单位代码： 密 级：



硕 士 学 位 论 文



论文题目： 室内外三维环境感知系统技术研究

1220055726

学号

姓名

导 师

学 科 专 业

研 究 方 向

申请学位类别

论文提交日期

朱海鹏

高浩

电子信息

工程硕士

二零二三年四月

Research on Indoor and Outdoor 3D Environmental Perception System Technology

Thesis Submitted to Nanjing University of Posts and Telecommunications for the Degree of

Master of Engineering



By

Haipeng Zhu

Supervisor: Prof. Hao Gao

April 2023

南京邮电大学学位论文原创性声明

本人声明所呈交的学位论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得南京邮电大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

本人学位论文及涉及相关资料若有不实，愿意承担一切相关的法律责任。

研究生学号：\_\_\_\_\_\_\_\_\_\_\_ 研究生签名：\_\_\_\_\_\_\_\_\_\_\_\_ 日期：\_\_\_\_\_\_\_\_\_\_\_\_

南京邮电大学学位论文使用授权声明

本人承诺所呈交的学位论文不涉及任何国家秘密，本人及导师为本论文的涉密责任并列第一责任人。

本人授权南京邮电大学可以保留并向国家有关部门或机构送交论文的复印件和电子文档；允许论文被查阅和借阅；可以将学位论文的全部或部分内容编入有关数据库进行检索；可以采用影印、缩印或扫描等复制手段保存、汇编本学位论文。本文电子文档的内容和纸质论文的内容相一致。论文的公布（包括刊登）授权南京邮电大学研究生院办理。

非国家秘密类涉密学位论文在解密后适用本授权书。

研究生签名：\_\_\_\_\_\_\_\_\_\_\_\_ 导师签名：\_\_\_\_\_\_\_\_\_\_\_\_ 日期：\_\_\_\_\_\_\_\_\_\_\_\_\_

摘要

多源传感器融合感知与三维重建一直以来都是无人车、机器人、智能化设备以及VR/AR等领域的研究热点。其中包括多传感器的融合、高精度的地图重建、鲁棒的环境感知以及准确的路径规划等工作。随着多传感器融合感知与三维重建相关技术的发展，以及我国对于航天领域探索的不断深入，对航天领域无人车的多传感器融合感知与重建系统的研究越发重要。然而，目前多源传感器融合感知以及三维重建还面临着传感器种类少、时空同步精度低、复杂且未知地形的重建信息少、小物体的3D感知不鲁棒以及数据依赖等问题。本文基于航天探索背景的二代火星车以及火星场地形，结合火星车搭载的传感器种类多、火星场地形复杂等特点，并针对现有方法所存在的问题做出了以下工作：

（1）多源传感器时空一致表征：针对现有多传感器融合算法存在的融合传感器种类较少、时空同步精度较低等问题，本文提出了一种用于多源传感器融合的时空同步方法。该方法包括机器人、智能驾驶汽车等相关领域中大于5种的常见传感器的时间同步与空间同步。本文从各类传感器的软件与硬件方面同时着手，首先进行多种传感器的时间同步，因为各个传感器的初始采集数据会存在一定的时间差，所以要将各类传感器与系统时钟保持同步，用以保证能够获取到同一时刻的传感器数据；其次进行多种传感器的空间同步，该部分主要进行各种传感器的内参标定与外参标定，用以确定传感器自身参数以及传感器之间的空间位置关系。高精度的时空同步结果也是多源传感器融合感知与三维重建的保障。

（2）基于多源信息的多层地图构建：针对现有地图重建算法在夜晚等恶劣环境下对复杂的未知地形重建感知能力差的问题，本文提出了一种基于多源传感器信息的多层地图构建方法。该方法包括三维地图重建和多层地图构建两个步骤。首先，基于多视图的MVE算法进行三维地图的重建；然后，通过多传感器空间同步所得到的热成像仪、多光谱相机与RGB相机的内参以及相对位置关系，进行多层地图的构建；最终构建出一个不仅包含颜色，还包含热和光谱信息的多源多层点云地图，为路径规划和环境感知等任务提供更加准确和全面的地图信息。

（3）基于双目视觉的三维环境感知：针对现有的三维环境感知算法中存在的对小物体的3D感知不鲁棒和数据依赖等问题，本文提出了一种基于双目立体视觉的2D感知+3D映射的三维环境感知方法。首先，在双目相机的左目图像上进行基于SOLO算法的实例分割，实现在2D上的物体感知；其次，利用基于深度学习的双目立体匹配网络BGNet恢复左目相机坐标系下的视差图，并将视差图转为深度图和点云；最后，将2D物体映射到3D点云中，并对物体的3D点云加以约束，获得最终的感知结果。实验表明，使用双目立体视觉和实例分割算法，能够有效提升对小物体的3D感知的鲁棒性，并且不需要任何的3D标注，使得感知结果更加准确和可靠。

关键词: 火星场，多源传感器融合，三维重建，环境感知，双目立体视觉

Abstract

Multi-source sensor fusion perception and 3D reconstruction have always been research hotspots in the fields of unmanned vehicles, robots, intelligent devices, and VR/AR. These include multi-sensor fusion, high-precision map reconstruction, robust environment perception, and accurate path planning. With the development of multi-sensor fusion perception and 3D reconstruction related technologies, and the continuous deepening of my country's exploration of the aerospace field, the research on the multi-sensor fusion perception and reconstruction system of unmanned vehicles in the aerospace field is becoming more and more important. However, there are still challenges in multi-sensor fusion perception and 3D reconstruction, such as limited types of sensors, low accuracy of spatiotemporal synchronization, lack of reconstruction information for complex and unknown terrains, non-robust of 3D perception for small objects, and data dependency. Therefore, this paper is based on the second-generation Mars rover and the terrain of the Mars field in the background of space exploration, combined with the variety of sensors on the rover and the terrain of the Mars field is complex and other characteristics, and the following work has been done to address the problems existing in the existing methods:

(1) Multi-source sensor spatio-temporal consistent representation: Aiming at the problems of limited sensor types and low spatio-temporal synchronization accuracy in existing multi-sensor fusion algorithms, this paper proposes a spatio-temporal synchronization method for multi-source sensor fusion. The method includes time synchronization and space synchronization of more than five common sensors in related fields such as robots and intelligent driving cars. This paper starts from the software and hardware of various sensors at the same time. Firstly, the time synchronization of various sensors is carried out. Because there will be a certain time difference in the initial data collection of each sensor, it is necessary to keep all sensors synchronized with the system clock to ensure that sensor data can be obtained at the same time; Secondly, the spatial synchronization of multiple sensors is carried out. This part mainly performs internal and external reference calibration of various sensors to determine the parameters of the sensor itself and the spatial position relationship between the sensors. High-precision spatio-temporal synchronization results are also the guarantee for multi-sensor fusion perception and 3D reconstruction.

(2) Multi-layer map construction based on multi-source information: Regarding the problem of poor perception ability for complex unknown terrain reconstruction in adverse environments such as night-time in existing map reconstruction algorithms, this paper designs a multi-source sensor information-based multi-layer map reconstruction method. The method first reconstructs the three-dimensional map based on the multi-view MVE algorithm; Then, constructs a multi-layer map through the internal reference and relative position relationship of the thermal imager, multispectral camera and RGB camera obtained by multi-sensor space synchronization; Finally, a multi-source multi-layer point cloud map containing not only color but also thermal and spectral information is formed, which provides more accurate and comprehensive map information for tasks such as path planning and environmental perception.

(3) 3D environment perception based on binocular vision: To address the problems of non-robust 3D perception of small objects and data dependence in existing 3D environment perception algorithms, this paper proposes a 3D environment perception method based on 2D perception + 3D mapping using binocular stereo vision. Firstly, instance segmentation based on the SOLO algorithm is performed on the left image of the binocular camera to achieve object perception in 2D; Secondly, use the binocular stereo matching network BGNet based on deep learning to restore the disparity map in the coordinate system of the left camera, and convert the disparity map into a depth map and point cloud; Finally, 2D objects are mapped to 3D point clouds, and constraints are applied to the 3D point cloud of the objects to obtain the final perception result. And this method not only allows for the perception of smaller objects but also does not require any 3D annotation. The experiments showed that using binocular stereovision and instance segmentation algorithms can effectively improve the robustness of 3D perception for small objects, and no 3D annotation is needed, making the perception results more accurate and reliable.

Key words: Mars terrain, Multi-source sensor fusion, 3D reconstruction, Environment perception, Binocular stereo vision.

目录

[摘要 I](#_Toc130892710)

[Abstract III](#_Toc130892711)

[第一章 绪论 1](#_Toc130892712)

[1.1 研究背景与意义 1](#_Toc130892713)

[1.2 国内外研究现状 2](#_Toc130892714)

[1.2.1 多传感器融合研究现状 2](#_Toc130892715)

[1.2.2 三维重建研究现状 4](#_Toc130892716)

[1.2.3 三维环境感知研究现状 6](#_Toc130892717)

[1.3 本文的研究内容 7](#_Toc130892718)

[1.4 本文的章节安排 8](#_Toc130892719)

[第二章 相关背景知识介绍 11](#_Toc130892720)

[2.1 相机成像模型 11](#_Toc130892721)

[2.1.1 针孔相机模型 11](#_Toc130892722)

[2.1.2 相机畸变模型 12](#_Toc130892723)

[2.2 多视图几何模型 13](#_Toc130892724)

[2.2.1 三维空间的刚体运动 13](#_Toc130892725)

[2.2.2 对极几何 14](#_Toc130892726)

[2.3 双目立体视觉模型 15](#_Toc130892727)

[2.4 本章小结 16](#_Toc130892728)

[第三章 多源传感器时间一致表征 17](#_Toc130892729)

[3.1 时间一致表征系统总体设计 17](#_Toc130892730)

[3.2 激光雷达的时间同步 18](#_Toc130892731)

[3.3 视觉相机的时间同步 20](#_Toc130892732)

[3.4 惯性测量单元IMU的时间同步 22](#_Toc130892733)

[3.5 本章小结 24](#_Toc130892734)

[第四章 多源传感器空间一致表征 25](#_Toc130892735)

[4.1 空间一致表征系统总体设计 25](#_Toc130892736)

[4.2 传感器内参标定 26](#_Toc130892737)

[4.2.1 视觉相机内参标定 26](#_Toc130892738)

[4.2.2 惯性测量单元IMU内参标定 29](#_Toc130892739)

[4.3 多传感器外参联合标定 30](#_Toc130892740)

[4.3.1 视觉相机与视觉相机之间联合标定 30](#_Toc130892741)

[4.3.2 视觉相机与激光雷达之间联合标定 33](#_Toc130892742)

[4.3.3 视觉相机与IMU之间联合标定 36](#_Toc130892743)

[4.4 本章小结 38](#_Toc130892744)

[第五章 基于多源信息的多层地图构建 39](#_Toc130892745)

[5.1 重建算法设计方案 39](#_Toc130892746)

[5.2 基于MVE算法的多层地图重建 40](#_Toc130892747)

[5.2.1 运动恢复结构SFM 40](#_Toc130892748)

[5.2.2 多视图立体视觉MVS 42](#_Toc130892749)

[5.2.3 表面重建FSSR 43](#_Toc130892750)

[5.2.4 多层地图映射 44](#_Toc130892751)

[5.3 实验结果与分析 45](#_Toc130892752)

[5.4 本章小结 48](#_Toc130892753)

[第六章 基于双目视觉的三维环境感知 49](#_Toc130892754)

[6.1 感知系统算法设计方案 49](#_Toc130892755)

[6.2 基于双目视觉的目标空间定位 50](#_Toc130892756)

[6.2.1 基于SOLO网络的实例分割算法 50](#_Toc130892757)

[6.2.2 基于BGNet网络的双目立体匹配算法 51](#_Toc130892758)

[6.3 实验结果与分析 54](#_Toc130892759)

[6.3.1 实例分割算法实验结果 54](#_Toc130892760)

[6.3.2 双目空间定位算法实验结果与分析 57](#_Toc130892761)

[6.4 本章小结 59](#_Toc130892762)

[第七章 总结与展望 61](#_Toc130892763)

[7.1 本文工作总结 61](#_Toc130892764)

[7.2 未来工作展望 62](#_Toc130892765)

[参考文献 63](#_Toc130892766)

[附录1 攻读硕士学位期间撰写的论文 67](#_Toc130892767)

[附录2 攻读硕士学位期间申请的专利 68](#_Toc130892768)

[附录3 攻读硕士学位期间参加的科研项目 69](#_Toc130892769)

[致谢 70](#_Toc130892770)

# 绪论

1.1 研究背景与意义

随着工业4.0的到来，越来越多的智能制造[1]、智能机器人、智能驾驶等技术在不断的发展和完善，也正逐步的改变着我们的生活。这些工业设备、机器人、自动驾驶汽车等都对环境的感知尤为依赖。例如工业设备需要根据实际应用场景使用单目相机、双目相机、线扫激光等传感器；室内清洁、服务型机器人多以单线雷达进行地图重建，并使用RGBD相机进行场景的感知和避障；对于自动驾驶汽车而言，它作为一种特殊类型的移动类机器人，其工作的环境特殊且复杂，需要多线激光雷达、视觉相机、惯性测量单元、轮速计、GPS等多种传感器结合的方式进行场景的重建和环境感知。尤其对于复杂、恶劣环境，以及未知环境的探索，将会对多传感器的融合感知技术有更高的需求。例如航天对于火星、月球等环境的探索，都需要高精度的三维地图[2]，高精度的机器人定位和高精度且鲁棒的环境感知系统，这都离不开各种传感器之间的融合与协作。

在2020年的7月份，长征五号遥四运载火箭将火星任务车“祝融号”发射升空。其为天问一号任务火星车，重达240公斤，高度将近2米。火星车主要依赖轨道器预先开展约三个月的对地观测，以便对预选着陆区的详细勘测[3]。之后火星车与轨道器分离，并且利用反推火箭以及降落伞协同的方式在火星表面安全着陆。着陆后，火星车将通过搭载的导航相机、多光谱相机、激光雷达、磁场探测仪、气象测量仪、成分探测仪等传感器按计划开展巡视区域的环境感知、火星表面移动和科学探测。同时轨道器也持续运行在轨道中，与火星表面的火星车形成空地一体协同探测，并为火星车的巡视提供稳定的中继通信和保障。

对于早期的室内外无人化设备、机器人等智能体多以一个或两个传感器进行物体的识别以及周围环境感知，甚至最初只能实现二维平面内的一些定位、识别和感知功能。这也导致以前的许多智能化、自动化设备动作机械化，容易受干扰，且不鲁棒。随着多传感器融合技术以及三维重建、三维环境感知等相关技术的发展使得各类智能体，机器人包括智能驾驶汽车逐渐的发展起来，也能实现更多精细化的功能。

早期的室内外重建，大多使用单线激光雷达或单目相机用于室内，多线激光雷达用于室外的地图重建，并在此地图的基础上进行机器人导航。但是由于单一传感器对于大场景地图重建中存在的累积漂移现象，使得基于多传感器融合的即时定位与建图逐渐的成为了主流。基于多传感器融合的定位与建图在没有回环的情况下也能有很好的效果，所重建出来的地图信息也更加丰富，并且基于多传感器的环境感知，也更加精准。由此可见，对于复杂地形环境，尤其是针对像火星、月球等航天领域的探索中，多传感器的配准融合、多维地图的重建和三维环境的感知显得尤为重要。

总而言之，如何对于复杂环境下进行高质量、多信息的三维地形重建与物体感知是本文的主要目标。针对火星、月球等航天领域的环境探索，如何在未知环境下，进行鲁棒性较高的地图高精度、高维度的重建与感知是目前科技发展与航天领域中需要不断探索与研究的问题。本文将从多传感器的时空一致表征，火星场多层次地图的重建，双目视觉三维环境感知三个方面展开研究，针对航天院火星车多传感器系统以及实验室搭建的多传感器系统，从软硬件方面入手并对现有算法进行改进与应用。

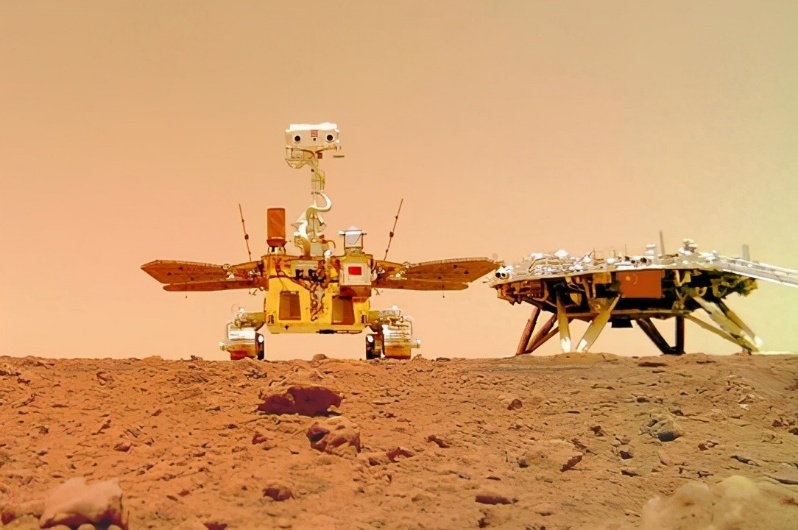


图1.1 “祝融号”火星车图像

1.2 国内外研究现状

1.2.1 多传感器融合研究现状

随着深度学习实用性的显着发展和5G通信技术的超高速信息传输速率将突破数据壁垒车联网传输，智能驾驶技术正在逐渐成为影响交通的关键技术之一。对于智能驾驶系统而言，传感器是其感知外界环境的关键，并且其协作性能直接决定了自动驾驶车辆的安全性。

多传感器融合（Multi-sensor Fusion, MSF）技术[4]，是利用现有计算机相关知识与技术，综合利用所获得不同传感器的信息和数据并以一定的准则进行自动分析和融合，避免了感知限制和不确定性，形成对环境更全面的感知和识别或目标，并提高外部感知能力系统。目前，多传感器融合技术已广泛应用于故障检测领域[5][6], 遥感[7], 人体健康监测[8]、机器人系统[9][10]、人机交互[11]、目标识别和跟踪[12][13]、即时定位与建图[14] 和高级驾驶员辅助系统[15]等领域。

近年来在智能驾驶中多传感器融合的不同策略也会有不同的融合感知效果，常用的传感器包括毫米波雷达、单线和多线激光雷达、RGB摄像头、深度相机、热成像仪、超声波、GPS、IMU 和轮速计等。各类传感器都有各自的优点与不足，因此在自动驾驶系统中各个传感器通常会有不同任务的划分[16]，也会相互配合完成一项任务，进而达到高效、高精度的效果。目前，几种主要传感器组合系统包括相机-激光雷达（Camera-LiDAR，CL）系统、毫米波雷达-相机（Radar-Camera，RC）系统以及毫米波雷达-相机-激光雷达（Radar-Camera- LiDAR，RCL）系统。根据现有数据图1.2表明，使用最多的传感器组合是RC系统，因为这种组合不仅可以获得周围物体的距离信息，而且具有优异的分辨率。同样，激光雷达和相机的组合系统即可以获得颜色信息又可以获得深度信息，并且有人在此基础上又加入了毫米波雷达，来提高安全冗余度。虽然视觉相机可以获得目标的许多信息，但其缺点也是很明显的，例如对光照敏感且只有2D信息。双目相机与深度相机的可以获得图像的深度信息，但其在精度和抗噪声方面还有待提升。毫米波雷达相较于相机和激光雷达具有更长的波长，在雨雪天气以及雾天有不错的效果。美中不足的是，毫米波雷达比较容易受到一些杂波的干扰。激光雷达除了可以昼夜连续工作外，还可以提供高精度的三维点云数据。综上所述，为了应对各种恶劣工作条件，较好的解决方案是采用多传感器融合技术。

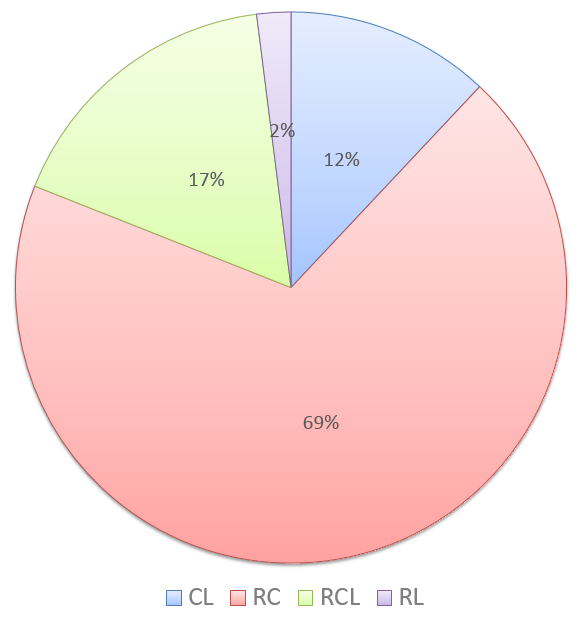


图1.2 常见传感器组合使用比例

目前常见的多传感器融合主要分为两种。一种是如图1.3的前融合算法，另一种是如图1.4的后融合算法。前融合指的是每个传感器都是融合体的一部分，将所有传感器依据他们之间的关联性融合成类似于一个传感器包含所有初始信息的形式，最终输出融合结果；后融合算法又称松耦合算法，它的本质是将所有传感器独立开来，每个传感器进行相应的识别、感知算法，最后将所有感知结果进行一定的汇总，后融合的算法是传感器间相互独立的，不存在传感器之间的约束。



图1.3 前融合算法框图



图1.4 后融合算法框图

1.2.2 三维重建研究现状

目前在过去几年中，GPU硬件的进步促进了三维重建系统的发展。此前，Pollefeys [17]等人、Agarwal[18]等人和Furukawa[19]等人的大规模努力从非结构化照片集重建了城市场景的部分。最先进的算法所依赖的持续优化理论基础不断加强和拓宽，如Chambolle和Pock[20]、以及Goldluecke [21]等人的工作，已成为机器人和计算机视觉应用的基础。硬件和理论结合在一起，使其能够构建如图1.5所示的创建大规模3D密集重建的系统。然而，许多密集三维重建系统的最新技术很少考虑可扩展性在自动驾驶或检查等地图应用中的实际应用。最普遍的方法是最近的着眼于小规模的重建开发，如Klingesmith [22]等人、Engel [23]等人以及Schöps [24]等人在手机和平板电脑上的工作。

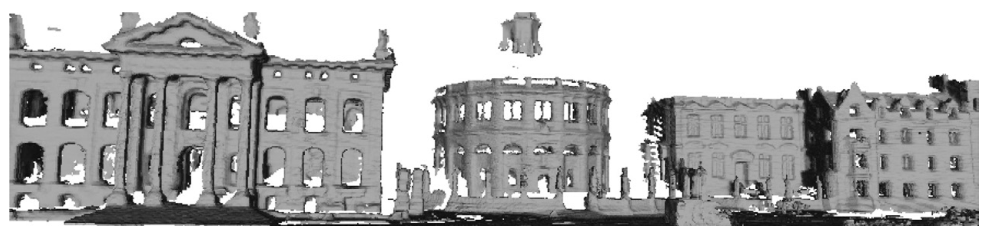


图1.5立体视觉或激光数据高效生成多公里尺度环境的密集模型

一些研究人员研究了更适用于自动驾驶汽车的数据和传感器。最值得注意的是，2010年谷歌发布了一篇学术文章，详细介绍了他们的“树视图”应用程序，该应用程序利用激光和相机数据创建世界各地城市的密集3D重建[25]。然而，他们的算法过度强调激光数据，并假设所有深度图仅包含分段平面对象。2013年，谷歌提出了一种仅使用相机传感器的运动方法替代结构[26]。Xiao和Furukawa[27]提出了一种利用激光和图像输入对大型室内环境进行建模的系统，但他们修改的曼哈顿世界假设限制了重建，无法对除垂直和水平平面以外的任何东西进行建模。Bok等人[28]在现有技术的基础上，使用相机和激光创建大规模3D地图。他们的最终重建是稀疏的，仅使用相机进行里程测量和环路闭合，而不是作为额外的深度传感器来改善密集的重建。

在RGB-D相机广泛使用之前，DTAM[29]提出了一种使用单眼相机生成高质量深度图的方法。在相机的焦平面前面构建成本体积，并使用来自连续图像帧的2D正则化深度估计来不断更新成本体积。最终的重建提供了精细的细节，但该系统将单眼相机的范围限制为近场重建。2010年，第一台商品RGB-D相机发布。RGB-D相机为图像中的每个像素提供厘米级的精确深度测量：在第一台设备中，分辨率为640×480，频率为30Hz。Curless和Levoy[30]通过Kinect Fusion[31]系统扩展了他们的工作，以利用这种高频和高质量的深度图流。利用截断有符号距离函数（TSDF），深度观测值存储在体素网格中，其中每个体素在表面前面时包含一个正数，在表面后面包含一个负数。求解零值水平集会生成原始曲面的密集模型。因此，Kinect Fusion可以为大约7m³的工作空间实时生成前所未有的高质量密集3D重建。

与Kinect Fusion（体素网格固定在空间中的一个位置）不同，Kintinuous[32]试图通过允许体素网格随相机移动来扩展重建场景的大小。先前观察到的区域的连续流作为网格流式传输到磁盘，但如果再次观察到该区域，则可以重新加载到GPU中。该系统理论上无限扩展了重建工作空间的大小。然而，它不能利用距离RGB-D相机3米以外的传感器观测，因为它仍然基本上基于传统的固定尺寸体素网格。

Nießner的哈希体素网格[33]（HVG）也扩展了重建的大小，但只在观察到表面的区域分配体素。这样删除的内存浪费了存储可用空间。当表面远离传感器时，与GPU和硬盘之间的数据流相结合时，重建的大小基本上没有限制。这种实施方式将传感器范围限制在4米，因为这接近Kinect相机的最大有效范围，但该范围可以很小地扩展。Whelan等人[34]的解决方案利用滚动循环缓冲作为体积重建数据结构。这是一种有趣的方法，它允许局部体积区域在摄影机在环境中移动时进行虚拟平移。在深度学习环境中，正确选择有效的体积数据结构也获得了关注，因为它会影响3D任务的分辨率，包括3D对象分类、方向估计和点云标记[35]。

三维重建作为计算机视觉领域的一个活跃研究领域，在虚拟现实、逆向工程和机器人视觉等领域有着重要的应用。

1.2.3 三维环境感知研究现状

自从计算机视觉出现[36][37]以来，3D环境重建和物体6D姿态估计一直是一个核心问题。先进的机器人经常出现在科幻电影中，这提高了人类对未来机器人的期望。我们希望机器人能像我们一样感知周围的世界。然而，现实世界的机器无法实现这种智能。我们的人类视觉系统可以快速轻松地推断3D对象的属性，并理解周围环境的组成。相比之下，机器人视觉传感器捕捉环境的图像，但输入设备不具有感知能力或认知能力；它不了解世界。因此，迄今为止，赋予机器这些能力是一项具有挑战性的任务。

计算机视觉系统最重要的目标之一是从复杂环境中估计物体的6D姿态（3D平移和旋转）。精确的6D姿态估计技术可以驱动与机器人操纵、自动驾驶、增强现实等相关的各种新兴技术领域。为了在杂乱场景中感知准确的6D姿势，大量研究致力于建立高效的3D物体感知系统。想法包括了利用对象几何信息的基于特征的方法[38][39]以及使用深度神经网络的基于学习的方法[40][41][42]。无论使用哪种范式，它们都依赖于高质量的对象模型。换句话说，优越的物体模型是准确的6D姿态估计的先决条件。

在3D计算机视觉的早期，它主要处理多面体物体[43][44]。为了识别细粒度几何形状，一些方法[45][46]直接从3D点云估计物体的6D姿态。他们利用点特征描述来计算对象模型和场景点云之间的一组对应关系，并通过随机样本一致性（RANSAC）算法稳健地估计6D姿态。除了基于点云的方法外，基于图像的方法也取得了重大进展。Hinterstoisser等人[38]提出了一种模板匹配方法Linemod，该方法利用颜色梯度和表面法线来描述对象轮廓和内部信息。在[39]中，通过从CAD 3D模型自动生成模板数据集，对其进行了进一步改进。然而，由于缺乏丰富的纹理，这些作品无法处理杂乱环境中的无纹理对象，以检测可区分的特征进行匹配。

最近，深度卷积网络技术的出现，特别是基于CNN的类别检测器[47][48][49]，已经显示出用于对象检测和对象分割的优异结果。受2D物体检测的显著进步的启发，越来越多的工作将深度学习用于6D姿态估计。对于RGB导出的6D姿态估计，大多数方法遵循类似的范式：首先，它们采用神经网络来检测与目标对象相关的八个3D边界框角点。他们执行透视n点（PnP）算法来计算方向和平移。然而，该范式在无纹理对象的检测精度低和昂贵的后处理步骤方面存在严重缺陷[50]。

与2D图像相比，点云更接近原始的3D几何形状，并具有可靠的深度距离以精确定位对象。Qi等人提出了一种端到端网络PointNet[51]，它直接将点云作为输入，并对输入点云进行分类和分割。在[52]中，提出了PointNet的改进版本，使网络能够学习不同规模的局部结构。虽然这些方法仍处于类别级别，但Li提出了PointRCNN[53]，它可以在实例级别从原始点云实现3D对象检测。然而，一个典型的场景模型包含超过100k个点；训练这样的网络需要高的计算和存储器要求。此外，当前基于点云的学习技术利用公开可用的数据集进行训练和测试。尽管数据集可以加快研究过程，但如何生成适用于真实世界对象的通用系统仍然具有挑战性。

在自动驾驶系统中，环境感知[54]包括物体定位、离线障碍物和道路映射、移动障碍物跟踪以及交通信号检测和识别。在现在的研究中，关注较多的是自动驾驶系统中的多对象三维感知问题。通常，环境感知的作用是检测物体，并从城市场景中的传感器数据中获取其定位信息。

先前对自动驾驶多目标感知的研究大致可分为手工特征算法和深度学习[55]方法。手工制作的特征算法在很大程度上依赖于专业知识和技能来提取代表性特征；因此，多传感器融合方法已成为一种有前途的方法，它使用多模态输入数据（即图像和点云数据），并使用各种融合策略组合其各自的特征图，以实现更准确和鲁棒的检测[56][57][58][59]。MV3D[56]使用紧凑的多视图表示对稀疏的3D点云数据进行编码，并结合来自多个视图的区域特征，用于3D空间中的对象检测。截头体PointNet[57]框架利用成熟的2D对象检测器来生成区域建议，并在几何关系的约束下精确估计3D边界框。AVOD[58]在自动驾驶场景中聚集用于对象检测的视图。在所提出的网络中，从点云和RGB图像数据生成共享特征图，然后通过区域建议网络融合多模态特征图以获得可靠的3D对象建议。该方法对于城市场景中的目标检测是有效的。

1.3 本文的研究内容

在上一小节中，主要简述了目前国内外包括多传感器融合、三维重建以及三维环境感知等相关技术的研究现状。但是通过大量的调研发现目前针对多传感器融合感知以及三维重建还存在着诸多问题。具体的，本文将针对以下几个关键问题展开研究。

（1）多传感器融合的时空一致表征

针对的问题：多传感器融合算法融合传感器种类少、时空同步方法复杂、精度低。

为了解决以上问题，本文提出了一种用于多传感器融合的时空同步方法。该方法包括机器人、智能驾驶汽车等相关领域中大于5种的常见传感器的时间同步与空间同步。本文从各类传感器的软件与硬件方面同时着手，首先进行多种传感器的时间同步，因为各个传感器的初始采集数据会存在一定的时间差，所以要将各类传感器与系统时钟保持同步，用以保证能够获取到同一时刻的传感器数据；其次进行多种传感器的空间同步，该部分主要进行各种传感器的内参标定与外参标定，用以确定传感器自身参数以及传感器之间的空间位置关系。高精度的时空同步结果也是多传感器融合感知与三维重建的保障。

（2）基于多源信息的多层地图构建

针对的问题：现有地图重建算法在夜晚等恶劣环境下对复杂的未知地形重建感知能力差。

为了解决以上问题，本文提出了一种基于多源传感器信息的多层地图构建方法。该方法包括三维地图重建和多层地图构建两个步骤。首先，基于多视图的MVE算法进行三维地图的重建；然后，通过多传感器空间同步所得到的热成像仪、多光谱相机与RGB相机的内参以及相对位置关系，进行多层地图的构建；最终构建出一个不仅包含颜色，还包含热和光谱信息的多源多层点云地图，为路径规划和环境感知等任务提供更加准确和全面的地图信息。

（3）基于双目视觉的三维环境感知

针对的问题：对小物体的3D感知不鲁棒、数据依赖。

为了解决以上问题，本文提出了一种基于双目立体视觉的2D感知+3D映射的三维环境感知方法。首先，在双目相机的左目图像上进行基于SOLO算法的实例分割，实现在2D上的物体感知；其次，利用基于深度学习的双目立体匹配网络BGNet恢复左目相机坐标系下的视差图，并将视差图转为深度图和点云；最后，将2D物体映射到3D点云中，并对物体的3D点云加以约束，获得最终的感知结果。实验表明，使用双目立体视觉和实例分割算法，能够有效提升对小物体的3D感知的鲁棒性，并且不需要任何的3D标注，使得感知结果更加准确和可靠。

1.4 本文的章节安排

针对多源传感器融合感知以及三维重建任务当前面临的问题。本文主要从多源传感器时空一致表征；基于多源信息的多层地图构建；基于深度学习的双目立体视觉感知三个方面着手。基于航天院二代火星车及火星场，进行三个方面的实践应用与研究。本文各章节组织结构如图1.6所示。



图1.6 章节组织结构

本文由七个章节组成，各个章节的内容和相关结构安排如下：

第一章绪论部分，主要介绍了目前关于多传感器融合在智能驾驶等无人车领域的应用，以及三维重建和三维环境感知的国内外研究现状。简单介绍了基于航天探索领域的相关应用背景，并且对本文的研究内容与章节安排进行了概述。

第二章进行相关基础知识与原理的介绍。首先进行单目相机的小孔成像模型、相机畸变模型的介绍；其次对三维重建中较为关键的多视图几何模型，包括刚体的运动模型、对极几何等原理进行了介绍；最后介绍了三维环境感知中常见的双目立体视觉模型的相关原理。

第三章主要从多种传感器的软硬件着手，进行多源传感器融合的前期准备工作，即多源传感器的时间一致表征。重点是实现多类传感器的时间一致，保障这些传感器能采集到同一时刻的数据。

第四章主要进行的多源传感器的空间一致表征。重点是实现多类传感器的内参标定、畸变校正以及传感器之间的外参标定。将各类传感器统一到一个坐标系下，为后续的多源传感器融合、感知、重建等算法提供良好的工作准备。

第五章首先使用RGB图像序列实现基于MVE算法的火星车附近场景的三维重建，再利用第三章与第四章的时空一致的相关数据，将热成像仪、多光谱成像仪的数据映射到已经建好的三维地图中，实现一个多源多层的点云地图。

第六章介绍两个深度学习网络，分别是基于深度学习的SOLO图像分割网络和基于深度学习的双目立体匹配网络BGNet进行视差图的恢复。通过2D的感知和3D的映射，再对物体的3D点云加以约束，最终实现基于双目视觉的特定目标的三维空间定位。

第七章，总结与展望。主要是对本文工作与相关实践内容的一个总结，分析了自身工作中的问题以及相关算法的不足，并对未来工作的相关工作进行了展望。

# 相关背景知识介绍

本章将介绍三维视觉里面的相关的基础知识、数学表达及原理。首先进行单目相机的小孔成像模型、相机畸变模型的介绍；又对三维重建中较为关键的多视图几何模型，包括刚体的运动模型、对极几何等原理进行了介绍；最后介绍了三维环境感知中常见的双目立体视觉模型的相关原理与算法。

2.1 相机成像模型

2.1.1 针孔相机模型

最早的照相设备以及现在很多人在使用的相机。其在成像上都是采用的针孔成像的原理，只不过对于这个针孔，已经逐渐的由简单的物镜变成了较为复杂且高质量的相机镜头，对于成像平面的位置变成了电子成像原件或者是感光胶片之类的成像载体。

如图2.1为理想的针孔相机的成像模型[60]。其中针孔o与像平面之间的距离称之为焦距f。对于理想的针孔相机模型，物体的成像没有畸变，且所成的像是一个倒像，这样计算机处理计算时需要一个负号。因为，通常研究的如图2.1中的虚拟像平面。

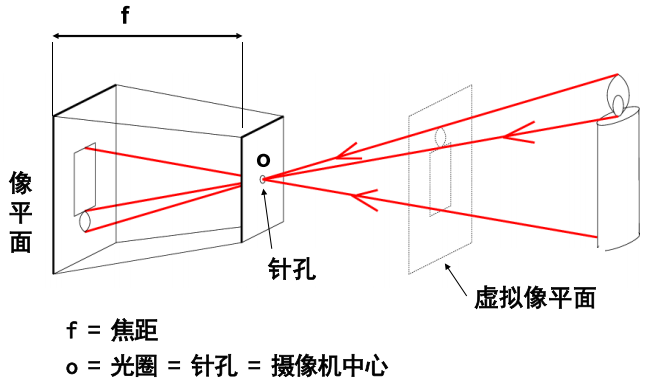


图2.1 针孔相机模型

下面为一个三维点到二维点的映射过程。设O-x-y-z为相机坐标系，其中O为针孔，x-y平面与成像平面平行，z轴指向物体。在此相机坐标系下，设空间中的某一物体点，该三维空间点，经过针孔成像在像平面--上成像点为，其中针孔到像平面的距离为f。所以最终可以通过图2.2所示相似三角形关系可得：

将式子整理可得：

上述式子即根据相似三角形原理，描述了三维点到二维成像平面的映射关系，最终可得到空间点P在相机成像O-u-v平面的成像坐标为

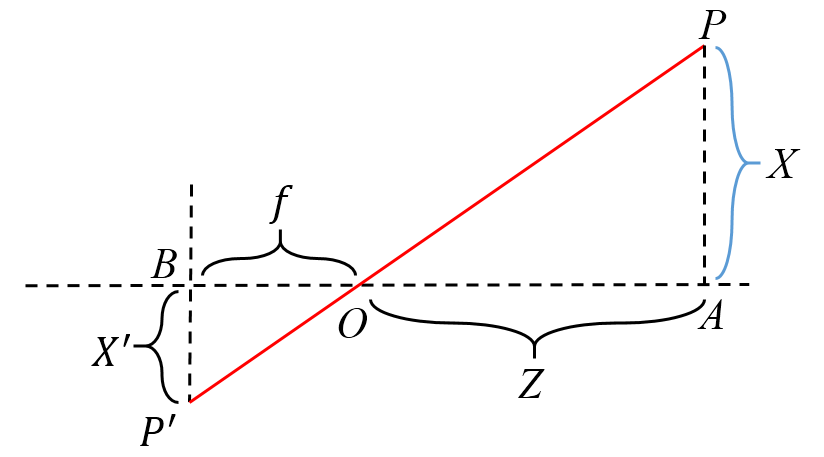


图2.2 相似三角形成像图

2.1.2 相机畸变模型

上述的针孔成像模型是理想状态下的线性成像模型，随着相机镜头的出现，现在也更多的是非线性的透镜成像。但是由于镜头的安装误差以及制作过程中工艺的误差，通常在成像平面上与理想模型计算所得坐标不一致，这也就是透镜畸变[61]。

透镜的畸变主要包括径向畸变和切向畸变。其中径向畸变主要是因为镜头工艺不理想，在成像平面上产生径向失真；切向畸变主要是由于安装的过程中无法使镜头与像平面平行，导致切向失真。

（1）径向畸变

相机的径向畸变主要是由于镜头制作工艺的误差带来的影响，使得像素点无法理想的到达成像平面对应位置，产生径向失真。常见的径向畸变失真有“肥胖”桶型和“瘦小”枕型两种失真。径向畸变的数学模型可由如下公式近似表示：

其中分别对应x、y方向的径向畸变分量，为径向畸变系数。

（2）切向畸变

相机的切向畸变主要是由于镜头安装的过程中，人为的无法确保镜面与像平面平行，产生切向失真。切向畸变的数学模型可由如下公式近似表示：

其中分别对应x、y方向的切向畸变分量，为切向畸变系数。r为理想像素点距离畸变中心点的像素长度，即。

所以，对于一个镜头，通常需要5个参数来描述畸变以及进行畸变的矫正。

2.2 多视图几何模型

2.2.1 三维空间的刚体运动

机器人在三维空间中的位姿变化，一般会描述为一个刚体的运动。在三维的空间里面，通常用来描述刚体运动的方式主要是旋转和平移。对于平移来说，相对较好理解和描述，通常用一个平移向量来描述刚体三个自由度的平移量；对于旋转而言，描述的方式就比较多了，例如旋转矩阵、四元数等。对于旋转不同的表示方法来说，他们之间也都有着各自的优势和劣势，通常根据实际需要来选择具体的表示方式。下面为几种常见的三维空间中刚体旋转的表示方式：

（1）旋转矩阵

对于一个向量a，其在两个坐标系下的表示是相等的，可知：

再将左右两边同乘一个进行等式化简，可得：

由上可得向量经过矩阵R的转换，可以得到向量a，也即是一个向量在两个坐标系下的表示是相等的。上述的矩阵R描述了旋转，所以被称为旋转矩阵。此矩阵又是一个正交矩阵，且行列式为1。最终可将旋转矩阵的集合定义为如下的一个特殊正交群：

（2）旋转向量和欧拉角

对于上面介绍的旋转矩阵而言，其需要9个量，难免会显得有些冗余。旋转向量和欧拉角是更为紧凑的表示方式。对于任意的旋转而言，都是可以由一个旋转轴外加一个旋转角度来进行描述。所以，使用一个旋转轴为**n**，旋转角度为的向量来进行描述，这个向量就称为旋转向量。除了旋转向量，欧拉角的描述方式也非常的紧凑，只需要3个量来描述旋转，欧拉角的描述方式在无人机等领域应用较多，一般通常使用roll-pitch-yaw分别表示绕XYZ轴的旋转角度。所以，使用即可描述任意旋转。

旋转向量与旋转矩阵之间还可以还可以通过罗德里格斯公式进行相互的转化，公式如下：

（3）四元数

虽然已经有了紧凑的旋转向量和欧拉角，但是他们存在一个非常严重的问题，那就是万向锁问题，也即是它们存在奇异性。一个三维的向量是不可避免奇异性的，所以才引入了四元数来描述旋转，它即紧凑，又没有奇异性，是一个常用的描述旋转的方式。四元数由一个实部、三个虚部组成，定义如下：

其中***i***，***j***，***k***为三个虚部，他们之间满足如下关系：

四元数到旋转矩阵的转换公式如下：

旋转向量到四元数的转换公式如下（其中设={}，）：

由上述公式可知，在实际的计算中，如果趋近于0，则会导致解的不稳定，此时需要使用其他方法进行旋转矩阵到四元数的求解。

2.2.2 对极几何

对极几何一般指的是现实中的两个相机之间，像平面不平行，并且以两个相机之间的基线作为轴的几何约束关系。这种约束关系只与两个相机的内参以及两个相机之间的相对位置有关，与空间中实际物体无关。对极几何关系图如图2.3所示，表示两个相机拍摄同一场景。其中*O*--P称为极平面，*O、*分别为左右相机的坐标系原点，连线为基线。极平面与两个相机的像平面分别相交于*l*和，P点对应的两个投影点分别为*x*和，基线与像平面的交点称为极点，分别为e和。

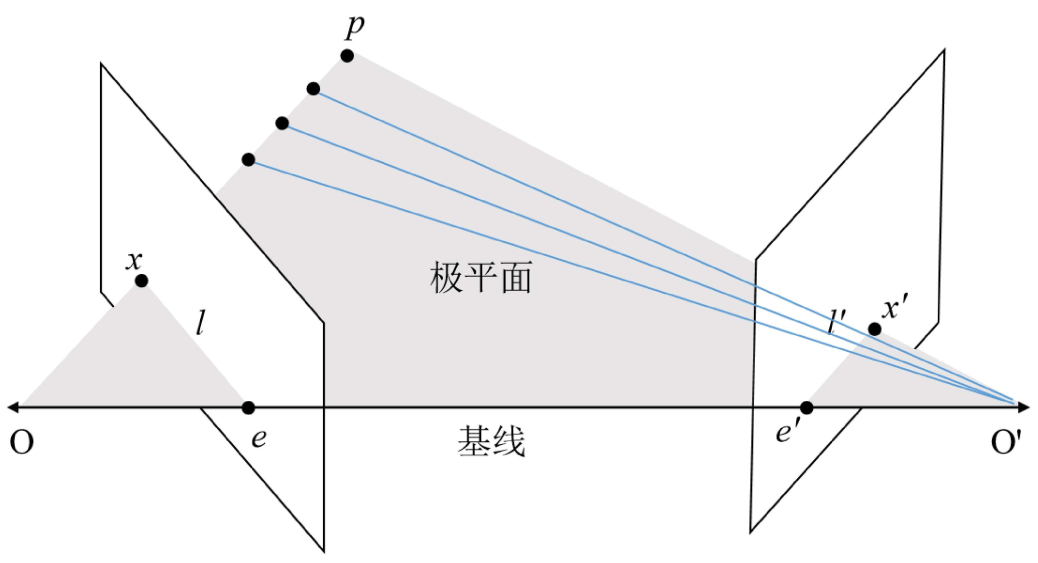


图2.3 对极几何关系图

对于P点在左相机的投影点x，其在右相机上的投影一定在极线上，在P点深度未知的情况下，极线上任意位置都有可能。这就是对极几何的极线约束，通过此几何约束，在进行左右视图之间的匹配时，对于左图的某一点，仅仅只需要在有图的极线上进行检索，大大的减小的搜索范围和计算量，同时也加速的算法的匹配速度。

2.3 双目立体视觉模型

在立体视觉中，双目相机可以根据两个相机之间的视差进行三维信息的恢复。相较于单目相机，其可以恢复三维信息；相较于深度相机，其又相对简单，成本低。双目视觉常被用来做双目测距、三维重建、深度恢复等工作。

为了简单的描述双目立体视觉模型，减小计算量，通常建立如图2.4所示的两个相机光轴平行的理想双目立体视觉模型。其中P为三维空间中一点，--与--分别为左右相机成像平面，空间点P在两个相机成像点为和。左右相机坐标系分别为---和---，并且将左相机坐标系作为世界坐标---。设P在左相机成像下的像素坐标为，在右相机成像下的像素坐标为。并且如图所示的两个相机的基线间距为b，相机的焦距为f。

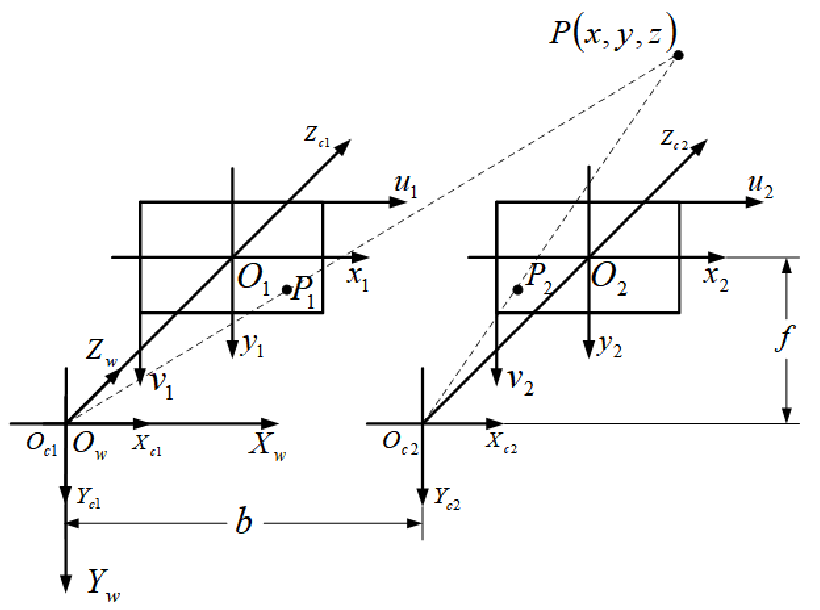


图2.4 双目相机立体成像模型

通过计算可得、之间的距离dis = b – ()。

所以，由图中映射关系，并且结合相似三角形，可建立如下式子：

将式中dis进行替换，并进行整理可得：

上述公式中的f为相机的焦距，b为双目相机的基线距离，均可通过相机内参标定获得。所以，只需要获得双目相机的视差，即，就可以计算出对应点的深度信息。

2.4 本章小结

本章首先进行了相机成像模型的介绍，包括对针孔相机模型的介绍以及利用相似三角形进行成像原理的描述，并进行了相机透镜的径向畸变与切向畸变模型；然后介绍了三维重建中十分重要的多视图几何模型，其中先介绍了三维空间中刚体的运动模型，重点介绍了描述旋转的旋转矩阵、旋转向量、欧拉角和四元数，然后通过几何原理介绍了对极几何中的几何约束；最后介绍了三维环境感知中常见的双目立体视觉模型，通过对三维空间中的一点在两个相机中的投影点，并利用相似三角形原理进行了双目测距的数学推导。

# 多源传感器时间一致表征

多源传感融合是提高自动驾驶、机器人等无人设备性能与精度的非常实用的办法。在多源传感器融合系统中，对于各种各样的传感器，如果无法获取到同一时刻的传感器数据，那么将无法在后续的算法中进行融合，即便个别传感器之间可以融合，也需要复杂的运动矫正。所以在多传感器融合、多传感器感知等系统中，首先需要做的工作就是进行多传感器的时间同步，即确保所有传感器能同一时刻进行数据的采集。

3.1 时间一致表征系统总体设计

时间同步的任务是在多传感器的数据进行融合前消除时间上的不同步，各个传感器的初始采集数据会存在一定的时间差，把这个时间差尽量减小，甚至消除，实现各传感器采集的数据在时间上统一即为时间同步[62]。本章将从软硬件相结合来实现多线激光雷达、彩色相机、鱼眼相机、深度相机、热成像仪、多光谱相机和IMU的时间同步，总体结构示意图如图3.1。



图3.1 时间同步系统总体示意图

该系统主要包含使用激光雷达授时的方式，实现激光雷达与系统UTC时间一致；利用触发信号进行各类相机的硬触发实现相机之间的时间一致；通过测量激光雷达与相机采集数据的时间差调整相机的触发信号达到激光雷达与相机之间时间一致；再通过高频的IMU数据与激光雷达之间进行就近时间戳插值对齐的方式实现IMU与激光雷达时间一致。

3.2 激光雷达的时间同步

激光雷达内部时钟，默认是以激光雷达上电时刻开始从0走时。所以激光雷达的时间同步主要是采用授时的方式，即将系统的UTC时间同步到激光雷达内部，以此来改变激光雷达内部时钟，实现激光雷达与系统UTC时间同步。激光雷达时间同步方案示意图如下：



图3.2 激光雷达时间同步方案示意图

激光雷达时间同步方案如图3.2所示。其中硬件方面主要由Xavier工控机，通过USB转RS232串口与激光雷达连接，用于发送GPRMC报文，并且通过Xavier工控机的一个IO口实现PPS信号的发送；其中软件方面主要由Xavier工控机端运行模拟GPS数据(GPRMC报文数据)发送模块和PPS信号发送模块。

之所以采用模拟的GPRMC数据，主要原因是本文的研究背景中不允许使用GPS之类的传感器，但是激光雷达的授时同步又需要此类报文数据，所以只能由工控机通过串口编程，模拟一串GPRMC格式的报文(包含系统UTC时间)，发送到激光雷达端。

对于GPRMC报文数据帧格式包含14个字段，具体如下：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| F0 | F1 | F2 | F3 | F4 | F5 | F6 | F7 | F8 | F9 | F10 | F11 | F12 | F13 |

其中对于每一个字段：

F 0：帧头，$GPRMC

F 1：UTC时间, 格式hhmmss.ssss，代表时分秒.毫秒

F 2：状态 A/V:分别代表定位成功/失败

F 3：纬度，格式ddmm.mmmmmm

F 4：纬度 N/S

F 5：经度，格式dddmm.mmmmmm

F 6：经度 E/W

F 7：速度

F 8：方位角

F 9：UTC日期 DDMMYY 日月年

F 10：磁偏角

F 11：磁偏角方向E/W

F 12：模式，A =自动，D =差分，E =估计，AND =无效数据

F 13：校验和

如下例子，则表示的系统UTC时间为21-03-24日的9:32分17秒：

“$GPRMC,093217.119,A,2237.496474,N,11356.089515,E,0.0,225.5,240322,2.3,W,A\*23”

通过PPS和模拟GPS数据（GPRMC数据）的组合就能够实现对激光雷达时钟同步。激光雷达在接收Xavier工控机发送模拟GPS信号后，以下一个PPS信号的上升沿作为时间同步基准。通过编程实现Xavier发送模拟GPS信号功能，如下图所示，具体要求为：激光雷达串口要求串口波特率：9600；数据位：8；无校验；停止位:1；PPS信号要求为标准的TTL信号，脉宽大于 200ns。通过修改激光雷达驱动程序让激光雷达能够接收GPRMC数据帧提供的标准时间信号。使用RS232协议进行串口通信模拟GPRMC数据帧发送如下图所示：

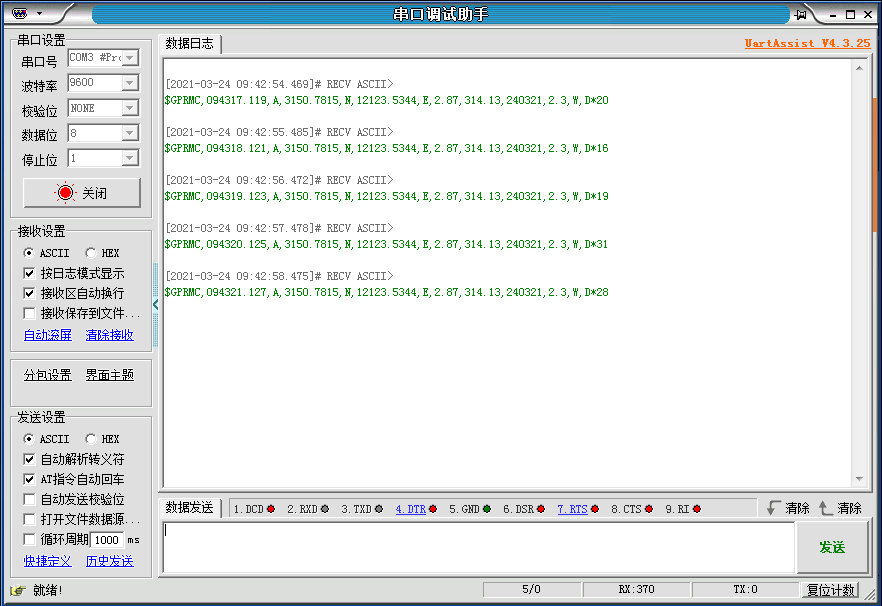


图3.3 串口助手发送GPRMC数据帧

3.3 视觉相机的时间同步

由于涉及视觉传感器较多，而且每一帧的图像数据量较大，无法采用串行软触发的方式进行串行图像采集。所以本文中使用多相机的硬件触发，以此来实现相机之间的时间同步。视觉相机之间的时间同步示意图如图3.1视觉相机部分。

相机的采集模式一般分为内触发和外触发两种模式，其中内触发主要靠相机内部给出的信号进行图像的采集，外触发主要由外部信息例如软件信号、硬件信号来触发图像的采集。本文中使用的相机外部触发中的硬件触发，一般是使用外部设备通过相机的I/O接口相连，通过触发脉冲信号给到相机，来触发采集图像，其内部原理是直接对相机内部寄存器进行的读写。本文使用的U3相机背面接口图以及相机I/O接口定义如下：

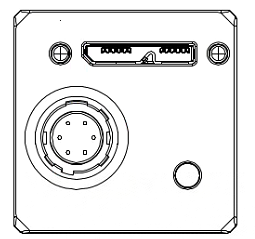


图3.4 U3相机背面接口图

接口中6pin的电源及I/O输入口对应的引脚定义如下表：

表3.1 管脚信号定义

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 管脚 | 信号 | I/O类型 | 说明 |
| 1 | 12V电源 | 输入 | 电源输入 |
| 2 | Opt-Iso In | 输入 | 光耦隔离输入 |
| 3 | GPIO | 输入/输出 | 可配置I/O口 |
| 4 | Opt-Iso Out | 输出 | 光耦隔离输出 |
| 5 | I/O Ground | 输入 | 触发-，信号地 |
| 6 | GND | 输入 | 电源地 |

具体的，硬件方面主要由彩色相机、鱼眼相机、热成像仪、多光谱相机均与Xavier工控机连接，并将各类相机的外部硬件触发引脚统一接到Xavier工控机上的IO口；软件方面需要配置相机采集帧率与激光雷达帧率相同，本文为20HZ，并且设置触发模式为上升沿硬触发，采集帧率20HZ目的是为了与激光雷达帧率相同，便于后续的相机与激光雷达之间的时间同步，另外需要在Xavier工控机上运行触发模块，实现IO口方波信号的输出，并利用上升沿来触发相机采集图像。如图3.5为虚拟逻辑分析仪查看的工控机通过IO输出的方波信号：

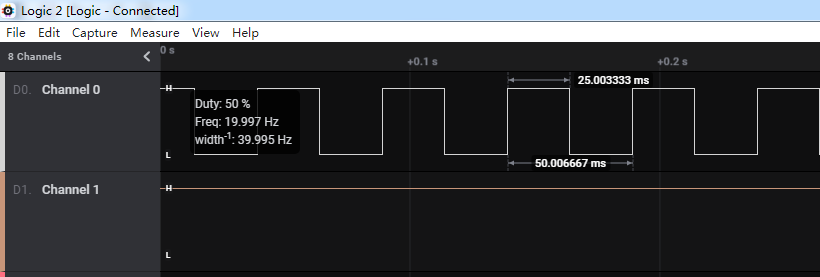


图3.5 工控机IO口输出的方波信号

针对Xavier工控机发出指令到触发信号产生这段时间，由于Xavier是ARM架构，ARM一条指令的执行时间是一个时钟周期，通常小于1纳秒，因此该部分时间误差可以忽略不计。对于相机之间的时间同步效果如下图：

|  |  |  |
| --- | --- | --- |
|  |  |  |
| 1. 彩色相机 | 1. 鱼眼相机 | 1. 多光谱相机 |

图3.6 彩色相机、鱼眼相机、多光谱相机时间同步效果图

由于手机秒表最高精度只到10ms级别，为了验证时间同步的精度，本文还使用了高精度的1ms级别的秒表进行相机外触发时的拍摄，以此说明此方法能够达到1ms以内的精度。其中彩色导航相机与鱼眼避障相机的高精度秒表拍摄图如下：

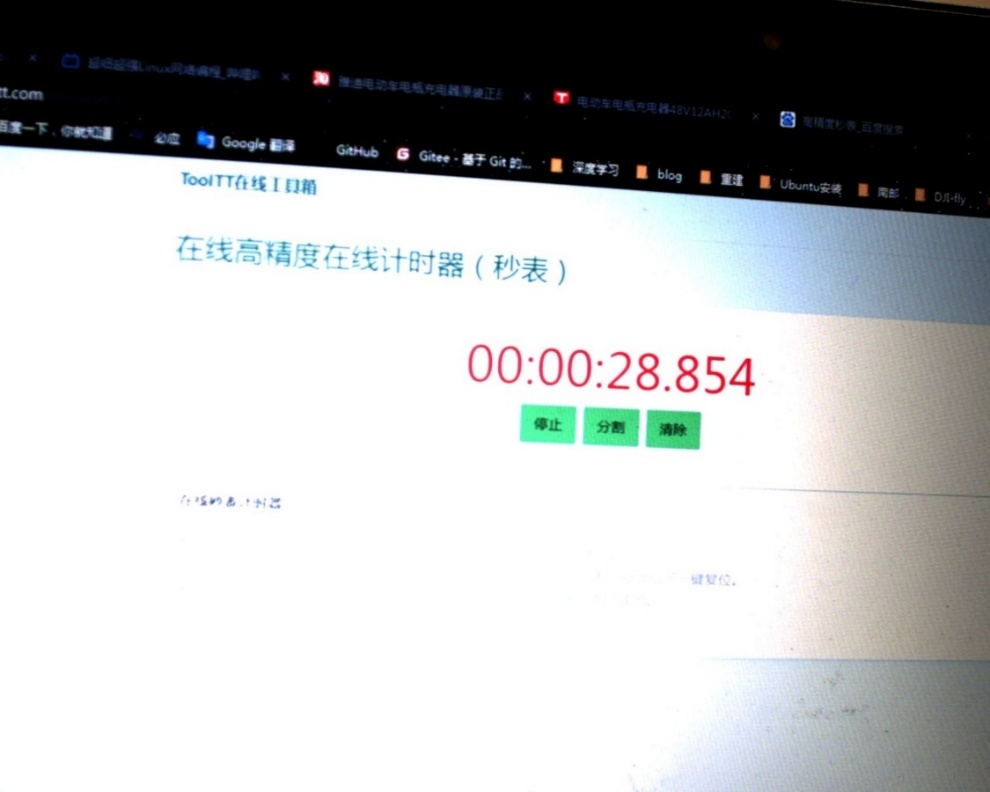
 

图3.7 彩色相机与鱼眼相机时间同步效果图

由上述触发模块实现了相机之间的外部硬触发，达到了相机之间的毫秒级时间同步。但是对于相同帧率的相机与激光雷达之间难免会存在一定的相位帧差，如图3.8所示。对于这种情况，本文通过测量t1、t2之间的时间差，进行相机外触发模块的信号就行对应时间的补偿，即可实现激光雷达与各类相机之间的时间同步。

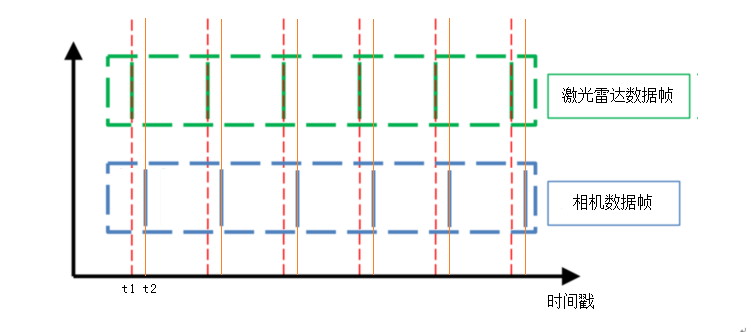


图3.8 雷达与相机时间戳相位帧差

3.4 惯性测量单元IMU的时间同步

有关IMU的时间同步，同样可以采用3.2小节的与激光雷达相同的授时同步，也可以采用3.3小节中与相机外部硬件触发的同步方式。但是以上两种方式，对于IMU来说，显得较为昂贵。主要原因是，常用的IMU大多数是MPU6050或者MPU9250为主的较为便宜的只有6轴数据输出的传感器，且这种IMU已经能满足正常需求，但是对于支持授时同步或者支持外部触发的IMU来说，价格将贵上好几倍甚至几十倍。所以本文中的时间同步方案，出于对成本的考虑，并且考虑到IMU的高频率采集数据的特点，采用的最直接的就近时间戳对齐的方式。

下面以激光雷达与IMU进行举例。如图3.9为雷达与IMU的数据帧示意图：

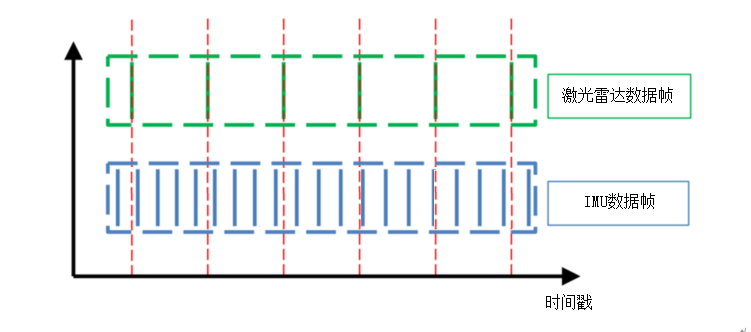


图3.9 激光雷达与IMU数据帧示意图

可以看到IMU的采集频率非常高，一般的IMU采集频率为200HZ，但是一般的激光雷达采集帧率只有20HZ甚至更低。所以，对于IMU的时间同步，采用就近时间戳对齐的方式无疑是最省时省力的。激光雷达与IMU数据帧之间就近时间戳对齐示意图如下：

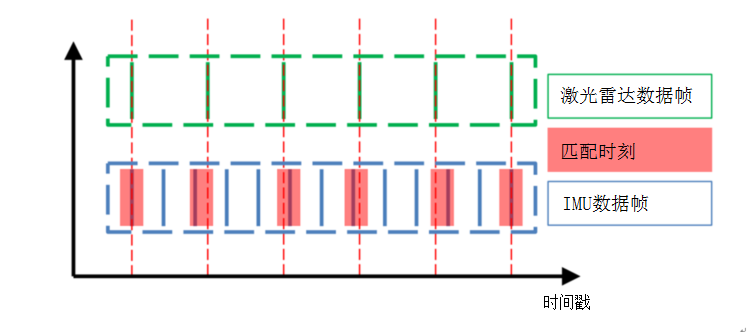


图3.10 激光雷达与IMU数据帧就近时间戳对齐示意图

下图为使用实验室传感器采集的激光雷达数据和IMU数据。数据帧示意图如下：

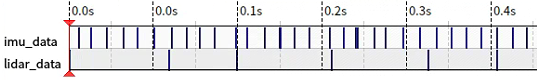


图3.11 雷达与IMU实际采集数据帧

通过就近时间戳对齐后的数据帧示意图如下：

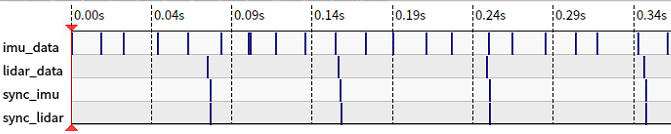


图3.12 雷达与IMU实现就近时间戳对齐

其中由上到下的4个topic分别为，时间戳对齐后的IMU数据帧、时间戳对齐后的雷达数据帧、时间戳对齐前的IMU数据帧、时间戳对齐前的雷达数据帧。

由图3.12可以看出时间戳对齐后的IMU数据和雷达数据已经达到了毫秒级的精度，已经满足本文中的对应的时间同步指标。但是由上图也可以看到，在经过就近时间戳对齐后的IMU与雷达数据之间还是会存在一定的时间相位差，所以，对于更高要求的时间同步系统而言，还可以在本文方案的基础上对高频的IMU数据进行插值，以此来提高时间同步的精度。对于本文来说毫秒级的精度已经满足对应指标，所以没有对IMU数据进行插值处理。

以上所述，便实现了IMU与激光雷达的时间同步。在3.2小节中实现了激光雷达与系统时钟的同步，在3.3小节中实现了激光雷达与相机之间的时间同步，在本小节中又实现了激光雷达与IMU之间的时间同步。至此，本文中的多种传感器已经完成了时间同步，且时间同步精度达到了毫秒级。

3.5 本章小结

本章主要围绕多传感器融合系统中的时间同步方法，基于各类传感器的软硬件特性，从三个方面实现了多种传感器的时间同步。其一，通过PPS信号+模拟GPRMC信号，使用Xavier工控机实现激光雷达的授时同步，实现激光雷达与系统UTC时间同步；其二，使用Xavier工控机发送相机外触发信号，实现各类相机之间的时间同步，并通过计算相机与激光雷达之间的数据帧的相位差来调整相机的触发信号，实现相机与激光雷达之间时间同步；其三，由于IMU采集频率很高的特性，使用就近时间戳对齐的方式实现IMU与激光雷达之间的时间同步。综上所述，本文中的所有传感器实现了时间同步，且时间同步精度达到了毫秒级。

# 多源传感器空间一致表征

在多源传感器融合系统中，除了前面第三章介绍过的多源传感器时间一致之外，还需要进行多源传感器的空间一致表征。传感器的空间一致主要是进行传感器之间的相对位置关系的确定，也即是传感器的外参标定。对于不同传感器之间的外参标定来说，还需要提前进行传感器的内参标定，也即是传感器自身参数的标定。对于多源传感器融合系统来说，传感器的内外参的标定精度也会直接影响到后面的融合感知结果。所以，传感器的内外参标定显得尤为重要。

4.1 空间一致表征系统总体设计

多个传感器的空间一致表征是进行数据的高效融合的基本需求，良好的一致表征是多传感器融合的基础，在平台上装了多个/多种传感器，而它们之间的坐标关系是需要确定的。本文主要用内外参标定来确定它们的固有参数以及它们之间的位置关系，完成多传感器空间一致性表征，总体框架示意图如图4.1所示。



图4.1 传感器标定总体框架示意图

该部分主要分为传感器的内参标定和外参标定。内参标定为相机内参、IMU内参的标定；外参标定为相机与相机之间的标定、相机与IMU之间的标定、相机与激光雷达之间的标定。

4.2 传感器内参标定

传感器的内参标定在本文中主要为相机的内参标定和IMU的内参标定。相机的内参标定需要标定出各个光学相机的固有参数，例如焦距、主点偏移、畸变系数等；IMU的内参标定需要标定出其静置状态下的零偏和白噪声。

4.2.1 视觉相机内参标定

视觉相机的内参标定，较为成熟的便是张正友标定法[63]。对于标定工具，可以使用OpenCV标定、使用Halcon标定、使用Matlab标定、使用Kalibr工具箱标定等。张正友标定法的具体步骤流程图如图4.2所示：



图4.2 张正友标定法步骤

其中步骤1中可以使用A4纸打印棋盘格贴在平面上，但是此方法标定精度不是特别高，当然也可购买高精度的标定板。特殊的，对于本文中的热成像仪[64]而言，需要特定的能发热的棋盘格标定板，在本文中是通过导热性较好的铝板和隔热材料进行自制的简易发热标定板，也达到了不错的标定精度。

其中步骤3中的提取棋盘格角点，具体算法步骤为：先进行图像的灰度化，其次进行角点检测，最后为了提高精度在角点位置进行亚像素角点检测。

对于普通光学相机和热成像仪拍摄标定板图像分别如图4.3和图4.4：

|  |  |  |
| --- | --- | --- |
| E:\Hyper\公司项目\张正友标定\Calib\1617700762.622108.png | E:\Hyper\公司项目\张正友标定\Calib\1617700770.376519.png | E:\Hyper\公司项目\张正友标定\Calib\1617700771.720387.png |
| E:\Hyper\公司项目\张正友标定\Calib\1617700782.146869.png | E:\Hyper\公司项目\张正友标定\Calib\1617700787.504745.png | E:\Hyper\公司项目\张正友标定\Calib\1617700789.643445.png |

图4.3 光学相机拍摄标定板图像

|  |  |  |
| --- | --- | --- |
| F:\My_Halcon_Pro\07\1.jpg | F:\My_Halcon_Pro\07\3.jpg | F:\My_Halcon_Pro\07\5.jpg |
| F:\My_Halcon_Pro\07\13.jpg | F:\My_Halcon_Pro\07\16.jpg | F:\My_Halcon_Pro\07\17.jpg |

图4.4 热成像仪拍摄自制标定板图像

下面是本文中用到的视觉传感器内参具体标定流程介绍：

关于相机内参标定，棋盘格的拍摄有一定的要求。首先是需要保证拍摄的图像数量大于4张，且少于20张较好；其次需要保证拍摄到的棋盘格所占图像空间的1/4至3/4最好；另外还需要保证多张图像的棋盘格覆盖过整个图像视野，且拍摄的棋盘格具有角度倾斜。

对于普通光学相机，可以直接进行图像的二值化，然后进行棋盘格角点提取并进行标定。但是对于热成像仪拍摄的棋盘格需要先进行图像预处理，即将棋盘格部分提取出来，在进行角点的提取和内参的标定。如下图所示即为热成像仪拍摄棋盘格的图像预处理结果：

|  |  |
| --- | --- |
| F:\My_Halcon_Pro\07\1.jpg | F:\My_Halcon_Pro\save\1.bmp |
| F:\My_Halcon_Pro\07\32.jpg | F:\My_Halcon_Pro\save\32.bmp |

图4.5 热成像仪拍摄棋盘格预处理结果

各类光学相机内参标定结果以及畸变系数分别如表4.1和4.2：

表4.1 相机内参标定结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 相机种类 | fx | fy | cx | cy |
| 彩色相机 | 610.9 | 610.2 | 325.9 | 228.5 |
| 鱼眼相机 | 537.3 | 537.3 | 543.4 | 500.8 |
| 深度相机 | 384.6 | 384.6 | 319.5 | 240.3 |
| 热成像仪 | 548.6 | 541.3 | 321.5 | 268.9 |
| 多光谱相机 | 1449.2 | 1448.6 | 638.9 | 490.3 |

表4.2 相机畸变系数

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 相机种类 | k1 | k2 | p1 | p2 | k3 |
| 彩色相机 | 0.137251 | -0.268334 | 0.001417 | -0.000494 | 0 |
| 鱼眼相机 | 6.8391e-02 | 8.93718e-04 | -2.8611e-04 | -8.0824e-06 | 0 |
| 深度相机 | 0.000287 | 8.327e-05 | -0.000458 | 0.000675 | 0 |
| 热成像仪 | -3.564e-01 | 1.660e-01 | -5.323e-03 | 2.431e-03 | 0 |
| 多光谱相机 | -0.102196 | 0.167307 | 0.000881 | 3.774e-05 | 0 |

4.2.2 惯性测量单元IMU内参标定

惯性测量单元IMU[65]是由陀螺仪和加速度计构成的，其中陀螺仪有3个轴，加速度计也有3个轴，最理想的情况下，陀螺仪和加速度计各自的3个轴是正交的。但是对于常见的一般IMU来说，这种情况是不存在的。

准确的IMU数据对目前较为普遍的多传感器融合VIO，例如VINS[66]、ORB3[67]等算法起到非常关键的作用，能使定位的精度更高，视觉里程计算法更鲁棒。所以在使用IMU之前，都需要对IMU进行内参标定，也即是标定出IMU自身的一些参数，也叫参数辨识。这些参数主要包括加速度计和陀螺仪各自的零偏、标度因数和安装误差。

下面是对于标定参数的分析：

（1）零偏：零偏就是IMU的输出比输入多了一个常量误差。

加速度计的零偏一般可以表示为：

陀螺仪的零偏一般可以表示为：

（2）标度因数误差：标定因数误差也称为刻度误差，也即标准输出和输入的比值不是1的时候，就需要标定。

加速度计的标度因数一般可以表示为：

陀螺仪的标度因数一般可以表示为：

对于常见的VIO、LIO算法中的IMU来说，需要标定的噪声一般分为两种：零偏（Bias）和白噪声（White Noise），它们都有连续时间模型和离散时间模型两种形式，两种方式也都可以进行相互转换。有的VIO算法需要的是连续型的噪声参数、有的需要的则是离散型的噪声参数。对于常用的标定工具，例如imu\_utils和kalibr\_allan，其中imu\_utils工具包输出的是连续型的，kalibr\_allan输出的是离散型的。

由于后续需进行IMU与相机的联合标定，所以此处需要介绍一下Kalibr工具所需的IMU噪声参数的形式。Kalibr标定工具需要的噪声参数形式属于连续时间模型，具体参数如下表：

表4.3 IMU噪声参数格式

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | YAML文件元素 | 符号 | 单位 |
| 陀螺仪“White Noise” | gyroscope\_noise\_density |  |  |
| 加速度计“White Noise” | accelerometer\_noise\_density |  |  |
| 陀螺仪“random walk” | gyroscope\_random\_walk |  |  |
| 加速度计“random walk” | accelerometer\_random\_walk |  |  |
| IMU采集频率 | update\_rate |  | Hz |

关于IMU内参的标定较为简单，本文中使用的是港科大开源的imu\_utils工具包进行的内参标定。标定步骤操作简答，只需要录制IMU在静置状态下的2个小时以上的数据即可进行标定。使用imu\_utils工具包进行IMU内参标定的结果如下图所示：

表4.4 IMU内参标定结果

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 标定参数 | | 单位 | x-axis | y-axis | z-axis | avg-axis |
| 陀螺仪 | Noise | rad/s | 1.396347e-04 | 3.199011e-05 | 1.806569e-05 | 6.323017e-05 |
| Walk | 1.023083e-06 | 2.587342e-07 | 1.308692e-07 | 4.708955e-07 |
| 加速度计 | Noise | m/s^2 | 4.627721e-03 | 5.061922e-03 | 6.699715e-03 | 5.463119e-03 |
| Walk | 1.820176e-04 | 2.204612e-04 | 3.008543e-04 | 2.344443e-04 |

4.3 多传感器外参联合标定

上述4.2小节主要介绍了视觉相机的内参标定和惯性测量单元IMU的内参标定，其目的是为了标定出对应传感器的自身参数、噪声等。本小节主要在前面内参标定完成的基础上进行的传感器之间的外参标定，也即是求解出传感器之间的相对位置关系，以便后续的融合感知等系统的使用。

4.3.1 视觉相机与视觉相机之间联合标定

视觉相机之间的外参标定，主要应用在三维重建、双目立体视觉匹配等系统中。本文中的多类视觉相机，包括彩色相机、灰度相机、热成像仪和多光谱相机之间的标定都可以看做是双目相机的标定，也即是分别进行彩色相机-灰度相机之间外参的标定、彩色相机-热成像仪之间外参的标定、彩色相机-多光谱相机之间外参的标定。

双目相机标定[68]的主要目的是确定两个相机之间的相对位置关系，以便用于后续的三维重建，双目测距等系统。但是，一般两个不同的相机或是双目相机的成像平面是不平行的，所以需要进行双目相机标定来将光轴不平行的双目相机模型图转换成图2.4所示的，两个相机光轴平行的双目视觉模型。

下面将介绍如图4.6所示的双目相机标定的算法原理：



图4.6 双目标定算法原理

对于双目相机的标定，主要是需要求解出两个相机之间的相对位置关系，即相对旋转和相对平移。假设为所拍摄的标定板上某一点在空间中的坐标；、分别为左右两个相机成像的世界坐标系下的坐标；、为点相对左目相机的旋转和平移；、为点相对右目相机的旋转和平移。所以，有如下公式：

将4.5式进行化简可得：

对4.6式进行两式相减，消去可得：

对4.7式两侧同乘并进行化简可得：

对于4.8式中，令、，则有：

所以，上式就是右目相机相对于左目的空间变换，R为旋转矩阵，T为平移向量。又因为旋转矩阵是正交矩阵，所以。所以对于4.9式中的R、T可以表示为：

对于上式中的、、、这几个参数在进行单目相机标定的时候即可获得。所以，带入式，通过拍摄的许多组双目标定板图像，即可利用最小二乘法，也可进行奇异值分解的方法，进行最小化误差，最终可得到最佳的旋转矩阵。

下面是本文中用到的视觉传感器外参具体标定流程介绍：

关于相机外参标定，棋盘格的拍摄有一定的要求。采集数据为棋盘格图像数据，采集图像数据要保证棋盘格角点的清晰，最好保持相机的采集帧率不要过高，防止图像模糊，不易于提取角点。采集图像总数量最好为20张左右，并且需要所有数据中的棋盘格尽量覆盖图像所有区域，且拍摄的棋盘格具有角度倾斜。其中双目相机外参标定靶标位置示意图如下：

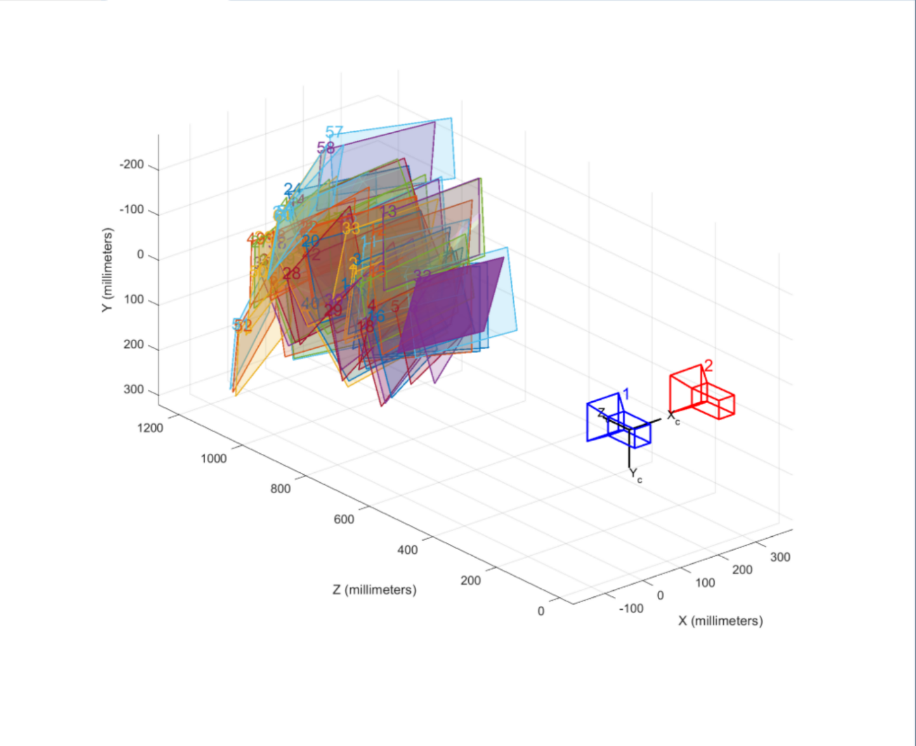


图4.7相机外参标定靶标位置示意图

最终进行两两相机之间外参标定，彩色相机与灰度相机、分别有彩色相机与深度相机、彩色相机与热成像仪、彩色相机与多光谱相机。最终将标定结果如下所示：

表4.5 相机之间外参标定结果

|  |  |  |
| --- | --- | --- |
| 外参标定 | 旋转矩阵R | 平移向量T |
| T\_color\_gray |  |  |
| T\_color\_depth |  |  |
| T\_color\_heat |  |  |
| T\_color\_multi |  |  |

为了验证双目相机标定算法的有效性及精度，本文使用标定好的双目相机的内外参，拍摄一帧棋盘格图像，并将其恢复到3D空间中，计算其中两个棋盘格角点与真实棋盘格大小的误差。本文中实验所用棋盘格大小为25mm，经过双目恢复两3D点之间距离计算，证明了本文标定算法的有效性与精度。

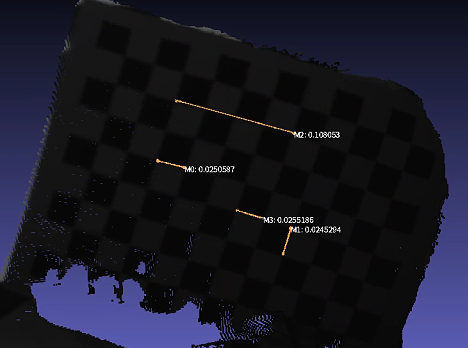


图4.8双目测距标定精度验证

4.3.2 视觉相机与激光雷达之间联合标定

激光雷达与相机之间的联合标定[69]在常见的多传感器融合感知系统中应用非常多，所以使用一个精度较高的标定方法是取得高精度结果不必可少的一步。常见的激光雷达与相机之间联合标定方法有手动法和自动法，手动法即为人工手动的选取点云和图像中的对应的特征点，然后进行求解计算；自动法则为算法自身进行点云和图像特征点的提取，但这种方法精度一般，也有进行线面特征等进行关联，然后3D配准进而增加标定精度的方案。

常用的标定工具有Autoware和Apollo工具箱[70]。其中Autoware是一个由日本开源的一个自驾相关的平台工具，其包含了相机的标定，相机与雷达的联合标定，所需的数据较为简单，但是需要在标定过程中由人工来选取激光点云中的3D特征点与图像中的2D特征点进行匹配，进而求解出外参；Apollo是百度公司推出的阿波罗自动驾驶平台，其也包含相机与雷达之间的联合标定，但是其标定对于数据的需求较高，不仅需要普通的数据，有时还需要IMU以及里程计等传感器的相关数据，并且其还需要一个粗略的外参才能进行外参的求解与优化。所以最常见的还是使用Autoware来进行相机与激光雷达之间的外参标定。

下面将介绍如图4.9所示的lidar-cam联合标定的算法原理：

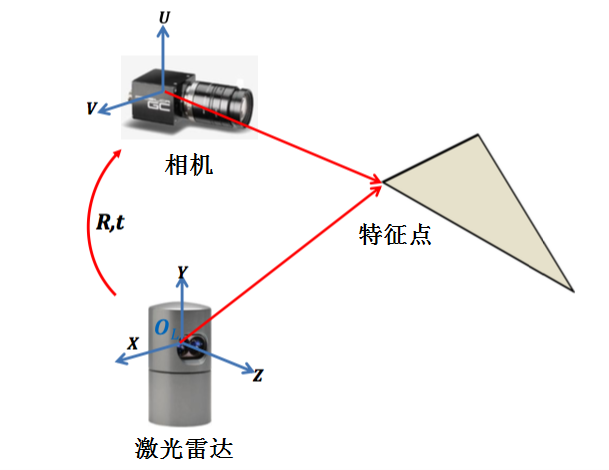


图4.9 lidar-cam联合标定示意图

在如图4.9所示的lidar-cam联合标定示意图中，相机坐标系为-U-V，激光雷达坐标系为-X-Y-Z，他们之间相差一个旋转变换对应的旋转矩阵R和平移变换对应的平移向量t，这就是需要标定的量。假设某一特征点的三维坐标为,此点映射到二维像素平面点为，则由映射关系可得如下式子：

其中、、、均为相机参数，也即是相机的内参，、分别为相机XY方向的焦距，、是像平面的主点坐标。其中R为旋转矩阵、t为平移向量。其中的M为变换矩阵，具体可表示为：

对4.11进行矩阵展开可得：

将4.13式进行变量提取，再转换为矩阵乘积形式可得：

对于4.14式，可以通过不同姿态下标定板平面，或者通过激光雷达的三维点云特征点与图像中对应的二维图像特征点，可以得到一系列的线性方程，进而可求解出M矩阵，也即是所求的外参标定参数。

下面是本文中用到的视觉传感器与激光雷达之间外参标定流程介绍：

本文中，使用的视觉与激光雷达之间外参标定的算法为Autoware工具箱中的标定算法，其原理同前面所介绍。但需要手动的进行激光雷达中3d特征点点的选取，以及图像中与其对应的2d特征点的选取。标定场景选取如下图：

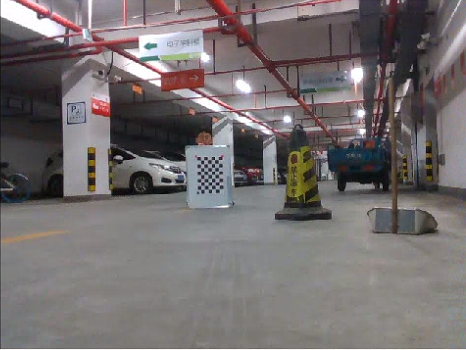
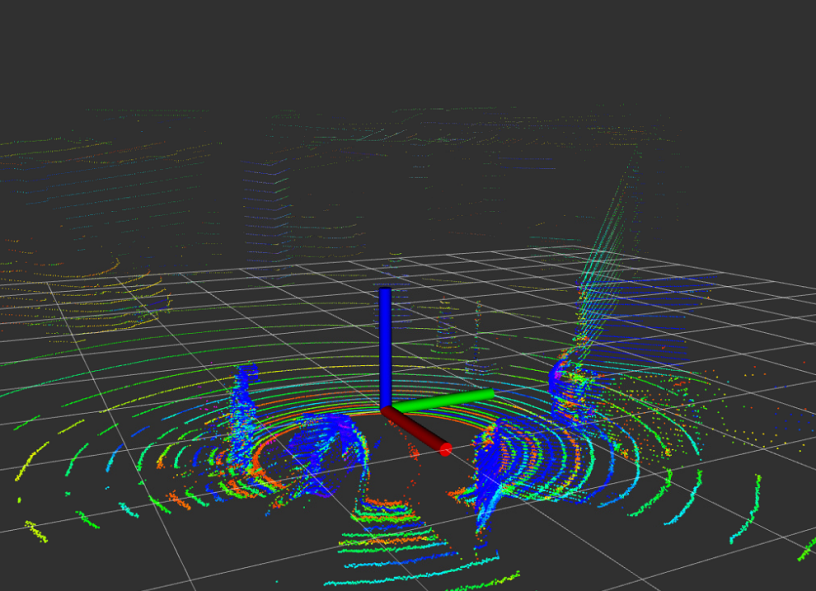
 

图4.10 彩色相机原始图像与激光雷达原始点云

使用Autoware进行相机与激光雷达的外参标定，需要首先在rviz界面中进行手动选择3d点，然后在图像中选取对应的2d像素点，共要9对数据即可进行标定。结果如下图所示：

表4.6 相机与IMU标定结果

|  |  |  |
| --- | --- | --- |
| 外参标定 | 旋转矩阵R | 平移向量T |
| T\_lidar\_cam |  |  |

相机与激光雷达标定后可以进行点云到图像的投影可视化，如下图所示即为滤除噪点后的激光雷达点云投影到彩色图像可视化结果：

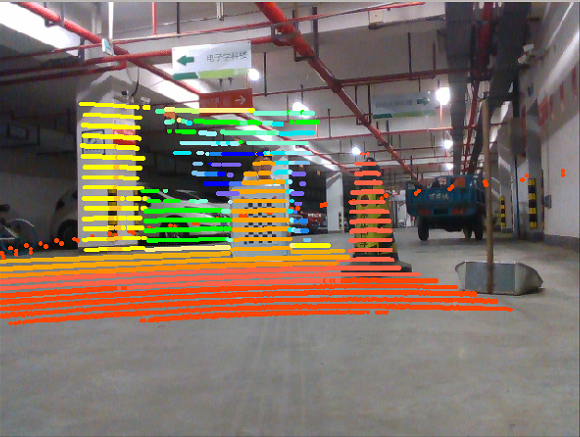


图4.11 激光点云到彩色图像的投影可视化

4.3.3 视觉相机与IMU之间联合标定

在现有的多传感器融合SLAM方案中，视觉与惯导的组合一直十分被认可，主要也因其成本低、精度高等优点。对于这两种传感器的组合，之所以有比较好的融合效果，这是因为他们之间可以做到信息的互补：例如相机在快速运动时会产生模糊，在光照环境不好的条件下视觉算法会失效，但有了IMU的加入，由于IMU的高频数据采集的特性，且不受光照等环境的影响，可以大大的弥补视觉相机的不足。相机与IMU配合如图4.12，可以获得相机较为丰富的图像数据，又能通过视觉回环检测来有效的修正IMU长时间工作带来的累积误差。

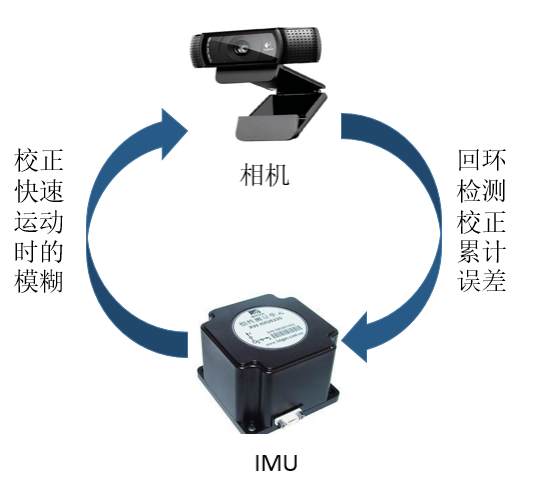


图4.12 相机与IMU融合示意图

使用Kalibr进行相机-IMU之间的外参标定，一般分为三个步骤：

（1）粗略进行相机与IMU之间的时延估计：

由于传感器的传输延时和触发延时等情况的存在，导致如下图所示的传感器的时间戳存在时间差，即：

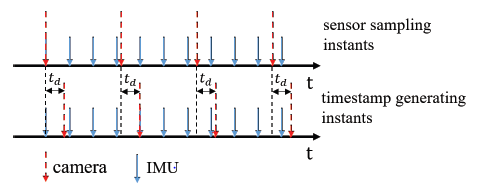


图4.13 相机与IMU之间时间差

在前面的4.2.1小节，以及进行过相机内参的标定。现在已知了图像每一帧的3d与2d之间的对应，所以就可以计算出相机每一帧的离散位姿。通过这些离散的位姿就可以进行连续的B-spline构造，进而获得任意时刻的相机位姿。

这样就可以利用相机的样条曲线获得任意时刻下相机对应的旋转角速度，同时对于IMU来说，其陀螺仪又可以测量出角速度。所以先忽略噪声的影响，相机与IMU之间只相差一个旋转。所以利用相机与IMU数据的两个曲线的相关性，可以粗略估计相机与IMU之间的时延，并且利用此相关性，可以将时延的误差缩小至2个IMU数据周期内。

（2）获得cam-imu之间的初始旋转，以及重力加速度、陀螺仪偏置等初值：

同样的，还可以利用角速度测量的关系，再进行一个优化问题的构造：

通过上述问题的优化，就可以获得相机与IMU之间的旋转，并且也能够获得陀螺仪偏置初始值。

假设在整个标定的过程中，IMU的加速度计平均加速度为零，并且在忽略加速度计的偏置与噪声的同事，可以获得重力加速度在参考坐标系下的表示如下：

（3）进行大优化，具体包括相机的重投影线误差和IMU的测量误差与噪声：

该步骤的大优化，主要是调整所有的待优化的量，使得其观测误差最小。前面两步也为该步骤提供了一个不错的初始值。待优化的误差项主要包括相机拍摄标定板的重投影误差、IMU自身的测量误差和噪声。

下面是本文中用到的视觉传感器与IMU之间外参具体标定流程介绍：

本文中使用的cam-imu外参标定的工具为kalibr工具箱。首先需要进行bag包的录制。其中需要相机一直能拍摄到棋盘格，以便优化时进行重投影误差；并且相机的拍摄频率不能太快；最重要的是需要给IMU足够激励，即使得IMU在xyz各个方向都有移动和旋转。

对于kalibr工具箱的输入需要有上述录制的bag包以及相机与IMU各自的内参文件。通过kalibr工具箱进行cam-imu之间的外参标定结果如下：

表4.7 相机与IMU标定结果

|  |  |  |
| --- | --- | --- |
| 外参标定 | 旋转矩阵R | 平移向量T |
| T\_cam\_imu |  |  |

4.4 本章小结

本章主要围绕多源传感器融合系统中的空间一致表征，从两个方面实现了多源传感器的空间同步。分别为传感器各自的内参标定和传感器之间的外参标定。其一，对于本文中的传感器各自的内参标定，主要包括所有视觉相机的内参以及IMU的内参；其二，对于本文中的传感器之间的外参标定，主要包括视觉相机之间的外参标定，相机与激光雷达之间的外参标定，相机与IMU之间的外参标定。通过测量传感器的安装位置与标定结果的对比，实现了多源传感器的空间一致。

# 基于多源信息的多层地图构建

三维重建发展至今，相关技术取得了巨大的成就。不管是在游戏开发、遗迹保存、医学三维CT等领域，其都有着及其广泛的应用。在进十几年的发展过程中，不论是基于单目相机的三维重建，还是基于RGB-D相机的三维重建算法层出不穷，并且基于深度学习的许多端到端三维重建网络在这几年发展的也越来越好。各种算法都有其自身的优势和缺点，例如基于深度学习的端到端的三维重建网络虽然发展火热，但是此类算法需要相机的已知位姿，无法做到直接由一系列图像到三维点云的重建，并且针对火星此类场景，由于火星车的慢速移动且无回环，很难通过传统的SLAM算法给出精确的位姿。所以，基于航天院火星场的研究背景，本文基于多视图环境（Multi-View Enviroment，MVE）三维重建算法，并且进行了热和光谱信息在三维地图的叠加，实现多层地图的构建，以便后续的感知和路径规划等工作。

5.1 重建算法设计方案

本文基于航天院火星场的场景进行三维地图重建，并在三维地图上加入热、光谱信息，形成一个高维的多层地图。主要使用的传感器有RGB相机、热成像仪和多光谱成像仪。具体算法框图如图5.1所示，首先，利用RGB图像序列通过SFM[72]进行相机位姿的计算；然后，利用已知的相机参数通过MVS[73]进行各个视角下的深度图恢复；最后通过FSSR[74]算法进行物体表面的恢复；在完成了上述的重建算法之后，利用已知的RGB相机与热成像仪、多光谱之间的外参，可以进行热和光谱信息到三维彩色地图的映射，最终形成多层地图。



图5.1 多源多层地图三维重建算法框图

5.2 基于MVE算法的多层地图重建

本文中的多源多层地图重建，主要基于航天院火星场的场景进行实验，传感器主要包括RGB相机、热成像仪和多光谱相机。本文采用MVE算法[71]，实现一个端到端的三维地图重建，最后再将热和光谱信息向地图中映射，生成多层地图。MVE算法，即多视图环境，它的输入是场景的图像，最终输出的是三角形网格，该算法主要由从运动恢复结构、多视图立体重建、密集点云生成和曲面重建构成。具体算法流程示意图如图5.2所示。

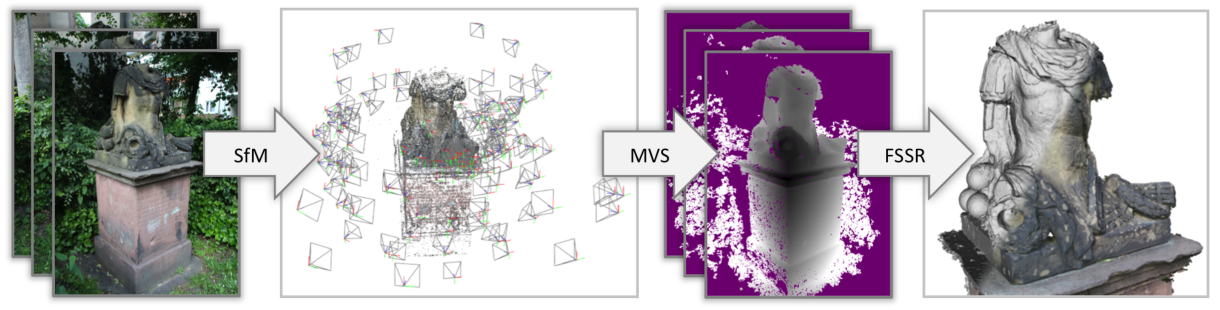


图5.2 多视图重建MVE算法流程示意图

5.2.1 运动恢复结构SFM

从运动恢复结构的SFM算法可以说是计算机视觉和三维重建中非常高的成就之一。本质上来说，SFM是从非结构化的一组图像中的稀疏对应关系重建相机的参数和稀疏点云。其恢复的相机参数主要包括相机位姿。对于常见的增量式SFM算法可以分为如下几步：

（1）特征点提取与匹配：

特征点提取：对输入图像进行特征提取需要考虑特征对于旋转、噪声、光照以及图像尺寸等易变的信息的不变性。常用是具有尺度不变性和旋转不变性的SIFT描述子，该描述子对于上述提到的一些易变信息具有较强的鲁棒性，非常适合用来提取具有尺度变换或是带有旋转的图像特征点信息。

特征点匹配：通过上述特征点提取，可以获得两幅图像的SIFT特征点描述子，又由于两幅图像之间的对应点收到相机模型之间的极线约束，所以通过这些约束可以去除掉许多的错误匹配。对于两两匹配的图像，常以欧氏距离进行匹配成功与否的判断，如果距离小于某一阈值则认为匹配成功，但是此种匹配方式依然会存在较多的误匹配，还需要进行误匹配的剔除。针对误匹配点的剔除常见方法有两种，第一，剔除掉最近距离与次近距离的比值较大的；第二，对匹配点采用RANSAC算法计算其对应的基础矩阵F，若不满足基础矩阵，也需要进行剔除。

（2）相机位姿估计，并重建三维坐标点；

相机相对位姿求解：在进行了第一步的特征点匹配后，即可获得两张图像中的多对2D-2D匹配特征点。可以通过8点法估计本质矩阵E。有了本质矩阵E，即可将其进行SVD分解出两个相机之间的旋转矩阵R和平移向量t，但是SVD分解会得到四组可能的解，需要通过检查旋转矩阵R的正确性，并找到能尽可能把2D点投影到相机坐标系下的3D点的R和t，对应的解则为最可能的解。

重建三维坐标点：在已知了两个相机之间的变换关系以及成对的匹配点坐标之后，便可以通过三角测量来恢复出匹配点在三维空间中的坐标。最后再将这些恢复出的三维空间点向两个视角进行重投影，计算与原来2D坐标的距离，以此来说明三角化误差。

计算第三个视角到世界坐标系的变换矩阵：由前面的两个视角的特征点匹配计算出的视角二相对于视角一，也即是世界坐标系的变换矩阵。在加入第三个视角后，使其与第二帧进行特征点匹配，在这些匹配点中，有一部分是第一帧与第二帧已经匹配成功且已恢复三维坐标的点，所以这种情况下就可以根据PnP算法进行求解第三个相机到世界坐标系的变换矩阵。对于加入的更多图像帧，同上方法即可求得新的一帧可恢复出的三维点云，与全局点云进行融合即可。具体流程如图5.3所示。

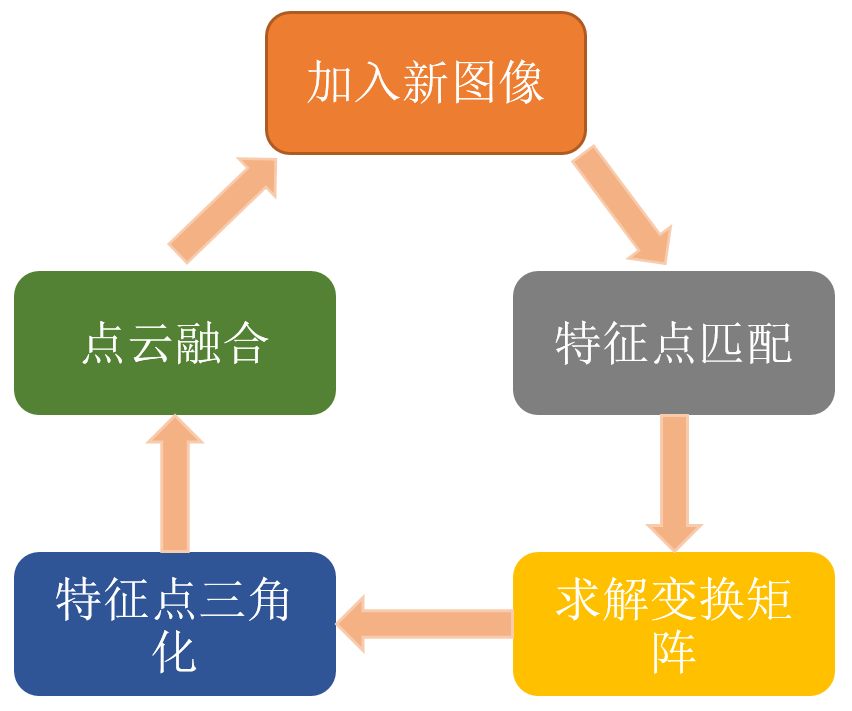


图5.3 SFM恢复三维点云流程图

（3）进行BA优化

在随着上述不断的新图像加入到算法中的时候，累积误差会逐渐增大，可能会导致最终的重建误差较大甚至是完全偏离重建目标。为了重建过程中相机位姿以及重建的三维点云的准确，需要进行点云和相机位姿的优化。常用的方法为光束平差法（Bundle Adjustment），其本质上是一个非线性的优化算法，对于BA的求解，Ceres Sover库具有很高的效率。

5.2.2 多视图立体视觉MVS

对于前面所介绍的SFM算法，其无法得到密集点云的原因主要是因为SFM算法中是有特征点匹配所得到的三维点，这些特征点天生就是稀疏的。所以还需要使用MVS进行稠密点云重建。整体思路是，利用输入的多帧图像，选取一帧为参考帧，通过这些图像恢复参考帧图像的深度图和法向量信息。据图流程如下图所示：



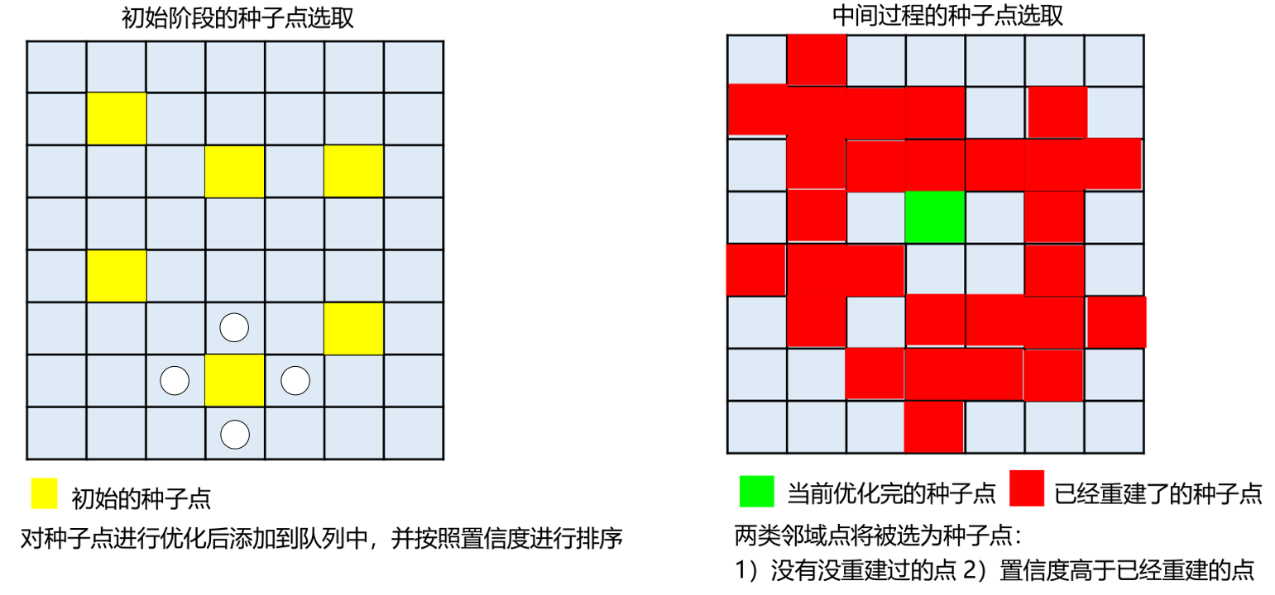
图5.4 MVS算法整体流程

（1）预处理部分：

1. 稀疏3D点云重建常用前面介绍的SFM进行特征点提取和匹配，并通过匹配点恢复相机参数以及生成稀疏的特征点点云。
2. 全局视角图像序列的选取：一般通过一些特定的指标，从输入的一系列图像中选出一帧作为参考图像。

（2）深度图生成：

1. 种子点选取：从已经重建好的3D点云中，筛选出能投影到至少一帧的全局视角图像的点。
2. 种子点深度与法向量恢复：重建好的3D点云中的点，向参考图像上投影，该点到参考图像坐标系原点的距离就位该点的对应像素的初始深度，且该点与参考图像坐标系原点所形成的向量即为初始法向量。然后再进行深度和法向量的优化即可。
3. 其余像素点深度恢复：经过对上述的种子点计算并优化后结果的置信度，将置信度从高到低进行队列的创建和优化，并且再将这些点周围的一些点也加入到队列中。如图5.5中(a)图所示，将黄色格子的种子点深度与法向量作为与其接壤的四个像素点的初始深度和法向量并进行优化。如图5.5中(b)图所示，优化完的种子点需要再与周围已经优化过的点置信度进行大小比较，如果当前点置信度较高，则把当前点的信息赋值给比较点。



(a)初始阶段种子点选取 (b)中间过程的种子点选取

图5.5 种子点更新图

（3）三维稠密点云：

该部分较为简单，前面已经得到了每一帧图像的深度图，接下来就可以根据相机位姿与相机内参将各个点投影到三维空间，最终获得稠密的3D点云。

5.2.3 表面重建FSSR

经过前面的算法，已经可以得到一个较为稠密的3D点云结构，但是其点与点之间还是会有较大的空间，无法很好的表达出物体的表面信息。所以在得到稠密点云后，还需要进行物体的表面重建操作来使得模型更加真实。表面重建需要构建出一种网格模型，也即是一种网格化的三维拓扑结构，由空间中的三个点组成一个面片的方式进行三个点的链接。其中MVE算法中所使用的就是基于浮动尺度的表面重建FSSR算法。该算法能够使用几乎无参数的方法从采集的样本数据有效地重建高质量的网格，即使对于大型数据或是有噪声的数据集也是如此。

从概念上说，FSSR方法基于从输入样本重构隐式函数F。F具有空间连续尺度（浮动尺度），即，由F表示的表面细节的尺度如输入样本的尺度所定义的那样连续变化。然后，定义了F的离散、尺度自适应采样，并提取与F的零级集合相对应的等值面。隐函数F被构造为紧支撑基函数的和。但与例如径向基函数[75]，或光滑有符号距离重建不同，FSSR的方法不需要解决全局问题，在计算上易于处理，并且在给定样本的情况下，可以容易地评估隐式函数。这种紧凑的支持导致了一种方法，即重建开放网格，并在数据过于稀疏的区域留下空洞，无法进行可靠的重建。这对于无法完全捕获的场景（例如室外场景）非常有用。

现在该方法已被封装为可直接调用的算法，其输入需要是一个ply类型的点集文件，具体的ply文件内点集的必须要有的属性如下表：

表5.1 FSSR算法输入点集属性要求

|  |  |  |  |
| --- | --- | --- | --- |
| 点云属性 | PLY 属性 | 类型 | 是否需要? |
| 3D坐标 | x,y,z | float 3-vector | 需要 |
| 法向量 | nx, ny, nz | float 3-vector | 需要 |
| 尺度值 | value | single float value | 需要 |
| 置信度 | confidence | single float value | 可选 |
| 颜色 | red, green, blue | uchar 3-vector | 可选 |

该FSSR算法工具以ply文件作为输入，在八叉树层次结构上对饮食函数进行采样，并提取等值面，最后再将其存储在ply网格中。对于生成的网格也会存在一些置信度较低的顶点、小的孤立组件以及来自等值面的退化三角形，这些也都将会被清理掉。

5.2.4 多层地图映射

由上述MVE算法，可以得到一个基于RGB图像序列的三维地图，也即是该地图中的3D点仅包含x、y、z坐标，nx、ny、nz法向量以及r、g、b颜色属性。但是基于航天探索背景研究，需要更多的地图信息，以便于适应不同情况下的环境重建以及环境感知。所以，本章使用时空同步过的RGB相机、热成像仪以及多光谱成像仪的数据构建多源多层的地图。具体的，为了解决部分RGB信息的缺失造成的感知问题，以便于火星车可通过性的分析，本文使用映射的方式进行多层地图的构建，如图5.6所示。

由于需要进行多层地图的构建，一个3D点云需要包含多层信息，也即是除了要包含上述提到的xyz与rgb信息之外，还需要包含热成像仪的热红外信息以及多光谱成像仪的光谱信息。所以需要从新定义既包含pcl::PointXYZRGB又包含Point\_irInfo、Point\_multiInfo的数据结构。

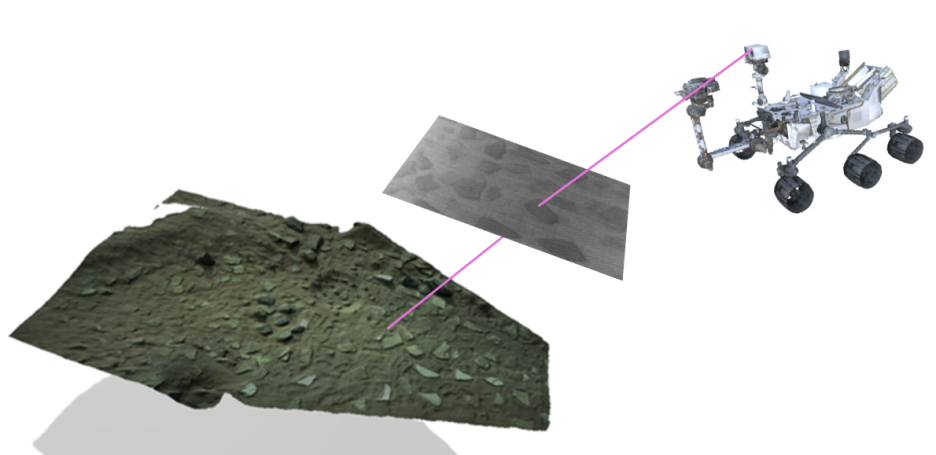


图5.6 多层地图构建示意图

对于已经时空同步过的传感器，在已有三维地图的情况下，将热红外信息与多光谱信息映射到点云地图中就比较简单了，通过如下图所示的坐标空间变换即可进行信息的映射：



图5.7 坐标系转换流程图

此处设世界坐标下的一点坐标为，相机坐标为，像素坐标为(u,v)。则上述5.7图对应的坐标系间矩阵转换关系如下：

其中为相机内参，为相机外参。

5.3 实验结果与分析

经过前面5.2小节的算法介绍，本小节将基于航天院火星场以及火星车上搭载的传感器进行多层地图的重建。最终形成一个包含颜色、热、光谱等多维信息的多源多层点云地图。为后续多维语义地图的划分、环境的感知以及可通过性的分析提供良好的地图。

|  |  |  |
| --- | --- | --- |
|  | ··· |  |

图5.8 多视角重建原始图像序列

如上图5.8所示图像序列为航天院火星车真实采集的室内模拟火星场的数据，该数据为火星车向前缓慢移动时采集。通过原始RGB图像序列进行基于MVE算法的三维重建，得到的火星场三维地图如图5.9所示。其中图5.10所示内容分别为图5.9中红色以及绿色框内的重建细节展示，由此可以看出重建效果非常不错。

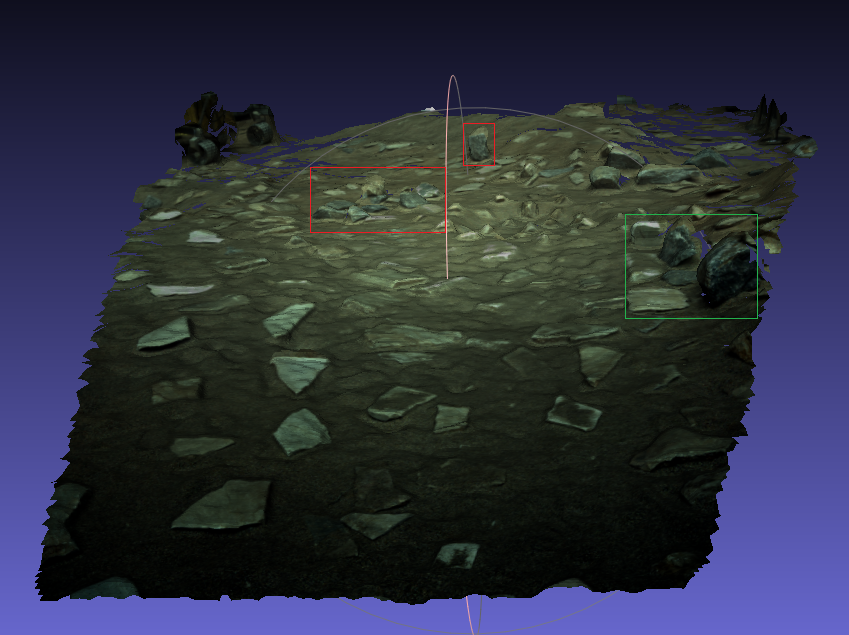


图5.9 彩色地图重建结果

|  |  |
| --- | --- |
|  |  |
| （a）红框内重建细节 | （b）绿框内重建细节 |

图5.10 三维地图重建细节

对于时空同步的传感器数据，使用与构建三维彩色地图同帧的热红外图像以及多光谱图像，并依次向彩色地图上映射，形成多层地图。以第一帧图像为例，如下图所示：

|  |  |  |
| --- | --- | --- |
|  |  |  |
| (a) RGB图像 | (b)热红外图像 | (c)光谱图像(非可见光谱段) |

图5.11 对应帧传感器数据

由于各个传感器分辨率不同、焦距不同以及所用镜头不同，所以，对应的各自视野大小也不一样。其中RGB相机的视野最大，热红外的视野次之，光谱图像的视野最小。对于多光谱图像，由于其有多个非可见光谱段的图像，所以本文仅以一个谱段进行实验与验证。所以在彩色地图之上映射多帧热红外信息后的彩色与热红外的多层地图如图5.12所示。由于标准的点云可视化都是三维的XYZRGB类型，所以对于本文中自定义的多层地图数据的可视化，只能是将对应位置点云颜色更改来进行可视化。



图5.12 彩色与热红外组成的多层地图

由上图可以看出热成像仪的视野范围相比RGB图像小很多，在彩色点云的基础上，只有中间一部分覆盖了热红外信息。对应的热红外映射细节如下图5.13所示。

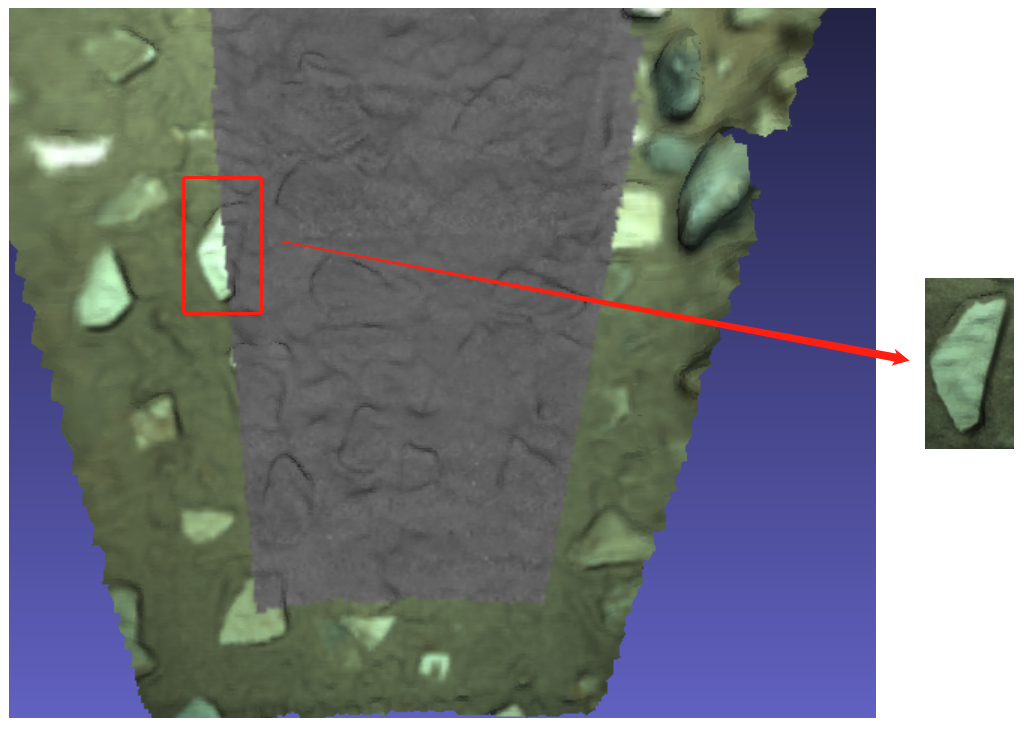


图5.13 热红外映射结果，右图为红框内石块的原色

由于多光谱相机所拍摄的图像均为非可见光谱段的数据，所以无法进行细节的可视化观察。但在本文的实验中，光谱图像的映射方式同热红外图像映射原理相同，所以，此处仅可视化一帧数据来表示光谱图像的映射。

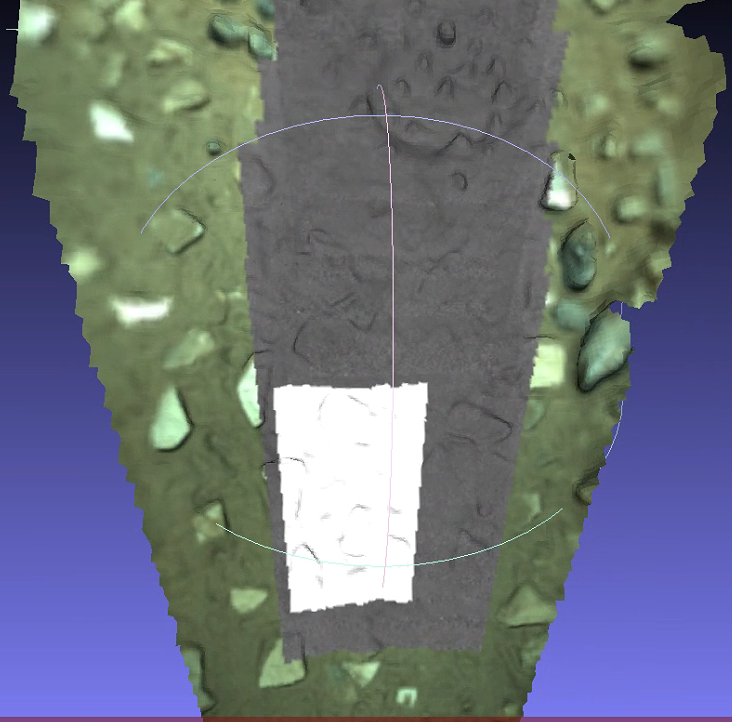


图5.14 多源多层地图可视化结果

5.4 本章小结

本章主要基于航天院火星场，以及火星车所搭载的RGB相机、热红外仪和多光谱相机，通过MVE算法首先实现了三维RGB地图的重建，其次进行了新的点云数据结构的构建并利用第三章与第四章所实现的时空同步后的传感器数据分别将热红外图像与光谱图像映射到三维RGB地图上，最终生成多源多层点云地图。该多源多层地图，可通过不同层的信息进行相关环境感知，并且还可以进行地图的可通过性的分析与路径规划。

# 基于双目视觉的三维环境感知

在三维环境感知领域，常见的算法大多是基于激光雷达进行的端到端的3D点云目标检测。现有的3D点云检测算法和分割算法大都停留在对公开数据集或者车、人、建筑物等大物体的检测。但是对于一些小物体，例如果园的树枝、水果、石头等，以及类似于航天探索领域的砂石等物体便难以鲁棒的进行检测和分割。初次之外，现有的算法大都依赖激光雷达的点云数据，这也就导致了高昂的成本，并且对于3D点云的网络训练数据的标注费时费力。因此，本章采用了成本较低的双目视觉系统来进行三维环境的感知和特定目标的空间定位。

6.1 感知系统算法设计方案

本文主要研究背景为航天探索相关的火星场场景，使用成本较低的双目视觉系统进行目标的三维空间定位。具体算法分为两步，如图6.1所示。首先，通过双目相机获取的图像，使用BGNet立体匹配网络[76]进行训练，并进行视差图估计，再将视差图通过相机内参转换成深度图进而恢复出3D点云。其次，使用SOLO分割网络[77]实现对特定目标2D位置信息的获取，从而估计目标物体的在左目相机中的空间位置信息，即物体在左目相机中的三维空间坐标。因为无人车在行进过程中触碰到物体即会发生危险，目标物体的空间位置信息定义为：距离左目相机光心最近的物体表面三维点的空间坐标。



图6.1 双目感知系统算法

6.2 基于双目视觉的目标空间定位

本文中的基于双目视觉的三维环境感知系统主要以火星场、仿真软件中的模拟场景进行实验。通过双目立体匹配网络进行左目相机下的深度图恢复，并使用左目图像进行对应物体的分割，最终在多个3D物体点云中再进行特定物体的剔除与筛选。

6.2.1 基于SOLO网络的实例分割算法

实例分割算法在这些年发展的越来越好，但是同时也存在了很大的挑战性。实例分割相较于语义分割更加细致，一个语义类别包含的实例个数可能是多个。在SOLO算法之前，现有的方法基本要么遵循Mask R-CNN[78]一类的算法先检测后分割测量，要么先进行类别的预测，再进行聚类、关系亲合场等方式来进行实例的划分。SOLO算法的效果和优势包括：端到端，无需任何的后处理；在COCO数据集上达到了Mask R-CNN的效果；不需要bbox标注信息，只需要mask的标注信息。SOLO与Mask R-CNN算法相比，不同之处如图6.2所示：

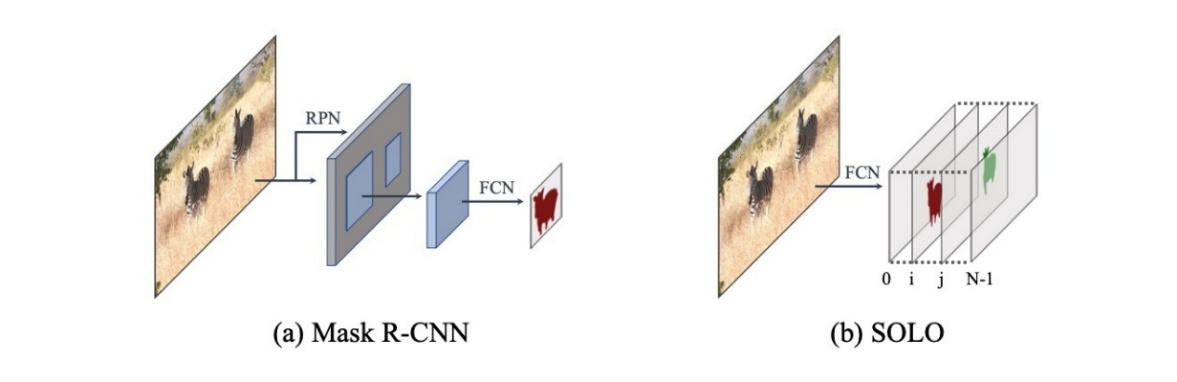


图6.2 SOLO与Mask R-CNN算法对比图

SOLO实例分割算法的目的是确定图像中是否有对应的实例，有的话就返回实例的mask。该算法的核心思路是将实例分割的问题转化成类别预测和实例 mask 生成的两个问题，具体的：首先将图像分成S×S个格子，然后分别通过分类分支和mask分支，如果对应目标的中心在某个格子里，则对应格子需要输出实例类别和分割实例。具体算法如图6.3所示：

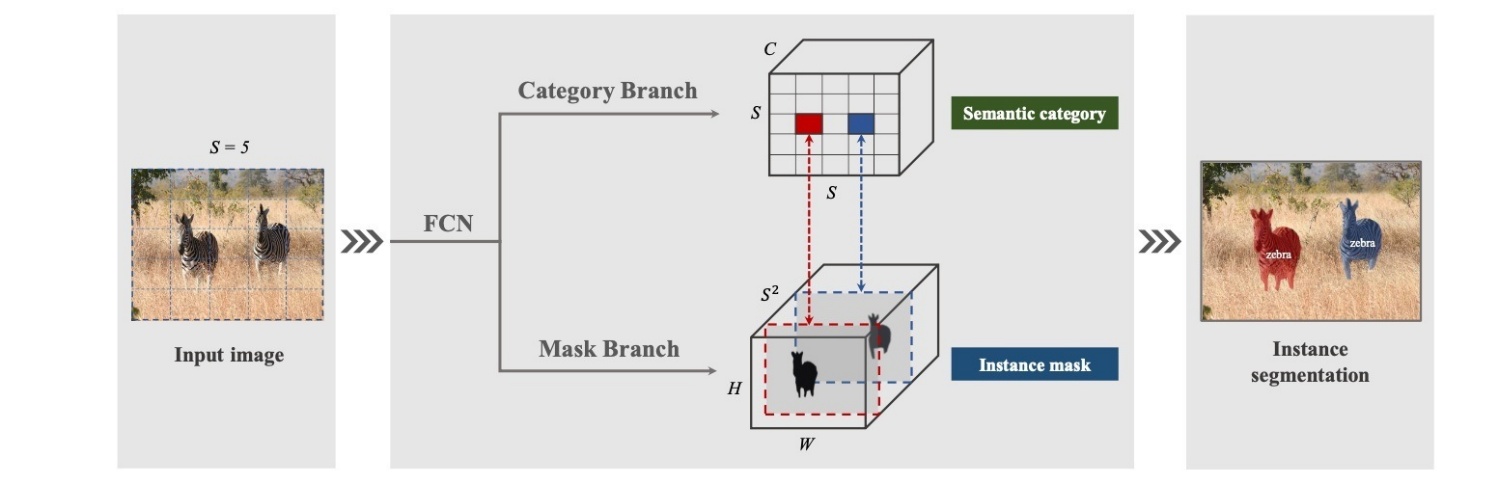


图6.3 SOLO算法网络框图

实例类别：如图6.3所示中的类别分支，对于输入图像中的每一个grid网格，该算法都会进行C维的预测输出，这里的C为类别个数，每一维代表一个类别的置信度得分。所以类别分支的输入是S×S个格子的数据，输出则为S×S×C的多维数据，这样每个格子就只会属于一个实例。

实例掩码：如图6.3所示中的mask分支，对于输入图像I中的S×S个格子，最多会进行次预测，将这些mask以一个3D的tensor进行输出，也即是有3个通道。则mask分支的输出维度为：。并且类别分支和mask分支之间也具有一对一的对应关系，例如某一个cell位置为(i,j)，那么对应的mask分支中的第i\*S+j个mask。

网络架构：该算法的backbone使用FPN，FPN可以在每一个通道上产生大小不同的特征图（通道数通常为256的倍数），FPN输出的特征图作为类别分支head和mask分支head的输入。同一个head的参数都是共享的，但是最后的1×1卷积参数的没有共享的。

关于SOLO网络的学习：网络对于每一个格子都需要给定预测的目标的概率，如果某一格子内落入了任一的gt mask的中心，就会被分为正例。对于每一个gt mask的正样本，都会设定一个二值分割mask，由于mask分支的输出是，所以此处共有个grid，对每个正样本，其对应的binary mask都会被标记。其中，训练过程中的损失函数公式定义如下：

上式中的是类别分支的分类损失，是用于预测mask的损失函数：

如果网格对应的类别索引是从左到右、从上到下的，则上式中i=|k/S|，j=k%S，则为正样本的个数，以及分别对应的类别和对应的mask。代表指数函数，如其内部式子成立则为1，否则为0。

关于SOLO网络的预测过程：（1）图像经过FPN的backbone网络以及分类分支和mask分支，得到(i,j)位置的类别得分以及对应位置的mask；（2）使用特定的阈值来进行类别得分的过滤；（3）对前得分前500对应的mask进行NMS操作；（4）最后将目标类别的mask进行二值mask输出。

6.2.2 基于BGNet网络的双目立体匹配算法

三维重建和三维感知中立体匹配是十分关键的步骤。给定双目图像数据，立体匹配的目的是找到左目图像和右目图像的像素之间建立密集的对应关系。现有的传统的双目立体匹配方法针对火星表面地形无纹理区域和反射表面弱的缺陷，很难找到准确的对应点，并且计算效率很低。所以，本文采用比较新的基于深度学习的立体匹配网络BGNet。通过双目相机和TOF相机制作的火星场双目数据集以及现有公开数据集，利用端到端的BGNet网络进行训练，计算出视差图，再将视差图利用相机内参进行转换成深度图，以便进行后续的空间定位。现有方法针对火星表面地形无纹理或纹理差区域和反射表面弱的缺陷，很难找到准确的对应点。双目图像通过BGNet网络转化为深度图的流程如下图：



图6.4 基于BGNet的双目立体匹配

实时的、高精度的立体匹配算法是目前三维感知和三维重建中研究的热点，在例如机器人，自动驾驶，AR等领域都有很广泛的应用。虽然相关的立体匹配网络的发展也在逐步的发展，但是在速度和精度上依然存在比较大的发展空间。

BGNet算法使用了一种基于可学习的双边网格的代价空间上采样模块（Cost volume Upsampling in the learned Bilateral Grid, CUBG），该模块结构如图6.5所示。基于这个具有边缘保持特性的上采样模块，再通过没有参数的切片层（slicing layer），网络就可以从低分辨率的代价空间获得高质量的高分辨率代价空间。使用这种方式，可以将代价聚合在低分辨率中执行，大大的降低了时间消耗。该算法中的此模块也可以应用到许多现有的其他立体匹配网络当中（如GCNet[79]，GANet[80]等），并且其在保持不错的精度条件下依然可以取得最高将近30倍的加速，这是可微双边网格首次在立体匹配网络中的应用。此外BGNet网络在一些公开数据集，例如KITTI数据集[81]下能够达到30-40fps，并且精度非常高，超过了现有的大部分实时立体匹配网络。

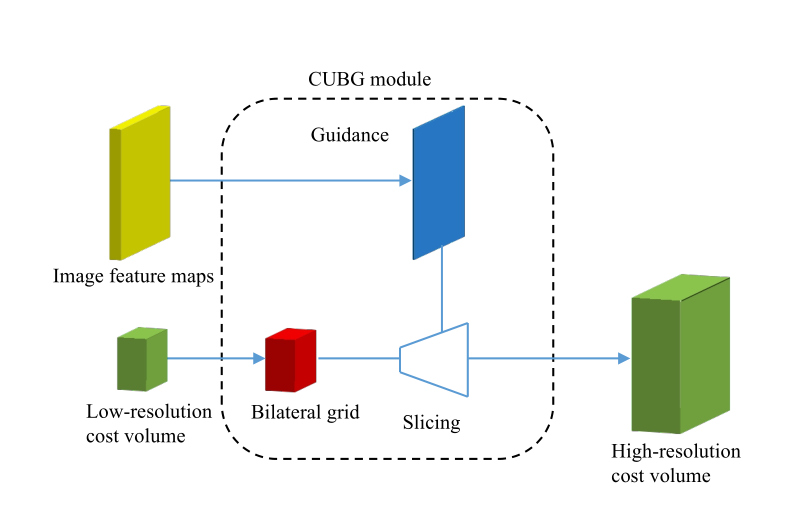


图6.5 CUBG模块结构

BGNet网络就是基于上述的CUBG模块实现的一个高效的端到端的立体匹配网络，其在KITTI数据集下可以实时运行。该网络主要有特征提取、成本聚合、CUBG和视差优化几个模块组成。具体的BGNet网络结构如下图：

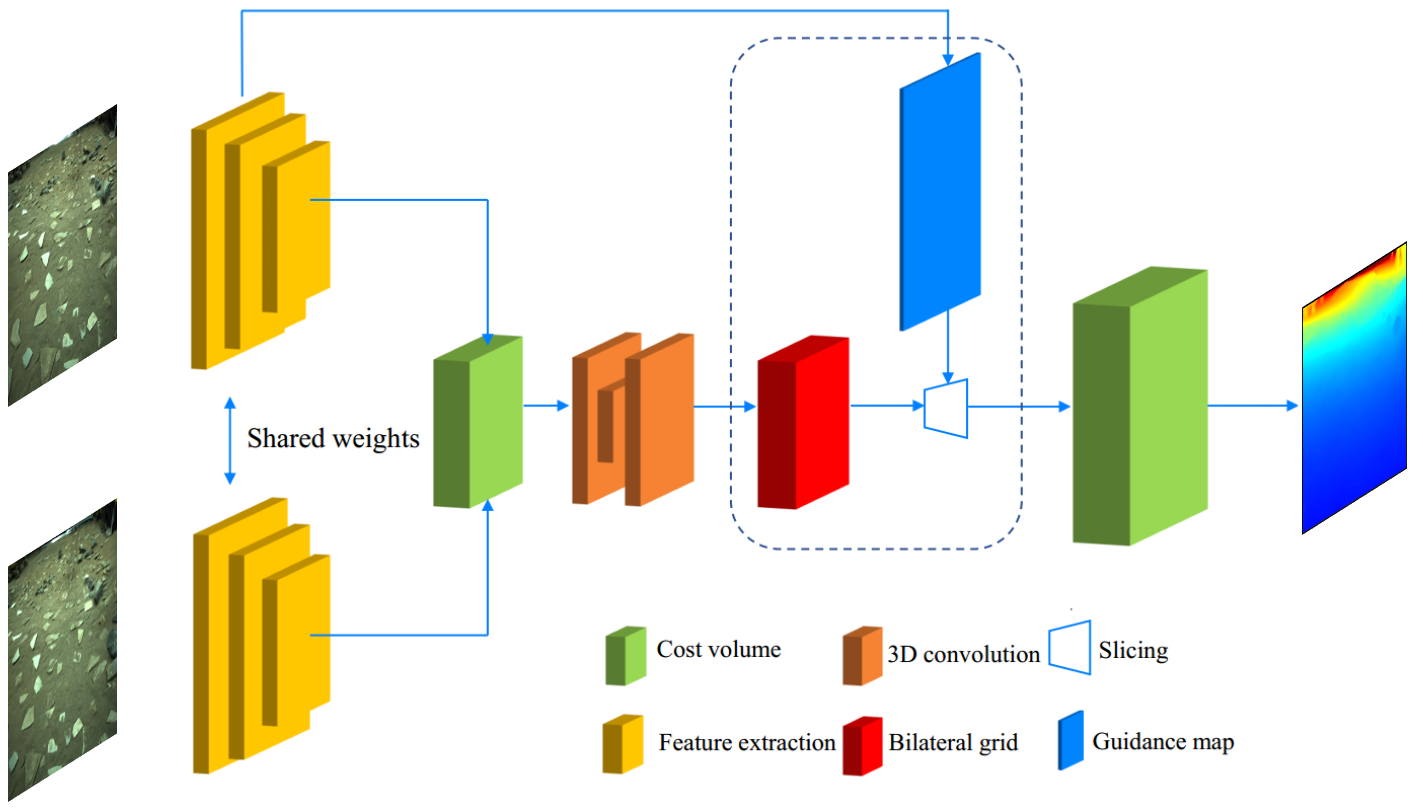


图6.6 BGNet网络结构

对于算法中的特征提取模块：采用类ResNet的架构来进行图像特征的匹配。首先，对于开始的三层，分别使用的步长为2、1、1且kernel的大小的3×3的卷积核进行图像的下采样。其次，使用步长分别为1、2、2、1的连续的四个残差层，并且使用1/8的分辨率生成一元特征。其中遵循两个沙漏特征的网络用以扩大感受野。最后，把所有的1/8分辨率的特征图都进行连接，最终生成一个352通道的特征图，并用来生成成本量。

对于算法中的成本聚合模块：在经过了上述的特征提取后，就为成本聚合构建了一个分组相关成本量，该成本量结合了相关联以及串联量的有点。并且，考虑到效率，算法中只使用了一个沙漏的架构来进行成本量的过滤。具体来说，首先使用两个3D卷积进行代价量通道数的降低，再使用类U-Net结构的3D卷积网络进行代价聚合以降低计算成本。

对于算法中的视差回归的定义公式如下：

对于算法中使用的损失函数的公式定义如下：

其中：

经过上述端到端的BGNet深度学习网络，可以得到左目相机坐标系下的视差图。通过相机内参可以将视差图转换为左目相机坐标系下的深度图，进而可以恢复出3D点云来进行相关的环境感知以及目标空间定位。视差图转化为深度图的公式较为简单，同2.3小节中介绍的双目立体视觉模型中的式（2.14）。

为了验证本文中的方案在三维空间中定位的准确性，本文还采用了TOF相机作为真值来进行对比分析，即将TOF相机获得的深度图转换到左目相机坐标系，并与双目生成的左目相机坐标系下的深度图进行定位误差分析。设TOF相机内参为，其坐标系下的一点坐标为[x,y,z]，那么该点转换到左目相机坐标系的转换，在已知左目相机内参矩阵为K的情况下，可以根据如下公式获得左目相机下对应的深度图：

其中u、v对应左目相机坐标系下的x、y坐标，d对应深度。

6.3 实验结果与分析

经过上述小节的算法介绍，为了验证本文中的双目视觉目标空间定位系统的效果，对许多场景进行了对应的分割算法与立体匹配算法，并进行了空间中特定目标的三维空间定位，并以TOF的数据为真值进行了定位的精度分析。实验场景主要包括航天院火星场，以及仿真软件内的模拟火星环境。

6.3.1 实例分割算法实验结果

本小节主要进行分割算法的实验，针对的场景包括航天院火星场、仿真软件里的模拟环境以及NASA的部分航天数据。具体的软硬件配置表如下：

表6.1 分割算法软硬件配置表

|  |  |
| --- | --- |
| 类别 | 配置 |
| 系统 | Ubuntu18.04 |
| CPU | Intel i7-12700 |
| 显卡 | Nvidia 2080ti \* 2 |
| 搭载相机 | 工业相机，分辨率2048\*2048 |
| 实验场景 | 航天院火星场、仿真场景、NASA航天数据 |

对于航天院火星场一期、二期的场地进行分割检测，主要检测的对象包括较大的岩石（rock）以及较为扁平的基岩（bedrock）。

|  |  |
| --- | --- |
|  |  |
| （a）效果图1 | （b）效果图2 |
|  |  |
| （c）效果图3 | （d）效果图4 |

图6.7 航天院火星场一期场地分割结果

对于航天院火星场一期的场地，场景主要为平整的沙地，以及少部分的岩石和基岩组成，作为最初的算法实验场地。但是为了更加模拟出火星的场景和后续的三维重建等工作，进行了场地的更换。二期的场景包括岩石、基岩、细沙、沙坑、斜坡、土堆、山脊等多种环境。

图 6.8 航天院火星场二期场地分割结果

另外还在仿真软件中进行了场景的搭建，进行了更多环境的设置，包括有沙坑（bunker），矿石（ore），岩石（rock）和突起（protuberance）。

|  |  |
| --- | --- |
|  |  |
| （a）效果图1 | （b）效果图2 |
|  |  |
| （c）效果图3 | （d）效果图4 |

图6.9 仿真环境地形分割结果

针对NASA的数据集，由于数据量较小，所以分割结果可能存在一定的误差。其中rock为岩石，bedrock为基岩，soil为土壤，shadow为阴影。

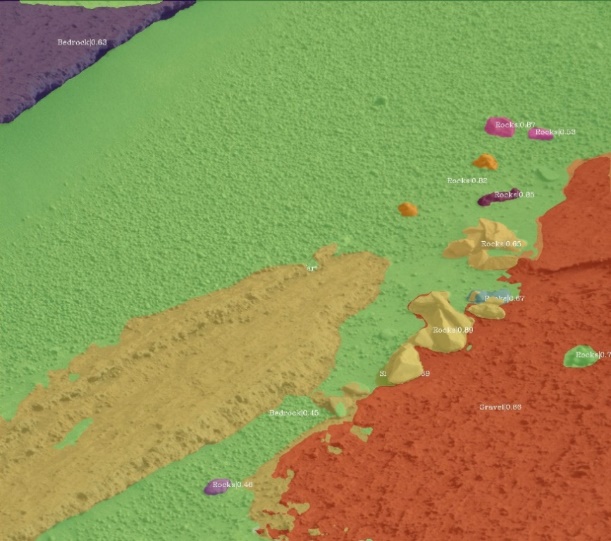
 

图6.10 NASA数据分割结果

通过上述不同场景的实验，训练数据只用了几百张图像，可以看出该算法的分割效果是非常好的。这也为后续的利用双目恢复出的三维信息进行空间定位提供了良好的条件。

6.3.2 双目空间定位算法实验结果与分析

本小节主要进行双目立体匹配与物体的空间定位算法的实验，针对的场景主要为航天院火星场的数据。具体的软硬件配置表如下：

表6.2 双目定位算法软硬件配置表

|  |  |
| --- | --- |
| 类别 | 配置 |
| 系统 | Ubuntu18.04 |
| CPU | Intel i7-12700 |
| 显卡 | Nvidia 2080ti \* 2 |
| 搭载相机 | 工业相机\*2，分辨率2048\*2048；TOF相机，分辨率640\*480 |
| 实验场景 | 航天院火星场 |

基于航天院火星场的场景，利用6.2.2中介绍的BGNet双目立体匹配网络进行左目相机坐标系的视差图恢复，再进行视差图到深度图的转换。下图所展示的即为网络输出并转化得到的深度图：

|  |  |  |
| --- | --- | --- |
| C:/Users/绿油油如你/AppData/Local/Temp/picturecompress_20211123193124/output_38.pngoutput_38 | C:/Users/绿油油如你/AppData/Local/Temp/picturecompress_20211123193124/output_39.pngoutput_39 | C:/Users/绿油油如你/AppData/Local/Temp/picturecompress_20211123193124/output_40.pngoutput_40 |
| （a）左目相机图像 | （b）左目视差图 | （c）左目深度图 |

图6.11 立体匹配网络输出及转化结果

通过上述算法生成的深度图，结合左目相机内参即可生成3D场景点云，再结合分割算法检测出的特定目标，进行3D点云与2D目标的映射，便可以计算出特定目标的空间坐标。例如在火星车行进过程中导致无法移动或者会对火星车造成损害的地形和物体便属于本文中需要进行空间定位的目标，例如岩石、陨石坑等。

要保证火星车安全的行进，则需要周围障碍物的三维位姿信息，而在图片中的二维位姿不带深度信息，无法从2D图片中获得物体的绝对尺度和位置，没有办法有效避免碰撞，具有一定的安全隐患。因此，场景的3D感知在火星车的移动，路径规划等方面都发挥着极其重要的作用，对于火星表面的目标探索有着重要的意义。所以本文使用分割+双目立体匹配的方案进行火星车周围高风险目标的定位，整体框架图为：



图6.12 双目视觉空间定位算法流程框图

通过获取物体的空间坐标，计算物体的方向包围盒，获取物体的质心以及高度进行高风险目标定位。为检测算法可行性，分割出高风险物体进行空间定位效果如下图：

|  |  |  |
| --- | --- | --- |
| C:/Users/绿油油如你/AppData/Local/Temp/picturecompress_20211123193124/output_42.pngoutput_42 | C:/Users/绿油油如你/AppData/Local/Temp/picturecompress_20211123193124/output_105.jpgoutput_105 |  |
| (a) 原图 | (b) 分割结果 | (c) 高风险目标 |

图6.13 高风险目标定位结果

本文基于航天院定义的高风险指标进行的实验，对应的指标在图6.13中对应的是岩石高度超过15cm则被认为是高风险。上述图6.13中图(c)的两块被框起来的岩石被定义为高风险。由于距离越远，误差带来的影响越大，所以本文中只考虑5m以内的物体。对于图(c)两个石头通过算法获得的距离与高度分别为rock1:距离4.18m、高度18.82cm；rock2：距离3.62m、高度25.59cm。

为了验证双目空间定位的精度，本文还以TOF相机的原始点云下的对应的高风险目标的深度作为真值，算法生成的深度图作为预测值，进行对应的误差分析。定位误差计算示意如下图所示：

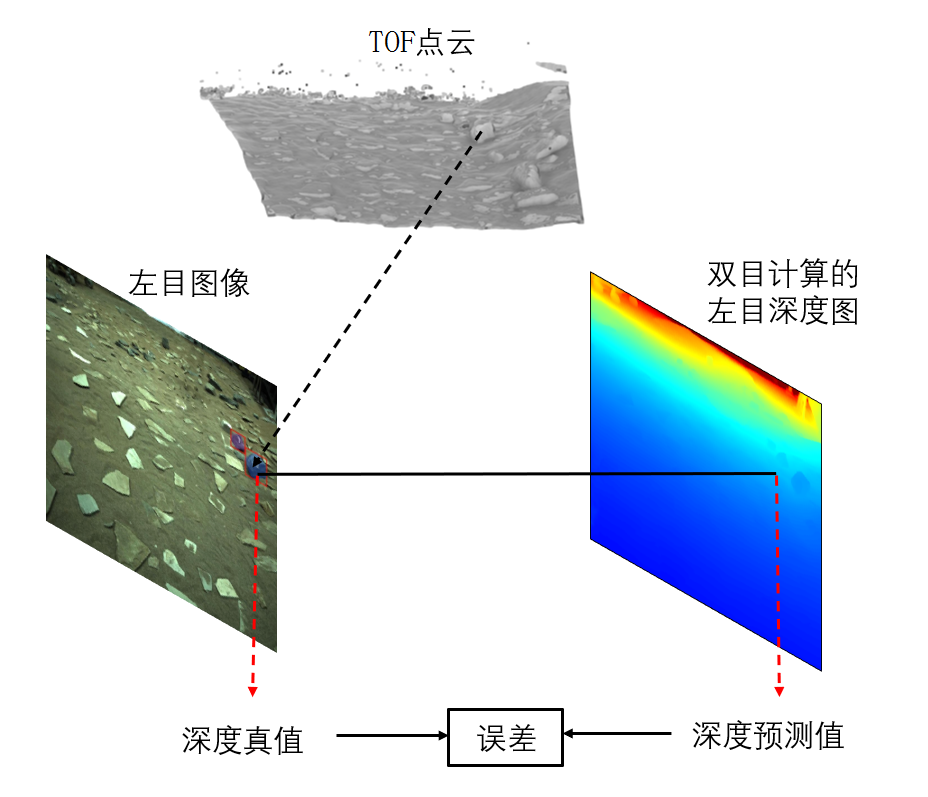


图6.14 定位误差计算示意图

在高风险目标区的分割区域内采样部分像素点，它们对应的真实距离为 预测距离, 然后计算这些像素点的平均真实距离和平均预测距离。最后采用以下公式计算定位误差：

利用上述定位误差计算方法，对图6.13时刻的图像以及高风险目标进行误差分析结果如下表所示：

表6.3 高风险目标定位误差分析

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 物体编号 | 高度(cm) | 距离预测值(m) | 距离真实值(m) | 定位误差 |
| 1 | 18.82 | 4.181 | 4.183 | 0.048% |
| 2 | 25.59 | 3.618 | 3.620 | 0.055% |

6.4 本章小结

本章主要基于航天院火星场的场景以及双目视觉系统进行三维空间环境的感知。与常见的端到端的激光点云检测算法不同的是，本文通过2D分割+3D映射，实现基于两步走的方式完成对三维环境中特定目标的定位。具体的，首先在左目图像上进行二维的图像分割，然后利用双目立体匹配网络生成的深度图，将左目图像分割得到的二维目标映射到三维空间中，并进行一定条件的高风险目标的判断，最终以两步走的方式实现三维空间环境的感知。此种方式相较于端到端的点云检测、分割网络的优势在于，训练时不需要大规模的进行3D标注，并且可以实现小物体的三维感知。本文使用的方案识别精度较高，并且对于航天院火星场数据进行目标定位误差为毫米级。

# 总结与展望

7.1 本文工作总结

针对目前多传感器融合、三维重建以及三维环境感知等相关技术所存在的一些问题。本文基于航天探索背景的二代火星车以及火星场地形，结合火星车搭载的传感器种类多、火星场地形复杂等特点进行研究，主要工作可以概括为以下几点：

（1）多源传感器时空一致表征：

针对现有多传感器融合算法存在的融合传感器种类较少、时空同步精度较低等问题，本文提出了一种用于多源传感器融合的时空同步方法。该方法包括机器人、智能驾驶汽车等相关领域中大于5种的常见传感器的时间同步与空间同步。本文从各类传感器的软件与硬件方面同时着手，首先进行多种传感器的时间同步，因为各个传感器的初始采集数据会存在一定的时间差，所以要将各类传感器与系统时钟保持同步，用以保证能够获取到同一时刻的传感器数据；其次进行多种传感器的空间同步，该部分主要进行各种传感器的内参标定与外参标定，用以确定传感器自身参数以及传感器之间的空间位置关系。高精度的时空同步结果也是多源传感器融合感知与三维重建的保障。

（2）基于多源信息的多层地图构建：

针对现有地图重建算法在夜晚等恶劣环境下对复杂的未知地形重建感知能力差的问题，本文提出了一种基于多源传感器信息的多层地图构建方法。该方法包括三维地图重建和多层地图构建两个步骤。首先，基于多视图的MVE算法进行三维地图的重建；然后，通过多传感器空间同步所得到的热成像仪、多光谱相机与RGB相机的内参以及相对位置关系，进行多层地图的构建；最终构建出一个不仅包含颜色，还包含热和光谱信息的多源多层点云地图，为路径规划和环境感知等任务提供更加准确和全面的地图信息。

（3）基于双目视觉的三维环境感知：

针对现有的三维环境感知算法中存在的对小物体的3D感知不鲁棒和数据依赖等问题，本文提出了一种基于双目立体视觉的2D感知+3D映射的三维环境感知方法。首先，在双目相机的左目图像上进行基于SOLO算法的实例分割，实现在2D上的物体感知；其次，利用基于深度学习的双目立体匹配网络BGNet恢复左目相机坐标系下的视差图，并将视差图转为深度图和点云；最后，将2D物体映射到3D点云中，并对物体的3D点云加以约束，获得最终的感知结果。实验表明，使用双目立体视觉和实例分割算法，能够有效提升对小物体的3D感知的鲁棒性，并且不需要任何的3D标注，使得感知结果更加准确和可靠。

7.2 未来工作展望

本文基于航天院火星场以及火星车为背景，进行了一些多传感器融合感知与三维重建的工作。在此基础上，还可以从以下几个方面进行更进一步的研究：

（1）多源传感器时空一致表征

当前本文对于传感器的标定工作主要是通过多种不同的标定工具对于不同的传感器进行的标定，此种方式，对于较多的传感器需要进行多种工具的安装与标定，操作较为麻烦，未来可以进行相关标定算法的开发，同时进行多种传感器的内外参标定，这样能极大的节省前期工作的时间。

（2）多源多层地图重建

针对火星车移动缓慢的特点，在本文的第五章工作中使用的三维重建算法，为离线算法，此种重建算法不错，但是需要较长的时间进行大范围的重建。今后的工作中，可以进一步探索适用于不管高速还是低速行驶的无人车的实时稠密重建算法。为更多环境的重建提供可能。

（3）三维环境感知

本文中采用基于双目视觉的3D点云恢复以及感知，相较于目前的端到端点云检测网络确实不需要进行大量的3D标注工作，并且可以进行小物体的感知。但是本文中采用的是两个网络的结合，这也使得算法整体时间并不可观。所以，未来可以进行针对此类问题的端到端算法的研究，进一步的提升算法性能。

# 

# 参考文献

1. Xu L D, Xu E L, Li L. Industry 4.0: state of the art and future trends[J]. International Journal of Production Research, 2018, 56(8): 2941-2962.
2. 龙霄潇, 程新景, 朱昊, 等. 三维视觉前沿进展[J]. 中国图象图形学报, 2021, 26(06): 1389–1428.
3. Wu B, Dong J, Wang Y, et al. Landing Site Selection and Characterization of Tianwen-1 (Zhurong Rover) on Mars[J]. Journal of Geophysical Research: Planets, 2022, 127(4): e2021JE007137.
4. Khaleghi B, Khamis A, Karray F O, et al. Multisensor data fusion: A review of the state-of-the-art[J]. Information Fusion, 2013, 14(1): 28-44.
5. Song L, Yan R. Bearing fault diagnosis based on Cluster-contraction Stage-wise Orthogonal-Matching-Pursuit[J]. Measurement, 2019, 140: 240-253.
6. Ankarao V, Sowmya V, Soman K P. Multi-sensor data fusion using NIHS transform and decomposition algorithms[J]. Multimedia Tools and Applications, 2018, 77: 30381-30402.
7. Siachalou S, Mallinis G, Tsakiri-Strati M. A hidden Markov models approach for crop classification: Linking crop phenology to time series of multi-sensor remote sensing data[J]. Remote Sensing, 2015, 7(4): 3633-3650.
8. Lin K, Li Y, Sun J, et al. Multi-sensor fusion for body sensor network in medical human–robot interaction scenario[J]. Information Fusion, 2020, 57: 15-26.
9. Wang P, Liu C, Wang Y, et al. Advanced pedestrian state sensing method for automated patrol vehicle based on multi-sensor fusion[J]. Sensors, 2022, 22(13): 4807.
10. Mao X, Li W, Lei C, et al. A brain–robot interaction system by fusing human and machine intelligence[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2019, 27(3): 533-542.
11. Gan S, Zhuang Q, Gong B. Human-computer interaction based interface design of intelligent health detection using PCANet and multi-sensor information fusion[J]. Computer Methods and Programs in Biomedicine, 2022, 216: 106637.
12. Zhu Y, Liang S, Gong M, et al. Decomposed POMDP optimization-based sensor management for multi-target tracking in passive multi-sensor systems[J]. IEEE Sensors Journal, 2022, 22(4): 3565-3578.
13. Carvalho H, Vale A, Marques R, et al. Remote inspection with multi-copters, radiological sensors and SLAM techniques[C].EPJ Web of Conferences. EDP Sciences, 2018, 170: 07014.
14. Saponara S. Sensing and connection systems for assisted and autonomous driving and unmanned vehicles[J]. Sensors, 2018, 18(7): 1999.
15. Ayyasamy S. A Comprehensive Review on Advanced Driver Assistance System[J]. Journal of Soft Computing Paradigm, 2022, 4(2): 69-81.
16. 咸宝金, 陈松涛. 智能移动机器人多传感器信息融合及应用研究[J]. 宇航计测技术, 2010 (2): 41-44.
17. Pollefeys M, Nistér D, Frahm J M, et al. Detailed real-time urban 3d reconstruction from video[J]. International Journal of Computer Vision, 2008, 78: 143-167.
18. Agarwal S, Furukawa Y, Snavely N, et al. Building rome in a day[J]. Communications of the ACM, 2011, 54(10): 105-112.
19. Furukawa Y, Curless B, Seitz S M, et al. Towards internet-scale multi-view stereo[C] IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010: 1434-1441.
20. Chambolle A, Pock T. A first-order primal-dual algorithm for convex problems with applications to imaging[J]. Journal of Mathematical Imaging and Vision, 2011, 40: 120-145.
21. Goldluecke B, Strekalovskiy E, Cremers D. The natural vectorial total variation which arises from geometric measure theory[J]. SIAM Journal on Imaging Sciences, 2012, 5(2): 537-563.
22. Klingensmith M, Dryanovski I, Srinivasa S S, et al. Chisel: Real Time Large Scale 3D Reconstruction Onboard a Mobile Device using Spatially Hashed Signed Distance Fields[C]. Robotics: Science and Systems. 2015, 4(1).
23. Engel J, Stückler J, Cremers D. Large-scale direct SLAM with stereo cameras[C]. IEEE/RSJ International Conference on Intelligent Robots and Systems, 2015: 1935-1942.
24. Schöps T, Sattler T, Häne C, et al. 3D modeling on the go: Interactive 3D reconstruction of large-scale scenes on mobile devices[C]. International Conference on 3D Vision, 2015: 291-299.
25. Anguelov D, Dulong C, Filip D, et al. Google street view: Capturing the world at street level[J]. Computer, 2010, 43(6): 32-38.
26. Klingner B, Martin D, Roseborough J. Street view motion-from-structure-from-motion[C].IEEE International Conference on Computer Vision. 2013: 953-960.
27. Xiao J, Furukawa Y. Reconstructing the world’s museums[J]. International journal of computer vision, 2014, 110: 243-258.
28. Bok Y, Choi D G, Kweon I S. Sensor fusion of cameras and a laser for city-scale 3D reconstruction[J]. Sensors, 2014, 14(11): 20882-20909.
29. Richard A, Newcombe R A, Steven L. Dense tracking and mapping in real-time[C]. IEEE International Conference on Computer Vision, 2011, 2320: 2327.
30. Curless B, Levoy M. A volumetric method for building complex models from range images[C]. Computer Graphics and Interactive Techniques. 1996: 303-312.
31. Newcombe R A, Izadi S, Hilliges O, et al. Kinectfusion: Real-time dense surface mapping and tracking[C]. IEEE International Symposium on Mixed and Augmented Reality, 2011: 127-136.
32. Whelan T, Kaess M, Fallon M, et al. Kintinuous: Spatially extended kinectfusion[J]. 2012.
33. Nießner M, Zollhöfer M, Izadi S, et al. Real-time 3D reconstruction at scale using voxel hashing[J]. ACM Transactions on Graphics, 2013, 32(6): 1-11.
34. Handa A, Whelan T, McDonald J, et al. A benchmark for RGB-D visual odometry, 3D reconstruction and SLAM[C]. IEEE International Conference on Robotics and Automation, 2014: 1524-1531.
35. Riegler G, Osman Ulusoy A, Geiger A. Octnet: Learning deep 3d representations at high resolutions[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 3577-3586.
36. Guo M H, Xu T X, Liu J J, et al. Attention mechanisms in computer vision: A survey[J]. Computational Visual Media, 2022, 8(3): 331-368.
37. Le N, Rathour V S, Yamazaki K, et al. Deep reinforcement learning in computer vision: a comprehensive survey[J]. Artificial Intelligence Review, 2022: 1-87.
38. Hinterstoisser S, Holzer S, Cagniart C, et al. Multimodal templates for real-time detection of texture-less objects in heavily cluttered scenes[C]. International Conference on Computer Vision, 2011: 858-865.
39. Hinterstoisser S, Lepetit V, Ilic S, et al. Model based training, detection and pose estimation of texture-less 3d objects in heavily cluttered scenes[C]. Asian Conference on Computer Vision, 2013: 548-562.
40. Kehl W, Manhardt F, Tombari F, et al. Ssd-6d: Making rgb-based 3d detection and 6d pose estimation great again[C]. IEEE International Conference on Computer Vision, 2017: 1521-1529.
41. Tekin B, Sinha S N, Fua P. Real-time seamless single shot 6d object pose prediction[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 292-301.
42. Wang C, Xu D, Zhu Y, et al. Densefusion: 6d object pose estimation by iterative dense fusion[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 3343-3352.
43. Besl P J, Jain R C. Three-dimensional object recognition[J]. ACM Computing Surveys (CSUR), 1985, 17(1): 75-145.
44. Bhanu B, Ho C C. CAD-based 3D object representation for robot vision[J]. Computer, 1987, 20(08): 19-35.
45. Tombari F, Di Stefano L. Object recognition in 3d scenes with occlusions and clutter by hough voting[C]. Image and Video Technology, 2010: 349-355.
46. Buch A G, Kraft D, Kamarainen J K, et al. Pose estimation using local structure-specific shape and appearance context[C]. IEEE International Conference on Robotics and Automation, 2013: 2080-2087.
47. Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing Systems, 2015, 28.
48. Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
49. Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]. Proceedings of the European Conference on Computer Vison, 2016: 21-37.
50. Peng S, Liu Y, Huang Q, et al. Pvnet: Pixel-wise voting network for 6dof pose estimation[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 4561-4570.
51. Qi C R, Su H, Mo K, et al. Pointnet: Deep learning on point sets for 3d classification and segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 652-660.
52. Qi C R, Yi L, Su H, et al. Pointnet++: Deep hierarchical feature learning on point sets in a metric space[J]. Advances in Neural Information Processing Systems, 2017, 30.
53. Shi S, Wang X, Li H. Pointrcnn: 3d object proposal generation and detection from point cloud[C] . IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 770-779.
54. Badue C, Guidolini R, Carneiro R V, et al. Self-driving cars: A survey[J]. Expert Systems with Applications, 2021, 165: 113816.
55. Dong S, Wang P, Abbas K. A survey on deep learning and its applications[J]. Computer Science Review, 2021, 40: 100379.
56. Chen X, Ma H, Wan J, et al. Multi-view 3d object detection network for autonomous driving[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1907-1915.
57. Qi C R, Liu W, Wu C, et al. Frustum pointnets for 3d object detection from rgb-d data[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 918-927.
58. Ku J, Mozifian M, Lee J, et al. Joint 3d proposal generation and object detection from view aggregation[C]. IEEE/RSJ International Conference on Intelligent Robots and Systems, 2018: 1-8.
59. Liang M, Yang B, Chen Y, et al. Multi-task multi-sensor fusion for 3d object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 7345-7353.
60. Juarez-Salazar R, Zheng J, Diaz-Ramirez V H. Distorted pinhole camera modeling and calibration[J]. Applied Optics, 2020, 59(36): 11310-11318.
61. Weng J, Cohen P, Herniou M. Camera calibration with distortion models and accuracy evaluation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(10): 965-980.
62. 宋俊男, 朱世强, 原崧育, 等. 地外探测设备多传感器硬件时间同步方法研究[J]. 空间控制技术与应用, 2022, 47(6): 19-26.
63. Zhang Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
64. 王谭, 王磊磊, 张卫国, 等. 基于张正友标定法的红外靶标系统[J]. 光学精密工程, 2019, 27(8): 1828-1835.
65. Samatas G G, Pachidis T P. Inertial Measurement Units (IMUs) in Mobile Robots over the Last Five Years: A Review[J]. Designs, 2022, 6(1): 17.
66. 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.
67. Campos C, Elvira R, Rodríguez J J G, et al. Orb-slam3: An accurate open-source library for visual, visual–inertial, and multimap slam[J]. IEEE Transactions on Robotics, 2021, 37(6): 1874-1890.
68. Xu G Y, Chen L P, Gao F. Study on binocular stereo camera calibration method[C]. International Conference on Image Analysis and Signal Processing, 2011: 133-137.
69. 刘今越, 唐旭, 贾晓辉, 等. 三维激光雷达-相机间外参的高效标定方法[J]. 仪器仪表学报, 2022, 40(11): 64-72.
70. Raju V M, Gupta V, Lomate S. Performance of open autonomous vehicle platforms: Autoware and Apollo[C]. International Conference for Convergence in Technology, 2019: 1-5.
71. Fuhrmann S, Langguth F, Goesele M. Mve-a multi-view reconstruction environment[C]. Graphics and Cultural Heritage, 2014: 11-18.
72. Schonberger J L, Frahm J M. Structure-from-motion revisited[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 4104-4113.
73. Furukawa Y, Curless B, Seitz S M, et al. Towards internet-scale multi-view stereo[C] . IEEE Conference on Computer Vision and Pattern Recognition, 2010: 1434-1441.
74. Fuhrmann S, Goesele M. Floating scale surface reconstruction[J]. ACM Transactions on Graphics, 2014, 33(4): 1-11.
75. Carr J C, Beatson R K, Cherrie J B, et al. Reconstruction and representation of 3D objects with radial basis functions[C]. Computer Graphics and Interactive Techniques. 2001: 67-76.
76. Xu B, Xu Y, Yang X, et al. Bilateral grid learning for stereo matching networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 12497-12506.
77. Wang X, Kong T, Shen C, et al. Solo: Segmenting objects by locations[C]. Proceedings of the European Conference on Computer Vision, 2020: 649-665.
78. He K, Gkioxari G, Dollár P, et al. Mask r-cnn[C]. IEEE International Conference on Computer Vision, 2017: 2961-2969.
79. Kendall A, Martirosyan H, Dasgupta S, et al. End-to-end learning of geometry and context for deep stereo regression[C]. IEEE International Conference on Computer Vision, 2017: 66-75.
80. Zhang F, Prisacariu V, Yang R, et al. Ga-net: Guided aggregation net for end-to-end stereo matching[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 185-194.
81. Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? the kitti vision benchmark suite[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2012: 3354-3361.

# 附录1 攻读硕士学位期间撰写的论文

1. Zhu S, Pun C M, **Zhu H**, et al. An artificial bee colony algorithm with a balance strategy for wireless sensor network[J]. Applied Soft Computing, 2023: 110083. (SCI二区：IF=8.263)

# 附录2 攻读硕士学位期间申请的专利

1. 高浩，**朱海鹏**.一种复杂地形下多传感器配准融合系统及方法，202111130555.4，2021.9.
2. 高浩，**朱海鹏**，葛森林，宋杰. 一种用于多传感器融合的时间同步方法、系统及装置，2023100917857，2023.2.

# 附录3 攻读硕士学位期间参加的科研项目

1. 国家重点实验室开放课题：基于火星车和直升机的天地一体火星表面环境感知与语义建模。编号：JBK219003 课题主要参与人
2. 江苏省研究生科研与实践创新计划：复杂环境下地图构建与感知无人系统。项目编号：SJCX21\_0289 课题负责人
3. 科研项目：大范围多视角多层次系统级地图融合感知与验证系统研制。课题主要参与人

# 致谢

拙笔艰涩难成章，鸣谢之辞有心长。

回顾前路，思绪良多。首先要感谢的就是我的导师高浩教授，在研究生的这几年里，不管是生活上还是实践上，都给与了我许多机会与帮助，对我的课题也是给与了很大的支持。此处还需要感谢航天院的胡海东老师，正是高老师给的机会，我才能有许多的实践动手的机会，才能在与高老师、胡老师关于项目沟通中学到很多的技术细节以及整理和完善材料的流程和规范。在此非常感谢高老师和胡老师的指导与帮助。

感谢420实验室的伙伴们的陪伴，尤其感谢王远师兄和张頔师兄在我项目遇到困难时给予的指导，感谢徐飞易师兄对我日常学习和毕业论文上的指导和帮助。

感谢我的父母，没有你们的培养我也无法成人。从初入学堂的交不起学费，到现在的我即将毕业。感谢你们的对我生活上、经济上、学习上的帮助与支持。你们永远是我的牵挂和动力。

至此，我的学生生涯马上结束，也即将开启新的篇章。最后感谢一下自己读书二十载，一直砥砺前行的不放弃，迎接新的明天。