

硕士学位论文

火星表面多漫游器协同构建导航地图 方法研究

REASEARCH ON MARS MULTI-ROVERS COOPERATIVELY CONSTRUCTING NAVIGATIONAL MAPS

顾远凌

哈尔滨工业大学

2018 年 6 月

国内图书分类号: V448.224

国际图书分类号: 629.783

学校代码: 10213

密级: 公开

工学硕士学位论文

火星表面多漫游器协同构建导航地图 方法研究

硕士研究生: 顾远凌

导师: 张泽旭教授

申请学位: 工程硕士

学科: 航天工程

所在单位: 飞行器动力学与控制研究所

答辩日期: 2018年6月

授予学位单位: 哈尔滨工业大学

Classified Index: V448.224

U.D.C: 629.783

Dissertation for the Master Degree in Engineering

**RESEARCH ON MARS MULTI-ROVERS
COOPERATIVELY CONSTRUCTING
NAVIGATIONAL MAPS**

Candidate :	Gu Yuanling
Supervisor :	Prof.Zhang Zexu
Academic Degree Applied for :	Master of Engineering
Speciality :	Aerospace Engineering
Affiliation :	Spacecraft Dynamics and Control Institute
Date of Defence :	June, 2018
Degree-Conferring-Institution :	Harbin Institute of Technology

摘 要

火星与地球所处的宇宙环境十分相似,对火星开展探测,在寻找地外生命、研究地球大气与磁场演变等方面具有重要意义。利用漫游器对火星进行探测,是人类目前为止最有效的探测火星的手段。然而由于缺少火星表面的导航地图,目前火星漫游器的行进路线主要依靠地面科研人员根据火星漫游器拍摄的图像设计,大大降低了火星漫游器的探测速度。此外,对于某个火星漫游器拍摄的图像,也只用于其自身的定位导航,对后续其它漫游器经过相同区域时的自主导航没有帮助,造成了信息的浪费。

针对上述问题,本文提出一种多漫游器协同构建导航地图的方法,主要研究内容如下。

首先,研究基于双目视觉的 SLAM (Simultaneous Localization and Mapping, 同时定位与地图构建) 方法,对单个火星漫游器进行自主定位,并构建火星表面的稀疏特征点地图,该地图可以用于其它漫游器经过相同区域时的自主定位。

其次,研究基于词袋模型的重叠区域搜索算法。利用火星漫游器拍摄的图像构建字典,利用字典计算图片的相似性评分,进而利用阈值法抑制较低的相似性评分,以最终寻找到连续多帧的相似图像。

第三,研究不同漫游器下的稀疏特征点地图与漫游器位姿协同拼接方法。利用重叠区域拍摄的连续多帧的图片,将不同漫游器构建的稀疏特征点地图以及漫游器的位姿估计拼接,并对结果进行非线性优化。

最后,研究面向后续漫游器导航应用的八叉树地图构建方法。利用漫游器位姿估计的结果和漫游器拍摄的双目图像,构建火星表面稠密点云地图,对稠密点云进行优化处理并利用稠密点云地图构建八叉树地图,该地图可以表示空间的占据情况,用于后续漫游器经过相同区域时的自主路径规划。

本文对研究内容进行了仿真实验,验证所提出方法的有效性。

关键词: 火星漫游器; 同时定位与地图构建; 词袋模型; 非线性优化; 八叉树地图

Abstract

As Mars is very similar to the environment in which the Earth is located, it is of great significance to explore Mars in search of extraterrestrial life and to study the evolution of the Earth's atmosphere and magnetic fields and so on. The use of Mars rover to detect Mars is the most effective means for human detection so far. However, due to the lack of navigational maps on the surface of Mars, the current Mars rover's travel route mainly relies on the design of the Earth's personnel and can not achieve autonomous path planning, which greatly reduces the detection speed of the Mars rover. In addition, the images captured by the Mars rover are used only for the navigation of a single rover, which does not provide help when other Mars rovers pass through the same area, resulting in waste of information.

Based on these issues, this paper proposes a Mars multi-rovers collaborative construction method of navigational map. The main research content is as follows.

First, a SLAM (Simultaneous Localization and Mapping) method based on stereo vision is studied. Using this method, a Mars rover is autonomously positioned and the sparse-points map of the surface of Mars is constructed, which can be used for autonomous positioning of other rovers when passing through the same area.

Then, this paper studies the overlapping region search algorithm based on the bag of word, including the method of building a dictionary using the images captured by the Mars rovers, calculating the similarity scores of pictures, suppress the lower similarity scores, and finally the similar images of consecutive frames are found.

Afterwards, this paper studies the method of stitching the sparse-points map constructed by different Mars rovers and the pose and position estimation of these rovers using the similar images of consecutive frames, and the results are nonlinearly optimized.

Finally, this paper studies the octree map construction method for follow-up Mars rovers navigation applications, including the method of constructing the dense points map on Mars surface, optimizing the point cloud, and using the dense points map to construct the octree map which can represent the space occupancy situation.

This article has carried on the simulation experiment to each chapter content, demonstrated that the method proposed in each part is accurate and effective.

Keywords: Mars rover; simultaneous localization and mapping; word of bag; nonlinear optimization; octree map

目 录

摘 要	I
ABSTRACT	II
第 1 章 绪 论	1
1.1 课题背景和意义	1
1.1.1 课题来源	1
1.1.2 研究目的及意义	1
1.2 国内外的研究现状及分析	2
1.2.1 火星漫游器导航方法研究现状	2
1.2.2 SLAM 方法研究现状	4
1.2.3 多机器人协同构建地图方法研究现状	5
1.2.4 空间三维地图研究现状	6
1.3 本文主要研究内容与结构	7
第 2 章 基于双目视觉的火星漫游器 SLAM 算法	9
2.1 引言	9
2.2 相机模型	9
2.2.1 相机成像模型	9
2.2.2 双目相机测距模型	10
2.2.3 图像畸变的校正	11
2.3 特征点提取与匹配	12
2.3.1 特征点检测方法	12
2.3.2 特征点表示方法	13
2.3.3 特征点匹配方法	14
2.4 漫游器位姿求解方法	15
2.5 仿真实验与分析	17
2.5.1 基于漫游器实际拍摄图像的特征点提取与匹配实验	17
2.5.2 基于地面模拟场地图像的漫游器 SLAM 仿真实验	18
2.5.3 基于计算机模拟图像的漫游器 SLAM 仿真实验	22
2.6 本章小结	23

第 3 章 基于词袋模型的重叠区域搜索算法.....	24
3.1 引言	24
3.2 基于词袋模型的字典生成算法	24
3.2.1 词袋模型的概念	24
3.2.2 字典的构建方法	25
3.3 基于字典的重叠区域搜索算法	26
3.3.1 图片描述向量的生成与比较	26
3.3.2 重叠区域查找流程	28
3.4 仿真实验与分析	29
3.4.1 基于漫游器实际拍摄图像的相似图片搜索实验	29
3.4.2 基于地面模拟场地图像的重叠区域搜索实验	31
3.5 本章小结	33
第 4 章 稀疏特征点地图及漫游器位姿拼接算法	34
4.1 引言	34
4.2 稀疏特征点地图及漫游器位姿拼接原理	34
4.3 坐标转换关系初始估计算法	36
4.4 三维点云与相机位姿联合优化方法	37
4.5 仿真实验与分析	41
4.5.1 基于地面模拟场地图像的稀疏特征点地图及位姿拼接实验	41
4.5.2 基于计算机模拟图像的稀疏特征点地图及位姿拼接实验	44
4.6 本章小结	46
第 5 章 稠密点云地图与八叉树地图的构建算法	47
5.1 引言	47
5.2 稠密点云地图构建算法	47
5.2.1 基于极线搜索与块匹配的双目图像匹配	47
5.2.2 稠密点云地图的构建	50
5.3 八叉树地图构建算法	51
5.4 仿真实验与分析	52
5.4.1 基于地面模拟场地图像的八叉树地图构建实验	52
5.4.2 基于计算机模拟图像的八叉树地图构建实验	54

5.5 本章小结	55
结 论	56
参考文献	57
攻读硕士学位期间发表的论文及其它成果	61
哈尔滨工业大学学位论文原创性声明和使用权限	62
致 谢	63

第1章 绪 论

1.1 课题背景和意义

1.1.1 课题来源

本课题来源于国家自然科学基金面上项目“火星表面大范围漫游探测协同导航与控制方法研究（项目批准号：61374213）”。

1.1.2 研究目的及意义

火星距地球最近距离大约为 0.37 个天文单位，自转周期为 24.6229h，几乎与地球相同，公转周期为 687d，约为地球的 2 倍。火星与地球所处的宇宙环境十分相似，对火星开展探测，在寻找地外生命、研究大气与磁场演变、研究地理变迁等方面具有重要意义^[1]。2004 年 ESA 发现火星南极存在固态水^[2]，2015 年 NASA 公布证据表明火星表面现今仍有液态水流动^[3]，这更极大引起了世界各航天大国对于火星探测的兴趣。从上世纪六十年代至今，人类共发射了 43 颗火星探测器^[4]，我国也计划于 2020 年左右发射自行研制的火星探测器^[5]。

火星漫游器可以对火星进行近距离探测，是迄今为止人类对火星最有效的探测工具。1996 年美国发射“索纳杰”号火星漫游器，对火星进行了 83 天的有效探索，获得了火星表面的图像资料，并对火星大气成分进行了分析^[6]。在此之后，美国先后发射了“勇气”号、“机遇”号、“好奇”号火星车，拍摄了一系列火星地表图像，获取了大量的火星土壤、岩石、大气成分的分析数据。但是，这些单漫游器各自独立进行探测，存在着探测范围有限的缺点。目前探测范围最大的“机遇”号火星车，在 10 年的时间里移动了 43.87 公里，创下了地外无人探测器移动记录^[7]，但是这仍然远远无法满足未来火星大范围探测任务的需求。

限制漫游器探测范围的因素主要有两点。一是单个漫游器携带大量科学探测载荷，重量较大，而为漫游器提供动力的电池功率有限，因此漫游器行进速度较慢。以“好奇”号火星漫游器为例，其在平坦沙地上最大的行驶速度也仅为 130m/h^[8]。二是由于目前漫游器行进时的路径规划仍然主要依靠地面科研人员根据前一个行进阶段漫游器获取的图像信息来确定，不能实现火星车的自

主路径规划,因此漫游器在每个行进阶段的移动距离很有限。而对于这些图像信息,目前只是使用在单个漫游器的导航方面^[9],对后续其他漫游器经过相同区域时的自主导航没有任何帮助,这样就造成了有效信息的浪费。

基于此,本课题提出了一种多漫游器协同导航的方法。对于在行进区域有重叠的多个漫游器,首先利用双目视觉 SLAM (Simultaneous Localization and Mapping, 同时定位与地图构建) 方法,对每个漫游器进行自主定位,并构建火星表面的稀疏特征点地图。然后根据不同漫游器拍摄的图像,寻找多个漫游器行进区域的重叠部分,利用重叠区域中的特征点,对多个漫游器构建的稀疏特征点地图进行拼接,形成更大范围的稀疏特征点地图,用于后续其他漫游器的自主定位。最后利用漫游器拍摄的双目图像和漫游器位姿信息,构建大范围稠密点云地图,并据此构建表示空间占据情况的八叉树地图,为后续其他漫游器的自主路径规划提供信息。

利用这种方法,可以在前期利用较少安装或不安装科学载荷、重量轻、行进速度较快的多个漫游器进行快速大范围探测,构建火星表面大范围导航地图。后续到达的安装精密科学载荷、重量大、行进速度较慢的漫游器,可以利用构建好的导航地图进行自主的定位与路径规划,较快速地行驶到感兴趣的目标点。

1.2 国内外的研究现状及分析

1.2.1 火星漫游器导航方法研究现状

美国、欧盟和前苏联都曾向火星发射过漫游器进行探测,但是前苏联和欧盟的漫游器探测任务均以失败告终,只有美国的“索纳杰”号、“勇气”号、“机遇”号和“好奇”号漫游器,在火星表面成功执行了探测任务。

“索纳杰”号火星车携带有三个相机,用以观察周围地形,并将图像发回地面,由地面科学家设计移动路线。在行进过程中,“索纳杰”号利用车载陀螺仪和里程计数据更新位置信息^[10]。“索纳杰”号漫游器在火星执行任务时长为 83 个火星日,共行驶了 100 多米^[11]。

“勇气”号、“机遇”号和“好奇”号漫游器在着陆后,通过无线电系统与地球测控站以及火星轨道器进行通信,根据多普勒信号进行无线电测控定位,成功获取了火星车的初始位置^[12]。在行进过程中,有三种模式可供切换:第一种模式是盲行模式:在平坦、无大障碍物的地区,火星车将相机获取的图像传回地面,由地面科学家制定行驶路线,火星车根据行驶指令行驶,利用加速度计和陀螺仪以每秒 8 次的频率更新位姿信息;第二种模式是自主避障和视觉

测程模式：漫游器利用相机获取的图像自主选择前进路线，每行驶一段距离就停止，根据相机获取的图像差异判断行进距离；第三种模式与第二种模式类似，只是视觉测程的间隔更短，每隔半个漫游器的长度就停下来进行视觉测程。盲行模式存在的问题是当车轮发生打滑时会产生较大的误差，而视觉测程则由于车载计算机计算速度的限制，需要较长的处理时间。在漫游器停止时，还利用太阳图像确定方位角，修正位姿估计的误差^[13]。

美国计划于 2020 年发射的火星漫游器“Mars 2020”号，将携带多达 23 个相机，包括 9 个工程相机、7 个科学相机和 7 个着陆辅助相机。其工程相机中包含 2 个导航相机和 6 个避障相机，可以利用双目图像获取周围环境信息，进行自主导航定位^[14]。



图 1-1 “Mars 2020”号火星漫游器

我国首个地外漫游器“玉兔”号，采用航迹推算法和视觉测程法联合进行定位导航^[15]。首先利用基于惯性测量元件的航迹推算法，估计漫游器的运动情况，作为视觉测程法的初始估计值。然后利用漫游器双目相机在不同位置拍摄的月面站点的图像提取特征点进行匹配，根据特征点的位置关系，结合航迹推算结果进行光束平差优化，最终计算得到漫游器的运动轨迹^[16]。

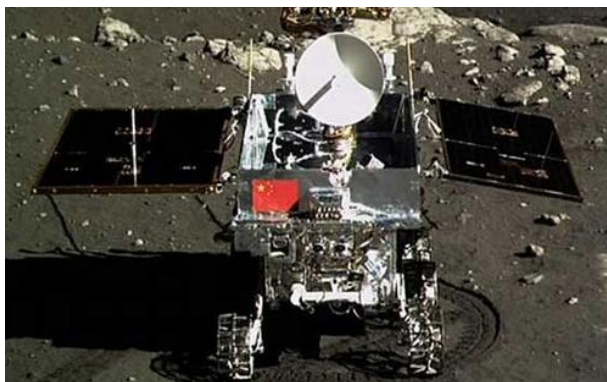


图 1-2 “玉兔”号月球漫游器

1.2.2 SLAM 方法研究现状

自 1986 年提出以来^[17], SLAM 一直是机器人领域的热点研究问题。

从获取环境信息的传感器上看, 早期的 SLAM 方法使用的主要是声纳、激光雷达等测距传感器^[18], 使得硬件设备成本较高, 在一定程度上限制了 SLAM 方法的应用。21 世纪以来, 随着计算机性能的大幅提升, 视觉 SLAM 方法凭借其成本较低、可获取信息量大、特征易提取等优点, 开始逐渐成为研究热点^[19]。目前已有的视觉 SLAM 方案中, 利用的传感器有单目相机、双目相机和深度相机这三种。2007 年, A. J. Davison 提出的 MonoSLAM, 是世界上第一个可以实时使用的单目 SLAM 系统^[20]。利用单目相机的 SLAM 系统设备简单, 但是由于单目相机无法获取深度信息, 因此在单目 SLAM 系统中, 必须借助于其他传感器的测量数据或者某些先验条件进行初始化, 否则构建的地图以及估计的运动轨迹将与实际值存在未知比例的缩放。2014 年, 由 M. Labbe 和 F. Michaud 提出的 RTAB-MAP 方案, 利用深度相机, 实现了包括实时构建稠密地图在内的完整 SLAM 过程^[21]。利用深度相机的 SLAM 方案, 获取深度信息准确可靠, 算法简单, 但是受限于深度相机的发展, 测量距离一般小于三十米, 目前只能使用在室内小范围环境下。利用双目相机的 SLAM 方案, 通过双目视差获取物体的深度信息, 设备简单, 测量距离长, 是一种很有效的大环境下 SLAM 方案, 其难点是双目匹配算法复杂, 且需要较大的计算量。J. Stephen^[22]、T. Lemaire^[23]、D. Schleicher^[24]等学者对此进行了研究, 提出了可满足实时性要求的双目 SLAM 系统。

从后端优化方法上看, 早期的 SLAM 方法主要是基于概率估计的滤波方法。Smith 和 Cheeseman 提出了使用扩展卡尔曼滤波器的 SLAM 方案^[25], Montemerlo 和 Thrun 提出了使用粒子滤波器的 SLAM 方案^[26]。使用扩展卡尔曼滤波器的 SLAM 方案, 当地图变大时, 特征点增多, 计算量会急剧增大。据此学者们提出了各种优化的方法, 例如利用 SEIF^[27]或 UKF^[28]的 SLAM 方案。而使用粒子滤波的方法, 容易出现粒子退化的问题, 学者们也对此进行了研究, 如 Montemerlo 将扩展卡尔曼滤波器与粒子滤波器结合, 改进粒子的提议分布^[29]。随着高效求解方式的出现, 利用所有观测信息对地图及位姿的估计进行整体非线性优化的方法逐渐成为 SLAM 方案主要采用的方法。在此研究方向上, Lu 与 Milios 最早提出了基于图优化的 SLAM 方法^[30], 用节点表示机器人的位姿, 用节点之间的相连的边来表示位姿之间的约束关系, 通过调整节点的位置使得各节点能最好地满足边的约束关系。

对于火星漫游器而言,激光雷达功耗较高、设备较重,而视觉传感器功耗低、设备简单,更适合作为导航传感器。单目相机存在尺度不确定的问题,深度相机观测距离太短,因此适合火星漫游器导航定位的只有双目相机。考虑到火星漫游器活动范围较广,基于非线性优化的方法要优于滤波的方法^[31],更适合火星漫游器解决 SLAM 问题。

1.2.3 多机器人协同构建地图方法研究现状

目前国外学者对于多机器人构建地图的研究主要提出了两种方案。

第一种方案是每个机器人单独构建地图,当地图有重叠区域时将地图拼接在一起。这种方案既可以在线实时拼接地图,也可以离线拼接地图,对于机器人的初始位置没有要求,也不需要机器人之间进行实时通信,因此系统实现方案较为简单。这种方案的研究较为充分,有很多学者提出了自己的方法。Gutmann 和 Konolige 提出 LRGC (Local Registration Global Correlation) 方法^[32],将多个机器人构建的几何特征地图或拓补地图拼接在一起。Dedeoglu 和 Sukhatme 提出基于路标的地图拼接方法^[33],将不同机器人构建的拓补地图拼接在一起。Fox 等人提出基于贝叶斯滤波和马可夫模型的地图拼接方法^[34]。潘薇提出了一种在机器人会和时测量彼此之间的相对距离和角度,以此将多机器人构建的地图拼接在一起的方法^[35]。

第二种方案是已知每个机器人在全局坐标系下的初始位置,所有机器人同时构建同一幅地图。这种方案不需要机器人的行驶路径有重叠部分,机器人的行动更加灵活,但是需要确定机器人的初始位置,在地图拼接过程中需要机器人之间进行通信,系统更为复杂。Fenwick 等人最早对这种方案进行了研究,他们提出了一种使用扩展卡尔曼滤波器同时估计所有机器人位姿与路标点位置的方法^[36],但是这种方法的计算量十分巨大。苑全德提出了一种基于粒子滤波器的多机器人构建地图方法^[37],每个机器人单独使用粒子滤波器进行 SLAM 过程,同时将自己观测到的环境信息广播给其他机器人,并接受来自其他机器人的观测信息,据此对自己所构建的地图进行扩展。

第二种方案将所有漫游器的定位与地图的构建合并为一个大的 SLAM 问题,计算量较大,并不适合火星表面大范围探测的条件下使用。而第一种方案只需要不同的漫游器的路径有重叠区域即可。构建火星大规模地图,选用第一种方案较为合适。

1.2.4 空间三维地图研究现状

当前 SLAM 方法中使用的空间三维地图主要是特征点地图或几何特征地图,利用图像中的特征点或者几何特征确定机器人在地图中的位置。这两种地图可以用于定位,并且所需存储空间较小,但由于其稀疏性,不能表示空间的占据情况,无法用于自主路径规划。据此,一些学者提出了构建稠密地图的方法。

Andreas 等人提出了一种利用双目图像构建三维稠密点云地图的方法^[38],该方法构建的地图利用大量三维点云表示环境信息,有利于地图中的物体识别以及人机交互,但是会占据很大的存储空间。Mason 等人利用单目相机和激光雷达获取空间中的三维稠密点云,在此基础上构建稠密栅格地图^[39],将空间划分为二维栅格,估计每个栅格的被物体占据的概率。这种方法构建的地图容易构建、保存,也易于在路径规划中使用,但是需要使用激光雷达获取距离信息,而且地图占据的存储空间仍然较大。在其基础上 Hornung 等人设计了一种三维占有率栅格地图 OctoMap^[40],这种地图利用八叉树的形式对空间进行三维分割,如果某个父节点下面的所有子节点都是被物体占据的状态或都是未被物体占据状态,则将所有子节点合并,只用父节点对空间占据状态进行表示,因此具有很好的压缩性,地图存储空间大大缩小。

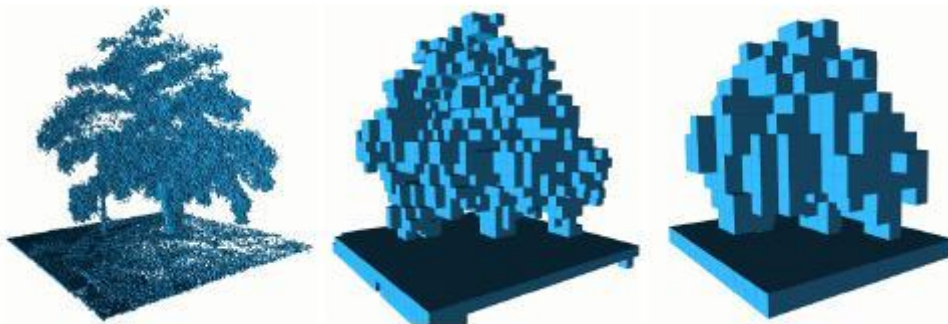


图 1-3 八叉树地图

此外还有学者提出了利用节点与节点间连通路描述环境的拓补地图^[41],这种地图存储空间小、结构紧凑,利于进行自主路径规划,但是只适用于简单环境下,并不适用于火星表面这种复杂的环境。语义地图^[42]通过对环境中的物体进行识别、标记,描述环境信息,有利于进行人机交互,但是也不适用于火星表面这种无人造物体的环境。

稀疏的特征点地图即可满足火星漫游器的自主定位,但是要想为后续的漫

游器提供路径规划的指导信息，就需要使用稠密地图。稠密的三维点云地图和稠密三维栅格地图虽然能显示空间的结构，但是在大范围环境下地图的存储量过于庞大。利用八叉树分割空间的三维占有率地图则可在占据较小的存储空间的前提下，表示大范围地形的空间占据情况，适合火星漫游器自主路径规划时使用。

1.3 本文主要研究内容与结构

本文提出一种多漫游器协同构建火星表面导航地图的方法。对于在行进区域有重叠的多个漫游器，首先利用双目视觉 SLAM 方法，对每个漫游器进行自主定位，并构建火星表面的稀疏特征点地图；然后根据漫游器拍摄的图像，寻找多个漫游器行进区域的重叠部分，利用重叠区域中的特征点，对多个漫游器构建的稀疏特征点地图进行拼接，形成更大范围的稀疏特征点地图，用于后续其他漫游器的自主定位。最后利用漫游器拍摄的双目图像和漫游器位姿信息，构建大范围稠密点云地图，并据此构建表示空间占据情况的八叉树地图，为后续其他漫游器的自主路径规划提供信息。

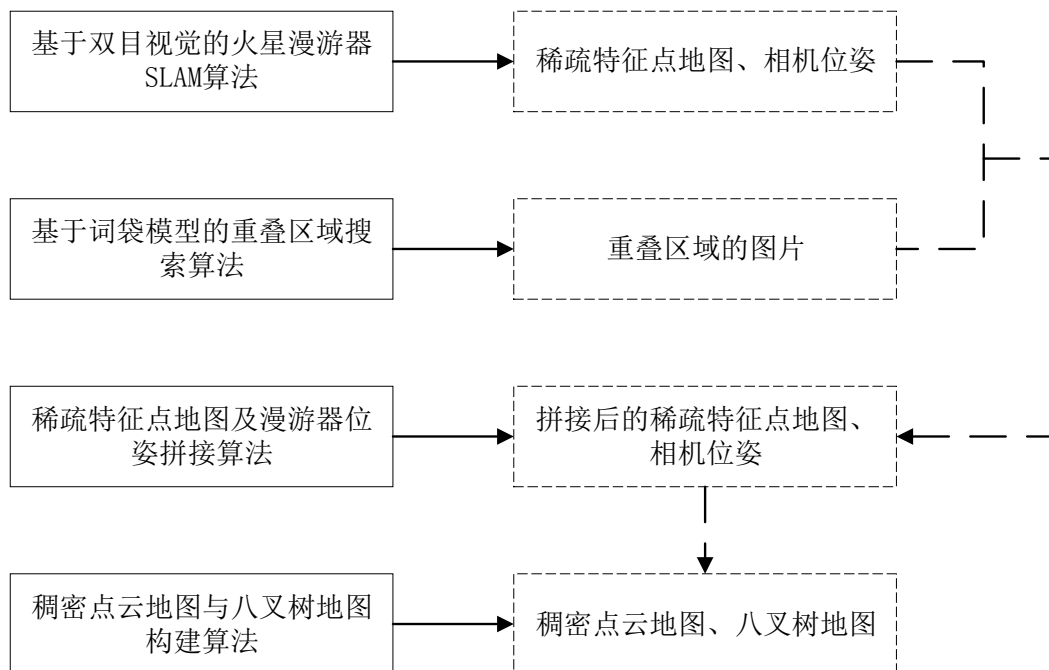


图 1-4 论文结构框图

(1) 基于双目视觉的火星漫游器 SLAM 算法

建立双目相机模型，从双目图像中提取特征点并利用相机模型建立三维点，对前后帧图像获取的特征点进行匹配，剔除误匹配点对，利用奇异值分解法求

解相机位姿，进而得到漫游器位姿。最后利用“勇气”号火星车拍摄的图像以及双目相机在模拟场地采集的图像对算法进行仿真验证。

（2）基于词袋模型的重叠区域寻找算法

研究利用词袋模型建立字典、根据字典生成图片的描述向量、根据描述向量查找相似图片的方法，并利用“好奇”号火星车拍摄的图像对算法进行仿真验证。

（3）稀疏特征点地图及漫游器位姿拼接算法

研究利用重叠区域的双目图像拼接两幅特征点地图的方法。首先利用多组重叠区域图像提取特征点、建立三维点云，然后计算特征点地图坐标变换关系，进而对两幅特征点地图以及漫游器的位姿估计进行拼接。最后利用双目相机在模拟场地采集的图像以及电脑仿真图像对算法进行仿真验证。

（4）稠密点云地图与八叉树地图的构建算法

研究利用双目图像构建稠密点云地图的方法和利用稠密点云地图构建八叉树地图的方法，并利用双目相机在模拟场地采集的图像以及电脑仿真图像对算法进行仿真验证。

第2章 基于双目视觉的火星漫游器 SLAM 算法

2.1 引言

单漫游器准确实现自主定位与地图构建 (SLAM)，既是其执行探测任务的基础，也是实现地图拼接与构建八叉树的基础。双目相机以其功耗小、可视距离远、获取信息量大等优点，适用于火星漫游器的导航。本章研究利用双目相机实现火星漫游器自主定位与地图构建的方法，并利用“勇气”号火星车拍摄的图像、双目相机在模拟场地采集的图像以及计算机模拟图像对算法进行了仿真验证。

2.2 相机模型

2.2.1 相机成像模型

以相机光心 O 为原点，通过光心、垂直于成像平面指向相机前方的射线为 z 轴， y 轴向下， x 轴根据按右手系确定，建立相机坐标系 $Oxyz$ 。相机焦距为 f 。三维世界中的点 P 经过光心投影到成像平面点 P' 上，如图 2-1 所示。点 P 坐标为 (X, Y, Z) ，点 P' 坐标为 (X', Y', Z') ，则根据三角形相似关系有：

$$\frac{Z}{f} = \frac{X}{-X'} = \frac{Y}{-Y'} \quad (2-1)$$

以所成图像右上角 O' 为原点， u 轴与相机坐标系 x 轴方向相同， v 轴与 y 轴方向相同，建立像素坐标系 $O'uv$ 。像素坐标系与成像平面之间有一定缩放。设像素坐标与成像平面相比，在 u 轴方向缩小了 a 倍，在 v 轴方向缩小了 b 倍，所成图像在 u 轴方向上的长度为 $2c_x$ ，在 v 轴方向上的长度为 $2c_y$ 。则 P' 点在像素坐标系下的坐标为：

$$\begin{cases} u = a \cdot X' + c_x = -f \cdot a \cdot \frac{X}{Z} + c_x = f_x \cdot \frac{X}{Z} + c_x \\ v = b \cdot Y' + c_y = -f \cdot b \cdot \frac{Y}{Z} + c_y = f_y \cdot \frac{Y}{Z} + c_y \end{cases} \quad (2-2)$$

将上式转换成矩阵形式为：

$$Z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (2-3)$$

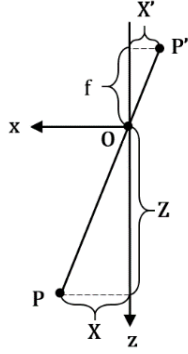


图 2-1 相机成像模型

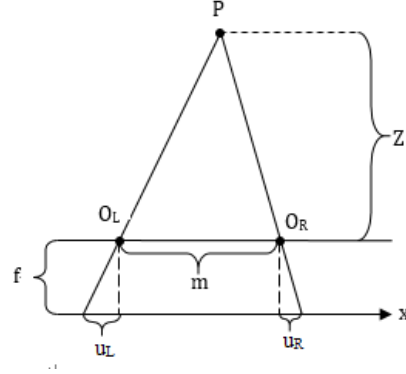


图 2-2 双目相机测距模型

2.2.2 双目相机测距模型

双目相机获得景物深度信息的模型如图 2-2 所示。设左、右两相机的光心分别为 O_L 和 O_R 。 O_L 、 O_R 之间的距离为 m ，称为基线长度。两相机焦距均为 f 。三维世界中的点 P 在两成像平面上的投影（即点 P 在像素坐标系下的坐标）分别为 u_R 和 u_L 。点 P 在相机坐标系下的深度为 Z 。根据三角形相似关系有：

$$\frac{Z+f}{Z} = \frac{m-u_L+u_R}{m} \quad (2-4)$$

整理得：

$$Z = \frac{f \cdot m}{u_R - u_L} = \frac{f \cdot m}{d} \quad (2-5)$$

式(2-5)中， d 是同一个世界点在左右相机成像平面上投影坐标的偏差，称为视差，其最小值为一个像素。因此在相机焦距一定的条件下， Z 存在最大值，即双目相机可以测量的深度是有一定范围的。基线长度 m 越大， Z 的最大值越大，双目相机可以测量的最大深度越远。

对于图像上的点 P ，根据式(2-5)得到点 P 的深度信息后，利用式(2-3)，即可计算点 P 所对应的三维世界中的点的坐标。

2.2.3 图像畸变的校正

由于相机 CCD 元件的制造误差、镜头的制造误差、光线在透镜中的折射、透镜光轴安装的偏差等原因，相机实际拍摄的图片与针孔模型所成的理想图像之间存在一定的误差，称为图像的畸变。图像的畸变又可根据像素偏离的方向分为径向畸变和切向畸变，如图 2-3 所示。

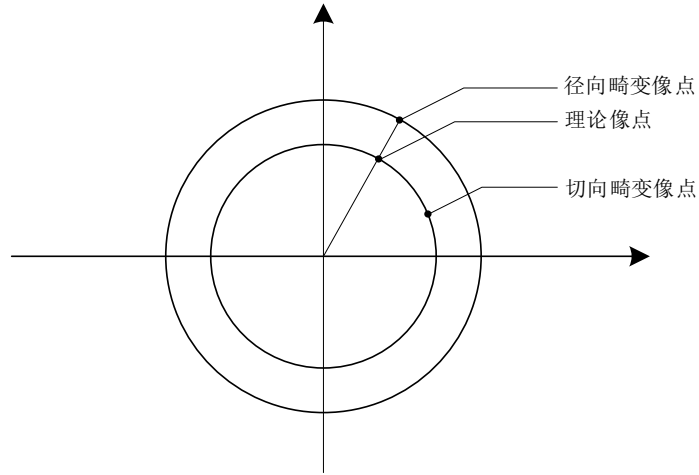


图 2-3 图像畸变示意图

径向畸变可以由如下多项式建立畸变模型：

$$\begin{cases} x_{disorted} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \\ y_{disorted} = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \end{cases} \quad (2-6)$$

式中， $[x \ y]^T$ 是点在相机归一化平面的坐标； $[x_{disorted} \ y_{disorted}]^T$ 是畸变后点的坐标； $r = \sqrt{x^2 + y^2}$ ， $[k_1 \ k_2 \ k_3 \ k_4]$ 是相机径向畸变系数。对于除鱼眼镜头外的其他相机，利用 $[k_1 \ k_2]$ 两个系数就能很好地描述径向畸变。

对于切向畸变，可以使用另外两个参数 $[p_1 \ p_2]$ 建立畸变模型：

$$\begin{cases} x_{disorted} = x + 2p_1 xy + p_2(r^2 + 2x^2) \\ y_{disorted} = y + p_1(r^2 + 2y^2) + 2p_2 xy \end{cases} \quad (2-7)$$

式中， $[x \ y]^T$ 是点在相机归一化平面的坐标； $[x_{disorted} \ y_{disorted}]^T$ 是畸变后点的坐标； $r = \sqrt{x^2 + y^2}$ ， $[p_1 \ p_2]$ 是相机切向畸变系数。

相机的径向、切向畸变系数可以通过相机标定的过程获得。对于相机标定方法的研究最早可追溯至上世纪 70 年代，至今已发展得十分完善，典型的标定方法如 DLT 方法、RAC 定标算法、张正友标定法等都能很精确地获取

相机畸变参数。利用这些参数建立畸变模型，即可对畸变的图像进行校正。

对畸变图像的校正过程为：

(1) 将像素坐标系下的点利用相机内参矩阵转换到相机归一化坐标系中：

$$\begin{bmatrix} x_{disorted} \\ y_{disorted} \\ 1 \end{bmatrix} = K^{-1} \begin{bmatrix} u_{distorted} \\ v_{distorted} \\ 1 \end{bmatrix} \quad (2-8)$$

(2) 利用畸变模型，求解方程组计算校正后的坐标 $[x, y]^T$ ：

$$\begin{cases} x_{disorted} = x + 2p_1xy + p_2(r^2 + 2x^2) + x(k_1r^2 + k_2r^4 + k_3r^6) \\ y_{disorted} = y + p_1(r^2 + 2y^2) + 2p_2xy + y(k_1r^2 + k_2r^4 + k_3r^6) \\ r = \sqrt{x^2 + y^2} \end{cases} \quad (2-9)$$

(3) 利用内参矩阵将校正后的坐标转换到像素坐标系下：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2-10)$$

如此即完成了对畸变图像的校正。本文在相机获取图像后，首先进行畸变校正，然后再进行后续的处理。后续章节中使用的图像都默认是已经过校正的图像。

2.3 特征点提取与匹配

2.3.1 特征点检测方法

漫游器的自主定位与地图构建依赖于特征点快速、准确的检测与识别。

为了保证特征点检测的快速性，采用 FAST 角点检测方法。这种方法的思想是：如果有一个像素与周围像素的灰度值差异较大，则这个像素对应着一个特征点，如图 2-4 所示。它的检测过程如下：

- (1) 设定检测的阈值 T ；
- (2) 在图像中选取像素点 P ，得到其灰度值 I_p ；
- (3) 以 P 为圆心，3 个像素为半径，得到一个由 16 个像素点组成的圆；
- (4) 判断圆上是否有连续的 N 个点 (N 通常取 12，称为 FAST-12)，它们的灰度值均大于 $I_p + T$ 或均小于 $I_p - T$ ，如果是，则像素点 P 为一个特征点；
- (5) 重复 (2) 到 (4) 步，判断图像上的每一个点是否为特征点。

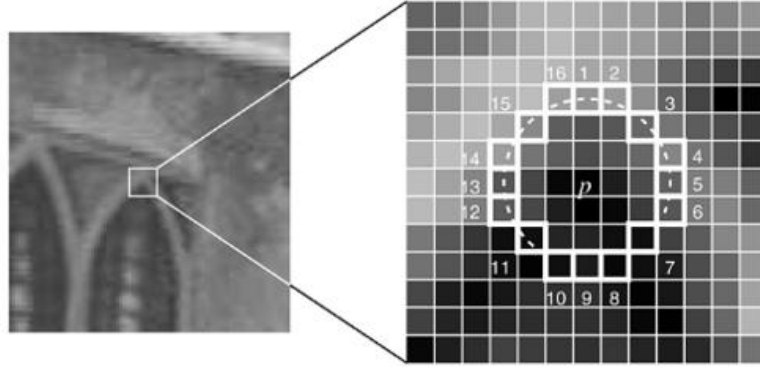


图 2-4 FAST 角点示意图

通过这种方式检测到的特征点，可能在某个区域特别集中，因此还需要通过非极大值抑制的过程，在一定区域内只保留 Harris 响应值最大的特征点。

此外，由于在检测特征点的时候只考虑了以 3 个像素为半径的圆，在距离发生改变时，得到特征点也会发生变化，因此需要构建图像金字塔，对于同一图像进行不同层次的降采样，在不同分辨率的图像上分别进行检测，使特征点获得尺度不变性，在不同距离上均能被准确识别。

2.3.2 特征点表示方法

在检测到特征点之后，需要对每个特征点计算描述子，用来区分不同的特征点。

BRIEF 描述子是一种二进制描述子，在特征点周围，按照一定的模式选取 N 组点对，对于每组点对中 A 、 B 两个点，提取它们的灰度值 I_A 、 I_B ，并进行比较：

$$T = \begin{cases} 0, I_A \leq I_B \\ 1, I_A > I_B \end{cases} \quad (2-11)$$

利用这 N 组点对的比较结果，形成一个长度为 N 的二值码串，即为此特征点的描述子。描述子与选择点对时的基础方向有关，当图像发生旋转时，基础方向发生变化，对于特征点的描述就会发生改变。针对这一问题，在得到特征点后，可以利用灰度质心法确定生成描述子时的基础方向，使特征点获得方向不变性。具体做法是：

(1) 首先以特征点 O 为中心选取一个图像块 B ，计算图像块的灰度矩：

$$m_{pq} = \sum_{x,y \in B} x^p y^q I(x,y) \quad (2-12)$$

其中—— $p, q = \{0, 1\}$ 。

(2) 通过图像块的灰度矩可以找到图像块的灰度质心 C ：

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (2-13)$$

(3) 连接 \overline{OC} ，并以此作为基础方向，选取点对并形成描述子。当图像发生旋转时，同一个特征点的基础方向不发生改变，描述子将基本保持不变，特征点仍能被准确识别。

2.3.3 特征点匹配方法

基于特征点的视觉 SLAM 方法，根据相机在不同时刻对相同的特征点观测计算相机位姿的变化。因此需要对不同图像中的特征点进行匹配，寻找不同图片中拍摄到的相同特征点。

对两幅图像中的特征点进行匹配，最简单的方法是暴力匹配，即对图像 A 中的每一个特征点，分别用图像 B 中的每一个特征点与其进行比较，寻找图像 B 中与之最为相似的特征点，若两个特征点的描述子相差小于一定阈值，则将这两个特征点视为匹配特征点。这样得到的匹配结果中会存在一些误匹配的情况，需要进行剔除。具体做法是：

- (1) 取图像 A 中的特征点 S_i ，计算图像 B 中每一个特征点与 S_i 的汉明距；
- (2) 取图像 B 中与 S_i 汉明距最小的特征点 P_i ，若 S_i 与 P_i 汉明距小于阈值，则将 S_i 与 P_i 视为一对匹配点；
- (3) 重复 (1)、(2) 步直至遍历图像 A 中的每一个特征点；
- (4) 将图像 A、B 所有匹配特征点之间最小的汉明距设为 d ，剔除汉明距大于 $2d$ 的匹配点对。

也可根据实际情况调整匹配点对和剔除误匹配点对的阈值条件。

对于左、右眼图像的匹配点对，可以利用对极约束剔除误匹配。

假设在左眼相机坐标系下存在三维点 $P_1 = [X_1 \ Y_1 \ Z_1]^T$ ，其在左眼图像中对应的像素点为 $p_1 = [u_1 \ v_1 \ 1]^T$ ，在右眼相机坐标系下对应的三维点为 $P_2 = [X_2 \ Y_2 \ Z_2]^T$ ，在右眼图像中对应的像素点为 $p_2 = [u_2 \ v_2 \ 1]^T$ 。根据 2.2.1 节介绍的相机成像模型，有：

$$Z_1 p_1 = K P_1 \quad (2-14)$$

$$Z_2 p_2 = K P_2 \quad (2-15)$$

因为双目相机的成像平面在同一平面上，因此同一个三维点在两个相机坐标系下的深度相同，即有：

$$Z_1 = Z_2 \quad (2-16)$$

左、右眼相机的相对位置是已知的，即两相机坐标系的坐标系转换关系 R 、 t 是已知的，即：

$$P_2 = R P_1 + t \quad (2-17)$$

将式(2-14)、(2-15)带入到(2-17)中可得：

$$Z_1 R K^{-1} p_1 + t = Z_2 K^{-1} p_2 \quad (2-18)$$

两边同时左乘 t^\wedge ，有：

$$t^\wedge R K^{-1} p_1 = t^\wedge K^{-1} p_2 \quad (2-19)$$

对式(2-19)两边再同时左乘 $(K^{-1} p_2)^T$ ，可得：

$$(K^{-1} p_2)^T t^\wedge R K^{-1} p_1 = 0 \quad (2-20)$$

即：

$$p_2^T K^{-T} t^\wedge R K^{-1} p_1 = 0 \quad (2-21)$$

计算从左、右眼图像中获取的匹配点对是否在给定的误差下满足式(2-21)，剔除不满足的点对，即可剔除左、右眼图像特征点的误匹配。

2.4 漫游器位姿求解方法

在得到不同时刻相机拍摄图像的匹配特征点后，利用双目相机模型可以将匹配点还原到三维世界中。在 t_1 时刻得到相机坐标系下 n 个三维世界中的点的坐标： $P_1' \dots P_n'$ 。

经过一段较短的 Δt 时间，漫游器发生了位置、姿态的变化，相机坐标系也发生了旋转与平移，在 t_2 时刻得到相机坐标系下上述 n 个三维世界中的点的坐标： $P_1 \dots P_n$ 。

假设相机固定在漫游器上，不发生移动，则求解漫游器位姿变化的问题可以转化为寻找旋转矩阵 R 和平移向量 t ，使得：

$$\forall i, P_i = R P_i' + t \quad (2-22)$$

求解这个问题，首先构造误差函数 J ：

$$J = \frac{1}{2} \sum_{i=1}^n \|P_i - (RP_i' + t)\|^2 \quad (2-23)$$

设两组点的质心分别为 P 和 P' ，则有：

$$P = \frac{1}{n} \sum_{i=1}^n (P_i) \quad (2-24)$$

$$P' = \frac{1}{n} \sum_{i=1}^n (P_i') \quad (2-25)$$

将式(2-24)、(2-25)代入到式(2-23)中得：

$$J = \frac{1}{2} \sum_{i=1}^n \|P_i - P - R(P_i' - P')\|^2 + \|P - RP' - t\|^2 \quad (2-26)$$

要获得最小的误差函数 J ，只需要先找到 R 使得第一项最小，然后找到 t 使得第二项为 0 即可。设每个点的去质心坐标为：

$$q_i = P_i - P \quad (2-27)$$

$$q_i' = P_i' - P' \quad (2-28)$$

则式(2-26)中第一项可以写为：

$$\frac{1}{2} \sum_{i=1}^n \|q_i - Rq_i'\|^2 = \frac{1}{2} \sum_{i=1}^n q_i^T q_i + q_i'^T q_i' - 2q_i^T Rq_i' \quad (2-29)$$

式(2-29)中只有第三项与 R 有关，只需寻找到使其最小的 R 即可使误差函数达到最小。对式(2-29)第三项进行变换，可得：

$$\sum_{i=1}^n -q_i^T Rq_i' = \sum_{i=1}^n -tr(Rq_i' q_i'^T) = -tr\left(R \sum_{i=1}^n q_i' q_i'^T\right) \quad (2-30)$$

设：

$$W = \sum_{i=1}^n q_i' q_i'^T \quad (2-31)$$

对 W 进行奇异值分解 (SVD)，可得：

$$W = UDV^T \quad (2-32)$$

其中， D 为奇异值对角阵，对角线元素从大到小排列。 U 和 V 为对角阵。

则使式(2-30)最小的 R 为：

$$R = UV^T \quad (2-33)$$

令式(2-26)第二项为 0 可得 t ：

$$t = P - RP' \quad (2-34)$$

至此，计算得到了漫游器在 t_2 时刻相对于 t_1 时刻的姿态旋转矩阵 R 和平移向量 t 。

在导航过程中，提取、匹配双目图像中的特征点，得到每个特征点的深度，然后匹配两帧左眼图像中的特征点，利用相机模型将特征点还原到三维空间中，最后利用奇异值分解法求解相机位姿变化。将相机位姿变化进行积分，即可得到漫游器每一时刻的位姿状态。

2.5 仿真实验与分析

2.5.1 基于漫游器实际拍摄图像的特征点提取与匹配实验

美国“勇气”号火星漫游器携带有双目导航相机、全景相机和避障相机，在长达七年多的工作过程中，获取并向地球传回了大量的火星地表图像。图 2-5 是漫游器双目导航相机获取的一组图片，图 2-5 (a) 是左眼相机获取的图像（图像编号：2n225260056radasnuf0006l0c1），图 2-5 (b) 是右眼相机获取的图像（图像编号：2n225260056radasnuf0006r0c1）。

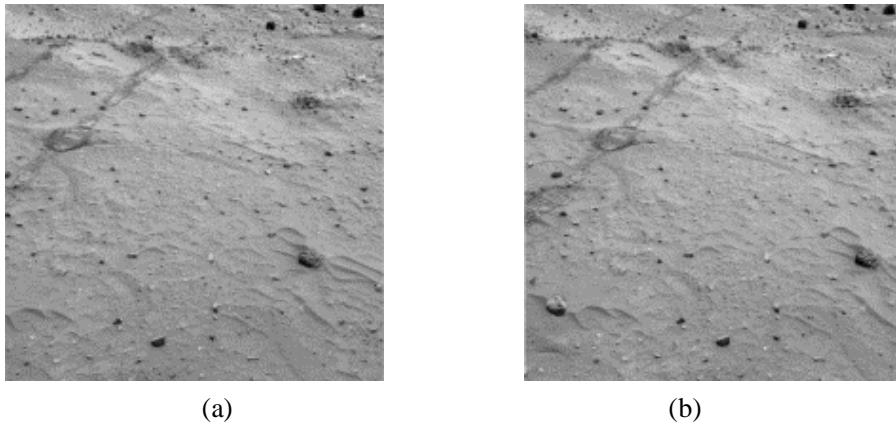


图 2-5 “勇气”号漫游器双目相机获取的图片

利用 2.3 节的方法提取并匹配左、右眼图像中的特征点，如图 2-6 所示。剔除误匹配后，剩余的特征点匹配如图 2-7 所示。可以看出，本章介绍的特征点提取与匹配方法可以有效提取出火星表面图像的特征点并剔除误匹配。

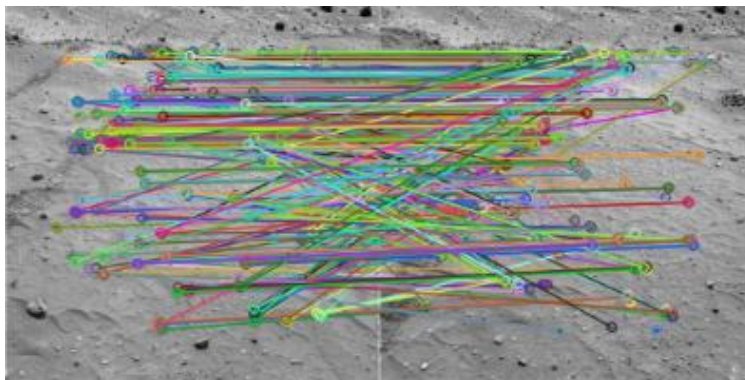


图 2-6 两幅图像所有匹配的特征点

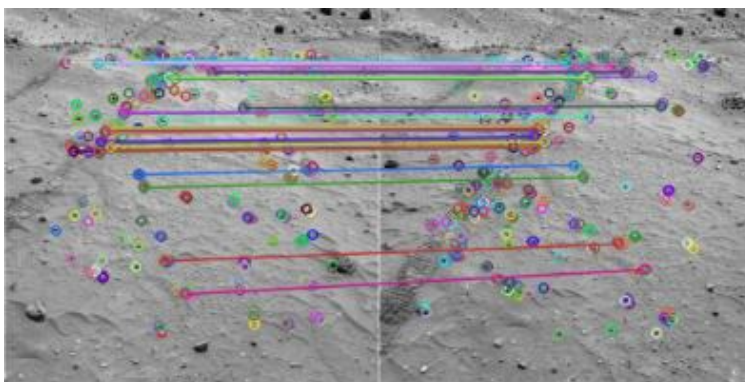


图 2-7 利用汉明距阈值剔除误匹配后的匹配结果

2.5.2 基于地面模拟场地图像的漫游器 SLAM 仿真实验

由于缺少长路径上火星漫游器连续拍摄的双目图像数据，本节采用地面模拟场地采集的图像对漫游器自主定位与稀疏特征点地图构建的算法进行仿真验证。

地面模拟场地的环境以沙土为主，其中间杂有石块分布，与火星表面相类似。相机采用双目灰度相机，图片采集频率为 10Hz，所采集的图像宽度为 752 像素，高度为 480 像素。仿照美国“勇气”号、“机遇”号等火星漫游器的双目导航相机的朝向，采用斜向下的角度进行拍摄。实验所用电脑配备 i5-3210M 处理器，4G 内存。在采集图像时，利用 Lighthouse 室内定位技术，将跟踪器与相机固连，以 60Hz 的频率获取跟踪器的位姿信息，作为相机位姿的标准值，与实际位置的误差在 $\pm 2\text{mm}$ 以内。

首先利用张正有标定法，对双目相机进行标定，以获取相机的畸变系数，对图像进行校正。最终获得的相机的参数如下表所示。

表 2-1 左眼相机参数

参数	数值
焦距	467mm
畸变系数 K1	-0.365499
畸变系数 K2	0.103475
畸变系数 K3	-0.003296
畸变系数 P1	-0.001389
畸变系数 P2	0.000000

表 2-2 右眼相机参数

参数	数值
焦距	467mm
畸变系数 K1	-0.413433
畸变系数 K2	0.146088
畸变系数 K3	-0.003315
畸变系数 P1	0.001147
畸变系数 P2	0.000000

利用 2.2.3 节介绍的方法对图像畸变进行校正，其效果如图 2-8 所示。图 2-8 (a) 是校正前的图像，图 2-8 (b) 是校正后的图像。可以看出，畸变前弯曲的直线得到了很好的修正，畸变校正方法准确有效。双目相机在地面模拟场地拍摄的图像校正后如图 2-9 所示。图 2-9 (a) 是左眼图像，图 2-9 (b) 是右眼图像。



图 2-8 图像去畸变效果示意图

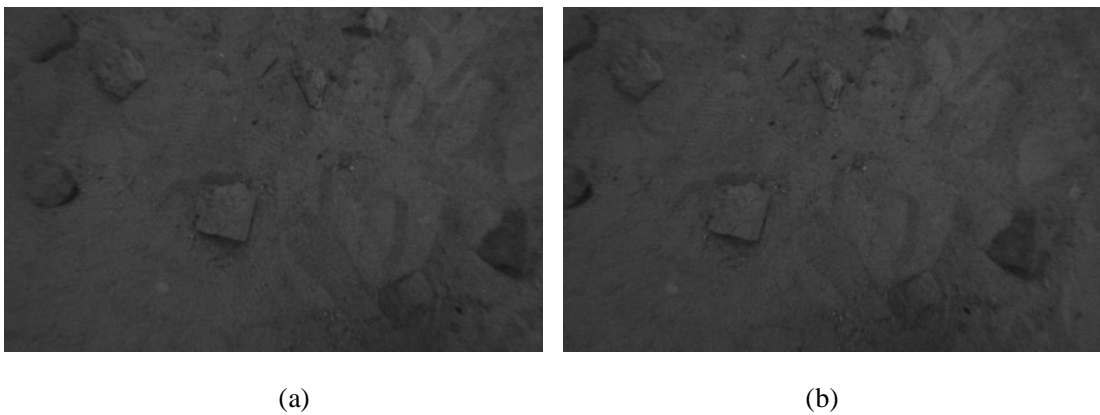


图 2-9 双目相机在地面模拟场地拍摄的图像

在地面模拟场地中，根据本章介绍的方法，以漫游器第一帧左眼相机坐标系为参考坐标系，利用双目相机拍摄图像，进行漫游器视觉自主定位与稀疏特征点地图构建的实验。图像处理速度为每秒 10 帧，共输入 253 幅图像进行仿真实验。

在第一帧图像的左眼相机坐标系下，绘制实验所得的稀疏特征点地图与运动轨迹如图 2-10 所示。横、纵坐标轴的单位为米。其中的负号是因为相机斜向下拍摄，根据坐标系的定义，向前行进为 y 轴负方向，向左行进为 x 轴负方向。

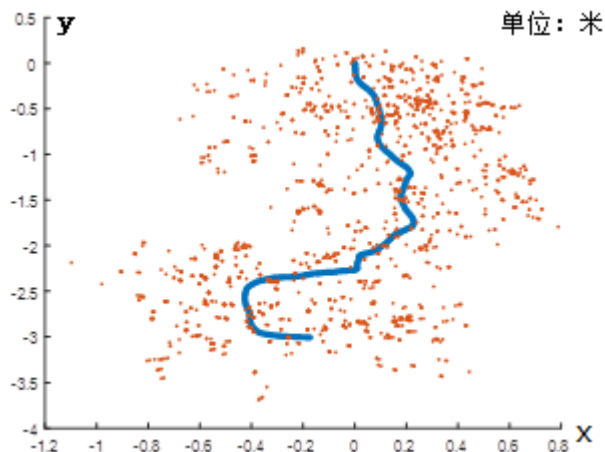


图 2-10 稀疏特征点地图与运动轨迹示意图

本次实验中，在第一帧图像的左眼相机坐标系下，漫游器初始位置为 $[0 \ 0 \ 0]^T$ ，末位置估计值为 $[-0.1728 \ -3.0051 \ 0.9520]^T$ ，漫游器自主定位得到的始末位置距离为 3.16 米。而由实际测量得到的漫游器始末位置距离为 3.18 米，误差为 0.02 米，占实际值的 0.63%。

由于两帧之间的时间间隔较小（约为 0.1s），漫游器位置变化不大，因此假设在两帧之间漫游器沿直线运动，计算漫游器的行进里程。根据跟踪器的测量值，漫游器实际行进里程为 3.70 米，由自主定位得到的行进里程为 3.60 米，误差为 0.10 米，占实际值的 2.7%。

选取地面试验场地的不同位置，进行多次实验，测得误差如表 2-3 所示。

表 2-3 漫游器自主定位误差

序 号	始末位置距离 实际值（米）	始末位置距离 估计值（米）	始末位置 距离误差	行进路程实 际值（米）	行进路程估 计值（米）	行进路 程误差
1	3.18	3.16	0.63%	3.70	3.59	2.97%
2	2.60	2.65	1.92%	4.00	3.89	2.75%
3	2.57	2.53	1.56%	2.80	2.71	3.21%
4	2.68	2.72	1.49%	3.50	3.42	2.29%

由上表可以看出，漫游器行进里程的误差平均在 3% 左右。在长距离探测时，考虑到误差的积累，每行进一定距离后可以使用无线电测程等精确定位方法，对漫游器位置进行校准，以满足火星漫游器探测任务的要求。

2.5.3 基于计算机模拟图像的漫游器 SLAM 仿真实验

由于地面模拟场地规模较小，漫游器的双目相机采用斜向下的角度进行拍摄。为了验证当双目相机平视行进方向时仍能利用本章介绍的方法进行火星漫游器的自主定位与稀疏特征点地图的构建，本节利用计算机模拟图像进行仿真实验。

首先利用 3DMax 软件制作火星表面的模拟场景，并绘制两条轨迹，如图 2-11 所示。在场景中添加两组双目相机，渲染得到火星漫游器沿轨迹行进时拍摄到的双目图像，如图 2-12 所示。红色路径长为 54.0000m，绿色路径长为 123.1942m，两条路径上分别均匀采集 1000 帧图像。

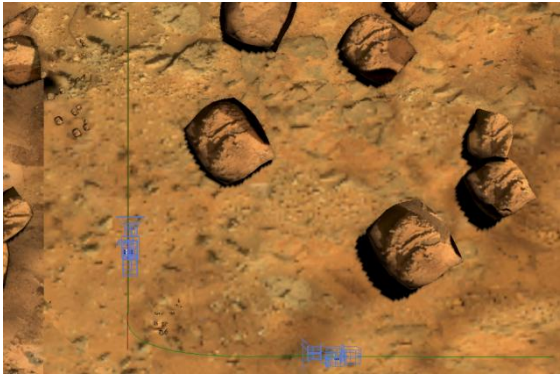


图 2-11 火星表面模拟场景

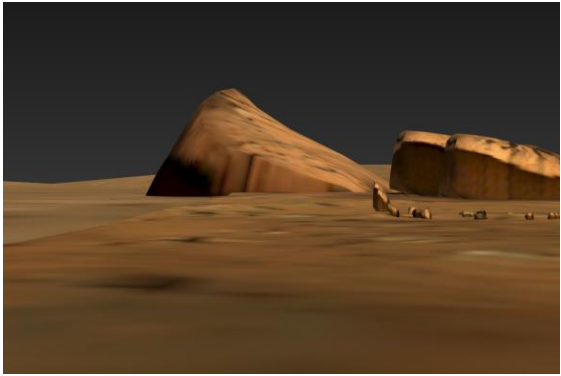


图 2-12 计算机模拟图像

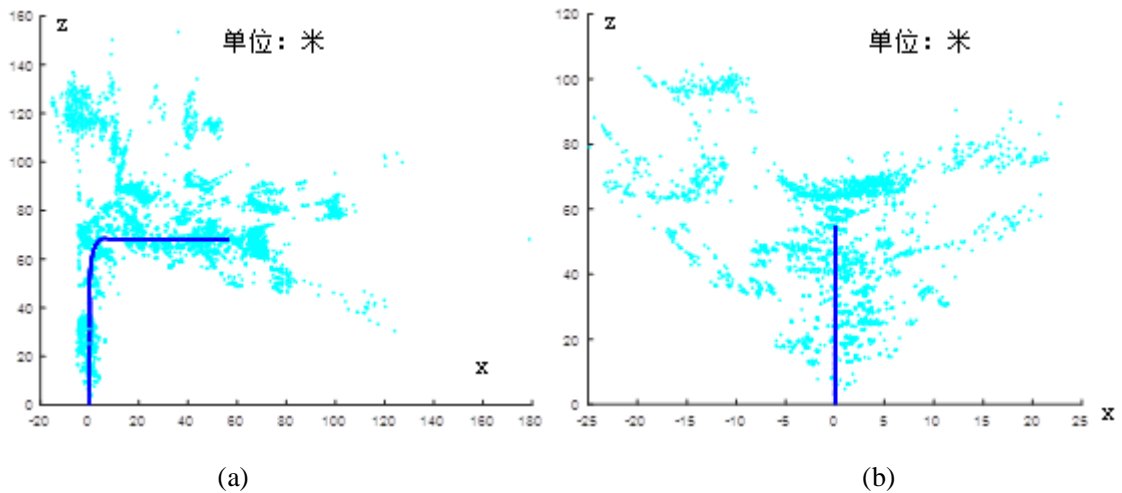


图 2-13 由模拟图像构建的稀疏特征点地图与轨迹

根据本章介绍的方法，利用两组图像进行漫游器的自主定位与稀疏特征点地图的构建，分别在其各自的第一帧左眼相机坐标系下绘制实验结果，如图

2-13 所示。图 2-13(a)是第一组图像计算的结果，对应图 2-11 中绿色的轨迹。图 2-13 是第二组图像计算的结果，对应图 2-11 中红色的轨迹。

测得误差如表 2-4 所示。由表可见，漫游器行进路程的误差平均在 2% 以下，与上节地面模拟场地的实验结果相近。漫游器相机视角与行进方向平行时，利用本章介绍的方法仍能获得较高的定位精度。

表 2-4 漫游器自主定位误差

序号	始末位置距 离实际值 (米)	始末位置距 离估计值 (米)	始末位置 距离误差	行进路程实 际值(米)	行进路程估 计值(米)	行进路 程误差
1	90.0000	88.4433	1.73%	123.1942	121.6789	1.23%
2	54.0000	54.9537	1.77%	54.0000	54.9537	1.77%

2.6 本章小结

本章研究了利用双目相机实现火星漫游器自主定位与地图构建的方法。

首先建立了相机模型，研究了相机成像模型以及双目相机测距模型，并且研究了利用畸变参数校正图像畸变的方法。然后研究了特征点提取与匹配的方法，包括从图像中提取 FAST 角点作为特征点、利用 BRIEF 描述子与灰度质心法生成具有旋转不变性的描述子、以及根据描述子之间的汉明距对特征点进行匹配并剔除误匹配的方法。最后研究了基于矩阵奇异值分解对漫游器位姿求解的方法。

本章利用“勇气”号火星车拍摄的图像验证了提取特征点、表示特征点、匹配特征点以及剔除误匹配的方法对于火星表面的图像是有效可行的；利用双目相机在模拟场地采集的图像以及计算机模拟图像对漫游器视觉自主定位以及稀疏特征点地图构建的整个过程进行了仿真验证，证明本章提出的方法可以用于火星表面漫游器的自主定位与稀疏特征点地图的构建。并对精度进行了分析。

第3章 基于词袋模型的重叠区域搜索算法

3.1 引言

本文利用两个漫游器行进路径上重叠区域的图像计算特征点地图与漫游器位姿的拼接关系。由于漫游器行进路径长、拍摄图片数量大，难以通过人工方式寻找重叠区域的图像。本章研究利用词袋模型寻找相似图片的算法，并利用“勇气”号火星漫游器拍摄的图像以及双目相机在模拟场地采集的图像对算法进行了仿真验证。

3.2 基于词袋模型的字典生成算法

3.2.1 词袋模型的概念

词袋模型（Bag of Word）最早是为了解决文档分类问题而提出的。对于一篇文档，忽略其中单词的顺序、句子的语法等因素，只根据字典，按顺序统计每种单词的出现次数，并依次将每种单词出现的次数记录下来，这样就可以形成一个离散的数值向量，用以描述这篇文档，并与其他文档进行比较。假设单词统计结果如图 3-1 所示，则文档的描述向量为（3，5，4，6）。

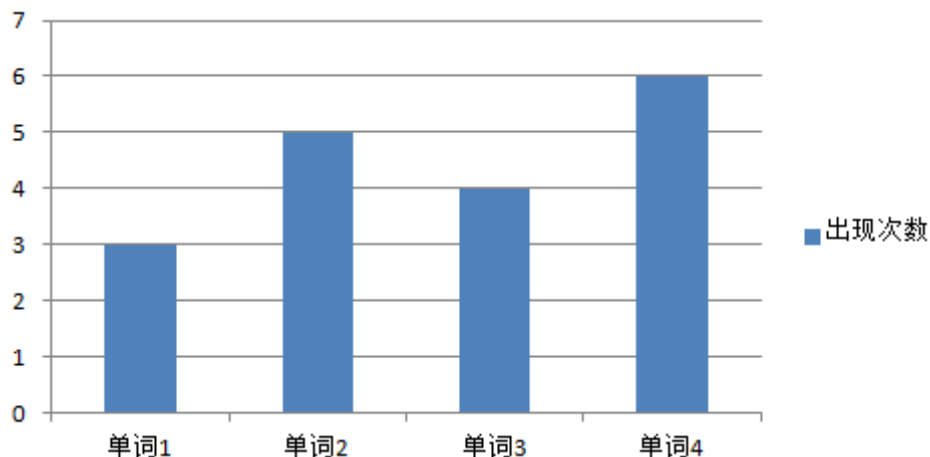


图 3-1 生成描述向量示意图

对于一幅图片，也可以利用这种思想，根据字典，按顺序统计每种特征点出现的次数，并依次将这些特征点出现的次数记录下来，形成一个离散的数值向量，用以描述这幅图片，并与其他图片进行比较。与文档分类的区别是，这

里使用的字典是利用大量图片提取的特征点构建的,并且由于难以完全穷尽所有特征点,对于待描述的图片,提取的每一个特征点需要依据某种判别方法寻找其在字典中最接近的单词。

3.2.2 字典的构建方法

字典是众多单词的集合,每一个单词代表一类特征点。生成字典时,首先选取图像作为训练集,对训练集的图像提取特征点,然后根据特征点的描述子,将这些特征点进行聚类,形成单词。

由于寻找重叠区域时,对准确性要求较高,而对实时性并没有太大要求,因此选择耗时较长,但在单幅图片中可以提取到更多数量的特征点、并且对特征点描述更为准确的 SIFT 特征点,用于构建字典以及区分图片。

在字典结构上,选择了 k-d 树的结构,如图 3-2 所示。首先利用 K-means 聚类方法把训练集图片提取到的所有特征点根据描述子的不同分为 k 个子节点,作为第一层结构。对于第一层的每个子节点,又利用 K-means 聚类方法继续分为 k 个节点,作为下一层结构。以此类推,分为 d 层,最后一层共包含 k^d 个节点,作为字典的单词。根据已知特征点查找单词时,利用特征点的描述子,从根节点开始,与每一层的节点进行比对,寻找其对应的子节点,直到进入 d 层,即可寻找到特征点对应的单词。可以看到,一方面,利用 k-d 树构建的字典最多可包含 k^d 个单词,保证了字典的丰富性;另一方面,输入一个特征点只需要比对 $k \cdot d$ 次即可找到对应的单词,查找过程简单高效。

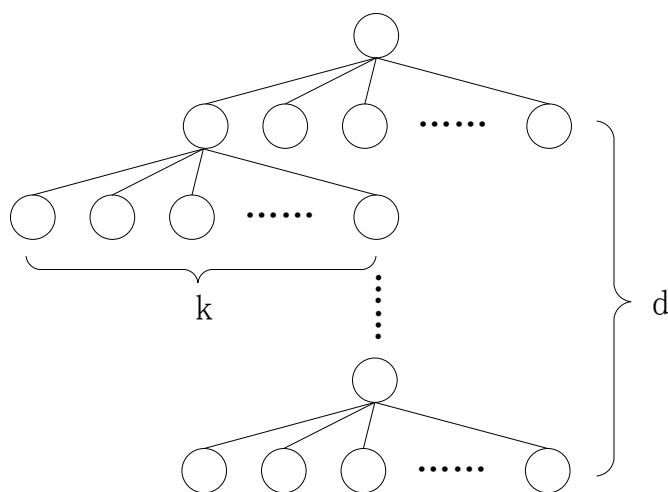


图 3-2 字典结构示意图

这样生成的字典,每个单词的权重相同,缺乏区分度。利用每个单词出现

的频率形成图片的描述向量同样也缺少区分度。为了使图像的描述效果更好，可以采用 TF-IDF 法，计算单词的权重，形成图片的描述向量。TF-IDF 法是一种文档检索方法。TF 是译频率，在生成图片的描述向量时计算，代表单词在文档中的区分度；IDF 是逆文档频率，在构建字典时计算，代表单词在字典中的区分度。

要使用 TF-IDF 法生成图片的描述向量，在生成字典后需要对字典的每一个单词计算逆文档频率 IDF。假设构建字典时一共用了 n 个特征点，其中有 m 个特征点对应于单词 w ，则单词 w 的逆文档频率 IDF_w 为：

$$IDF_w = \log \frac{n}{m} \quad (3-1)$$

对以上过程进行总结，构建字典的流程如图 3-3 所示。同时，根据词袋模型的原理，应选择与待区分图像拍摄环境相似的图片作为生成字典所用的训练集，并且训练集所含图片越多，生成的字典包含的单词越丰富、单词逆文档频率的计算越准确，字典的使用效果越好。

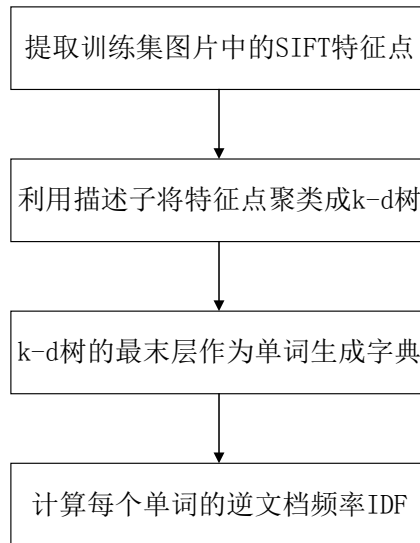


图 3-3 构建字典过程

3.3 基于字典的重叠区域搜索算法

3.3.1 图片描述向量的生成与比较

在完成字典的构建后，可以利用字典，将图片信息转化为数值向量。图片描述向量的生成过程如图 3-4 所示。

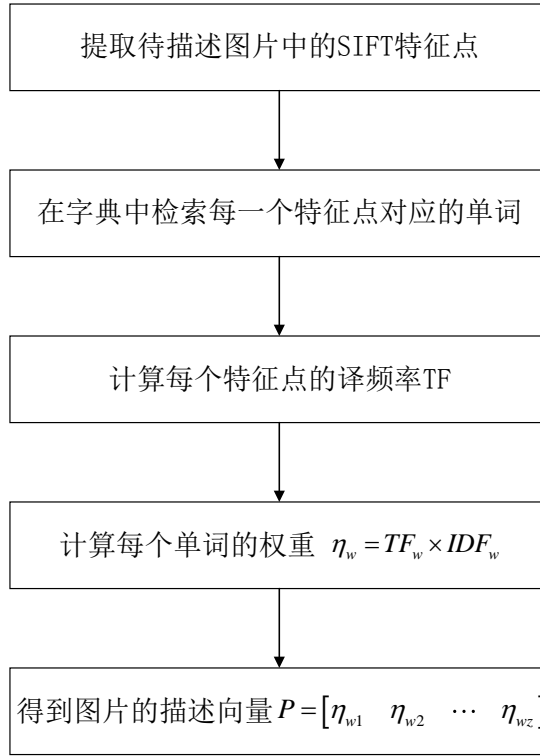


图 3-4 生成图片描述向量的过程

首先提取图像中的 SIFT 特征点，然后根据字典对图像中的每一个特征点进行检索，寻找特征点在字典中的对应单词。根据上节的叙述，为了提高单词的区分度，使图片描述效果更好，采用了 TF-IDF 法计算单词的权重，形成图片的描述向量。

假设图片 P 根据字典一共可以检索到 z 个单词，其中，单词 w 出现了 y 次，则单词 w 的译频率 TF_w 为：

$$TF_w = \frac{y}{z} \quad (3-2)$$

单词 w 的逆文档频率 IDF 在构建字典时根据式(3-1)计算得到，则单词 w 在图片的描述向量中总的权重 η_w 为：

$$\eta_w = TF_w \times IDF_w \quad (3-3)$$

按照这种方式，分别计算图片 P 的特征点检索到的每一个单词的权重 η_{w1} 、 η_{w2} 、 η_{w3} ……并最终形成图片的描述向量 V_P ：

$$V_P = [\eta_{w1} \ \eta_{w2} \ \cdots \ \eta_{wz}] \quad (3-4)$$

在生成图片的描述向量后，对两幅图片 P_A 、 P_B 相似性的比较就可以转化为对两幅图片的描述向量 V_A 和 V_B 进行比较，计算其相似性评分 $S(V_A - V_B)$ ：

$$S(V_A - V_B) = 1 - \frac{1}{2} \left| \frac{V_A}{|V_A|} - \frac{V_B}{|V_B|} \right| \quad (3-5)$$

3.3.2 重叠区域查找流程

寻找两个漫游器行进路径的重叠区域，实际上就是寻找两个漫游器所拍摄的图片中最相似的部分。利用前面介绍的基于词袋模型的方式寻找相似图片，首先需要构建字典。

根据前面章节的叙述，选择与待处理图像拍摄环境相似的图片作为生成字典所用的训练集，并且训练集所含图片越多，生成的字典包含的单词越丰富、单词逆文档频率的计算越准确，字典的使用效果越好。但是训练集数量越大，字典的构建越耗费时间。因此可以在两漫游器所拍摄图片中随机选取一部分作为训练集，按照 3.2.2 节介绍的方法构建字典。

利用多帧图像进行特征点地图的拼接，需要找到连续多帧相似的图像。如图 3-5 所示，首先利用字典对两个漫游器左眼相机拍摄的每一幅图片分别生成描述向量，然后将两组描述向量一一对比，计算其相似性评分。由于场景的相似性，可能存在众多的相似图片，可以利用阈值法对较低的相似性评分进行制。

由于字典的限制，相似性很高的图片也可能存在相似性评分的绝对值并不高的情况，但是与其它图像相比，其相似性评分的相对值仍将较高。因此不能简单使用一个固定的阈值对较低的相似性评分进行抑制，而应当使用一种相对阈值。

首先寻找到最高的相似性评分，然后将所有低于此数值一定百分比的图片间相似性评分都置为 0，这样就排除了很多相似性不那么高的图像。对于剩下的图像，寻找在时间上连续的多对图像，作为重叠区域的图片，用于后续地图拼接的计算。

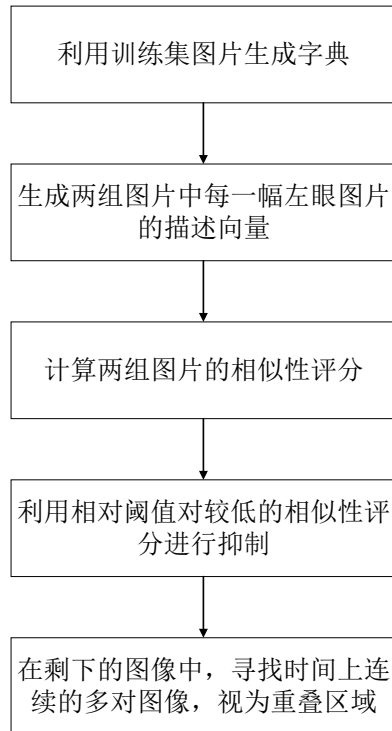


图 3-5 寻找连续多帧相似图像的过程

3.4 仿真实验与分析

3.4.1 基于漫游器实际拍摄图像的相似图片搜索实验

利用“勇气”号火星漫游器导航相机拍摄的一组图片进行重叠区域查找的实验。

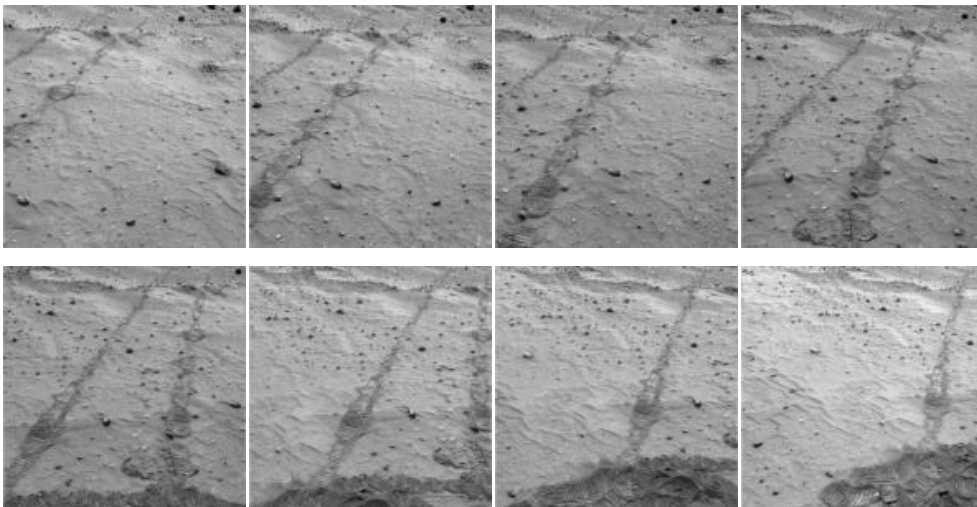


图 3-6 “勇气”号火星漫游器左眼导航相机所摄图片

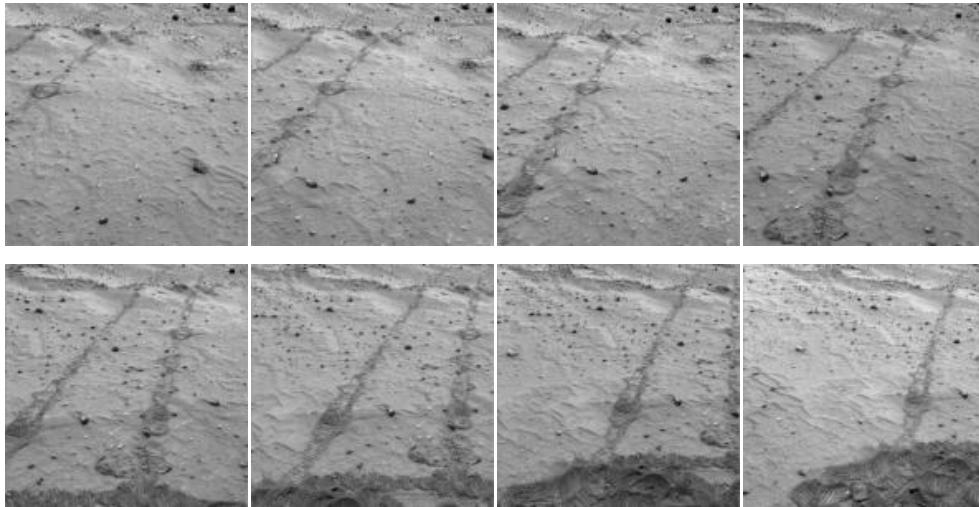


图 3-7 “勇气”号火星漫游器右眼导航相机所摄图片

首先利用漫游器左眼导航相机拍摄的图片作为训练集，按照 3.2 节介绍的方法构建字典。字典 k - d 树结构的参数为： $k=10$ ， $d=5$ 。得到一个包含 3191 个单词的字典。可以看出，由于图片数量较少，没有足够多的特征点， k - d 树并没有被填满。但是仍然可以利用这个字典对图片的相似性进行计算，查找最为相似的一对图像。

利用生成的字典，按照 3.3.1 节介绍的方法生成漫游器左、右眼导航相机拍摄的每张图片的描述向量，然后计算左、右眼图片的描述向量的相似性评分，如表 3-1 所示。

表 3-1 火星漫游器左、右眼图片相似性评分

	左眼 1	左眼 2	左眼 3	左眼 4	左眼 5	左眼 6	左眼 7	左眼 8
右眼 1	0.3504	0.1973	0.1000	0.0716	0.0776	0.0636	0.0430	0.0752
右眼 2	0.0956	0.2980	0.2449	0.0890	0.0436	0.0537	0.0549	0.0686
右眼 3	0.0841	0.1095	0.3378	0.1860	0.1039	0.0676	0.0578	0.0547
右眼 4	0.0709	0.0950	0.1045	0.3412	0.2648	0.0971	0.0583	0.0528
右眼 5	0.0451	0.0513	0.0572	0.1071	0.3320	0.2200	0.0970	0.0894
右眼 6	0.0764	0.0686	0.0619	0.0813	0.0966	0.3330	0.2405	0.0790
右眼 7	0.0571	0.0771	0.0480	0.0638	0.0642	0.1304	0.3357	0.2669
右眼 8	0.0630	0.0354	0.0380	0.0452	0.0411	0.0529	0.0940	0.2763

将表中数据画为灰度图，数值越高像素越亮，如图 3-8 所示。

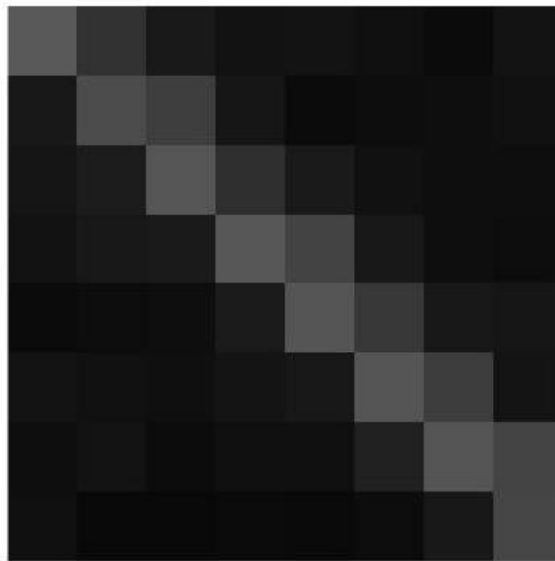


图 3-8 火星漫游器左、右眼图片相似性评分灰度图

可以看出，对角线上的元素最为明亮，即漫游器左、右眼对应的图片相似性评分最高，与实际情况相符。另外，右眼第 8 幅图片与左眼第 7、8 两幅图片的相似性评分都很高，这与观察实际图片得到的结果也是相符的。这个结果表明，利用本章介绍的方法，可以有效地在漫游器拍摄的图片中，寻找到最为相似的部分。

3.4.2 基于地面模拟场地图像的重叠区域搜索实验

由于缺少交叉路径上火星漫游器连续拍摄的双目图像数据，本节利用第二章实验中相机在模拟场地采集两组图像中左眼相机的图像进行重叠区域寻找的仿真实验，两组图像采集时相机路径有交叉部分，其中一组图像包含 180 幅图片，另一组包含 253 幅图片。实验所用电脑使用 i5-3210M 处理器，4G 内存，每张图片的大小为 752×480 ，

首先从所有图像中随机抽取 150 幅图片作为训练集样本，按照 3.2 节介绍的方法构建字典。字典 k-d 树结构的参数为：k=10，d=5。得到一个包含 24342 个单词的字典。

然后利用生成的字典，按照 3.3.1 节介绍的方法对两组图像中的每一幅图片分别生成描述向量，然后按照 3.3.2 节介绍的方法，将两组描述向量一一对比，计算其相似性评分，寻找连续多对相似性评分都大于一定阈值的图像，作为重叠区域的图片。将两组图像的相似性评分绘制为灰度图，数值越高像素越

亮，如图 3-9（a）所示。

图 3-9（a）中沿一条斜线分布的像素亮度都很高的区域即为重叠区域。有许多图像的相似性评分都很高，利用阈值法对其进行抑制。首先寻找到最高的相似性评分，然后将评分低于 0.6 倍此数值的相似性评分都置为 0。其结果如图 3-9（b）所示。

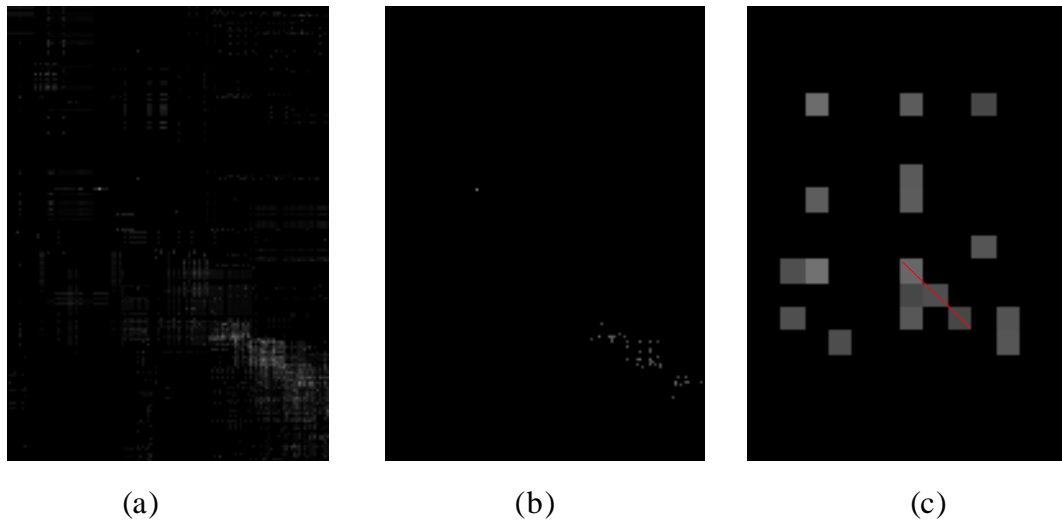


图 3-9 两组图像相似性评分灰度图

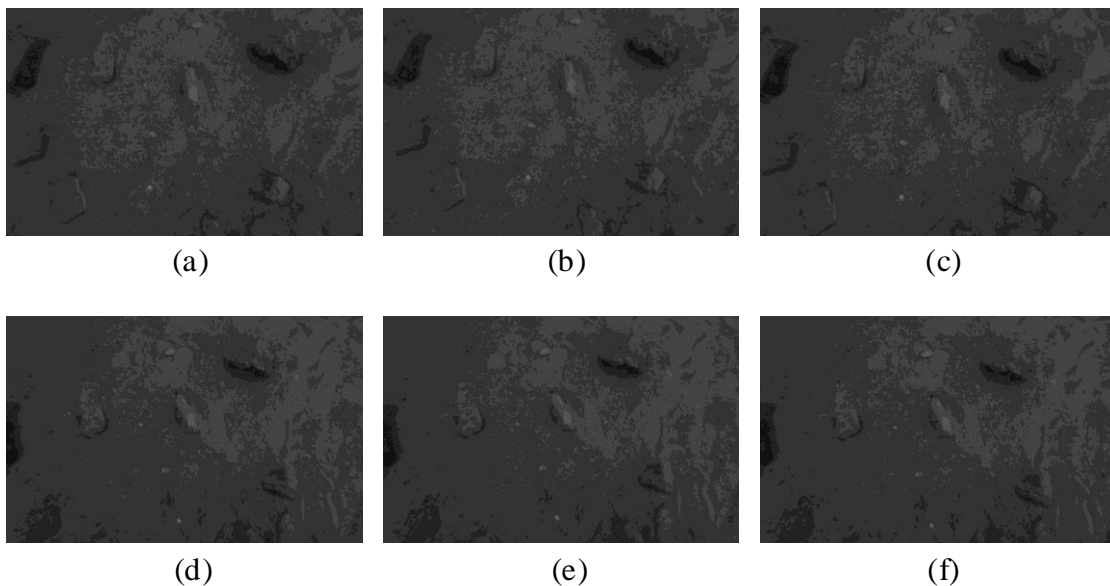


图 3-10 两组图像中的相似图片

将亮区局部放大，如下图所示。可以看到，在两组图像中有连续三对图片的相似性评分都满足阈值条件，如图 3-9（c）所示。这三对图片即可作为

漫游器行进轨迹的重叠区域。图 3-10 中 (a) (b) (c) 为第一组图像中的三幅图片, (d) (e) (f) 为第二组图像中的三幅图片。由图像可见, 地图重叠区域搜索成功。实验过程中, 构建字典共耗时 5min, 计算图片的相似性评分共耗时 3min。

由实验结果可见, 本章提出的重叠区域搜索方法准确有效, 且计算速度较快, 可以用于火星漫游器大范围探测时, 寻找不同漫游器行进路径中的重叠区域。

3.5 本章小结

本章研究利用词袋模型搜索不同漫游器行进路径上的重叠区域的算法。

首先介绍了词袋模型的概念, 然后研究了基于词袋模型利用训练集图片构建字典的方法。之后又研究了利用构建的字典生成图片描述向量、比较描述向量计算图片相似性评分、并最终寻找到重叠区域的方法。

本章利用“勇气”号火星漫游器拍摄的图像进行了构建字典、生成图片描述向量、计算图片相似性评分的实验, 证明了利用本章介绍的方法, 通过比较描述向量寻找火星表面相似图片的可行性。

利用双目相机在模拟场地采集的两组图像, 进行了重叠区域寻找的实验, 其结果证明了本章提出的基于词袋模型寻找火星漫游器行进路径上的重叠区域的方法准确高效。

第4章 稀疏特征点地图及漫游器位姿拼接算法

4.1 引言

通过两幅较小的稀疏特征点地图的拼接,可以构建大范围的稀疏特征点地图。而对漫游器位姿估计进行拼接,将其转换到相同的坐标系下,则是稠密点云地图构建的基础。

直接利用两个火星漫游器构建的稀疏特征点地图进行 ICP 配准拼接,由于特征点地图庞大,将导致计算量十分巨大。本章将稀疏特征点地图拼接问题转换为求取两漫游器在重叠区域的位姿变换关系,对此进行研究,并利用双目相机在模拟场地采集的图像以及计算机模拟图像对算法进行了仿真验证。

4.2 稀疏特征点地图及漫游器位姿拼接原理

在单漫游器视觉自主定位与稀疏特征点地图构建的过程中,得到的稀疏特征点地图建立在以漫游器第一帧左眼相机光心为原点、向右为 x 轴正方向、向下为 y 轴正方向的相机坐标系中。对两个漫游器构建的稀疏特征点地图进行拼接,实际上就是寻找两个漫游器拍摄第一帧图像时相机坐标系之间的转换关系,然后将其中一个相机坐标系中的稀疏特征点通过坐标变换转移到另一个相机坐标系中。

设漫游器 1 拍摄第一帧图像时的左眼相机坐标系记为 $O_1x_1y_1z_1$, 漫游器 2 拍摄第一帧图像时的左眼相机坐标系记为 $O_2x_2y_2z_2$ 。首先利用上一章介绍的词袋模型,寻找两漫游器拍摄的图片中连续多帧相似的图像,设漫游器 1 找到的是从第 J 帧开始的连续 n 帧图像,漫游器 2 找到的是从第 K 帧开始的连续 n 帧图像。

在单漫游器视觉自主定位过程中,已经得到了漫游器 1 拍摄第 J 帧图像时的左眼相机坐标系 $O_Jx_Jy_Jz_J$ 相对于 $O_1x_1y_1z_1$ 的坐标系转换关系 R_1 和 t_1 , 以及漫游器 2 拍摄第 K 帧图像时左眼相机坐标系 $O_Kx_Ky_Kz_K$ 相对于 $O_2x_2y_2z_2$ 的坐标系转换关系 R_2 和 t_2 。

利用漫游器 1 从第 J 帧开始的连续 n 帧图像和漫游器 2 从第 K 帧开始的连续 n 帧图像,构建三维点云,对三维点云的位置估计以及相机的位姿估计进行联合优化,得到漫游器 1 拍摄第 J 帧图像时的左眼相机坐标系 $O_Jx_Jy_Jz_J$ 相对于

漫游器 2 拍摄第 K 帧图像时的左眼相机坐标系 $O_K x_K y_K z_K$ 的坐标转换关系 R_{JK} 和 t_{JK} 。

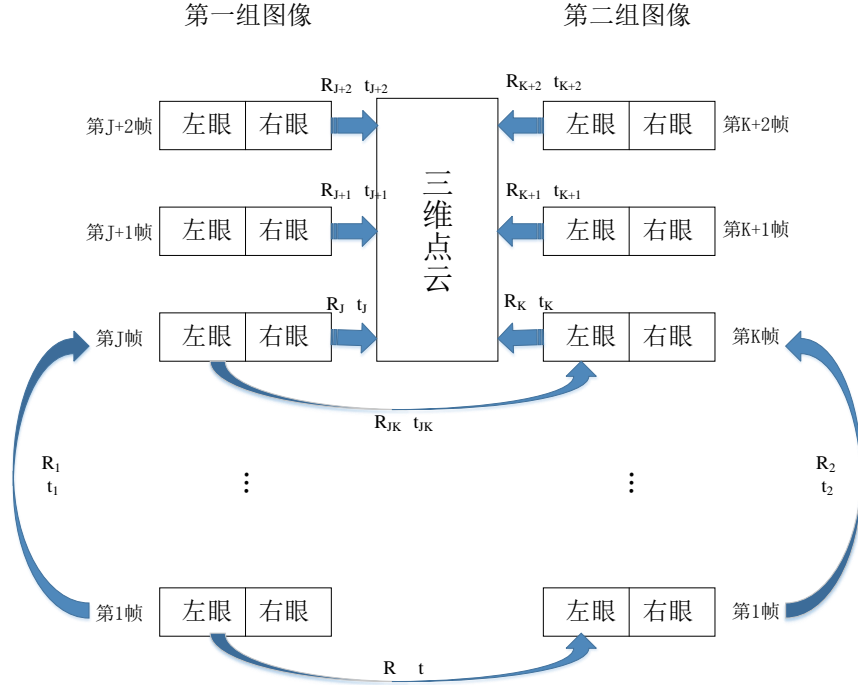


图 4-1 稀疏特征点地图拼接原理示意图

假设 $O_1 x_1 y_1 z_1$ 坐标系中的点 P_1 、 $O_2 x_2 y_2 z_2$ 坐标系中的点 P_2 、 $O_J x_J y_J z_J$ 坐标系中的点 P_J 、 $O_K x_K y_K z_K$ 坐标系中的点 P_K 在三维空间中是同一个点，根据上述坐标系转换关系有：

$$\begin{cases} P_1 = R_1 P_J + t_1 \\ P_2 = R_2 P_K + t_2 \\ P_J = R_{JK} P_K + t_{JK} \end{cases} \quad (4-1)$$

整理后可得：

$$P_1 = R_1 R_{JK} R_2^T P_2 - R_1 R_{JK} R_2^T t_2 + R_1 t_{JK} + t_1 \quad (4-2)$$

即：坐标系 $O_2 x_2 y_2 z_2$ 相对于 $O_1 x_1 y_1 z_1$ 的坐标变换 R 、 t 分别为：

$$R = R_1 R_{JK} R_2^T \quad (4-3)$$

$$t = -R_1 R_{JK} R_2^T t_2 + R_1 t_{JK} + t_1 \quad (4-4)$$

利用 R 、 t 可将两条轨迹中的稀疏特征点地图转换到同一个坐标系下，即完成了两漫游器构建的稀疏特征点地图的拼接。将漫游器位姿估计的结果转换

到同一个坐标系下的方法与之相同。

4.3 坐标转换关系初始估计算法

利用第 3 章介绍的方法,寻找两漫游器拍摄的图片中连续多帧相似的图像,设漫游器 1 找到的是从第 J 帧开始的连续 n 帧图像,漫游器 2 找到的是从第 K 帧开始的连续 n 帧图像。漫游器 1 拍摄第 J 帧图像时的左眼相机坐标系为 $O_J x_J y_J z_J$, 漫游器 2 拍摄第 K 帧图像时的左眼相机坐标系为 $O_K x_K y_K z_K$ 。以 $O_J x_J y_J z_J$ 为基准,将其他图片提取到的特征点添加到 $O_J x_J y_J z_J$ 中,同时估计其他图片拍摄时左眼相机的位姿,进而得到坐标系 $O_J x_J y_J z_J$ 相对于坐标系 $O_K x_K y_K z_K$ 的转换关系 R_{JK} 和 t_{JK} 的初始估计值。其过程如图 4-2 所示。

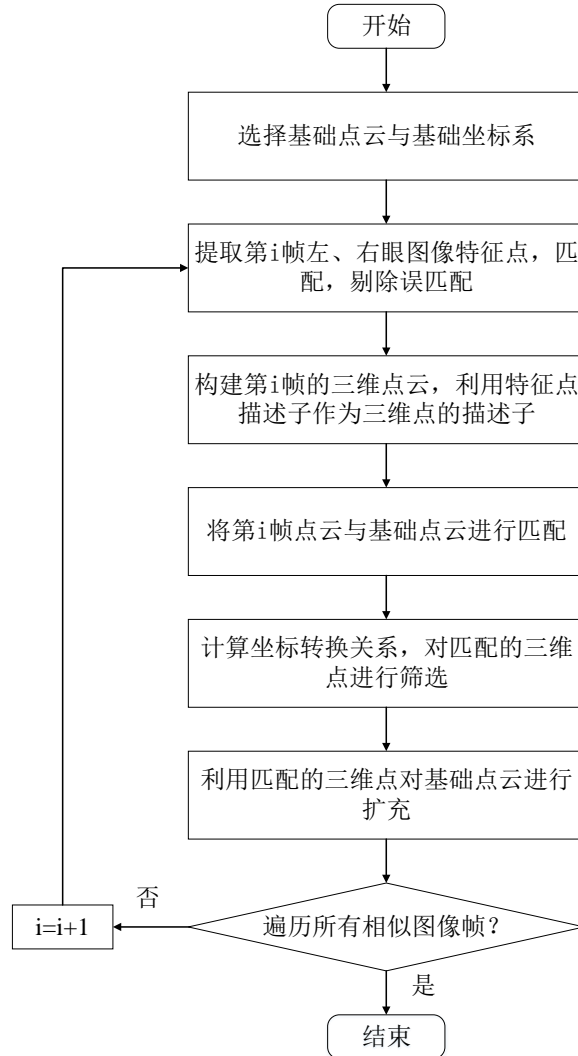


图 4-2 三维点云构建与坐标转换关系估计流程图

首先对每一帧的左、右眼图像提取特征点，并进行匹配。选择 SIFT 特征点提取方法以使特征点的检测更丰富、准确。利用暴力匹配的方法，对左、右眼图像中提取的特征点进行匹配，并利用对极约束剔除误匹配的点对。

然后利用双目相机模型，根据每一帧左、右眼图像中匹配的特征点，在每一帧的左眼相机坐标系下构建其观测到的三维点云，并利用左、右眼图像中特征点的描述子共同作为三维点的描述子。例如，左眼图像中描述子为 s_1 的点与右眼图像中描述子为 s_2 的点互为匹配点，利用双目相机模型根据这两个点构建的三维点描述子为 $\{s_1, s_2\}$ 。

以漫游器 1 第 J 帧图像构建的三维点云为基础点云，第 J 帧左眼相机坐标系为基础坐标系。依次用其他各帧图像构建的三维点云与基础点云进行匹配，计算每一对描述子的汉明距，将汉明距小于一定阈值的三维点视为候选的匹配点对，利用奇异值分解法，根据候选匹配点对计算各帧左眼相机坐标系相对于基础坐标系的坐标转换关系 \mathbf{R} 、 \mathbf{t} ，并利用 \mathbf{R} 、 \mathbf{t} 对候选匹配点对进行筛选，剔除误差大于给定阈值的匹配点对。

利用保留下的匹配点对，对基础点云的描述子进行扩展，例如基础点云中的三维点 \mathbf{p} 的描述子为 $\{s_1, s_2\}$ ，与其匹配的三维点 \mathbf{p}' 的描述子为 $\{s_3, s_4\}$ ，则扩展之后基础点云中的三维点 \mathbf{p} 的描述子为 $\{s_1, s_2, s_3, s_4\}$ 。对于未在基础点云中找到匹配点的三维点，保持其描述子不变，利用 \mathbf{R} 、 \mathbf{t} 将其坐标转换到基础坐标系下，添加进基础点云。

重复这一过程直到所有相似图像帧都处理完毕，保存基础三维点云、每一帧左眼相机坐标系相对于基础坐标系的转换关系以及每一帧对三维点的观测，在下一节进行优化处理。

4.4 三维点云与相机位姿联合优化方法

由于三维点的匹配是通过比较二维图像描述子实现的，会存在一些误匹配的情况，因此通过 4.3 节的方法得到的基础坐标系下的三维点云坐标以及每帧左眼相机坐标系相对基础坐标系的转换关系都存在误差，可以将其作为初始的估计值，进行联合优化。

假设基础坐标系中有点 \mathbf{P} ，其坐标为 $[X \ Y \ Z]^T$ 。某一帧的相机拍摄到了点 \mathbf{P} ，点 \mathbf{P} 在左眼相机拍摄的图片的像素坐标系下坐标为 $[u_p \ v_p]^T$ ，这一帧的左眼相机坐标系相对于基础坐标系的坐标转换关系是 \mathbf{R} 、 \mathbf{t} 。

根据 2.2 节介绍的相机模型, 可以将点 P 投影到左眼图片的像素坐标系中。
首先将点 P 在基础坐标系下的坐标 p 转换到左眼相机坐标系中, 记为 p' :

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ O & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4-5)$$

然后将 p' 投影到像素坐标系下:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{Z'} \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} \quad (4-6)$$

对点 P 的实际观测与投影之间存在误差 e :

$$e = \begin{bmatrix} u_p \\ v_p \end{bmatrix} - \begin{bmatrix} u \\ v \end{bmatrix} \quad (4-7)$$

考虑每一帧相机观测到的每一个三维点与投影之间的误差, 构建对相机位姿和三维点位置联合优化的代价函数:

$$f(x) = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n \|e_{ij}\|^2 \quad (4-8)$$

其中, m ——相机的数量;

n ——三维点的数量;

e_{ij} ——第 i 帧相机观测第 j 个三维点的误差;

x ——是待优化的变量。

利用李代数的形式表示相机位姿, 则式(4-5)可以写为:

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \exp(\xi^\wedge) \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4-9)$$

待优化的变量为:

$$x = [\xi_1 \quad \dots \quad \xi_m \quad p_1 \quad \dots \quad p_n]^T \quad (4-10)$$

其中, ξ_i ——李代数形式的第 i 帧相机位姿, 即第 i 帧左眼相机的相机坐标系相对于基础坐标系的转换关系;

p_j ——第 j 个三维点的位置, 即第 j 个三维点在基础坐标系下的坐标。

采用高斯牛顿法进行上述非线性优化问题的求解。首先将误差 e 进行一阶泰勒展开，则代价函数式(4-8)可以写为：

$$f(x) = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n \|e_{ij} + J_{ij} \cdot x\|^2 \quad (4-11)$$

式中， J_{ij} ——误差函数 e_{ij} 对待优化变量 x 的偏导数。

将式(4-11)展开可得：

$$\begin{aligned} f(x) &= \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n (e_{ij} + J_{ij} \Delta x)^T (e_{ij} + J_{ij} \Delta x) \\ &= \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n \left(\|e_{ij}\|_2^2 + 2e_{ij}^T J_{ij} \Delta x + \Delta x^T J_{ij}^T J_{ij} \Delta x \right) \end{aligned} \quad (4-12)$$

求式(4-12)关于 Δx 的偏导数，可得：

$$\frac{\partial f(x)}{\partial \Delta x} = \sum_{i=1}^m \sum_{j=1}^n (J_{ij}^T J_{ij} \Delta x + J_{ij}^T e_{ij}) \quad (4-13)$$

令式(4-13)右侧等于 0，有：

$$\sum_{i=1}^m \sum_{j=1}^n J_{ij}^T J_{ij} \Delta x = - \sum_{i=1}^m \sum_{j=1}^n J_{ij}^T e_{ij} \quad (4-14)$$

下面研究误差函数 e_{ij} 对待优化变量 x 的偏导数 J_{ij} 的计算方式。

根据 J_{ij} 的定义，有：

$$J_{ij} = \frac{\partial e_{ij}}{\partial x} = \begin{bmatrix} \frac{\partial e_{ij}}{\partial \xi_1} & \frac{\partial e_{ij}}{\partial \xi_2} & \dots & \frac{\partial e_{ij}}{\partial \xi_m} & \frac{\partial e_{ij}}{\partial p_1} & \frac{\partial e_{ij}}{\partial p_2} & \dots & \frac{\partial e_{ij}}{\partial p_n} \end{bmatrix} \quad (4-15)$$

由于误差函数 e_{ij} 只与第 i 帧相机和第 j 个三维点有关，因此式(4-15)实际为：

$$J_{ij} = \frac{\partial e_{ij}}{\partial x} = \begin{bmatrix} 0 & \dots & 0 & \frac{\partial e_{ij}}{\partial \xi_i} & 0 & \dots & 0 & \frac{\partial e_{ij}}{\partial p_j} & 0 & \dots & 0 \end{bmatrix} \quad (4-16)$$

式(4-16)中， $\frac{\partial e_{ij}}{\partial \xi_i}$ 是误差函数对相机位姿的偏导数，可以利用链导法则计算：

$$\frac{\partial e_{ij}}{\partial \xi_i} = \frac{\partial e_{ij}}{\partial p'} \frac{\partial p'}{\partial \xi_i} \quad (4-17)$$

其中， p' ——点 P 在第 i 帧左眼相机坐标系中的坐标。

根据式(4-6)和式(4-7)有：

$$\begin{cases} u = \frac{X'}{Z'} f_x + c_x \\ v = \frac{Y'}{Z'} f_y + c_y \end{cases} \quad (4-18)$$

$$e = \begin{bmatrix} u_p \\ v_p \end{bmatrix} - \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u_p - \frac{X'}{Z'} f_x + c_x \\ v_p - \frac{Y'}{Z'} f_y + c_y \end{bmatrix} \quad (4-19)$$

由式(4-19)可得式(4-17)的第一项为：

$$\frac{\partial e_{ij}}{\partial p'} = \begin{bmatrix} -\frac{f_x}{Z'} & 0 & \frac{f_x X'}{Z'^2} \\ 0 & -\frac{f_y}{Z'} & \frac{f_y Y'}{Z'^2} \end{bmatrix} \quad (4-20)$$

式(4-17)的第二项可以写为：

$$\frac{\partial p'}{\partial \xi_i} = \lim_{\delta \xi_i \rightarrow 0} \frac{\exp(\delta \xi_i^{\wedge}) \exp(\xi_i^{\wedge}) p - \exp(\xi_i^{\wedge}) p}{\delta \xi_i^{\wedge}} \quad (4-21)$$

对式(4-21)等号右侧的分子中 $\exp(\delta \xi_i^{\wedge})$ 一阶泰勒展开，有：

$$\exp(\delta \xi_i^{\wedge}) = I + \delta \xi_i \quad (4-22)$$

将式(4-22)代入到式(4-21)中，有：

$$\begin{aligned} \frac{\partial p'}{\partial \xi_i} &= \lim_{\delta \xi_i \rightarrow 0} \frac{\exp(\delta \xi_i^{\wedge}) \exp(\xi_i^{\wedge}) p - \exp(\xi_i^{\wedge}) p}{\delta \xi_i} \\ &= \lim_{\delta \xi_i \rightarrow 0} \frac{(I + \delta \xi_i^{\wedge}) \exp(\xi_i^{\wedge}) p - \exp(\xi_i^{\wedge}) p}{\delta \xi_i} \\ &= \lim_{\delta \xi_i \rightarrow 0} \frac{\delta \xi_i^{\wedge} \exp(\xi_i^{\wedge}) p}{\delta \xi_i} \\ &= \lim_{\delta \xi_i \rightarrow 0} \frac{\begin{bmatrix} \delta \phi_i^{\wedge} & \delta \rho_i \\ O & 0 \end{bmatrix} \begin{bmatrix} R_i p + t_i \\ 1 \end{bmatrix}}{\delta \xi_i} = \lim_{\delta \xi_i \rightarrow 0} \frac{\begin{bmatrix} \delta \phi_i^{\wedge} (R_i p + t_i) + \delta \rho_i \\ 0 \end{bmatrix}}{\begin{bmatrix} \delta \rho_i \\ \delta \phi_i \end{bmatrix}} \\ &= \begin{bmatrix} I & -(R_i p + t_i)^{\wedge} \\ O & O \end{bmatrix} \end{aligned} \quad (4-23)$$

将式(4-20)、式(4-23)代入式(4-17)中可得：

$$\frac{\partial e_{ij}}{\partial \xi_i} = \begin{bmatrix} -\frac{f_x}{Z'} & 0 & \frac{f_x X'}{Z'^2} \\ 0 & -\frac{f_y}{Z'} & \frac{f_y Y'}{Z'^2} \end{bmatrix} \begin{bmatrix} I & -(R_i p + t_i)^\wedge \\ O & O \end{bmatrix} \quad (4-24)$$

对于 $\frac{\partial e_{ij}}{\partial p_j}$ ，也可利用链导法则进行计算。

$$\frac{\partial e_{ij}}{\partial p_j} = \frac{\partial e_{ij}}{\partial p'} \frac{\partial p'}{\partial p_j} \quad (4-25)$$

其中， p' ——点 P 在第 i 帧左眼相机坐标系中的坐标。

$\frac{\partial e_{ij}}{\partial p_j}$ 的计算与前面式(4-20)相同。

根据式(4-5)可得：

$$\frac{\partial p'}{\partial p_j} = R_i \quad (4-26)$$

因此有：

$$\frac{\partial e_{ij}}{\partial p_j} = \begin{bmatrix} -\frac{f_x}{Z'} & 0 & \frac{f_x X'}{Z'^2} \\ 0 & -\frac{f_y}{Z'} & \frac{f_y Y'}{Z'^2} \end{bmatrix} R_i \quad (4-27)$$

利用上一节计算得到的三维点云以及相机位姿作为待优化变量 x 的初始值，每次迭代时，首先利用 x 计算得到每一项的 J_{ij} 和 e_{ij} ，然后求解方程组(4-14)得到待优化变量 x 的改变量 Δx ，对 x 进行调整，继续进行下一次迭代，直至代价函数小于一定阈值，或者 Δx 小于一定阈值为止，即完成了对相机位姿估计和三维点位置估计的联合优化。

4.5 仿真实验与分析

4.5.1 基于地面模拟场地图像的稀疏特征点地图及位姿拼接实验

本节利用第二章实验中相机在模拟场地采集两组图像进行稀疏特征点地图拼接的仿真实验。两组图像采集时相机路径有交叉部分，其中一组图像包含 180 幅图片，另一组包含 253 幅图片。

在第二章漫游器的自主定位与构建稀疏特征点地图的实验过程中，已经得

到了两组图像各自构建的稀疏特征点地图以及每一帧左眼相机坐标系相对于第一帧左眼相机坐标系的坐标转换关系。漫游器自主定位与稀疏特征点地图构建的结果如图 4-3 所示。

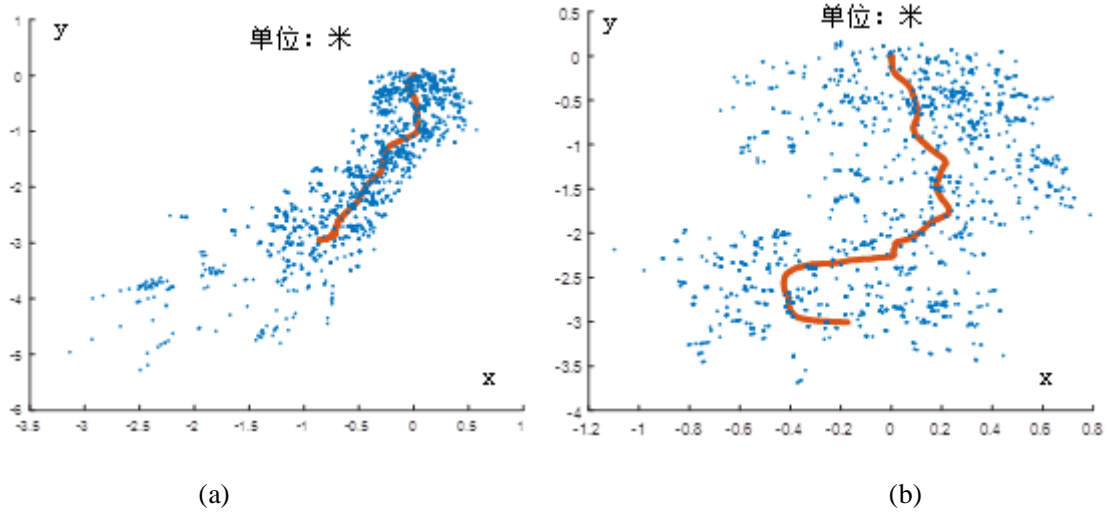


图 4-3 两组图像构建的稀疏特征点地图与轨迹

在第三章地图重叠区域搜索的实验中,已经找到了两组图片中连续的相似帧,如图 4-4 所示。

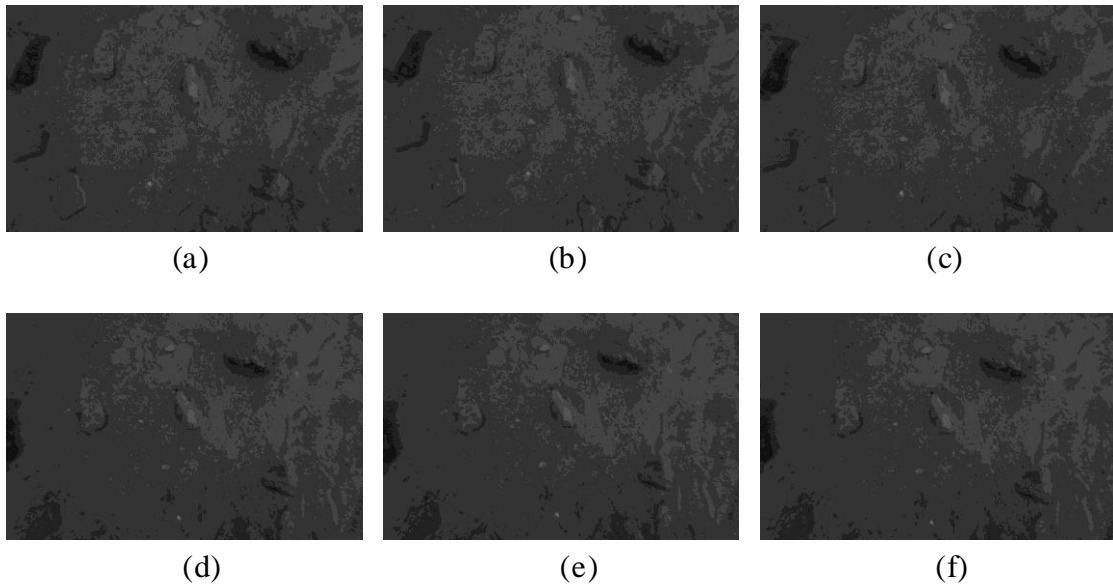


图 4-4 两组图像中的相似图片

利用重叠区域的左、右眼图片,根据本章介绍的方法计算两组图像构建的稀疏特征点地图的坐标转换关系,其结果如下表所示。

表 4-1 两稀疏特征点地图的坐标转换关系

参数名	参数值
相似图片首帧左眼相机坐标系间旋转矩阵 R	$\begin{bmatrix} 0.9901 & -0.0539 & 0.1299 \\ 0.0524 & 0.9985 & 0.0153 \\ -0.1305 & -0.0084 & 0.9914 \end{bmatrix}$
相似图片首帧左眼相机坐标系间平移向量 t	$[-0.0437 \quad -0.1333 \quad -0.0165]^T$
两稀疏特征点地图间旋转矩阵 R	$\begin{bmatrix} 0.9901 & 0.1323 & 0.0484 \\ -0.1276 & 0.9877 & -0.0897 \\ -0.0597 & 0.0825 & 0.9948 \end{bmatrix}$
两稀疏特征点地图间平移向量 t	$[-0.4140 \quad -0.2823 \quad -0.1471]^T$

将第二组图像得到的漫游器自主定位与稀疏特征点地图构建的结果转换到第一组图像所得的结果中，如下图所示。图中，红色的轨迹是由第一组图像得到的自主定位结果，绿色的点云是第一组图像得到的稀疏特征点地图；蓝色的轨迹是由第二组图像得到的自主定位结果，浅蓝色的点云是第二组图像得到的稀疏特征点地图。

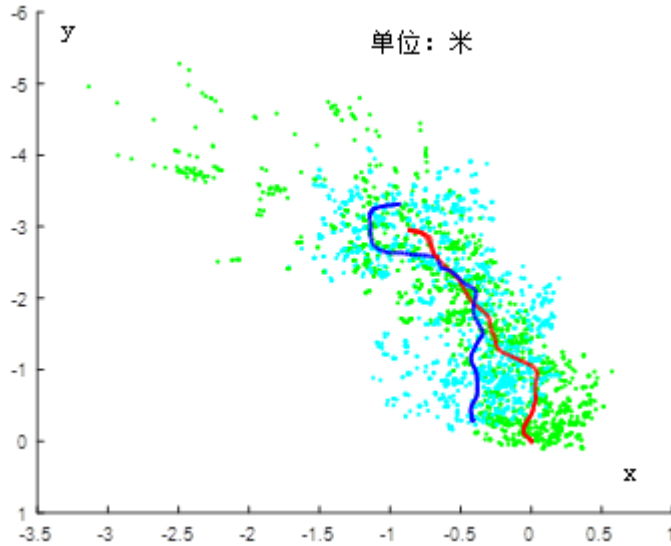


图 4-5 漫游器轨迹与稀疏特征点地图拼接结果

由于稀疏特征点在三维空间中的实际位置难以确定，而将漫游器位姿估计的结果转换到同一个坐标系下的过程与稀疏特征点地图的拼接过程相同，因此可用漫游器位置的误差来估计特征点地图拼接的误差。利用 lighthouse 室内定位技术获取两条轨迹初始位置之间的距离为 0.5395m，拼接后的估计值为

0.5322m，误差为 0.0073m，是实际距离的 1.35%。本章提出的稀疏特征点地图及漫游器位姿拼接算法具有较高的精度。

4.5.2 基于计算机模拟图像的稀疏特征点地图及位姿拼接实验

由于地面模拟场地规模较小，漫游器双目相机采用斜向下的角度进行拍摄。为了验证当双目相机平视行进方向时仍能利用本章介绍的方法进行稀疏特征点地图及位姿的拼接，本节利用第二章实验中渲染得到的火星表面的计算机模拟图像进行仿真实验。两组图像各包含 1000 帧，均匀采集自 54.0000m 和 123.1942m 长的两条路径中。

在第二章漫游器的自主定位与构建稀疏特征点地图的实验过程中，已经得到了两组图像各自构建的稀疏特征点地图以及每一帧左眼相机坐标系相对于第一帧左眼相机坐标系的坐标转换关系。漫游器自主定位与稀疏特征点地图构建的结果如图 4-6 所示。

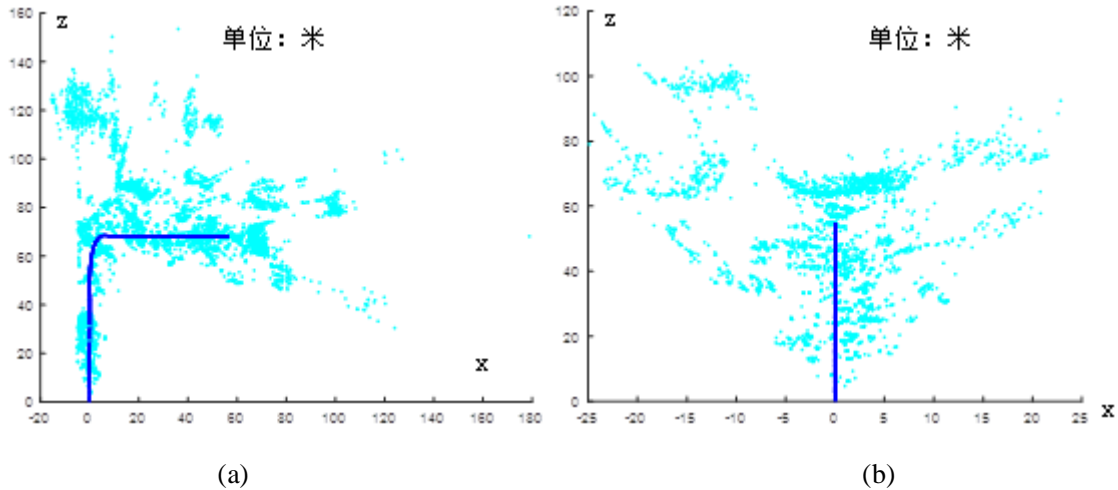


图 4-6 由模拟图像构建的稀疏特征点地图与轨迹

在模拟图像中随机选取 500 张图片作为训练集，利用第三章介绍的方法构建字典，然后寻找两组图像的左眼图片中连续多帧的相似图像，将这些帧视为地图重叠区域。相似图片的寻找结果如图 4-7 所示。

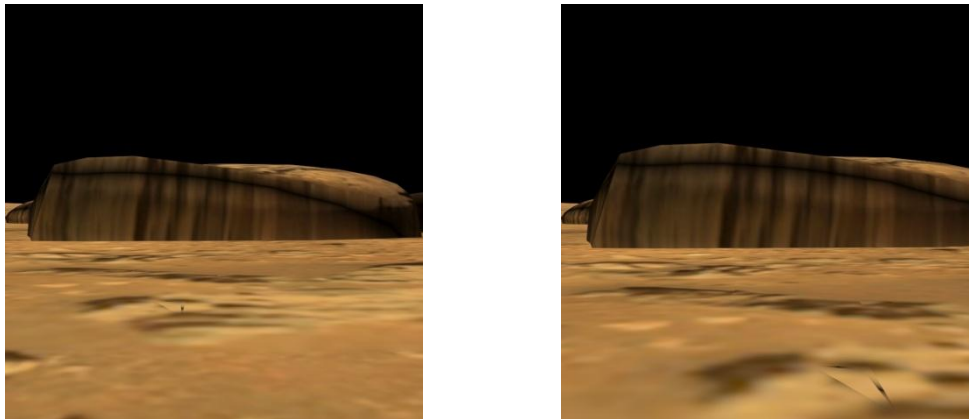


图 4-7 两组图像中的一对相似图片

利用相似图片计算两条路径基准坐标系之间的转换关系,对漫游器自主定位与稀疏特征点地图构建的结果进行拼接,如表 4-2 所示。

将第二组图像得到的漫游器自主定位与稀疏特征点地图构建的结果转换到第一组图像所得的结果中,如图 4-8 所示。图中,红色的轨迹是由第一组图像得到的自主定位结果,绿色的点云是第一组图像得到的稀疏特征点地图;蓝色的轨迹是由第二组图像得到的自主定位结果,浅蓝色的点云是第二组图像得到的稀疏特征点地图。

两条轨迹初始位置距离实际值为 69.3295m,拼接后的估计值为 69.1390m,误差为 0.1905m,是实际距离的 0.27%。

表 4-2 两稀疏特征点地图的坐标转换关系

参数名	参数值
相似图片首帧左眼相机坐标系间旋转矩阵 R	$\begin{bmatrix} 0.9999 & 0.0001 & 0.0014 \\ -0.0001 & 0.9999 & 0.0011 \\ -0.0014 & -0.0011 & 0.9999 \end{bmatrix}$
相似图片首帧左眼相机坐标系间平移向量 t	$[0.9810 \quad -0.0091 \quad -0.1424]^T$
两稀疏特征点地图间旋转矩阵 R	$\begin{bmatrix} -0.0013 & 0.0009 & -1.0000 \\ 0.0021 & 1.0000 & 0.0009 \\ 1.0000 & -0.0021 & -0.0013 \end{bmatrix}$
两稀疏特征点地图间平移向量 t	$[69.1390 \quad -0.1340 \quad -1.6748]^T$

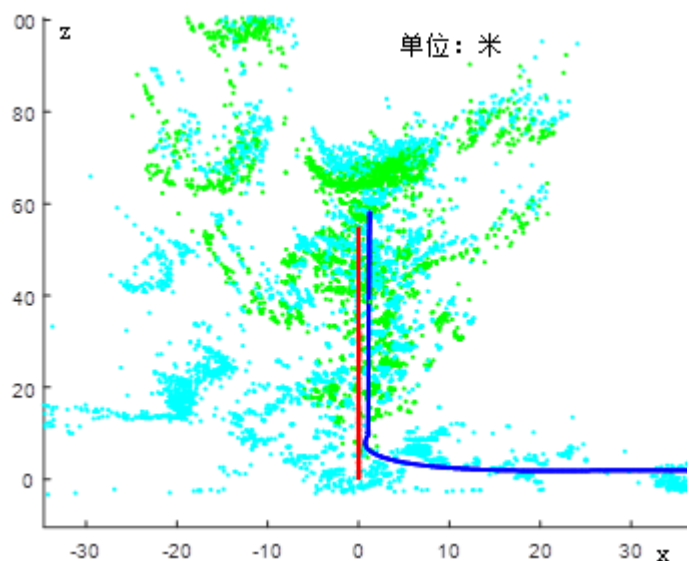


图 4-8 漫游器轨迹与稀疏特征点地图拼接结果

根据实验结果可以看到,本章设计的算法可以准确有效地利用漫游器在行进路线上重叠区域拍摄的图片实现稀疏特征点地图及漫游器位姿的拼接,具有较高的精度,可以满足探测任务的要求。

4.6 本章小结

本章将稀疏特征点地图及漫游器位姿拼接问题转换为求取两漫游器在重叠区域的位姿变换关系,首先介绍了稀疏特征点地图及漫游器位姿的拼接原理,然后研究了坐标转换关系的初始估计算法,利用此方法,根据重叠区域的图片构建三维点云并得到每一帧相机的位姿,进而得到两漫游器基准坐标系之间的转换关系的初始估计值。最后研究了利用非线性优化的方法对三维点云位置的估计与相机位姿的估计进行联合优化,进而得到两漫游器基准坐标系之间转换关系的优化值。

本章利用地面模拟场地采集的图像以及计算机模拟图像对算法进行了仿真实验,证明了本章提出的稀疏特征点地图拼接方法准确有效。

第5章 稠密点云地图与八叉树地图的构建算法

5.1 引言

大范围稀疏特征点地图已经可以满足火星漫游器的自主定位需求，但要想为后续经过相同区域的漫游器提供路径规划的指导信息，则需要使用稠密地图。稠密点云地图所需存储空间过大，不适合用于漫游器自主路径规划，而利用八叉树分割空间的三维占有率地图则可在占据较小的存储空间的前提下，表示大范围地形的空间占据情况。本章研究基于双目视觉构建火星表面稠密点云地图，以及利用稠密点云地图构建八叉树地图的方法，并利用双目相机在模拟场地采集的图像以及计算机模拟图像对算法进行仿真验证。

5.2 稠密点云地图构建算法

5.2.1 基于极线搜索与块匹配的双目图像匹配

为了获取稠密的点云，首先需要对左右眼图像进行匹配，获取尽可能多的像素点的深度。

假设空间中存在一个三维点 P ，其在左眼图像中投影的像素点为 $p_1 = [u_1 \ v_1 \ 1]^T$ ，在右眼图像中投影的像素点为 $p_2 = [u_2 \ v_2 \ 1]^T$ 。相机投影矩阵为 K ，则有：

$$Z_1 p_1 = K P_1 \quad (5-1)$$

$$Z_2 p_2 = K P_2 \quad (5-2)$$

式中， $P_1 = [X_1 \ Y_1 \ Z_1]^T$ ——像素点对应的三维点在左眼相机坐标系下的坐标；

$P_2 = [X_2 \ Y_2 \ Z_2]^T$ ——像素点对应的三维点在右眼相机坐标系下的坐标。

漫游器左、右眼相机之间的相对位姿是确定且已知的。记右眼相机的相机坐标系相对于左眼相机的相机坐标系的坐标旋转矩阵为 R ，平移向量为 t ，则有：

$$R P_1 + t = P_2 \quad (5-3)$$

将式(5-3)代入到式(5-2)中有：

$$Z_2 p_2 = K (R P_1 + t) = K R P_1 + K t \quad (5-4)$$

再将式(5-1)代入到式(5-4)中，可得：

$$Z_2 p_2 = K R K^{-1} Z_1 p_1 + K t \quad (5-5)$$

因为双目相机的成像平面在同一平面上，因此同一个三维点在两个相机坐标系下的深度相同，即有：

$$Z_1 = Z_2 = Z \quad (5-6)$$

将式(5-6)代入到式(5-5)中，有：

$$p_2 = K R K^{-1} p_1 + \frac{1}{Z} K t \quad (5-7)$$

在理想情况下，左、右眼相机的只是沿相机坐标系的 x 轴方向进行平移，因此有：

$$K R K^{-1} p_1 = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} \quad (5-8)$$

$$K t = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t_x \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} f_x t_x \\ 0 \\ 0 \end{bmatrix} \quad (5-9)$$

$$\begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} + \frac{1}{Z} \begin{bmatrix} f_x t_x \\ 0 \\ 0 \end{bmatrix} \quad (5-10)$$

即：

$$\begin{cases} u_2 = u_1 + \frac{1}{Z} f_x t_x \\ v_2 = v_1 \end{cases} \quad (5-11)$$

由式(5-11)可见，理想情况下，左、右眼图像中对应的像素点应该位于图片的同一行中。实际中，左、右眼相机的安装可能存在一定的误差，但是由于误差较小，仍然可以将旋转矩阵 R 视为单位阵，将平移向量 t 记为 $\begin{bmatrix} t_x & t_y & t_z \end{bmatrix}^T$ 。同时，由于安装误差相对于相机与拍摄物体之间的距离相比十分小，因此仍然可以近似认为式(5-6)成立。对于式(5-5)有：

$$\begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} + \frac{1}{Z} \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (5-12)$$

即：

$$\begin{cases} u_2 = u_1 + \frac{1}{Z}(f_x t_x + c_x t_z) \\ v_2 = v_1 + \frac{1}{Z}(f_y t_y + c_y t_z) \end{cases} \quad (5-13)$$

根据式(5-13)可得：

$$v_2 = \frac{f_y t_y + c_y t_z}{f_x t_x + c_x t_z} u_2 + v_1 - \frac{(f_y t_y + c_y t_z) u_1}{f_x t_x + c_x t_z} \quad (5-14)$$

由上式可以看出，对于左眼图像中的某个像素点，其在右眼图像中对应的像素点应该位于某条已知的直线上，具体的位置与三维点的深度有关。这条直线就称为极线。因此在寻找左、右眼图像中对应的像素点时，只需要沿着极线搜索即可，这样可以大大降低寻找匹配像素点时的计算量。

一条极线上可能存在若干个色彩或灰度相同的像素点，因此通过像素点之间的比较来获取双目图像的匹配并不可靠。本章采用块匹配的方法寻找双目图像的匹配像素。

对于左眼图像中待匹配的像素 p ，首先以 p 为中心，提取一个尺寸为 $w \times w$ 的像素块 A 。然后在右眼图像对应的极线上，以每一个像素 q_i 为中心，提取尺寸为 $w \times w$ 的像素块 B_i 。之后计算 A 与 B_i 的相似性评分 $S(A, B_i)$ ， $S(A, B_i)$ 越接近 1，说明两个像素块 A 与 B_i 越相似。

$$S(A, B_i) = \frac{\sum_{m=1}^w \sum_{n=1}^w A(m, n) B_i(m, n)}{\sqrt{\sum_{m=1}^w \sum_{n=1}^w A(m, n)^2} \sqrt{\sum_{m=1}^w \sum_{n=1}^w B_i(m, n)^2}} \quad (5-15)$$

寻找右眼图像极线上与 A 最相似的像素块 B_i ，并将相似性评分赋予两像素块的每一对对应像素点。对于左眼图像中的每个像素，都进行这样的搜索。由于像素块有重叠区域，寻找到的对应像素点对也会有重叠的部分，对于这样的部分，将其相似性评分相加。最后遍历左眼图像中的每个像素点，寻找右眼图像中与其相似性评分最高的像素点，作为对应像素点。如果左眼图像中有多个像素点对应着右眼图像中同一个像素点，则只取相似性评分最高的一对像素点，剔除其他的匹配。

在获取了两幅图像匹配像素点后，利用 2.2 节构建的相机模型，就可以根据像素点构建出三维空间中的点。

5.2.2 稠密点云地图的构建

在漫游器双目视觉自主定位的过程中，以拍摄第一帧图像时左眼相机坐标系为基础坐标系，获得了漫游器在整条路径上拍摄每帧图像时，左眼相机坐标系相对于基础坐标系的位姿变换。在稀疏特征点地图拼接的过程中，又得到了漫游器 1 与漫游器 2 基准坐标系之间的转换关系。根据这个关系，可以获得漫游器 2 拍摄每一帧图像时左眼相机坐标系相对于漫游器 1 基准坐标系的坐标转换关系。

利用 5.2.1 节的方法，可以获得漫游器每一帧左、右眼图像之间的匹配像素点。计算匹配像素点的视差，根据双目相机模型，可以得到匹配像素点的深度，进而可以利用单孔相机模型将这些像素点还原到每一帧的左眼坐标系中。附带点的 rgb 颜色特征，即得到了单帧图像构建的稠密的点云。

利用每一帧左眼相机坐标系相对漫游器 1 基准坐标系的坐标转换关系，将每一帧图像构建的稠密点云都转换到漫游器 1 的基准坐标系下，即得到了初步的火星表面大范围稠密点云地图。由于像素点的匹配等过程存在误差，并且各帧的稠密点云存在重复的部分，这样得到的稠密点云地图并不十分准确，需要进行一些优化处理。

本章对初步构建的稠密特征点地图进行了两项优化处理：去除离群的和降采样。

与其它点距离较远、以散点形式出现的点，通常为深度估计错误的点，这些点会影响稠密点云地图的准确性，降低后续八叉树地图构建的精度，需要对其进行剔除。对于稠密点云地图中的点 P，首先寻找与其距离最近的 n 个点，计算点 P 与这些点之间距离的标准差 σ ：

$$d = \frac{1}{n} \sum_{i=1}^n \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} \quad (5-16)$$

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} - d \right)^2} \quad (5-17)$$

式中： n——点的个数；

$[x_i \ y_i \ z_i]^T$ ——点 p 附近的点 P_i 的坐标；

$[x \ y \ z]^T$ ——点 p 的坐标；

d——点 p 与点 P_i 的距离。

将这 n 个点中与 P 的距离大于一个标准差的点视为离群点，进行剔除。

降采样则可以降低稠密点云地图的点云密度，并且可以将重叠的点合并。

这样做的好处是可以加快后续处理的速度，并且可以降低点云的可视化操作对计算机硬件的要求。首先将所有点云划分为大小相等的方格区域，称为单元格。

然后计算每个单元格内所有三维点的重心 $[x \ y \ z]^T$ ：

$$\begin{cases} x = \frac{1}{n} \sum_{i=0}^n x_i \\ y = \frac{1}{n} \sum_{i=0}^n y_i \\ z = \frac{1}{n} \sum_{i=0}^n z_i \end{cases} \quad (5-18)$$

式中： n ——单元格包含的点的个数；

$[x_i \ y_i \ z_i]^T$ ——单元格内的点 P_i 的坐标。

每个单元格只在重心位置留下一个点，其 **rgb** 颜色信息也按照重心求取的方法计算，这样即完成了稠密点云地图的降采样，使稠密点云地图中点的数量大大降低，而又保留了稠密点云地图的基本形貌特征。

5.3 八叉树地图构建算法

稠密点云地图可以用于火星表面地形的可视化研究，但是其所需存储空间过大，不适合用于火星漫游器的自主路径规划。而利用八叉树分割空间的三维占有率地图则可以在占据较小的存储空间的前提下，表示大范围地形的空间占据情况。

根据稠密点云构建八叉树地图的方法如下：

- (1) 将点云所在的空间平均分为八个等大的立方体区域；
- (2) 统计每个立方体区域内三维点的数量，如果三维点数量大于一定阈值，则将这个立方体区域视为占据状态；
- (3) 对每个划分得到的小立方体，继续将其平均分为八个等大的立方体区域；
- (4) 重复 (2) 至 (3) 步，直到立方体的尺寸小于给定的分辨率。

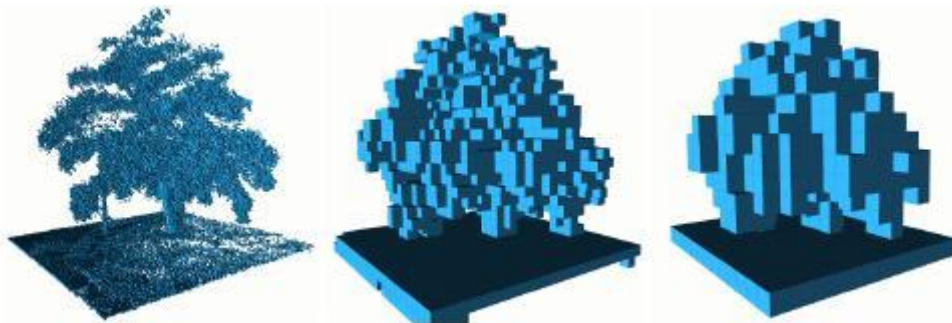


图 5-1 八叉树地图示意图

八叉树地图的一大优势在于节省存储空间。

在划分立方体、判断占据情况完成后，可以对立方体进行合并。假设一共划分了 n 层立方体，最大立方体为第 0 层，最末层为第 n 层，立方体 A 位于第 $(n-1)$ 层。如果由 A 划分的小立方体（位于第 n 层）的占据状态全部与 A 相同，则可以将这些小立方体合并，对地图进行压缩，只利用立方体 A 来表示此处空间的占据情况。

依次类推，如果第 $(n-2)$ 层立方体 B 所划分的位于第 $(n-1)$ 层的小立方体，以及这些小立方体所划分的位于第 n 层的更小的立方体，全部处于与立方体 B 相同的占据状态，则可以将立方体 B 以下的所有小立方体合并，对地图进行压缩，只用立方体 B 来表示此处空间的占据情况。

由于空间中的物体往往是连续的，因此可以有较多的立方体合并的情况，进而有效减小地图中包含的立方体数量，降低地图的存储空间。

5.4 仿真实验与分析

5.4.1 基于地面模拟场地图像的八叉树地图构建实验

利用双目相机在模拟场地采集的图像进行稠密点云地图与八叉树地图构建的实验。

首先利用双目相机在模拟场地采集图像，然后利用第二章介绍的方法进行漫游器自主定位与稀疏特征点地图构建，以拍摄第一帧图像时左眼相机坐标系为基准坐标系，得到拍摄每一帧图像时的左眼相机坐标系相对于基准坐标系的坐标转换关系。

然后利用 5.2.1 节介绍的方法，对每一帧的左、右眼图像进行匹配，利用双目相机模型构建每一帧的稠密点云，如图 5-2 所示。

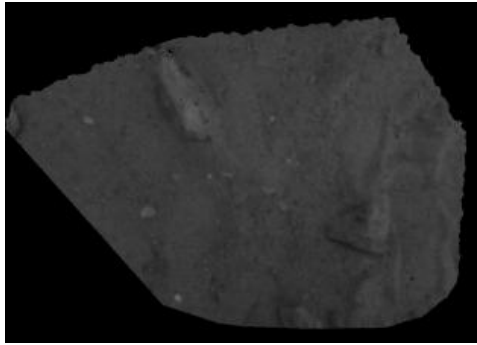


图 5-2 由单帧双目图像构建的稠密点云

之后利用 5.2.2 节介绍的方法, 根据每一帧的相机位姿将稠密点云都转换到基准坐标系下, 构建稠密点云地图, 并对其进行去除离群点和降采样的优化。去除离群点时, 参数 n 选取为 50; 降采样时, 选取单元格大小为 $3\text{mm} \times 3\text{mm} \times 3\text{mm}$ 。

得到的初始稠密点云地图包含 22739672 个点, 如图 5-3 所示。去除离群点后包含 22589742 个点, 共去除了 149930 个离群点, 如图 5-4 所示。降采样的结果如图 5-5 所示, 包含 1533772 个点, 仅是原来的 6.79%, 但是仍然保留了原有点云所展现出的地图形貌。

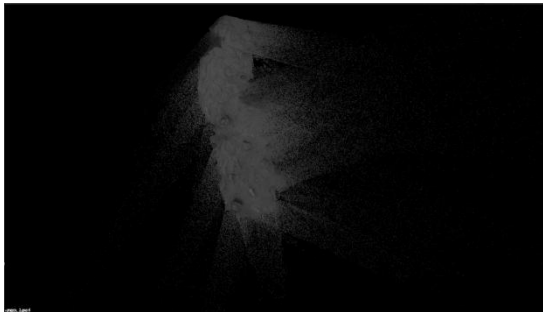


图 5-3 初始稠密点云地图

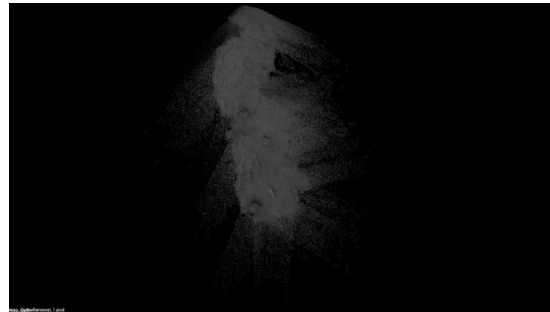


图 5-4 去除离群点后的稠密点云

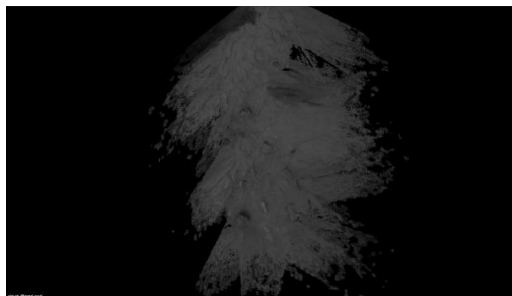


图 5-5 降采样后的稠密点云地图

最后利用 5.3 介绍的方法, 利用优化后的稠密点云地图构建八叉树地图,

如图 5-6 所示。



图 5-6 八叉树地图

稠密点云地图包含 1533772 个点，其大小为 36.4MB；八叉树地图最小单元格的大小为 0.0064m，其大小为 6.6MB。可见，八叉树地图有效地减小了地图的存储空间。

5.4.2 基于计算机模拟图像的八叉树地图构建实验

由于地面模拟场地规模较小，漫游器双目相机采用斜向下的角度进行拍摄。为了验证当双目相机平视行进方向时仍能利用本章介绍的方法进行稠密点云地图与八叉树地图的构建，利用计算机模拟图像进行仿真实验。

首先根据第二章介绍的方法，利用 3DMax 软件制作火星表面的模拟场景，并渲染得到漫游器在其中行进时双目相机拍摄的图像。然后利用计算机模拟图像进行漫游器自主定位与稀疏特征点地图构建，以拍摄第一帧图像时左眼相机坐标系为基准坐标系，得到拍摄每一帧图像时的左眼相机坐标系相对于基准坐标系的坐标转换关系。

利用 5.2.1 节介绍的方法，对每一帧的左、右眼图像进行匹配，并利用双目相机模型构建每一帧的稠密点云，如图 5-7 所示。



图 5-7 由单帧图像构建的稠密点云

之后利用 5.2.2 节介绍的方法，利用每一帧的相机位姿将稠密点云都转换到基准坐标系下，构建稠密点云地图，并对其进行去除离群点和降采样的优化。去除离群点时，参数 n 选取为 50；降采样时，选取单元格大小为 $3\text{mm} \times 3\text{mm} \times 3\text{mm}$ 。最终得到的稠密点云地图如图 5-8 所示。将火星表面模拟场景与稠密点云的形貌进行对比，可以看出，稠密点云基本与模拟场景一致，这既证明了本章提出的基于双目视觉的稠密点云地图构建方法正确有效，又进一步验证了第二章基于双目视觉的漫游器自主定位的准确性。



图 5-8 稠密点云地图（局部）

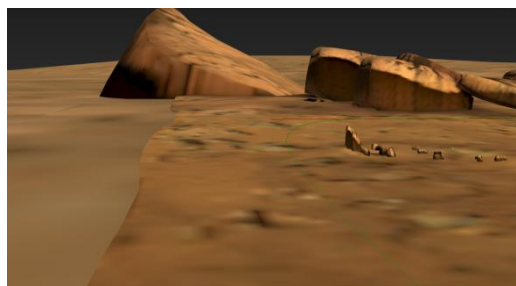


图 5-9 火星表面模拟场景（局部）

最后利用 5.3 节介绍的方法,利用优化后的稠密点云地图构建八叉树地图。构建的八叉树地图如图 5-10 所示,其中,图 5-10(a)的最小单元格尺寸为 0.32m,图 5-10(b)的最小单元格尺寸为 1.28m。优化后的稠密点云地图共包含 18529530 个点,其大小为 296.5MB;最小单元格尺寸为 0.32m 的八叉树地图,其大小为 223.9KB;最小单元格尺寸为 1.28m 的八叉树地图,其大小为 16.5KB。可见,八叉树地图有效地减小了地图的存储空间。

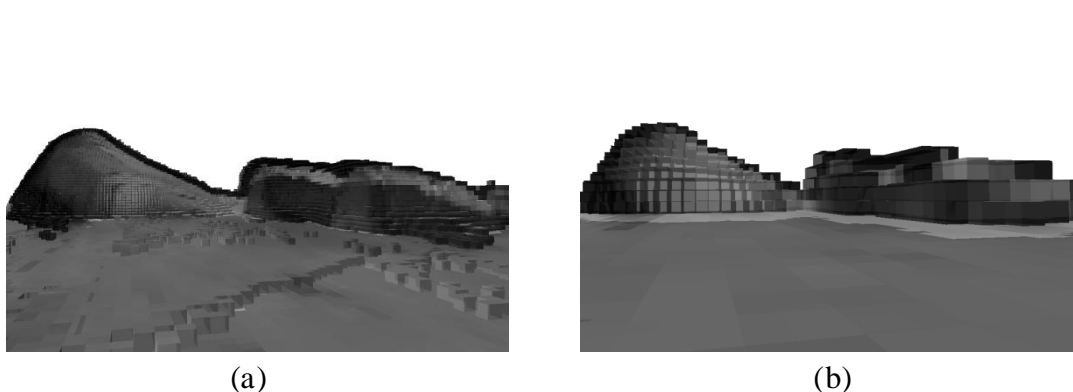


图 5-10 八叉树地图

5.5 本章小结

本章研究了基于双目视觉构建火星表面稠密点云地图,以及利用稠密点云地图构建八叉树地图的方法。

首先研究了利用极线搜索和块匹配的方法得到双目图像间的匹配像素的方法,然后研究了构建稠密点云地图的方法,并研究了对稠密点云地图进行剔除离群点、降采样等优化处理的方法。

之后,本章研究了利用稠密点云地图构建八叉树地图的方法。

最后利用双目相机在模拟场地采集的图像以及计算机模拟图像对算法进行仿真验证,证明本章提出的火星表面稠密点云地图与八叉树地图构建的方法准确有效。

结 论

火星表面大范围导航地图对于利用漫游器进行火星探测具有重要意义，特别是未来的火星地表探测任务，需要漫游器具有更加自主智能的路径规划能力。本文提出了一种多漫游器协同构建导航地图的方法，取得的研究成果如下。

(1) 给出了基于双目视觉的 SLAM 方法，能够对火星漫游器进行自主定位，并构建出火星表面的稀疏特征点地图。利用火星漫游器实际拍摄的图像进行了特征点提取与匹配的实验，实验结果表明，该算法可以用于火星表面环境；利用地面模拟场地图像和计算机模拟图像进行了仿真实验，实验结果表明，漫游器行进路程的误差低于 2%。

(2) 提出了基于词袋模型的火星漫游器行进路径上重叠区域的搜索算法，仿真实验结果表明，该算法具有很好的计算精度和计算速度，可以用于不同火星漫游器进行大范围探测时行驶路径中重叠区域的搜索，为后续导航地图协同拼接提供了支撑。

(3) 提出了不同漫游器下的稀疏特征点地图与漫游器位姿协同拼接方法。该方法能构建出大范围火星表面的稀疏特征点地图，并建立统一坐标系下的漫游器位姿估计，仿真结果证明了该算法的有效性。

(4) 给出了面向后续漫游器导航应用的八叉树地图构建方法，该方法能够建立存储空间较小的导航八叉树地图，可用于表示火星表面大范围探测。

论文后续还有以下方面的工作有待完善。

(1) 本文构建的稀疏特征点地图，忽略了火星表面环境的变化，只适用于短期内其他漫游器的自主定位。在后续的研究中，可以考虑对特征点进行筛选，只保留在石块、山谷等较长时间不发生改变的地形中提取的特征点。

(2) 在构建稠密点云时，由于双目图像的误匹配等原因可能存在一些误差较大的点。目前本论文对稠密点云的优化方法只有去除离群点和降采样，可以考虑采用更多的方法对稠密点云进行优化。

(3) 本论文只是提出了第一批漫游器经过后，利用其拍摄的图像构建导航地图方法，并未对后续漫游器经过相同路径时的导航过程进行研究。后续可以研究基于八叉树地图的自主路径规划方法，并对利用稀疏特征点地图进行导航的过程进行仿真。

参考文献

- [1] 史建魁, 刘振兴. 火星探测研究的概况及意义[A]. 中国地球物理学会. 中国地球物理学会第二十届年会论文集[C]. 中国地球物理学会, 2004: 1.
- [2] European Space Agency. Water At Martian South Pole[N], ScienceDaily, 2014-03-19
- [3] 刘园园. NASA 公布火星表面有液态水的“强有力”证据[N]. 科技日报, 2015-09-29 (001)
- [4] 士元. 火星探测——痛并快乐着[N]. 中国航天, 2011-11-11 (004) .
- [5] 付丽丽. 中国确定二〇二〇年左右发射火星探测器[N]. 科技日报, 2016-04-23 (003)
- [6] Wilt D M, Jenkins P P, Scheiman D A. Photodetector development for the Wheel Abrasion Experiment on the Sojourner microrover of the Mars Pathfinder mission[C]// Energy Conversion Engineering Conference, 1997. Iecec-97. Proceedings of the Intersociety. IEEE, 1997: 738-742 vol.1.
- [7] “机遇”号火星车破外星行驶距离纪录[J]. 科技创新导报, 2014, 11(22): 3.
- [8] 岳宗玉, 邸凯昌. 好奇心号巡视器及其特点分析[J]. 航天器工程, 2012, 21 (05): 110-116. [2017-09-06].
- [9] 邸凯昌. 勇气号和机遇号火星车定位方法评述[J]. 航天器工程, 2009, 18 (05): 1-5.
- [10] Jet Propulsion Laboratory. A Description of the Rover Sojourner[DB/OL]. <https://mars.jpl.nasa.gov/MPF/rover/descrip.html>, 1996/2017-09-08.
- [11] Jet Propulsion Laboratory. Mars Pathfinder Mission Status[DB/OL]. <https://mars.jpl.nasa.gov/MPF/ops/sol86.html>, 1997-10-01/2017-09-08
- [12] Li R, Di K, Matthies L H, et al. Rover localization and landing site mapping technology for 2003 Mars Exploration Rover mission[J]. Photogrammetric Engineering and Remote Sensing, 2004, 70 (1): 77-90
- [13] Trebi Ollennu A, Huntsberger T, Cheng Y, et al. Design and analysis of a Sun sensor for planetary rover absolute heading detection[J]. IEEE Trans. Robotics and Automation, 2001, 17 (6): 939-947
- [14] Cameras - Mars 2020 Rover
<https://mars.nasa.gov/mars2020/mission/rover/cameras/#CacheCam-Engineerin>

g-Cameras

- [15]吴伟仁, 周建亮, 王保丰, 刘传凯. 嫦娥三号“玉兔号”巡视器遥操作中的关键技术[J]. 中国科学: 信息科学, 2014, 44 (04): 425-440.
- [16]王保丰, 周建亮, 唐歌实, 等. 嫦娥三号巡视器视觉定位方法. 中国科学: 信息科学, 2014, 44: 452-460
- [17]R. C. Smith and P. Cheeseman, “On the representation and estimation of spatial uncertainty”, *International Journal of Robotics Research*, vol. 5, no. 4, pp. 56-68, 1986
- [18]R. Triebel, W. Burgard. Improving simultaneous mapping and localization in 3D using global constraints. *Proceedings of the 20th national conference on Artificial intelligence*. 2005: 1330-1335
- [19]林睿. 基于图像特征点的移动机器人立体视觉 SLAM 研究[D]. 哈尔滨工业大学, 2011.
- [20]A. Davison, I. Reid, N. Molton, and O. Stasse, “Monoslam: Real-time single camera SLAM”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 1052-1067, 2007
- [21]M. Labbe and F. Michaud, “Online global loop closure detection for large-scale multi-session graph-based slam”, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2661-2666, IEEE, 2014
- [22]J. Stephen. Real-time stereo visual SLAM. MSc Thesis, Heriot-Watt University. 2008
- [23]T. Lemaire, C. Berger, I. Jung, et al. Vision-based SLAM: stereo and monocular approaches. *International Journal of Computer Vision*. 2007, 74 (3): 323-364
- [24]D. Schleicher, L. M. Bergasa, R. Barea, et al. Real-time wide-angle stereo visual SLAM on large environments. *IEEE international Conference on Intelligent Robots and Systems*. 2007: 3878-3883
- [25]Smith R, Cheeseman P. On the representation and estimation of spatial uncertainty. *The International Journal of Robotics Research*, 1986, 5(4): 56-68
- [26]Montemerlo M, Thrun S, Koller D, Wegbreit B. FastSLAM: a factored solution to the simultaneous localization and mapping problem. *Proceedings of the National Conference on Artificial Intelligence*, Edmonton, Canada, 2002: 593-598
- [27]S. Thrun, Y. Liu, D. Koller, et al. Simultaneous localization and mapping with sparse extended information filters. *International Journal of Robotics Research*.

- 2004, 23 (7): 693-716
- [28]Julie S, Uhlman J, Durrant-Whyte H F. A new method for the nonlinear transformation of means and covariances in filters and estimators. IEEE Transactions on Automatic Control, 2000, 45 (3): 477-482
- [29]Montemerlo M. Fast SLAM: A Factored Solution to the Simultaneous Localization and Mapping Problem with Unknown Data Association [Ph. D. dissertation], Carnegie Mel-lon University, Pennsylvania, 2003
- [30]Lu F. Milios E. Globally consistent range scan alignment for environment mapping[J]. Autonomous robots, 1997, 4 (4): 333-349
- [31]H. Strasdat, J. M. Montiel, and A. J. Davison, "Visual slam: Why filter?", Image and Vision Computing, vol. 30, no. 2, pp. 65-77, 2012
- [32]Gutmann J S, Konolige K. Incremental mapping of large cyclic environments. In Proceedings of CIRA 99, Monterey, California, 1999: 318-325
- [33]Dedeoglu G, Sukhatme G S. Landmark-based Matching Algorithm for Cooperative Mapping by Autonomous Robots[C]. Distributed Autonomous Robotic Systems , Proceedings of the International Symposium on Distributed Autonomous Robotic Systems, Dars 2000, October 4-6, 2000, Knoxville, Tennessee, Usa. DBLP, 2000: 251-260.
- [34]Fox D, Ko J, Konolige K, et al. A Hierarchical Bayesian Approach to the Revisiting Problem in Mobile Robot Map Building[M]// Robotics Research. The Eleventh International Symposium. Springer Berlin Heidelberg, 2005: 551-558.
- [35]潘薇. 多移动机器人地图构建的方法研究[D]. 中南大学, 2009.
- [36]Fenwick J W, Newman P M, Leonard J J. Cooperative Concurrent Mapping and Localization[C]// IEEE International Conference on Robotics and Automation , 2002. Proceedings. ICRA. Washington, DC, 2002: 1810-1817 vol.2
- [37]苑全德. 基于视觉的多机器人协作 SLAM 研究[D]. 哈尔滨工业大学, 2016.
- [38]Andreas Geiger, Julius Ziegler and Christoph Stiller: StereoScan: Dense 3D Reconstruction in Real-time[C], 2011 IEEE Intelligent Vehicles Symposium(IV) Baden, Germany, June 5-9, 2011
- [39]Mason J, Ricco S, Parr R. Textured occupancy grids for monocular localization without features[C]/ IEEE International Conference on Robotics and Automation. Piscataway, NJ, USA: IEEE, 2011: 5800-5806.
- [40]A. Hornung, K. M. Wurm, M. Bennewitz, et al. OctoMap: An efficient

- probabilistic 3D mapping framework based on octrees[J]. *Autonomous Robots*, 2013, 34 (3): 189-206.
- [41]Boal J, Sánchez Miralles Á, Arranz Á. Topological simultaneous localization and mapping: a survey[J]. *Robotica*, 2014,32(5):803-821.
- [42]Nüchter A, Hertzberg J. Towards semantic maps for mobile robots[J]. *Robotics & Autonomous Systems*, 2008,56(11):915-926.

攻读硕士学位期间发表的论文及其它成果

- [1] 顾远凌, 张泽旭, 魏祥泉. 一种基于双目视觉的火星车自主导航方法. 中国宇航学会深空探测技术专业委员会第十四届学术年会, 成都, 2017.

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《火星表面多漫游器协同构建导航地图方法研究》，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：顾远凌 日期：2018年06月30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遵守有关保密规定，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：顾远凌 日期：2018年06月30日

导师签名：张洪成 日期：2018年06月30日

致 谢

逝者如斯，不舍昼夜。转眼间两年的硕士生活即将结束。回望在深空探测基础研究中心的这两年时光，我由衷地庆幸能与这么多优秀的师生在一起科研、生活，获得帮助与成长，更收获了一段难忘的人生经历。

首先要感谢我的恩师张泽旭教授。从刚踏入实验室起，张老师广博的知识就使我深深折服。从图像处理、导航制导到编队控制，张老师知识的触手深入到深空探测传统领域的各个角落。根基稳固而不囿于成见，张老师更开创性地将vr技术、神经网络等前沿技术引入到深空探测的研究中，不断碰撞出新的火花，激发学生的灵感，不断取得创新与突破。“要站在巨人的肩膀上，紧跟前沿的发展”，张老师的教导鼓励着学生们将目光放长远，将束缚思维的高墙打穿。成功之花由汗水灌溉，这是张老师给我的又一启迪。只要不出差，每天学生们路过张老师的办公室，一定能看见他在坐在办公桌前，或眉头紧锁苦苦思辨，或十指敲击书写报告。周末不休，小长假不休，就连春节放假，老师也是比学生们更晚离校、更早返回。以学为乐、以工作为乐、以创造为乐，厚积薄发，张老师为我树立了最好的榜样。“事事有回音，件件有着落”、制定进度安排、每天做总结、每月有回望，张老师在科研工作中时刻注意培养学生良好的工作习惯，训练团队协作能力，为将来学生走向工作岗位打下了坚实的基础。不止教书，更重育人，这是张老师最大的魅力所在。

我还要感谢崔祜涛老师、徐敏强老师、史雪岩老师、徐田来老师、李玉庆老师、梁廷伟老师、王日新老师、高晶波老师在学业上给予我的指导，各位老师们在毕设中提出的专业意见使我对论文不断丰富完善。

感谢张卓、郑博、张良、徐毅、苏宇师兄与胡茜师姐在学习和生活中给予我的关心与指导，感谢王笛、乔衍迪、杨名、马泽祥、张文政、张博连、董志强同学的帮助，感谢池雨灿师妹、李江、王洪亮、刘金龙师弟对我的启发，还要感谢迟旭、齐宗谕、陈秋实、刘有权对我在技术上的指导。感谢实验室的各位伙伴们，是你们共同营造了一个温馨、愉悦、进步的环境，让我快乐地度过这两年的时间。你们是我一生的朋友。

最后，我要感谢我的父亲、母亲，是你们在物质上与精神上的支持与鼓励，让我健康成长，你们是我坚实的后盾，让我敢于在外拼搏、闯荡。谢谢你们！

感谢身边每一个给予我帮助的人，我将珍藏这份感动，凝聚这股力量，在未来的生活中继续披荆斩棘，乘风破浪！