标定,然后设计了一种基 于梯度的评价指标来衡量图像细节的丰富程度,也就是得到了评价指标和曝光时



曝光时间**0.6ms** Fast特征数522

曝光时间3.1ms Fast特征数2977

曝光时间8ms Fast特征数2773

图 3-9 不同曝光时间的图像提取 FAST 角点的结果

间之间的函数关系,可以使用标准的梯度上升算法通过最大化评价指标来求解最 优曝光时间。为了进行梯度上升,需要计算评价指标相对于曝光时间的导数,可 以使用链式法则来进行计算。首先计算图像梯度相对于曝光时间的导数,然后计 算评价指标相对于图像梯度的导数。

由公式 (3-9) 可以求得梯度幅值 G(·) 对曝光时间 Δt 的导数如式 (3-16) 所示。

$$\frac{\partial G(\cdot)}{\partial \Delta t} = 2\nabla \mathbf{I}(\mathbf{u}, \Delta t)^{\top} \frac{\partial}{\partial \Delta t} [\nabla \mathbf{I}(\mathbf{u}, \Delta t)]$$
(3-16)

该式的第一项为图像的梯度,第二项可以用施瓦兹定理(梯度的导数等于导数的 梯度)重写为:

$$\frac{\partial}{\partial \Delta t} [\nabla \boldsymbol{I}(\mathbf{u}, \Delta t)] = \nabla \left[\frac{\partial}{\partial \Delta t} \boldsymbol{I}(\mathbf{u}, \Delta t) \right]$$
(3-17)

定义 $g = \ln f^{-1}$,其中f是相机响应函数,表示将曝光量 $E\Delta t$ 映射到图像亮度I;g为逆相机响应函数,表示将图像亮度I映射为对数曝光量 $\ln(E\Delta t)$ 。离线标定完成后,使用多项式拟合可以得到g和f的函数表达式以及其导数g'和f'。函数g和函数f有如式(3-18)所示的关系:

$$f^{-1} = e^{g}$$

$$[f^{-1}]' = [e^{g}]'$$

$$= e^{g} \cdot g'$$

$$= f^{-1} \cdot g'$$
(3-18)

公式 (3-17) 右侧梯度符号内实际上是相机响应函数相对于曝光时间的偏导数。因此对于一个亮度值为 *I_i* 的像素 *u_i*,其亮度值相对于曝光时间的偏导数可以由式 (3-19) 计算得到,

$$\frac{\partial I_i}{\partial \Delta t} = f' \left[f^{-1}(I_i) \right] V(\mathbf{u}_i) L(\mathbf{u}_i)
= \frac{V(\mathbf{u}_i) L(\mathbf{u}_i)}{\left[f^{-1} \right]'(I_i)}
= \frac{V(\mathbf{u}_i) L(\mathbf{u}_i)}{f^{-1}(I_i) g'(I_i)}
= \frac{1}{g'(I_i) \Delta t}$$
(3-19)

将式(3-19)插入到式(3-16)中即可得梯度幅值对曝光时间的导数:

$$\frac{\partial G(\cdot)}{\partial \Delta t} = 2 [\nabla \boldsymbol{I}(\cdot)]^{\mathsf{T}} \nabla \left[\frac{1}{g'(\boldsymbol{I}(\cdot))\Delta t} \right]$$
(3-20)

其中 g'(I(·)) 的含义是将 g' 应用在图像 I 的所有像素上。

由于可导百分比指标 *M*_{derivedPerc} 是图像所有像素的梯度幅值的加权和,故该指标相对于曝光时间的导数可以由下式得到:

$$\frac{\partial M_{\text{derivedPerc}}}{\partial \Delta t} = \sum_{i \in [0,S]} W_{i \text{ th}} \frac{\partial G_{i \text{ th}}}{\partial \Delta t}$$
(3-21)

主动曝光控制的目标是最大化下一副图像的可导百分比指标,给定一副图像 I 以及其相对应的曝光时间 Δt,则下一副图像的最优曝光时间 Δt_{next} 可以通过梯 度上升的方式计算得到:

$$\Delta t_{next} = \Delta t + \gamma \frac{\partial M_{\text{derivedPerc}}}{\partial \Delta t}$$
(3-22)

其中 γ 是步长大小,用来控制更新的幅度。主动控制相机的曝光时间为 Δt_{next} 可 以得到下一副图像,然后继续按照式(3-22)计算下下副图像所需的曝光时间,依 次类推即可实现主动曝光控制。

3.2.3 自适应光度矫正

视觉惯导里程计的前端通过对前后帧进行光流跟踪来实现稳定的特征匹配, 光流跟踪的一个前提假设是灰度不变假设^[42],即同一个空间点的像素灰度值在各 个图像中是不变的。如果灰度不变假设成立,那么对于*t*时刻位于(*x*,*y*)处的像素, 假设在*t* + *dt* 时刻它运动到(*x* + *dx*,*y* + *dy*)处,则有:

$$I(x + dx, y + dy, t + dt) = I(x, y, t).$$
(3-23)

但在实际使用中,灰度不变假设很可能不成立,特别是在野外复杂的光线环

境下,由于场景中的阳光和树木阴影的影响,相机会不断调整曝光参数,使得图像变亮或者变暗,从而破坏灰度不变假设。使用主动曝光控制的优点不仅仅在于可以控制曝光参数使图像中有更丰富的视觉特征,更重要的是可以利用相机响应函数和设置的曝光时间对采集的图像进行光度矫正,这对基于光流跟踪的视觉里程计前端是很重要的。

视觉里程计前端光流跟踪的目的在于对相邻帧的特征点进行匹配。已知参考 帧图像中待跟踪特征点周围的图像块 *P*,当前帧图像 *I* 和待跟踪特征点在当前帧 图像上的初始估计位置 *u*'。通过最小化公式(3-24)中的光度误差计算待追踪特征 点在当前帧的位置:

$$\arg\min_{\delta \mathbf{u}} \sum_{\Delta \mathbf{u} \in P} \left[P(\Delta \mathbf{u}) - I \left(\mathbf{u}' + \delta \mathbf{u} + \Delta \mathbf{u} \right) \right]^2$$
(3-24)

其中 Δu 在模板 P 内进行遍历。最终追踪上的特征点在当前帧的位置为 u' + δu。如 果已知参考帧和当前帧的曝光时间以及相机响应函数,则可以使用下式(3-25)进 行带有光度矫正的光流跟踪。

$$\arg\min_{\delta \mathbf{u}} \sum_{\Delta \mathbf{u} \in P} \left[f \left(\frac{\Delta t_c}{\Delta t_r} f^{-1}(P(\Delta \mathbf{u})) - \mathbf{I} \left(\mathbf{u}' + \delta \mathbf{u} + \Delta \mathbf{u} \right) \right]^2$$
(3-25)

其中 Δt_c 表示参考帧的曝光时间, Δt_r 表示当前帧的曝光时间, f 表示相机响应函数。实际使用中可以直接将整副图像进行光度矫正,但这样会增大计算量。

3.3 本章小结

本章介绍了一种主动曝光控制算法,用于解决野外环境的复杂光线变化对视 觉惯导里程计的影响。当相机看向场景中同时有阳光和树木阴影的 HDR 环境时, 由于相机的感光范围有限,无法完整记录整个场景的亮度信息,相机自带的自动 曝光算法会自动调节曝光参数使图像看起来更加自然,但这样会使图像过暗或过 亮,导致视觉信息丢失。为了解决这个问题,首先对相机响应函数进行离线标定, 然后利用一种梯度百分比指标对图像中含有的视觉信息量大小进行量化,最后求 出梯度百分比指标相对于曝光时间的导数,使用梯度上升的方法来更新相机的最 优曝光时间。为了解决由于曝光时间变化所导致的灰度不变假设不成立的问题,使 用每一帧图像的曝光时间和相机响应函数对图像进行光度矫正,在矫正后的图像 上使用光流跟踪。使用主动曝光控制算法,可以有效地保证相机采集的图片中含 有足够的视觉信息,防止由于野外光线导致的图像过暗或过亮问题。

第4章 两栖视觉惯导里程计

含有丰富视觉信息的图像结合惯性测量单元输出的加速度和角速度数据可以 构建视觉惯导里程计(Visual-Inertial Odometey)对机器人的位姿进行估计,实现 两栖机器人的定位。相对于移动机器人而言,两栖机器人在野外地形下的运动形 式更加复杂。当两栖机器人进行快速运动时,其视角会出现明显的上下俯仰运动, 这会影响视觉里程计前端光流跟踪的稳定性。为了解决两栖机器人俯仰运动和野 外起伏不定的地形对视觉惯导里程计的影响。本文提出了一种两栖视觉惯导里程 计(AmphiVIO),对视觉惯导里程计Vins-Mono^[30]的前端光流跟踪部分进行改进, 增加了新检测特征点和局部地图的匹配关系,通过额外的视觉约束来增加整个系 统的稳定性,防止定位丢失。整个 AmphiVIO 系统框图如图4-1所示,主要分为视 觉惯导预处理,初始化,局部地图匹配和紧融合的 VIO 优化四部分。本章首先介 绍 AmphiVIO 的预处理部分如何对相邻帧进行特征追踪并对 IMU 进行预积分,然 后介绍初始化模块如何利用预处理的数据为整个系统待估计的状态提供一个初始 值,然后介绍如何通过和局部地图匹配获得额外的视觉约束,最后介绍如何执行 一个基于滑动窗的紧融合非线性优化来估计所有的状态量实现两栖机器人的自主 视觉定位。



图 4-1 AmphiVIO 原理图

4.1 视觉惯导预处理

相机直接输出的是低频高维的图像数据, IMU 输出的是高频带噪声的三轴加

速度和三轴角速度数据。对于视觉图像数据,在连续图像帧中追踪特征点的位置,如果有特征追踪丢失,则在当前图像帧中检测一些新的特征点。对于 IMU 数据,则利用 IMU 预积分技术将连续两图像帧之间所有的 IMU 数据积分成一个测量值,便于后续处理。

每新来一副图像,首先使用3.2.3小节的方法对新图像进行光度矫正,然后使 用带金字塔的 KLT 稀疏光流^[43] 对前后帧的特征点进行跟踪。与此同时,为了保证 被跟踪的特征点维持在一个稳定的数目 Max_{cnt},当跟踪特征点数小于 Max_{cnt} 时, 需要检测新的特征点^[44]。由于后端计算资源有限,无法对每一帧图像都进行优化, 因此需要选取关键帧,将这些关键帧保持在滑动窗中,而将那些提供不了太多信 息的帧丢弃。关键帧的选取有两个准则,第一个是当前帧与上一帧的平均视差大 于某一个阈值 (10 个像素),则将该帧选为关键帧。由于旋转也可以造成视差,但 是特征点无法在纯旋转时候进行三角化,因此在计算视差时使用 IMU 估计的旋转 来补偿由于旋转造成的视差。第二个是如果追踪的特征点数量低于某一个阈值,则 将这一帧作为关键帧,防止追踪丢失。

一般而言, IMU 的频率比相机的频率高很多,这样相邻关键帧之间会存在大量的 IMU 测量数据,如图4-2所示。为了处理相邻关键帧之间大量的 IMU 测量数据,通常的做法是在世界坐标系下将这些数据积分成一个观测值。但使用这种方式会导致最后的积分项中包含上一时刻的状态量,当优化过程中状态量发生变化时,需要重新积分,增加了不必要的计算量。Lupton 等^[45]提出的预积分方法可以将积分的坐标系从固定的世界坐标系转换到了上一图像帧的相对坐标系。转化后只需要提前计算好积分量,在优化的过程中并不需要重新计算。使用预积分的策略减少了后端非线性优化的计算量,同时也提高了积分的精度。接下来介绍 AmphiVIO中使用的预积分方法,主要参考了 Vins-Mono^[30],关于 IMU 预积分理论更详细的介绍建议读者参考 Forster 等的文章^[28]。

从陀螺仪和加速度计获得的 t 时刻的原始数据 ût 和 ât 可以写成如下形式:

$$\hat{a}_{t} = a_{t} + b_{a_{t}} + R_{w}^{\nu} g^{w} + n_{a}$$

$$\hat{\omega}_{t} = \omega_{t} + b_{w_{s}} + n_{w}$$
(4-1)

加速度计测量的是除去重力后物体受到的合力 $a_t + R_w^b g^w$,同时还受到加速度计偏置 b_a 和测量噪声 n_a 的影响。陀螺仪测量的是物体的角速度 ω_t ,同时受到陀螺仪偏置 tb_w 和测量噪声 n_w 的影响。其中测量噪声服从高斯分布, $n_a \sim N(0, \sigma_a^2), n_w \sim$



图 4-2 相机和 IMU 频率示意图

 $N(0,\sigma_w^2)$,加速度计偏置和陀螺仪偏置被建模为随机游走噪声,其导数服从高斯分布:

$$\boldsymbol{n}_{b_a} \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{\sigma}_{b_a}^2\right), \quad \boldsymbol{n}_{b_w} \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{\sigma}_{b_w}^2\right)$$

$$\dot{\boldsymbol{b}}_{a_t} = \boldsymbol{n}_{b_a}, \quad \dot{\boldsymbol{b}}_{w_t} = \boldsymbol{n}_{b_w}$$
(4-2)

给定相邻的两个图像时刻 b_k 和 b_{k+1} 之间 IMU 的测量数据,则 IMU 在世界坐 标系下的位置 p,速度 v 和姿态 q 可以由加速度计和陀螺仪的测量值积分得到:

$$\boldsymbol{p}_{b_{k+1}}^{w} = \boldsymbol{p}_{b_{k}}^{w} + \boldsymbol{v}_{b_{k}}^{w} \Delta t_{k} + \iint_{t \in [t_{k}, t_{k+1}]} \left(\boldsymbol{R}_{t}^{w} \left(\hat{\boldsymbol{a}}_{t} - \boldsymbol{b}_{a_{t}} - \boldsymbol{n}_{a} \right) - \boldsymbol{g}^{w} \right) dt^{2}$$

$$\boldsymbol{v}_{b_{k+1}}^{w} = \boldsymbol{v}_{b_{k}}^{w} + \int_{t \in [t_{k}, t_{k+1}]} \left(\boldsymbol{R}_{t}^{w} \left(\hat{\boldsymbol{a}}_{t} - \boldsymbol{b}_{a_{t}} - \boldsymbol{n}_{a} \right) - \boldsymbol{g}^{w} \right) dt$$

$$\boldsymbol{q}_{b_{k+1}}^{w} = \boldsymbol{q}_{b_{k}}^{w} \otimes \int_{t \in [t_{k}, t_{k+1}]} \frac{1}{2} \boldsymbol{\Omega} \left(\hat{\omega}_{t} - \boldsymbol{b}_{w_{t}} - \boldsymbol{n}_{w} \right) \boldsymbol{q}_{t}^{b_{k}} dt$$

$$(4-3)$$

其中 Δt_k 表示的是相邻两帧 t_k 和 $t_k + 1$ 之间的时间间隔:

$$\Omega(\omega) = \begin{bmatrix} -\lfloor \omega \rfloor_{\times} \omega \\ -\omega^T & 0 \end{bmatrix}, \ \lfloor \omega \rfloor_{\times} = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}$$
(4-4)

由公式(4-3)可得,要获得 b_{k+1} 时刻的 IMU 的位置、速度和姿态,需要从 b_k 时刻的状态进行递推。如果在非线性优化的每一步迭代中更新了 b_k 时刻的状态,则需要将 b_k 到 b_{k+1} 时刻之间所有 IMU 的测量值积分才能得到 b_{k+1} 时刻的状态,这对有限的计算资源造成了不必要的负担。使用如公式(4-5)的预积分算法,通过将参考坐标系从世界坐标系变换到局部坐标系 b_k,则可以只积分和加速度测量值 û 有关的部分,避免重新积分。

$$\boldsymbol{R}_{w}^{b_{k}}\boldsymbol{p}_{b_{k+1}}^{w} = \boldsymbol{R}_{w}^{b_{k}} \left(\boldsymbol{p}_{b_{k}}^{w} + \boldsymbol{v}_{b_{k}}^{w} \Delta t_{k} - \frac{1}{2}\boldsymbol{g}^{w} \Delta t_{k}^{2}\right) + \boldsymbol{\alpha}_{b_{k+1}}^{b_{k}}$$

$$\boldsymbol{R}_{w}^{b_{k}}\boldsymbol{v}_{b_{k+1}}^{w} = \boldsymbol{R}_{w}^{b_{k}} \left(\boldsymbol{v}_{b_{k}}^{w} - \boldsymbol{g}^{w} \Delta t_{k}\right) + \boldsymbol{\beta}_{b_{k+1}}^{b_{k}}$$

$$\boldsymbol{q}_{w}^{b_{k}} \otimes \boldsymbol{q}_{b_{k+1}}^{w} = \boldsymbol{\gamma}_{b_{k+1}}^{b_{k}}$$

$$(4-5)$$

其中:

$$\alpha_{b_{k+1}}^{b_k} = \iint_{t \in [t_k, t_{k+1}]} \mathbf{R}_t^{b_k} \left(\hat{a}_t - b_{a_t} - n_a \right) dt^2$$

$$\beta_{b_{k+1}}^{b_k} = \int_{t \in [t_k, t_{k+1}]} \mathbf{R}_t^{b_k} \left(\hat{a}_t - b_{a_t} - n_a \right) dt$$

$$\gamma_{b_{k+1}}^{b_k} = \int_{t \in [t_k, t_{k+1}]} \frac{1}{2} \mathbf{\Omega} \left(\hat{\omega}_t - b_{w_t} - n_w \right) \gamma_t^{b_k} dt$$

(4-6)

除了加速度计和陀螺仪的偏置 \boldsymbol{b}_{a_t} 和 \boldsymbol{b}_{w_t} ,公式 (4-6)中的预积分项 $\boldsymbol{a}_{b_{k+1}}^{b_k}$, $\boldsymbol{\beta}_{b_{k+1}}^{b_k}$ 和 $\boldsymbol{\gamma}_{b_{k+1}}^{b_k}$,可以直接通过 IMU 的测量值积分得到,而不用考虑 b_k 和 b_{k+1} 时刻其他的状态量。当优化后的传感器偏置 \boldsymbol{b}_{a_t} 和 \boldsymbol{b}_{w_t} 变化较小时,可以利用 $\boldsymbol{a}_{b_{k+1}}^{b_k} \boldsymbol{\beta}_{b_{k+1}}^{b_k} \boldsymbol{\gamma}_{b_{k+1}}^{b_k}$ 相对于偏置的一阶近似来修正 $\boldsymbol{a}_{b_{k+1}}^{b_k} \boldsymbol{\beta}_{b_{k+1}}^{b_k} \boldsymbol{\gamma}_{b_{k+1}}^{b_k}$ 的值,无需重新积分。上式积分为连续形式,实际计算中需要使用数值积分进行离散化。离散化形式如式(4-7)所示,其中 i 表示 IMU 的测量时刻, δt 表示两个 IMU 测量 i和 i+1之间的时间间隔。

$$\hat{\boldsymbol{\alpha}}_{i+1}^{b_{k}} = \hat{\boldsymbol{\alpha}}_{i}^{b_{k}} + \hat{\boldsymbol{\beta}}_{i}^{b_{k}} \delta t + \frac{1}{2} \boldsymbol{R} \left(\hat{\boldsymbol{\gamma}}_{i}^{b_{k}} \right) \left(\hat{\boldsymbol{a}}_{i} - \boldsymbol{b}_{a_{i}} \right) \delta t^{2}$$

$$\hat{\boldsymbol{\beta}}_{i+1}^{b_{k}} = \hat{\boldsymbol{\beta}}_{i}^{b_{k}} + \boldsymbol{R} \left(\hat{\boldsymbol{\gamma}}_{i}^{b_{k}} \right) \left(\hat{\boldsymbol{a}}_{i} - \boldsymbol{b}_{a_{i}} \right) \delta t$$

$$\hat{\boldsymbol{\gamma}}_{i+1}^{b_{k}} = \hat{\boldsymbol{\gamma}}_{i}^{b_{k}} \otimes \begin{bmatrix} 1 \\ \frac{1}{2} \left(\hat{\boldsymbol{\omega}}_{i} - \boldsymbol{b}_{w_{i}} \right) \delta t \end{bmatrix}$$

$$(4-7)$$

在多传感器融合的状态估计中,对于每一个传感器源,除了该传感器提供的测量值外,往往还需要知道该传感器源的可信度,数学表示即为协方差。当将多 个时刻的 IMU 测量值预积分成一个值时,实际上测量误差也进行了累计,协方差 发生了传播,因此需要对协方差的传播过程进行建模。

由于一个旋转实际上只有三个自由度,当使用一个四维的旋转四元数 $\gamma_t^{b_k}$ 进行表示时出现了过参数化,给优化问题带来额外的约束,增加了求解的难度。因此可以定义其误差项为一个扰动项^[40],其中 $\delta \theta_t^{b_k}$ 是一个微小的三维扰动。

$$\boldsymbol{\gamma}_{t}^{b_{k}} \approx \hat{\boldsymbol{\gamma}}_{t}^{b_{k}} \otimes \begin{bmatrix} 1\\ \frac{1}{2}\delta\boldsymbol{\theta}_{t}^{b_{k}} \end{bmatrix}$$
(4-8)

使用微小的三维扰动来表示姿态估计误差,则可推导误差导数的连续形式如 下式(4-9)所示。

$\begin{bmatrix} \delta \dot{\boldsymbol{\alpha}}_{t}^{b_{k}} \\ \delta \dot{\boldsymbol{\beta}}_{t}^{b_{k}} \\ \delta \dot{\boldsymbol{\theta}}_{t}^{b_{k}} \\ \delta \dot{\boldsymbol{b}}_{t}^{b_{k}} \\ \delta \dot{\boldsymbol{b}}_{t}^{b_{k}} \end{bmatrix}$	=	$\begin{bmatrix} 0 \ \boldsymbol{I} & 0 \\ 0 \ 0 - \boldsymbol{R}_{t}^{b_{k}} \left[\hat{\boldsymbol{a}}_{t} - \boldsymbol{b}_{a_{t}} \right] \\ 0 \ 0 & - \left[\hat{\boldsymbol{\omega}}_{t} - \boldsymbol{b}_{w_{t}} \right]_{\times} \\ 0 \ 0 & 0 \\ 0 \ 0 & 0 \end{bmatrix}$	$ \begin{array}{ccc} 0 & 0 \\ \downarrow_{\times} - \boldsymbol{R}_{t}^{b_{k}} & 0 \\ \langle & 0 & -\boldsymbol{I} \\ 0 & 0 \\ 0 & 0 \end{array} \right \begin{bmatrix} \delta \boldsymbol{\alpha}_{t}^{b_{k}} \\ \delta \boldsymbol{\beta}_{t}^{b_{k}} \\ \delta \boldsymbol{b}_{a_{t}}^{b_{k}} \\ \delta \boldsymbol{b}_{a_{t}} \\ \delta \boldsymbol{b}_{w_{t}} \end{bmatrix} $	(4-9)
	+	$\begin{bmatrix} 0 & 0 & 0 & 0 \\ -\boldsymbol{R}_{t}^{b_{k}} & 0 & 0 & 0 \\ 0 & -\boldsymbol{I} & 0 & 0 \\ 0 & 0 & \boldsymbol{I} & 0 \\ 0 & 0 & 0 & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{n}_{a} \\ \boldsymbol{n}_{w} \\ \boldsymbol{n}_{b_{a}} \\ \boldsymbol{n}_{b_{w}} \end{bmatrix}$	$= \boldsymbol{F}_t \delta \boldsymbol{z}_t^{b_k} + \boldsymbol{G}_t \boldsymbol{n}_t$	(+-7)

k 时刻没有积分的预积分项协方差 $P_{b_k}^{b_k} = 0$,每两次 IMU 测量的时间间隔为 δt ,则积分后的预积分项协方差 $P_{b_{k+1}}^{b_k}$ 可以由下式 (4-10) 递推得到;

$$\boldsymbol{P}_{t+\delta t}^{b_{k}} = (\boldsymbol{I} + \boldsymbol{F}_{t}\delta t)\boldsymbol{P}_{t}^{b_{k}}(\boldsymbol{I} + \boldsymbol{F}_{t}\delta t)^{T} + (\boldsymbol{G}_{t}\delta t)\boldsymbol{Q}(\boldsymbol{G}_{t}\delta t)^{T} t \in [k, k+1]$$
(4-10)

其中 Q 是噪声 $\left(\sigma_{a}^{2}, \sigma_{w}^{2}, \sigma_{b_{a}}^{2}, \sigma_{b_{w}}^{2}\right)$ 构成的对角协方差矩阵。 由式 (4-5) 可以写出使用预积分算法后的 IMU 测量模型,

$$\begin{bmatrix} \hat{\boldsymbol{x}}_{b_{k+1}}^{b_k} \\ \hat{\boldsymbol{\beta}}_{b_{k+1}}^{b_k} \\ \hat{\boldsymbol{y}}_{b_{k+1}}^{b_k} \\ \boldsymbol{\hat{y}}_{b_{k+1}}^{b_k} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} \boldsymbol{R}_w^{b_k} \left(\boldsymbol{p}_{b_{k+1}}^w - \boldsymbol{p}_{b_k}^w + \frac{1}{2} \boldsymbol{g}^w \Delta t_k^2 - \boldsymbol{v}_{b_k}^w \Delta t_k \right) \\ \boldsymbol{R}_w^{b_k} \left(\boldsymbol{v}_{b_{k+1}}^w + \boldsymbol{g}^w \Delta t_k - \boldsymbol{v}_{b_k}^w \right) \\ \boldsymbol{q}_{b_k}^{w_{k}^w} \otimes \boldsymbol{q}_{b_{k+1}}^w \\ \boldsymbol{b}_{ab_{k+1}} - \boldsymbol{b}_{ab_k} \\ \boldsymbol{b}_{wb_{k+1}} - \boldsymbol{b}_{wb_k} \end{bmatrix}$$
(4-11)

其中预积分观测值 $\hat{a}_{b_{k+1}}^{b_k}$, $\hat{\beta}_{b_{k+1}}^{b_k}$, $\hat{\gamma}_{b_{k+1}}^{b_k}$ 和相对应的协方差矩阵 $P_{b_{k+1}}^{b_k}$ 可以分别由公式 (4-7) 和 (4-10) 递推得到。

4.2 AmphiVIO 系统初始化

单目紧融合视觉惯性里程计是一个高度非线性系统,系统中的状态量需要良好的初始化才能进行后续的估计。对单目相机而言,将相机的运动和场景的大小同时扩大两倍,相机拍摄的图像不会发生变化,因此无法直接从单目相机拍摄的图像中恢复尺度信息,尺度信息的恢复只能依赖IMU。本文参考Vins-Mono^[30]使用一种松耦合的方法来获得状态量的初始值,首先通过运动恢复结构(Struct from Motion, SFM)^[46]重构不带尺度的相机位姿和特征点位置。然后将IMU 积分得到的轨迹和

SFM 重构的轨迹进行对齐来恢复尺度,重力,速度和陀螺仪偏置。AmphiVIO 系 统初始化的示意图如图4-3所示。



图 4-3 初始化过程中视觉惯导对齐说明

在初始化过程中,为了使计算时间限制在一个合理的值上,维护一个固定大小(一般 10 个关键帧)的滑动窗口。初始化的第一步是进行 SFM 重构,通过判断 滑动窗中是否存在一帧 c_i 和最新帧 c_0 有足够的匹配点(一般超过 30 个)和足够的视差(一般超过 20 个像素)。如果存在,则利用对极几何恢复这两帧之间的旋转 $q_{c_i}^{c_0}$ 和不带尺度的平移 $\overline{p}_{c_i}^{c_0}$ 。接下来通过三角化来初始化这两帧匹配好的特征点的 3D 位置。基于这些三角化后的特征点,可以使用 perspective-n-point(PnP) 算法来估计滑动窗口中其他帧的位姿。最后,利用 Bundle Adjustment^[47]技术,通过最小 化所有特征点在滑动窗所有帧上的重投影误差来优化滑动窗所有帧的位姿 $q_{c_k}^{c_0}$ 和 特征点的位置。假定已知相机和 IMU 之间的外参数为 (p_c^b, q_c^b) ,则可以用式4-12将 所有的位姿从相机坐标系转换到 IMU 坐标系,也叫机器人坐标系。

$$\boldsymbol{q}_{b_{k}}^{c_{0}} = \boldsymbol{q}_{c_{k}}^{c_{0}} \otimes \left(\boldsymbol{q}_{c}^{b}\right)^{-1}$$

$$s \overline{\boldsymbol{p}}_{b_{k}}^{c_{0}} = s \overline{\boldsymbol{p}}_{c_{k}}^{c_{0}} - \boldsymbol{R}_{b_{k}}^{c_{0}} \boldsymbol{p}_{c}^{b}$$

$$(4-12)$$

其中 *s* 参数可以将视觉恢复的结构对齐到真实的物理尺度上, *k* 取遍滑动窗中的 所有帧。初始化过程的第二步是进行视觉惯导对齐,包括对陀螺仪偏置,速度,重 力矢量和尺度因子 *s* 的初始化。

视觉惯导对齐的第一步是初始化陀螺仪的偏置 b_w ,考虑滑动窗中的连续两帧 $k \ \pi k + 1$,可以从视觉 SFM 的结果中获得这两帧相对于 c_0 的旋转 $q_{b_k}^{c_0} \ \pi q_{b_{k+1}}^{c_0}$,从 IMU 预积分中可以获得这两帧之间的相对旋转 $\hat{\gamma}_{b_{k+1}}^{b_k}$ 。假设这两帧相对旋转的真实 值为 $\gamma_{b_{k+1}}^{b_k}$ 。IMU 预积分值由于没有考虑陀螺仪偏置的影响,实际上为真实值的一 阶近似,二者由式 (4-13) 中的雅克比矩阵联系起来。理论上,通过视觉获得两帧

之间的相对旋转应该和从 IMU 获得的两帧间的旋转相等,因此,通过最小化如下的误差函数可以估计陀螺仪的偏置。其中 *B* 表示滑动窗中的所有帧。

$$\min_{\delta b_{w}} \sum_{k \in \mathcal{B}} \left\| \boldsymbol{q}_{b_{k+1}}^{c_{0}-1} \otimes \boldsymbol{q}_{b_{k}}^{c_{0}} \otimes \boldsymbol{\gamma}_{b_{k+1}}^{b_{k}} \right\|^{2}
\boldsymbol{\gamma}_{b_{k+1}}^{b_{k}} \approx \hat{\boldsymbol{\gamma}}_{b_{k+1}}^{b_{k}} \otimes \begin{bmatrix} 1 \\ \frac{1}{2} \boldsymbol{J}_{b_{w}}^{\gamma} \delta \boldsymbol{b}_{w} \end{bmatrix}$$
(4-13)

在陀螺仪偏置初始化后,通过对滑动窗中所有 IMU 的预积分项进行重新积分,可以获得一个更加准确的 IMU 预积分测量值。然后初始化包括速度,重力矢量和 尺度因子在内的状态量 X_I。

$$X_{I} = \left[\boldsymbol{v}_{b_{0}}^{b_{0}}, \boldsymbol{v}_{b_{1}}^{b_{1}}, \cdots, \boldsymbol{v}_{b_{n}}^{b_{n}}, \boldsymbol{g}^{c_{0}}, s \right]$$
(4-14)

式中 $\nu_{b_k}^{b_k}$ ——第 k 帧时机器人的速度;

g^{c0}——在 c0 坐标系下表示的重力矢量;

s——将单目 SfM 结果恢复到真实尺度需要的尺度因子; 考虑滑动窗中的连续两帧 *b*_k 和 *b*_{k+1},公式(4-5)可以写成

$$\boldsymbol{\alpha}_{b_{k+1}}^{b_k} = \boldsymbol{R}_{c_0}^{b_k} \left(s \left(\overline{\boldsymbol{p}}_{b_{k+1}}^{c_0} - \overline{\boldsymbol{p}}_{b_k}^{c_0} \right) + \frac{1}{2} \boldsymbol{g}^{c_0} \Delta t_k^2 - \boldsymbol{R}_{b_k}^{c_0} \boldsymbol{v}_{b_k}^{b_k} \Delta t_k \right)$$

$$\boldsymbol{\beta}_{b_{k+1}}^{b_k} = \boldsymbol{R}_{c_0}^{b_k} \left(\boldsymbol{R}_{b_{k+1}}^{c_0} \boldsymbol{v}_{b_{k+1}}^{b_{k+1}} + \boldsymbol{g}^{c_0} \Delta t_k - \boldsymbol{R}_{b_k}^{c_0} \boldsymbol{v}_{b_k}^{b_k} \right)$$
(4-15)

将公式(4-12)和公式(4-15)整合成如下的线性测量模型:

$$\hat{\boldsymbol{z}}_{b_{k+1}}^{b_k} = \begin{bmatrix} \hat{\boldsymbol{\alpha}}_{b_{k+1}}^{b_k} - \boldsymbol{p}_c^b + \boldsymbol{R}_{c_0}^{b_k} \boldsymbol{R}_{b_{k+1}}^{c_0} \boldsymbol{p}_c^b \\ \hat{\boldsymbol{\beta}}_{b_{k+1}}^{b_k} \end{bmatrix} = \boldsymbol{H}_{b_{k+1}}^{b_k} \boldsymbol{X}_I + \boldsymbol{n}_{b_{k+1}}^{b_k}$$
(4-16)

其中

$$\boldsymbol{H}_{b_{k+1}}^{b_{k}} = \begin{bmatrix} -\boldsymbol{I}\Delta t_{k} & \boldsymbol{0} & \frac{1}{2}\boldsymbol{R}_{c_{0}}^{b_{k}}\Delta t_{k}^{2} & \boldsymbol{R}_{c_{0}}^{b_{k}} & (\boldsymbol{\overline{p}}_{c_{k+1}}^{c_{0}} - \boldsymbol{\overline{p}}_{c_{k}}^{c_{0}}) \\ -\boldsymbol{I} & \boldsymbol{R}_{c_{0}}^{b_{k}}\boldsymbol{R}_{b_{k+1}}^{c_{0}} & \boldsymbol{R}_{c_{0}}^{b_{k}}\Delta t_{k} & \boldsymbol{0} \end{bmatrix}$$
(4-17)

通过求解公式(4-18)所示的线性最小二乘问题,可以得到滑动窗中每一帧的速度, 以及重力向量和尺度参数。

$$\min_{X_{I}} \sum_{k \in \mathcal{B}} \left\| \hat{z}_{b_{k+1}}^{b_{k}} - \boldsymbol{H}_{b_{k+1}}^{b_{k}} X_{I} \right\|^{2}$$
(4-18)

通过将重力矢量旋转到和 z 轴对齐,可以获得世界坐标系 w 和相机坐标 c⁰ 之 间的旋转变换 q_{c0}^w。通过这个旋转变换将所有的状态量的参考坐标系从 (·)^{c⁰} 变换到 世界坐标系 (·)^w,然后利用尺度因子 s 将所有的平移量缩放到真实尺度上。至此, 初始化过程全部结束,这些初始化的量将会用在下一步滑动窗优化中。

4.3 基于局部地图匹配的视觉约束

两栖机器人相对于地面移动机器人而言,由于其使用足式机构而不是轮式机构,其运动形式会更加剧烈,再加上野外地形不平等原因,会导致其视角发生快速变化。图4-4是在野外环境下进行俯仰运动的示意图。图中展示了滑动窗中连续的10帧,可以看到,从第0帧到第4帧,相机视角逐渐从地面看向天空,从第5帧到第9帧,相机视角逐渐从天空看向地面,在当前帧,相机视角又看向前方,与滑动窗前面的帧看到相同的特征点。



图 4-4 两栖快速俯仰运动

由于光流跟踪只会对相邻帧中的特征点进行跟踪,即使当前帧和滑动窗中的 第1帧观测到了相同的场景点,也无法形成有效的特征匹配。图4-5对这个过程进 行了说明,假定滑动窗中存在三帧 *I*₁, *I*₂, *I*₃,这三帧可以是不连续的,其中特征 点的灰色深浅表示特征点在世界坐标系下的高度大小,用绿色小块表示追踪上的 特征点,用红色小块表示新检测的特征点。*I*₁帧相机视野朝向正前方,在图像的像 素位置 *u*₁₁,*u*₁₂,*u*₁₃,*u*₁₄,*u*₁₅ 处分别检测到场景中的特征点 *p*1,*p*2,*p*3,*p*4,*p*5。在 *I*₂ 帧 时,相机视角朝上,此时高度较小的特征点 *p*1,*p*2,*p*3 在 *I*₂ 帧中就不可见了,对这 三个特征点的追踪丢失了,特征点 *p*4,*p*5 在该帧的像素位置 *u*²₄*u*²₅ 同时在 *I*₂ 帧中 还检测到了高度较高的特征点 *p*6,*p*7。在 *I*₃ 帧中,相机视角再度朝前,此时再次观 测到特征点 *p*2,*p*3,但由于 *p*2,*p*3 在 *I*2 帧中已经追踪丢失了,因此在 *I*3 帧中,它们 被当做新的特征点进行检测。两栖机器人的这种俯仰运动特性会导致场景中一些相同的特征点反复地被删除和作为新的特征点检测。



图 4-5 俯仰运动下光流跟踪说明

为了解决上述问题,给每个特征点 *pi* 加入一个特征描述子 *di*,特征描述子一般为一个多维的向量,利用特征点周围的图像亮度以及梯度关系计算得到。利用特征描述子可以对场景中的每一个特征点唯一地进行标识。将滑动窗中所有已经 三角化的 N 个特征点 *pi* 及其描述子 *di* 构建一个局部地图 {*pi*_{i=1},*pi*_{i=1}}。如图4-6所 示,对图像 *I*₃ 先将其和上一幅图像 *I*₂ 进行光流跟踪,得到特征匹配 *u*'₃₇,*u*'₃₅,在光 流跟踪失败的区域检测新的特征点 *u*₃₂,*u*₃₃,*u*₃₈,并计算其描述子。利用计算的描述 子,将这些新检测的特征点和局部地图进行匹配,匹配成功的特征点 *u*'₃₂,*u*'₃₃ 将作 为局部地图特征点 *p*₂,*p*₃ 的一次观测,匹配失败的特征点 *u*₃₈ 则作为一个新的待初 始化特征点加入到局部地图中。这样就向滑动窗优化公式(4-20)中加入了新的视 觉约束项,有效提高了优化求解的稳定性。能这样做的原因是由于俯仰运动保证 了相同的场景会被滑动窗中的不同帧观测到。

为了实现上述功能,视觉惯导里程计前端的特征检测和描述子计算需要满足两个要求。第一是特征检测需要有较好的重复性,即确实能在图像 *I*₃ 上检测出特征 点 *p*2,*p*3 在该图像上的投影 *u*₃₂,*u*₃₃。第二是要求描述子的匹配足够准确,因为匹配



图 4-6 同局部地图进行特征匹配示意图

的结果会直接作为视觉观测加入到滑动窗优化中,并没有使用 RANSAC (Random sample consensus)^[48] 来剔除误匹配。

VinsMono 在前端光流跟踪中使用的是 FAST 角点^[49],该方法的原理如图4-7所示。FAST 角点检测的优点在于速度特别快,该方法直接将目标像素 *p* 和其周围邻域像素的亮度值进行比较,如果在周围 16 邻域上有连续 *n* 个像素点的亮度都比中心像素点的亮度大或者小,就可以认为像素 *p* 为一个特征点,实际使用中取*n* = 12,该标准定义为 FAST-12。该方法由于速度快,牺牲了一些性能上的考虑,并没有在多个尺度上对特征点进行检测,同时容易出现特征点扎堆的情况,即在一块区域上有很多特征点的响应都很强。这导致了基于 FAST 的特征点检测方法重复度不高。如图4-8所示,对于几乎是同一个场景拍摄的两幅图像,在左上角区域中可以看到,使用 FAST 角点并没有能检测到同样的特征点。

在 AmphiVIO 中使用 SIFT 特征点代替 FAST 角点实现重复度更高的检测。在 计算机视觉领域, SIFT (Sclae-invariant feature transform)^[50] 特征是一种广泛使用 的,非常稳定的特征检测算法。该方法的主要特点是在尺度空间中进行特征检测, 然后使用特征点邻域的梯度方向分布直方图来作为特征描述子。因为大尺度的特 征对图像的微小运动是不敏感的,这使得 SIFT 角点的重复性较好。同时由于在获 得了梯度直方图后,计算了梯度的主方向,故该描述子具有方向不变的特性。SIFT



图 4-7 FAST 角点检测原理示意图^[49]



图 4-8 FAST 角点检测重复度说明

特征的重复度说明如图4-9所示,可以看到,相对于基于 FAST 角点的检测方法, SIFT 特征的重复度更好,对于同一个场景基本能检测到相同的特征点。

SIFT 特征点相对于 FAST 角点的另一个优势是能检测更大尺度上的特征,如 图4-10所示。图4-10(a)所示的是 FAST 检测的角点,可以看到在蓝色框中框出来 的点位于黑色缺陷的边缘,而图4-10(b)所示的是 SIFT 检测的特征点,可以看到 该特征位于黑色缺陷的中心。大尺度的特征不仅保证了特征检测在图像变化时具 有良好的重复度,更有效抑制了由于快速运动时出现的运动模糊对图像特征检测 的影响。

然而直接将原始 SIFT 算法检测出的特征用在视觉惯导里程计中是不行的,这



图 4-9 SIFT 角点检测重复度说明





 的 如何有点
 b) SIFT 检测的大尺度特征

 图 4-10 Fast 和 Sift 检测特征的尺度说明

是因为其特征检测具有扎堆的特征,如图4-11(a)所示,扎堆的特征其实可以看做一个有效的特征,其他的特征点增加了无用的视觉观测约束。在 AmphiVIO 中使用了基于网格的特征删选方法,该方法可以获得均匀分散在整张图像上的特征点。网格删选算法的原理是首先设定一个较小的阈值从图像中检测足够多的特征点,然后将图像分成 *M*×*N* 的网格,对每个网格中选取其响应最大的那个作为最终的特征点。同时为了减少视觉惯导里程计后端优化的计算量,保证实时性,需要限制前端特征检测的数量不能超过一定值 max_{cnt} (一般为 200)。通过先对所有特征点的响应从大到小进行排序,依次判断该特征点是否为其所在网格中的最大响应特征点,当达到最大特征点数量 max_{cnt} 后即停止判断。使用该方法检测的特征点如图4-11所示,可以看到特征点的分布相对于原始方法更加分散。

为了实现上述非连续帧的特征匹配,在 AmphiVIO 的滑动窗管理策略中,对 于每个特征点 *p*_i,首先记录下每次视觉观测所在的帧相对于该特征点初次观测帧 之间的一个偏移,然后在滑动窗需要对最旧帧进行边缘化或删除最新帧时,更新



a) 原始特征检测出现的特征扎堆现象b) 使用网格删选方法后的特征点图 4-11 网格特征删选算法说明



图 4-12 使用 SIFT 特征匹配的滑动窗图像

每个特征点 *p*_i 的初始观测帧位置和其相对于初始观测位置的偏移,这样就能将跨越多帧的视觉观测联系起来。

利用 SIFT 对将当前帧新检测的特征点和滑动窗进行匹配的结果如图4-12所 示,在最新帧中,红色的点表示新检测的特征点,绿色的点表示从上一帧跟踪到 的特征点。可以看到,这些新增加特征点中的大部分都是在滑动窗前几帧中观测 到的,它们的描述子被存储在局部地图中。通过稳定的 SIFT 特征匹配可以找到了 新加的特征点和局部地图特征点之间的匹配关系,如蓝线所示。通过添加新的视 觉观测约束到滑动窗紧融合优化式(4-20)中,可以提高自主视觉定位的稳定性。

4.4 紧融合的 VIO 优化

在初始化完成之后,AmphiVIO 使用一个基于滑动窗的紧融合 VIO 优化来进行高精度鲁棒的状态估计。紧融合 VIO 优化的示意图如图4-13所示。



图 4-13 紧融合 VIO 优化示意图

在滑动窗中的状态量可以按照下式进行定义

$$\mathcal{X} = \begin{bmatrix} \boldsymbol{x}_0, \boldsymbol{x}_1, \cdots \boldsymbol{x}_n, \boldsymbol{x}_c^b, \lambda_0, \lambda_1, \cdots \lambda_m \end{bmatrix}$$
$$\boldsymbol{x}_k = \begin{bmatrix} \boldsymbol{p}_{b_k}^w, \boldsymbol{v}_{b_k}^w, \boldsymbol{q}_{b_k}^w, \boldsymbol{b}_a, \boldsymbol{b}_g \end{bmatrix}, k \in [0, n]$$
$$\boldsymbol{x}_c^b = \begin{bmatrix} \boldsymbol{p}_c^b, \boldsymbol{q}_c^b \end{bmatrix}$$
(4-19)

n——关键帧的总数量

λ₁——第 *l* 个特征点在初次观测到它的相机坐标系下的逆深度 使用紧融合的视觉惯导优化,也就是寻找一个最大后验估计*X**来最小化先验项和 所有测量残差项的范数和,该残差和的形式如公式(4-20)所示。

$$\min_{\mathcal{X}} \left\{ \left\| \boldsymbol{r}_{p} - \boldsymbol{H}_{p} \boldsymbol{\mathcal{X}} \right\|^{2} + \sum_{k \in \mathcal{B}} \rho \left(\left\| \boldsymbol{r}_{\mathcal{B}} \left(\hat{\boldsymbol{z}}_{b_{k+1}}^{b_{k}}, \boldsymbol{\mathcal{X}} \right) \right\|_{\boldsymbol{P}_{b_{k+1}}}^{2} \right) + \sum_{(l,j) \in C} \rho \left(\left\| \boldsymbol{r}_{C} \left(\hat{\boldsymbol{z}}_{l}^{c_{j}}, \boldsymbol{\mathcal{X}} \right) \right\|_{\boldsymbol{P}_{l}^{c_{j}}}^{2} \right) \right\}$$
(4-20)

$$\rho(s) = \begin{cases} 1 & s \ge 1\\ 2\sqrt{s} - 1 & s < 1 \end{cases}$$
(4-21)

其中 Huber 正则项按照公式(4-21)定义,该正则项的目的是防止因为误匹配而造成的大的残差项影响最后的优化方向,特别的,为了解决两栖机器人俯仰运动时带来的震动问题,也给 IMU 的测量项加入了正则项。式(4-20)中的 $r_{\mathcal{B}}(\hat{z}_{b_{k+1}}^{b_k}, X)$ 和 $r_C(\hat{z}_l^{c_j}, X)$ 分别表示 IMU 预积分和视觉测量的残差项, \mathcal{B} 表示 IMU 预积分项的个数, C表示视觉观测项的个数, $\{r_p, H_p\}$ 表示边缘化获得的先验项。接下来详细介绍 IMU 和视觉测量残差项以及边缘化先验项的构建。

考虑在滑动窗中两个连续图像帧之间所有的 IMU 测量值,利用 IMU 测量模型公式 (4-11),可以得到 IMU 测量残差项如下式所示:

$$\boldsymbol{r}_{\mathcal{B}}\left(\hat{\boldsymbol{z}}_{b_{k+1}}^{b_{k}},\boldsymbol{X}\right) = \begin{bmatrix} \delta\boldsymbol{\alpha}_{b_{k+1}}^{b_{k}} \\ \delta\boldsymbol{\beta}_{b_{k+1}}^{b_{k}} \\ \delta\boldsymbol{\beta}_{b_{k+1}}^{b_{k}} \\ \delta\boldsymbol{\theta}_{b_{k+1}}^{b_{k}} \\ \delta\boldsymbol{b}_{a} \\ \delta\boldsymbol{b}_{g} \end{bmatrix} = \begin{bmatrix} \boldsymbol{R}_{w}^{b_{k}}\left(\boldsymbol{p}_{b_{k+1}}^{w} - \boldsymbol{p}_{b_{k}}^{w} + \frac{1}{2}\boldsymbol{g}^{w}\Delta t_{k}^{2} - \boldsymbol{v}_{b_{k}}^{w}\Delta t_{k}\right) - \hat{\boldsymbol{\alpha}}_{b_{k+1}}^{b_{k}} \\ \boldsymbol{R}_{w}^{b_{k}}\left(\boldsymbol{v}_{b_{k+1}}^{w} + \boldsymbol{g}^{w}\Delta t_{k} - \boldsymbol{v}_{b_{k}}^{w}\right) - \hat{\boldsymbol{\beta}}_{b_{k+1}}^{b_{k}} \\ 2\left[\boldsymbol{q}_{b_{k}}^{w_{k-1}^{w}} \otimes \boldsymbol{q}_{b_{k+1}}^{w} \otimes \left(\hat{\boldsymbol{\gamma}}_{b_{k+1}}^{b_{k}}\right)^{-1}\right]_{xyz} \\ \boldsymbol{b}_{ab_{k+1}} - \boldsymbol{b}_{ab_{k}} \\ \boldsymbol{b}_{wb_{k+1}} - \boldsymbol{b}_{wb_{k}} \end{bmatrix}$$
(4-22)

其中 $[\cdot]_{xyz}$ 表示提取四元数的向量部分用于误差 $\delta \theta_{b_{k+1}}^{b_k}$ 的表示, $\left[\hat{a}_{b_{k+1}}^{b_k}, \hat{\beta}_{b_{k+1}}^{b_k}, \hat{\gamma}_{b_{k+1}}^{b_k} \right]^{\mathrm{T}}$ 表示 IMU 预积分项的测量值, 在残差项中同时还考虑了加速度计和陀螺仪的偏置。

在传统的针孔相机模型中,重投影误差是定义在归一化的图像平面上的,然 而这种模型并不能对广角和鱼眼相机的测量进行很好的建模。因此,考虑在单位 球面上定义视觉测量残差,在第 i 副图像上第一次观测到的第 l 个特征,该特征同 时在第 j 副图像上被观测到了,则可以定义视觉测量残差项为:

$$\boldsymbol{r}_{C}\left(\hat{\boldsymbol{z}}_{l}^{c_{j}},\boldsymbol{X}\right) = \begin{bmatrix} \boldsymbol{b}_{1} \ \boldsymbol{b}_{2} \end{bmatrix}^{T} \cdot \left(\hat{\boldsymbol{\mathcal{P}}}_{l}^{c_{j}} - \frac{\boldsymbol{\mathcal{P}}_{l}^{c_{j}}}{\|\boldsymbol{\mathcal{P}}_{l}^{c_{j}}\|}\right)$$
$$\hat{\boldsymbol{\mathcal{P}}}_{l}^{c_{j}} = \pi_{c}^{-1} \left(\begin{bmatrix} \hat{\boldsymbol{u}}_{l}^{c_{j}}\\ \hat{\boldsymbol{v}}_{l}^{c_{j}} \end{bmatrix}\right)$$
$$(4-23)$$
$$\boldsymbol{\mathcal{P}}_{l}^{c_{j}} = \boldsymbol{R}_{b}^{c} \left(\boldsymbol{R}_{w}^{b} \left(\boldsymbol{R}_{b_{i}}^{w} \left(\boldsymbol{R}_{c}^{b} \frac{1}{\lambda_{l}}\right) \left(\begin{bmatrix} \boldsymbol{u}_{l}^{c_{i}}\\ \boldsymbol{v}_{l}^{c_{i}} \end{bmatrix}\right) + \boldsymbol{p}_{c}^{b} - \boldsymbol{p}_{b_{j}}^{w}\right) - \boldsymbol{p}_{c}^{b}$$

其中 $[u_l^{c_i}, v_l^{c_i}]$ 表示编号为 l 的特征在第 i 帧图像中第一次被观测到时的像素坐标。 $[\hat{u}_l^{c_j}, \hat{v}_l^{c_j}]$ 则表示该特征点在第 j 帧图像中被观测到的像素坐标。 π_c^{-1} 表示反投影模 型,由于视觉残差的自由度为2,因此将残差向量投影到 $\hat{\mathcal{P}}_{l}^{c_{j}}$ 的切平面上, $\{b_{1}, b_{2}\}$ 为任意选取的一对正交基。

为了限制后端紧融合 VIO 优化的计算复杂度,需要将滑动窗中的某些 IMU 状态 *x*_k 和特征 *λ*_i 边缘化掉,同时将这些边缘化的状态转化为一个先验值加入到 式4-20中。当新来一个帧时,如果第二最新帧为关键帧,则将最新帧留在滑动窗 中,将最旧的一帧和其对应的测量量进行边缘化。如果第二最新帧不是关键帧,则 将最新帧对应的视觉测量直接丢弃,只将其 IMU 测量保留。利用这种边缘化策略,可以保证窗口中包含有空间上分散的关键帧,这可以使特征点三角化时有足够的 视差,同时保证加速度测量值有足够的激励。最后,所有和边缘化状态相关的测量值会通过舒尔补^[51] 构建成一个先验项,加入到紧融合 VIO 优化中。

4.5 本章小结

本章主要对两栖视觉惯导里程计进行了介绍。为了利用含有丰富视觉信息的 图像和 IMU 数据实现两栖机器人的自主定位,同时考虑到两栖机器人进行俯仰运 动所带来的视角快速变化问题,本文提出了一种 AmphiVIO 定位算法。该定位算法 首先对滑动窗中的连续帧进行光流跟踪来获得特征之间的匹配关系,并将图像帧 之间的所有 IMU 数据预积分成一个观测量,通过将视觉 SFM 构建的轨迹与 IMU 积分的轨迹进行对齐可以恢复尺度信息,然后对整个系统进行初始化。考虑到视 角快速变化会导致光流跟踪失效,对图像中跟踪失败的区域检测 SIFT 特征并计算 描述子,将检测到的 SIFT 特征和滑动窗口中的特征点进行匹配,如果匹配成功就 将其作为滑动窗中特征点的一个视觉观测加入到紧融合 VIO 优化中,否则将其作 为一个未初始化特征加入到滑动窗中。同时设计了一种滑动窗管理策略来实现基 于匹配的视觉测量,能够有效的在存在非连续观测的情况下实现最旧帧的边缘化 操作和最新帧的删除。最后考虑所有的 IMU 测量残差、视觉测量残差和边缘化先 验残差,执行一个紧融合的 VIO 优化获得最终的机器人位姿。

第5章 实验验证与结果分析

本章对本文提出的 AmphiActiveVIO 系统中各个模块的实际性能进行验证。

5.1 实验平台搭建



图 5-1 两栖机器人实验平台

两栖机器人在野外环境下的正面和侧面实拍图以及实验硬件平台如图5-1所示。其中:

 计算平台: Nvidia 的 TX2 开发板, GPU 带有 256 个 Nvidia Pascal 架构的 CUDA 核,用于 SIFT 特征的实时加速提取。CPU 为四核 ARM Cortex-A57 微处理器,用于视觉惯导里程计和最优视角的计算。

2)相机: Intel 生产的 Realsense D435i,具有全局快门,最大支持分辨率 1920x1080 的 30hz 彩色图像输出。同时还集成有一个消费级的惯性测量单元。

3) 三自由度云台: 2206/100T 无刷电机两只, 2805/100T 无刷电机一只。能实

现正负 25 度的 pitch 和 roll 角控制,正负 90 度的 yaw 角控制。同时可以实时反馈 相机相对于机器人本体的姿态角度。

4)两栖机器人平台:中科大开发的 AmphiHex-II 机器人^[8],可以有效地切换 三足和六足运动步态,实现野外环境下的高机动运动。



图 5-2 代码框架图

整个系统软件运行在 Ubuntu 16.04 上,使用 ROS(Robot Operating System) 作为 底层消息通信组件,其代码框架图如图5-2所示,主要有以下几个 ROS 包 (Package):

1) AmphiVIO 包: 该包运行视觉惯导里程计,订阅相机发布的图像和 IMU 数据,实时估计机器人的位姿,同时也发布当前帧检测到的特征点位置。

2) Exposure Controller 包:订阅相机实时采集的图像数据和光度标定发布的响应函数参数,计算下一帧合适的曝光时间,将其发布给 ROS 的动态参数控制组件, 实现对相机曝光时间的动态调整。

3) Information Filed 包: 该包利用当前帧检测到的特征点和机器人的位姿,构建全局特征点地图。然后从全局特征点地图中计算最大信息视角,并将其发布出去。

4) Active Localization 包:由最大信息视角,机器人的运动方向,综合计算得 到最优视角。

5) View Controller 包:和云台进行通信,用最优视角控制云台的转动。同时利用云台上的 IMU 反馈相机相对于机器人的姿态。

6) Photometric Calibration 包:通过离线光度标定得到相机的响应函数。

5.2 主动视角调整实验

主动视角调整实验主要是验证当机器人在场景中运动时,通过最大信息视角 算法2-1计算得到的最大信息视角确实看向场景中特征点较多的位置。图5-3所示的 是在 Euroc 数据集上进行测试的结果。其中洋葱色表示全局栅格地图,即在世界 坐标系下特征点的位置。当机器人在场景中运动时,对机器人周围区域计算其最 大信息视角如图中的箭头所示,可以发现,这些箭头基本都指向场景中特征点较 丰富的区域。



图 5-3 在 Euroc 数据集上测试最大信息视角算法

接下来在野外实际环境中对主动视角调整算法进行验证,当两栖机器人在野 外环境中进行如图5-4所示的爬坡运动时,两栖机器人逐渐从平地爬到坡上,其视 角会逐渐地看向天空,导致视野中的视觉信息丢失,无法检测到足够的特征点。

我们让两栖机器人使用搭载三自由度云台的主动视角和无三自由云台的固定 视角走同一条轨迹,结果如图5-5所示。左侧为主动视角的程序运行截图,其中洋 红色的点表示斜坡上被检测到的特征点,蓝色箭头表示两栖机器人的当前视角,红 色箭头表示的是最大信息视角。由于现在视野快要丢失,红色箭头指向的视角看 向当前位置的下方,控制相机视角向下转动,增加视野中的特征点。在主动视角



图 5-4 两栖机器人野外环境爬坡视角丢失示意图

调整下,机器人可以实现自主视觉定位,其定位轨迹如蓝色线所示。右侧是固定 视角的程序运行截图,使用固定视角,相机看向几乎没有稳定特征点的天空,此 时机器人无法实现自主视觉定位。



图 5-5 主动视角和固定视角对比实验结果

5.3 主动曝光控制实验

在野外环境中对主动曝光控制算法进行验证,首先验证主动曝光控制相对于

自动曝光控制对视觉惯导里程计前端光流跟踪的影响。结果如图5-6所示,其中第 一行为使用主动曝光控制拍摄的图像,第二行为对主动曝光图像进行特征检测的 结果。第三行为同样位置下使用自动曝光所拍摄的图像,第四行为对自动曝光图 像进行特征检测的结果。



图 5-6 主动曝光和相机自动曝光效果对比

从图5-6中可以看到,当相机视野看向 HDR 环境时,相机自带的自动曝光算 法会使图像看起来更加自然,但这导致了图像中有视觉纹理信息的前景过暗,无 法检测到足够的特征点。而主动曝光控制使得即使在 HDR 环境而言,相机采集的 图像中也能存在足够的视觉特征。另外,主动曝光控制也使得图像亮度变化更加 平滑,不会出现像自动曝光那样前后帧图像亮度变化很大的情况。图5-7 (a)所示 的是主动曝光算法所计算的最优曝光时间随时间变化的曲线,图5-7 (b)展示了主 动曝光对比实验所使用的设备。

为了验证主动曝光控制对自主视觉定位的影响,通过手持相机在野外 HDR 环境中进行运动,一共做了三组实验,三组实验使用主动曝光控制定位算法估计的





a) 设定的曝光时间随时间变化曲线

b) 用于野外主动曝光实验的两个相机

图 5-7 野外环境主动曝光曲线和实验设备

<u>表 5-1</u>	主动	曝光和自动	动曝光定位	这误差对比
古	决	<u> </u>	实验一	实验三

114	入型	入型二	入型一
主动曝光	2.713m	1.341m	0.887m
自动曝光	8.063m	2.332m	1.359m

轨迹和使用自动曝光控制定位算法估计的轨迹如图5-8所示。由于相机的实际运动 为从原点出发再回到原点的闭环运动,因此通过测量估计的轨迹终点距离原点的 位置作为定位误差,结果如表5-1所示,从表中可以得出使用主动曝光控制可以有 效提高自主视觉定位的精度。

5.4 两栖视觉惯导里程计实验

为了验证基于匹配的视觉惯导里程计算法 AmphiVIO 的正确性,首先在标准的公开无人机 VIO 数据集 Eurco^[52]上进行实验。选用 Euroc 中的 V1-03-diffcult 这个数据集上进行测试,该数据集为人手持着相机在室内不断晃动,其视角会出现明显的上下变化,其示意图如图5-9所示。

在该数据集上分别运行基于匹配的视觉惯导里程计 AmphiVIO 和 Vins-Mono, 图5-10(a) 展示了 AmphiVIO 沿着整个时间轴的位置误差,其 RMSE 误差为 0.2258m,整个轨迹的平均误差为 0.2129m。其误差的热力图曲线如图5-10(b) 所 示,其中颜色越红代表误差越大,绝对位置误差最大有 0.4m,其出现在相机快速 上下晃动且光照变化明显的地方。

图5-11展示的是 Vins-Mono 和 AmphiVIO 在该数据集上位置估计误差的对比 结果。其中图5-11(a) 是位置估计误差曲线的对比图, 蓝线为 Vins-Mono 的位

哈尔滨工业大学工学硕士学位论文



a) 实验一主动曝光和自动曝光轨迹对比



b) 实验二主动曝光和自动曝光轨迹对比



c) 实验三主动曝光和自动曝光轨迹对比 图 5-8 主动曝光和自动曝光三次实验轨迹对比图



图 5-9 Euroc 数据集 V1-03-diffcult 示意图





b) 轨迹各处误差的热力图

图 5-10 AmphiVIO 和真值比较得到的误差图

置估计误差,绿线为 AmphiVIO 的位置估计误差。从图中可以得到,基本上全程 AmphiVIO 的误差都要比 Vins-Mono 要小,其中在快速俯仰运动部分,Vins-Mono 由于特征追踪丢失,误差急速上升到 0.4m 左右,而 AmphiVIO 由于可以提供额外 的约束,成功地提高了后端优化的稳定性和准确度。

在如图5-12所示的野外环境中通过手持进行俯仰运动来测试 AmphiVIO 的性能。由于快速的俯仰运动,光流追踪的匹配较为稀疏,在图像边缘处会出现追踪 丢失的情况。但是在当前帧中由于看到和之前同样的场景点,可以对新检测的特 征点和局部地图进行匹配,将匹配结果作为一个视觉观测约束加入到后端优化中。

哈尔滨工业大学工学硕士学位论文



 a) 位置估计误差曲线对比
 b) 位置估计误差关键指标对比

 图 5-11 Vins-Mono 和 AmphiVIO 位置估计误差对比



图 5-12 野外俯仰实验示意图

在该数据集中分别运行 Vins-Mono 和 AmphiVIO,得到的轨迹如图5-13所示, 其中图5-13(a)显示了二者在三维空间中的轨迹,蓝色线为 Vins-Mono 估计的轨 迹,绿色线为 AmphiVIO 的轨迹,图5-13(b)为二者分别在 x 轴, y 轴和 z 轴上 的轨迹。从图中可以发现,二者估计的 z 轴呈现明显的上下俯仰的特点,同时二者 都能估计出大致的运动轨迹,但是存在一个明显的平移误差,这可能是由于野外 环境太相似,在平移运动时 AmphiVIO 中加入了误匹配导致的。



 a) 三维空间上轨迹对比
 b) 三个坐标轴上轨迹对比

 图 5-13 Vins-Mono 和 AmphiVIO 在野外环境下进行俯仰运动的轨迹

表 5-2 AmphiVIO 和 Vins-Mono+IMU 正则定位误差对比						
方法	x 轴误差	y 轴误差	z 轴误差	总误差		
AmphiVIO	-0.62m	0.12m	0.25m	0.313m		
Vins-Mono+IMU 正则	-3.45m	-2.36m	0.58m	4.219m		

为了分析两栖机器人在实际野外环境中运动时 AmphiVIO 的定位性能,记录下两栖机器人在同一野外环境运动时的传感器数据,分别运行 AmphiVIO, Vins-Mono 和 Vins-Mono+IMU 正则三种定位算法。实验结果如图5-14所示,其中红色轨迹为 Vins-Mono 的定位轨迹,绿色为 Vins-Mono 加 IMU 正则的定位轨迹,蓝色为本文提出的 AmphiVIO 的定位轨迹,每一小格的尺度 1m。

当两栖机器人由于自身运动特征进行俯仰运动,其运动形式会很剧烈,此时 IMU 的测量值很大,误差也很大,没有加入 IMU 正则项的 Vins-Mono 会直接定 位失效。当两栖机器人进行俯仰运动时,其视角基本一直观测前方,AmphiVIO 将当前帧的特征点和局部地图进行匹配来加入额外的视觉约束。由于缺少实际真 实定位值作为对比,将两栖机器人回到原点时定位距离原点的距离作为定位误差。 该实验中的最终定位误差如表5-2所示,AmphiVIO 的定位误差为 0.313m, Vins-Mono+IMU 正则的定位误差为 4.219m,实验证明 AmphiVIO 能有效解决由于两栖 机器人俯仰运动所带来的震动和特征追踪丢失问题,提高视觉定位的鲁棒性和精 度。

哈尔滨工业大学工学硕士学位论文



5.5 本章小结

本章主要为本文提出的 AmphiActiveVIO 系统在公开数据集和野外环境中性 能进行了实验验证。对于主动视角调整模块,首先在 Euroc 数据集上验证了其能准 确计算出当前机器人的最大信息视角。然后在野外环境中让两栖机器人进行实际 的爬坡运动,对比主动视觉定位和固定视角定位结果,实验证明使用主动视角定 位可以有效防止视野丢失。对于主动曝光控制模块,首先利用标定出来的相机响 应曲线,在野外 HDR 环境中进行主动曝光控制实验,相对于自动曝光算法而言, 主动曝光控制使得对于 HDR 环境而言,图像中也能存在足够的视觉特征,同时能 提高自主视觉定位的精度。最后对于两栖视觉惯导里程计模块,首先在公开数据 集 Euroc 上与 Vins-Mono 进行了对比,通过对比发现,增加额外约束后整个系统的 位置估计精度有所提高,从 Vins-Mono 的平均 0.259m 的位置误差降低到 0.206m, 特别是在进行俯仰运动时能有效抑制误差的增长。然后在两栖机器人野外环境的 实际运动中对 Vins-Mono, Vins-Mono+IMU 正则和 AmphiVIO 三种定位算法进行 了验证,通过实验发现,加入 IMU 正则项可以有效解决由于两栖机器人震动带来 的定位失效问题,同时通过匹配增加的视觉测量可以提高 AmphiVIO 的定位精度。

结 论

两栖机器人在野外环境下进行作业时,由于自身运动特性,野外崎岖地形和 光线变化会造成视觉信息缺失和相机视野丢失等问题。为了实现两栖机器人在野 外环境下稳定准确的自主定位。本文设计了一种两栖主动视觉惯导里程计(AmphiActiveVIO)系统。该系统以基于滑动窗的紧融合视觉惯导里程计为基础。但与 已有视觉惯导里程计的不同在于两个方面:第一为两栖,即考虑到两栖机器人特有 的俯仰运动特性,通过特征匹配来增加额外的视觉约束,提高后端优化稳定性;第 二为主动,这又有两方面的含义:第一是通过三自由度云台来主动调整相机视角 防止视野丢失;第二是通过主动设置相机的曝光时间,使相机捕获的图像中能包 含丰富的视觉信息。总的来说,本文完成的主要工作有:

(1)设计了一种两栖主动视觉惯导里程计(AmphiActiveVIO)系统。该系统能利用相机采集的图像和 IMU 的数据来实时估计机器人的位姿,同时通过主动控制相机的曝光时间和视角来保证图像中含有足够的视觉信息,从而减弱由于两栖机器人运动以及野外环境地形和光线变化对视觉定位带来的不利影响。

(2)提出了一种基于匹配的视觉惯导里程计,对每个新检测的特征点增加一个特征描述子,如果该特征点和局部地图中已检测的特征点存在匹配关系,则在滑动窗优化中新增加一个视觉约束项。该方法可以在视角快速变化时添加新的约束项,提高视觉惯导里程计的稳定性。

(3)提出了一种用于相机视角调整的最优视角计算算法,该算法利用已探索的场景特征点和当前两栖机器人的位置,通过信息理论计算出当前最大信息视角。然后结合机器人当前运动方向计算一个最优视角,在这个最优视角下能有效平衡机器人的定位和探索需求,保证稳定视觉定位的同时实现对场景最大限度的探索。

(4)野外复杂的光线变化会导致图像过暗或者过亮,无法提取足够的视觉信息。为了解决这个问题,本文使用了一种基于离线光度标定的主动曝光控制算法,该算法首先对一个固定场景采集多张不同曝光下的图片来离线标定相机的响应函数,然后通过最大化图像的可导百分比指标来计算下一时刻最合适的曝光时间,保证下一帧图像不会出现过曝光或欠曝光的情况。

但本文所提出的系统框架在实际使用时还有许多不足,需要进行如下的改进: (1) 在基于匹配的视觉惯导里程计中,本文使用了基于特征描述子的匹配方 法,但是这就要求了局部地图中的点在当前帧图像中能够检测到。但检测新特征 点时并没有考虑到场景中哪些特征点已经被存到局部地图中。一种更好的方法是 类似 SVO^[21] 对每个特征点保存其周围的图像块,当需要与当前帧进行匹配时,将 地图中的点投影到当前帧中,然后调整投影点的位置,使保存的图像块和当前图 像匹配准确。通过这种方法可以提供更多的匹配信息,但通过实验发现该方法要 求预测的下一帧位姿足够准确,否则投影位姿距离最终位置较远会导致无法成功 进行匹配。

(2) 在主动视角调整模块中,只计算了下一帧的最优视角,最好利用模型预测 控制,规划一条最优视角的轨迹,预测接下来多个帧的最优视角。

(3)现有的基于视觉惯导里程计的方法还是不够鲁棒,当两栖机器人运动剧烈时, IMU 的测量值误差很大时还是会导致视觉定位失败,下一阶段可以考虑基于 多传感器融合的自主定位。

参考文献

- [1] Mur-Artal R, Tardós 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.
- [2] Engel J, Koltun V, Cremers D. Direct Sparse Odometry[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(3): 611–625.
- [3] Ijspeert A J, Crespi A, Ryczko D, et al. From Swimming to Walking with a Salamander Robot Driven by a Spinal Cord Model[J]. Science, 2007, 315(5817): 1416– 1420.
- [4] Hirose S, Yamada H. Snake-like Robots[J]. IEEE Robotics and Automation Magazine, 2009, 16(1): 88–98.
- [5] Yu J, Ding R, Yang Q, et al. On a Bio-inspired Amphibious Robot Capable of Multimodal Motion[J]. IEEE/ASME Transactions on Mechatronics, 2012, 17(5): 847-856.
- [6] 王田苗, 仲启亮, 孟刚. 仿鳄鱼水陆两栖机器人机构优化设计与试验验证 [J]. 机 械工程学报, 2010, 46(13): 76-82.
- [7] Zhang S, Zhou Y, Xu M, et al. AmphiHex-I: Locomotory Performance in Amphibious Environments With Specially Designed Transformable Flipper Legs[J].
 IEEE/ASME Transactions on Mechatronics, 2016, 21(3): 1720-1731.
- [8] Zhong B, Zhang S, Xu M, et al. On a CPG-Based Hexapod Robot: AmphiHex-II With Variable Stiffness Legs[J]. IEEE/ASME Transactions on Mechatronics, 2018, 23(2): 542-551.
- [9] 尹咸阳, 边宇枢. 基于 GPS 的水陆两栖机器人导航系统的研究 [J]. 微型电脑应用, 2006, 22(7): 1-3.
- [10] Kaznov V, Seeman M. Outdoor Navigation with a Spherical Amphibious Robot[C]
 // Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. 2010: 5113-5118.
- [11] Ribas D, Ridao P, Neira J. Underwater SLAM for Structured Environments using an Imaging Sonar: Vol 65[M]. America: Springer, 2010.

- [12] Aulinas J, Lladó X, Salvi J, et al. Selective Submap Joining for Underwater Large Scale 6-DOF SLAM[C] // Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. 2010: 2552–2557.
- [13] Raibert M, Blankespoor K, Nelson G, et al. Bigdog, the rough-terrain quadruped robot[J]. IFAC Proceedings Volumes, 2008, 41(2): 10822-10825.
- [14] Bloesch M, Hutter M, Hoepflinger M A, et al. State estimation for legged robotsconsistent fusion of leg kinematics and IMU[J]. Robotics, 2013, 17: 17–24.
- [15] Reinstein M, Hoffmann M. Dead reckoning in a dynamic quadruped robot based on multimodal proprioceptive sensory information[J]. IEEE Transactions on Robotics, 2012, 29(2): 563-571.
- [16] Ma J, Bajracharya M, Susca S, et al. Real-time pose estimation of a dynamic quadruped in GPS-denied environments for 24-hour operation[J]. The International Journal of Robotics Research, 2016, 35(6): 631–653.
- [17] Davison A J, Reid I D, Molton N D, et al. MonoSLAM: Real-time Single Camera SLAM[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007(6): 1052–1067.
- [18] Klein G, Murray D. Parallel Tracking and Mapping for Small AR Workspaces[C] // Proceedings of the IEEE and ACM International Symposium on Mixed and Augmented Reality. 2007: 1–10.
- [19] 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.
- [20] Engel J, Schöps T, Cremers D. LSD-SLAM: Large-scale Direct Monocular SLAM[C] // Proceedings of the European Conference on Computer Vision. 2014: 834-849.
- [21] Forster C, Zhang Z, Gassner M, et al. SVO: Semidirect Visual Odometry for Monocular and Multicamera Systems[J]. IEEE Transactions on Robotics, 2016, 33(2): 249–265.
- [22] Labbe M, Michaud F. Online Global Loop Closure Detection for Large-scale Multisession Graph-based SLAM[C] // Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. 2014: 2661–2666.

- [23] Liu H, Zhang G, Bao H. Robust Keyframe-based Monocular SLAM for Augmented Reality[C] // Proceedings of the IEEE International Symposium on Mixed and Augmented Reality. 2016: 1–10.
- [24] Tan W, Liu H, Dong Z, et al. Robust Monocular SLAM in Dynamic Environments[C] // Proceedings of the IEEE International Symposium on Mixed and Augmented Reality. 2013: 209–218.
- [25] Kelly J, Sukhatme G S. Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration[J]. The International Journal of Robotics Research, 2011, 30(1): 56–79.
- [26] Leutenegger S, Lynen S, Bosse M, et al. Keyframe-based Visual-inertial Odometry using Nonlinear Optimization[J]. The International Journal of Robotics Research, 2015, 34(3): 314-334.
- [27] Mourikis A I, Roumeliotis S I. A Multi-state Constraint Kalman Filter for Visionaided Inertial Navigation[C] // Proceedings of the IEEE International Conference on Robotics and Automation. 2007: 3565–3572.
- [28] Forster C, Carlone L, Dellaert F, et al. On-Manifold Preintegration for Real-Time Visual-Inertial Odometry[J]. IEEE Transactions on Robotics, 2016, 33(1): 1–21.
- [29] Bloesch M, Burri M, Omari S, et al. Iterated Extended Kalman Filter Based Visualinertial Odometry Using Direct Photometric Feedback[J]. The International Journal of Robotics Research, 2017, 36(10): 1053 – 1072.
- [30] Qin T, Li P, Shen S. Vins-mono: A Robust and Versatile Monocular Visual-inertial State Estimator[J]. IEEE Transactions on Robotics, 2018, 34(4): 1004 – 1020.
- [31] Shim I, Lee J-Y, Kweon I S. Auto-adjusting Camera Exposure for Outdoor Robotics Using Gradient Information[C] // Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. 2014: 1011–1017.
- [32] Zhang Z, Forster C, Scaramuzza D. Active Exposure Control for Robust Visual Odometry in HDR Environments[C] // Proceedings of the IEEE International Conference on Robotics and Automation. 2017: 3894–3901.
- [33] Debevec P E, Malik J. Recovering High Dynamic Range Radiance Maps from Photographs[C] // Proceedings of the ACM SIGGRAPH 2008 classes. 2008: 31.
- [34] Feder H J S, Leonard J J, Smith C M. Adaptive Mobile Robot Navigation and Mapping[J]. The International Journal of Robotics Research, 1999, 18(7): 650– 668.
- [35] Davison A J, Murray D W. Simultaneous Localization and Map-building using Active Vision[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 865-880.
- [36] Roy N, Burgard W, Fox D, et al. Coastal Navigation-mobile Robot Navigation with Uncertainty in Dynamic Environments[C] // Proceedings of the IEEE International Conference on Robotics and Automation : Vol 1. 1999 : 35–40.
- [37] Ichter B, Landry B, Schmerling E, et al. Robust Motion Planning via Perceptionaware Multiobjective Search on Gpus[J]. arXiv preprint arXiv:1705.02408, 2017.
- [38] Zhang Z, Scaramuzza D. Beyond Point Clouds: Fisher Information Field for Active Visual Localization[C] // Proceedings of the IEEE International Conference on Robotics and Automation. 2019.
- [39] Lehmann E L, Casella G. Theory of Point Estimation[M]. America : Springer Science & Business Media, 2006.
- [40] Sola J. Quaternion Kinematics for the Error-state Kalman Filter[J]. arXiv preprint arXiv:1711.02508, 2017.
- [41] Kim S J, Pollefeys M. Robust Radiometric Calibration and Vignetting Correction[J].
 IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(4): 562 576.
- [42] Baker S, Matthews I. Lucas-kanade 20 years on: A Unifying Framework[J]. International Journal of Computer Vision, 2004, 56(3): 221–255.
- [43] Lucas B D, Kanade T, others. An iterative image registration technique with an application to stereo vision[C] // Proceedings of the 7th International Joint Conference on Artificial Intelligence. [S.l.]: Vancouver, British Columbia, 1981.
- [44] Shi J, others. Good Features to Track[C] // journal IEEE Conference on Computer Vision and Pattern Recognition. 1994: 593–600.
- [45] Lupton T, Sukkarieh S. Visual-inertial-aided Navigation for High-dynamic Motion in Built Environments without Initial Conditions[J]. IEEE Transactions on Robotics, 2011, 28(1): 61–76.

- [46] Snavely N, Seitz S M, Szeliski R. Photo Tourism: Exploring Photo Collections in 3D[J]. ACM transactions on graphics (TOG), 2006, 25(3): 835–846.
- [47] Triggs B, McLauchlan P F, Hartley R I, et al. Bundle Adjustment a Modern Synthesis[C] // Proceedings of the International Workshop on Vision Algorithms. 1999: 298-372.
- [48] Derpanis K G. Overview of the RANSAC Algorithm[J]. Image Rochester NY, 2010, 4(1): 2-3.
- [49] Rosten E, Drummond T. Machine Learning for High-speed Corner Detection[C]// Proceedings of the European Conference on Computer Vision. 2006: 430–443.
- [50] Lowe D G, others. Object Recognition from Local Scale-invariant Features[C]
 // Proceedings of the International Conference on Computer Vision: Vol 99. 1999: 1150-1157.
- [51] Sibley G, Matthies L, Sukhatme G. Sliding Window Filter with Application to Planetary Landing[J]. Journal of Field Robotics, 2010, 27(5): 587–608.
- [52] Burri M, Nikolic J, Gohl P, et al. The EuRoC Micro Aerial Vehicle Datasets[J]. The International Journal of Robotics Research, 2016, 35(10): 1157–1163.

攻读硕士学位期间发表的论文

- [1] Jiabi Sun, Jin Song, Haoyao Chen, Xiaopeng Huang, Yunhui Liu. Autonomous State Estimation and Mapping in Unknown Environments with Onboard Stereo Camera for MAVs[J]. IEEE Transactions on Industrial Informatics. (doi: 10.1109/TII.2019.2958183)
- Wu Y, Song J, Sun J, et al. Aerial Grasping Based on VR Perception and Haptic Control[C]. 2018 IEEE International Conference on Real-time Computing and Robotics (RCAR). 2018: 556–562.

哈尔派工业大学工学硕士学位论文

哈尔滨工业大学学位论文原创性声明和使用权限

学位论文原创性声明

本人郑重声明:此处所提交的学位论文《野外环境下两栖机器人的自主视觉 定位算法研究》,是本人在导师指导下,在哈尔滨工业大学攻读学位期间独立进 行研究工作所取得的成果,且学位论文中除已标注引用文献的部分外不包含他人 完成或已发表的研究成果。对本学位论文的研究工作做出重要贡献的个人和集体, 均已在文中以明确方式注明。

作者签名: 宋瑾

日期: 7070年1月7日

学位论文使用权限

学位论文是研究生在哈尔滨工业大学攻读学位期间完成的成果,知识产权归 属哈尔滨工业大学。学位论文的使用权限如下:

(1)学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文,并向国家图书馆报送学位论文;(2)学校可以将学位论文部分或全部内容编入有关数据库进行检索和提供相应阅览服务;(3)研究生毕业后发表与此学位论文研究成果相关的学术论文和其他成果时,应征得导师同意,且第一署名单位为哈尔滨工业大学。

保密论文在保密期内邀守有关保密规定,解密后适用于此使用权限规定。

本人知悉学位论文的使用权限,并将遵守有关规定。

作者签名: 宋瑾

日期:2020年1月7日

导师签名: 花花花

- 68 -

日期: 2020年 1月 7日

致 谢

时光荏苒,一转眼两年半的硕士生涯就要告一段落了,虽然心中仍有些许不 舍,但也到了必须要说再见的时候。在这短短的两年半的时间里,我在老师和同 学们的帮助下成长了许多。在这里要对曾经帮助过我的人表示衷心的感谢。

首先要感谢的是我的导师陈浩耀老师。研究生开始,我有幸能加入 NRSL 实验室,后来证明这是一个极其明智的选择,感谢老师能给当时我这个机会。研一 刚入学,老师就给了我们许多锻炼的机会,从各种项目,比赛再到论文。在这段之间中,我们在老师悉心地指导下科研能力相对于刚来实验室时提高了许多,为之 后做毕业设计打下了良好的基础。做毕设时,老师也给予了我很大的自由度,对 我提出的想法都鼓励我去实现,但也会及时纠正我错误的研究方向。感谢老师给 我们提供的良好的实验室学习氛围和生活氛围,让我们可以安心做研究。虽然我 们平时已经很忙了,但感觉老师比我们更忙,老师对什么事都认真负责的态度给 我树立了良好的榜样,成为我日后前进的方向。

然后要感谢 NRSL 实验室的同门师兄妹们,来到实验室后认识了许多小伙伴, 大家都非常的优秀,我们在一起度过了很多难忘的时光。特别要感谢同一届的孙 家弼,巫亚奇和朱芬芳同学,我们在一起完成了大大小小的各种项目和比赛,虽 然中间的过程很痛苦,但是最后成功的喜悦却是很难忘的。没有他们的帮助和鼓 励,我很难取得今天的成绩。和你们在一起学习,一起工作,还有一起玩耍的时光 是我这辈子最宝贵的回忆。

最后要感谢父母对我养育之恩,他们一直在背后默默地支持着我。父母的愿 望其实很简单,就是希望我能身体健康,过的开心。希望以后我能不辜负父母的 期望,也能好好的孝敬他们。

虽然我在学校的生活将要结束了,但是在学校所培养的探索问题和解决问题 的能力将会一直陪伴着我,成为我这辈子最宝贵的财富。