

本科毕业论文（设计）

基于 GPS 数据融合的开阔场景下双目视觉里
程计研究

RESEARCH ON BINOCULAR VISUAL
ODOMETRY IN OPEN SCENES BASED
ON GPS DATA FUSION

李明谦

哈尔滨工业大学

2024 年 5 月

密级：公开

本科毕业论文（设计）

基于 GPS 数据融合的开阔场景下双目视觉里程计研究

本 科 生	李明谦
学 号	8200880133
指 导 教 师	佟明斯 佟明斯
专 业	机器人工程
学 院	未来技术学院
答 辩 日 期	2024 年 5 月
学 校	哈尔滨工业大学

摘 要

随着无人机技术的迅猛进步，这种体积小、成本低的飞行器在军事和民用领域都显现出了巨大的潜力。它们可以用于军事侦察、气象监测，以及电力输电线路的巡检等任务，这些任务的完成都要求无人机自主定位具有较高的精确性和鲁棒性。因此使无人机具备更高的定位精度是扩展其应用领域的关键。然而现有的研究方法尚不足以解决无人机在开阔环境下由于飞行速度过快导致定位精度下降的问题，因此本文研究了基于 GPS 数据融合的开阔场景下双目视觉里程计的技术，主要包含特征点的检测与跟踪，双目视觉惯性里程计的实现以及双目视觉惯性里程计与 GPS 数据融合的方法。在特征点的检测与跟踪部分，利用 ORB 算法，Harris 角点检测算法和 FAST 角点检测算法实现了对无人机飞行姿态变化中特征点的检测与描述；在双目视觉惯性里程计的实现中，通过 IMU 预积分和非线性优化等操作实现了双目视觉里程计和 IMU 数据的融合；在双目视觉惯性里程计与 GPS 数据融合的部分中，利用扩展卡尔曼滤波技术和非线性最小二乘等方法实现了二者数据的融合和优化。实验结果表明，双目视觉惯性里程计和 GPS 融合后的数据在开阔场景下有较高的稳定性和精确性，能够满足无人机定位精度需求。通过本研究，为解决无人机视觉定位中的难题提供了新的解决方案，为无人机在复杂环境下的自主作业需求提供了技术支持。

关键词：GPS 数据融合；双目视觉里程计；开阔场景；无人机；定位精度

Abstract

With the rapid advancement of drone technology, this small, low-cost aircraft has shown great potential in both military and civilian fields. They can be used for tasks such as military reconnaissance, weather monitoring, and inspection of power transmission lines. The completion of these tasks requires UAVs to have high accuracy and robustness in autonomous positioning. Therefore, enabling drones to have higher positioning accuracy is the key to expanding their application areas. However, the existing research methods are not enough to solve the problem of reduced positioning accuracy due to too fast flying speed of UAVs in open environments. Therefore, this paper studies the technology of binocular visual odometry in open scenes based on GPS data fusion. Mainly It includes the detection and tracking of feature points, the implementation of binocular visual inertial odometry and the method of merging binocular visual inertial odometry and GPS data. In the feature point detection and tracking part, the ORB algorithm, Harris corner point detection algorithm and FAST corner point detection algorithm are used to detect and describe the feature points in the change of UAV flight attitude; in the implementation of binocular visual inertial odometry In the part, the fusion of binocular visual odometry and IMU data is achieved through operations such as IMU pre-integration and nonlinear optimization; in the fusion part of binocular visual inertial odometry and GPS data, extended Kalman filtering technology and nonlinear minimum are used Methods such as square multiplication realize the fusion and optimization of the two data. Experimental results show that the data after fusion of binocular visual inertial odometry and GPS has high stability and accuracy in open scenes, and can meet the accuracy requirements of UAV positioning. Through this research, a new solution is provided to solve the problems in the visual positioning of UAVs, and it provides technical support for the autonomous operation requirements of UAVs in complex environments.

Keywords: GPS data fusion, binocular visual odometry, open scenes, UAV, positioning accuracy

目 录

摘 要	I
ABSTRACT	II
目 录	III
第 1 章 绪 论	1
1.1 课题来源及研究的目的和意义	1
1.1.1 课题来源	1
1.1.2 研究的目的和意义	1
1.2 国内外在该方向的研究现状及分析	3
1.2.1 视觉里程计国内研究现状	3
1.2.2 视觉里程计国外研究现状	4
1.2.3 对国内外研究现状的分析	5
1.3 主要研究内容	6
1.3.1 基于 ORB 算法的特征点检测与跟踪的方法研究	6
1.3.2 基于扩展卡尔曼滤波的双目视觉惯性里程计的实现方法研究	7
1.3.3 基于双目视觉惯性里程计和 GPS 数据融合的方法研究	7
1.3.4 研究内容过程图	8
第 2 章 基于 ORB 算法的特征点检测与跟踪	9
2.1 角点检测	9
2.1.1 Harris 角点	9
2.1.2 FAST 算法与 Harris 算法结合的角度检测	10
2.2 特征点描述	12
2.3 双目相机特征点匹配	13
2.4 实验验证	14
2.4.1 实验设计	14
2.4.2 实验结果及分析	14
2.5 本章小结	15
第 3 章 基于扩展卡尔曼滤波的双目视觉惯性里程计	16
3.1 无人机状态估计	16
3.1.1 状态定义	16
3.1.2 代价函数	16
3.2 传感器因子	17
3.2.1 相机因子	17
3.2.2 IMU 因子	18

3.3 边缘化	21
3.4 非线性优化	22
3.5 实验验证	26
3.5.1 实验设计	26
3.5.2 实验结果及分析	27
3.6 本章小结	28
第 4 章 双目视觉惯性里程计与 GPS 的数据融合	29
4.1 GPS 定位原理	29
4.2 双目视觉惯性里程计与 GPS 数据融合	33
4.2.1 坐标转换	33
4.2.2 基于扩展卡尔曼滤波的数据融合	34
4.2.3 后端非线性优化	36
4.3 实验验证	37
4.3.1 实验设计	37
4.3.2 实验结果及分析	37
4.4 本章小结	39
结 论	40
参考文献	41
哈尔滨工业大学本科毕业论文（设计）原创性声明和使用权限	45
致 谢	46

第 1 章 绪 论

1.1 课题来源及研究的目的和意义

1.1.1 课题来源

本课题来源于智能控制系统研究所在研横向项目，该项目对无人机在开阔场景下的自主定位精度和鲁棒性有较高的要求，本课题针对该问题开展基于 GPS 数据融合的双目视觉里程计研究。

1.1.2 研究的目的和意义

在过去的数十年里，科技的浪潮席卷全球，其中尤以微型计算机和传感器技术的迅猛发展最为引人注目。在这些科技力量的共同推动下，无人机（Unmanned Aerial Vehicles，简称 UAVs）的理论框架和技术应用日益完善，成为了现代科技领域的一颗璀璨明珠。无人机的种类繁多，各具特色。按飞行方式划分，无人机可以大致分为多旋翼无人机、固定翼无人机和混合型无人机。每一种无人机都凭借其独特的设计和功能，在不同的领域发挥着不可替代的作用。随着无人机技术的不断进步，其体积逐渐缩小，成本日益降低，这为无人机在更广泛的领域里应用提供了可能。在军事领域，无人机凭借其小巧灵活、成本低廉的特点，成为了侦察和监视、目标定位和打击以及电子战等主多用途的重要工具。通过无人机的应用，可以极大地降低军人的风险，提高作战效率。在商业和农业领域，无人机被广泛应用于航拍摄影、地理测绘、精准农业等方面。在紧急救援中，它们可以用于搜索和救援任务，如寻找失踪人员或在自然灾害后进行灾害评估。此外，在环境监测领域，无人机也展现出了巨大的潜力，它们可以用于监测空气质量、土壤、植被、水质等多方面。通过无人机的应用，我们可以更加深入地了解自然环境的变化规律，为环境保护和生态平衡做出更大的贡献。在电影和媒体领域，无人机同样成为了航拍、电影制作和新闻报道的重要工具。最后，在物流和运输领域，无人机也展现出了巨大的潜力，它们可以用于紧急物资运输和远程地区物资供应等方面。无人机的应用可以极大地提高物流效率，降低运输成本，为人们的生活带来更大的便利，如图 1-1 所示。



图 1-1 无人机的应用

尽管无人机在多个领域中的运用已相当普及，但在电力输电线路巡检方面，其视觉定位仍面临两大挑战：

1、由于户外场地的广阔性质，其中可供相机系统识别的对象数量相对较少，这导致相机视觉感知和目标检测的任务变得更具挑战性。在这种环境下，相机系统可能会受到遮挡、光照变化和视角问题的影响，这些因素都可能对目标的准确识别产生负面影响。因此，无人机搭载的相机系统在飞行过程中会存在以下问题：缺乏足够的特定特征以供识别。这种特征不足的情况可能导致相机的识别精度降低，进而导致无人机自身的定位精度降低。

2、无人机在运动时具有较高的速度，这意味着在极短的时间内它将覆盖大量的地理区域。然而，仅仅依赖相机数据进行自身定位存在一些显著的挑战。一方面，相机的快速位置采集可能导致在连续图像帧之间出现明显的位置变化，这可能会对无人机的自身定位精度造成不利影响。另一方面，无人机本身的运动和加速度也会增加相机定位的不稳定性，因为快速的运动可能导致图像模糊或失真，进而干扰了准确的目标识别和特征匹配，从而导致无人机自身定位准确度下降。

经过对现有文献的深入研究和分析，当前所采取的方法仍不足以解决上述无人机在电力输电线路巡检中视觉定位所面临的挑战。针对这些问题，本文将双目视觉里程计和 GPS 数据进行融合，利用 GPS 在开阔场景下的稳定性提升无人机在开阔场景下自身定位的精准性和鲁棒性，满足无人机户外电力输电线路巡检的自主定位需求。

本次课题研究对于经济与环境效益具有积极的影响。在经济效益方面，无人机在执行紧急救援任务时可以节省人力成本，在精准农业方面可以提高农作物的生产效率，同时促进了该领域无人机定位技术的创新，从而打动经济增长。在环境效益方面，由于无人机采用电力驱动，可以减少对环境的污染，同时在监测生

态环境方面时，无人机可以减少人类活动对生态环境的干扰，有助于保护生态环境，提高自然资源的利用率。

1.2 国内外在该方向的研究现状及分析

视觉里程计作为机器视觉与机器人技术中的一项关键组成部分，其核心功能在于通过分析连续捕获的图像帧序列，来估算移动相机或机器人的相对运动状态，进而推导并描绘出其行进轨迹及定位数据。视觉里程计基于计算机视觉原理，通过识别并跟踪在连续图像帧中可见的特征点、边缘或区域，来测量相邻图像帧之间的相机位移。这些位移测量然后被积分或累积，以获得相机或机器人的运动轨迹和定位信息。

在视觉里程计（VO）的实践中，常见的三种方法^[1]是基于特征点的 VO、基于直接跟踪的 VO 以及基于混合半直接跟踪的 VO。按照相机的类型划分，视觉里程计技术可以进一步细分为单目视觉里程计、立体视觉里程计和 RGB-D 视觉里程计。在利用图像信息的维度上，VO 技术被归类为基于特征的方法和基于直接的方法。而在降低轨迹漂移的策略上，VO 技术则分为采用滤波器的方法和采用非线性优化法^[2]。这些方法的选择和应用，取决于具体的应用场景和需求，以及可用的硬件设备和计算资源。

1.2.1 视觉里程计国内研究现状

石宗英等来自清华大学的研究者提出了一种创新的无人机视觉定位技术，该技术基于街景图像进行实现^[3]，通过利用 Google 街景图像作为预先获取的环境信息，能够在城市环境中准确确定全球定位无人机（UAV）的位置。该方法独特之处在于仅依赖单目相机和视觉里程计估计，并分为两个核心步骤：首先是地点识别，随后是度量定位。地点识别算法通过匹配无人机上的相机图像与数据库中的地理标记图像来确定 UAV 的位置。度量定位方法通过跟踪图像序列中的特征、绝对尺度估计和轨迹优化来实现 UAV 的精确定位。该方法可以纠正视觉里程计的尺度漂移，并将 UAV 的姿态估计在真实世界坐标系中进行。该技术在监视、制图和交付服务等领域具有潜在的应用价值。杨绍武等学者提出了整合多摄像头测量数据的 SLAM（同时定位和地图构建）系统，用于无人机在未知复杂环境中的自主导航。该系统通过多摄像头视觉 SLAM 克服了稳健姿态跟踪和地图构建的难题，实现了高效的姿态跟踪和地图优化^[4]。王科、职恒辉、王文森和刘海涛等研究者则通过深度学习技术，解决了机器人移动过程中累积误差导致的位姿估计精度问题，

从而提高了视觉惯性里程计的准确性和鲁棒性^[5-8]。崔立志等人则提出了一种结合深度学习、Ghost 模块和 YOLOv5s 目标检测网络的 C3Ghost 模块。该方法通过去除一些变化的特征点，仅利用相对稳定的特征点估算无人机位姿，提高了定位的可靠性^[9]。李晋等研究者提出了一种结合视觉里程计和反向传播神经网络的行人轨迹预算算法，并结合扩展卡尔曼滤波技术，增强了行人航迹推算的稳定性和鲁棒性^[10]。李伟东等人提出了一种基于光流追踪的双目视觉里程计算法，用于实时追踪关键帧中的特征点，以实现高效的地图构建^[11]。在 GNSS 信号受到干扰的情境下，廖士楠等人设计了一套视觉惯性里程计系统，为多旋翼无人机提供辅助视觉导航，确保其稳定飞行在外部多变的环境中^[12]。闫河等人介绍了一种结合语义信息的视觉里程计不稳定特征点的去除方法，通过去除误差较大的特征点，提高了相机位姿估计的精确性和鲁棒性^[13]。张梦龙等人提出了一种融合 IMU 信息的点转移矫正视觉惯性里程计，有效减小了卷帘效应对定位精度的影响^[14]。田应仲等人则提出了一种结合点线特征的 VIO（视觉惯性里程计）方法，该方法在移动机器人定位过程中，通过融合点线特征，有效解决了点特征退化的问题^[15]。钟家跃深入研究了在未知环境下无人机利用双目视觉里程计的自主导航方法^[16]，并通过实验证明，在室外环境中采用 SIFT 算法^[17]的视觉里程计表现劣于采用 ORB 算法^[18]的视觉里程计。吕华龙则提出了一种无人机状态估计算法^[19]，该算法主要融合双目立体视觉、IMU 和 GPS 数据，显著提升了无人机状态估计的稳定性。

1.2.2 视觉里程计国外研究现状

针对无人飞行器在 GPS 信号不可用环境下的计算机视觉定位与制图方法，Cvišić 等人提出了一种名为 SOFT-SLAM 的方法^[20]。该方法主要基于双目视觉，通过特征跟踪获取场景深度信息，实现高效的定位和建图。该方法的优点在于可以在嵌入式计算机上实时运行，同时保留足够的计算资源用于控制和导航算法。该技术在 KIT 数据集上展现了比其他双目视觉 SLAM 技术更佳的性能。此外，该技术还可广泛运用于无人机自主导航、环境监测等领域。Cioffi 及其团队提出了一项方法，通过非线性优化算法，将 GPS 全球定位信息、视觉里程计数据和 IMU 惯性测量数据通过紧耦合的方式融合在一起。他们使用滑动窗口对最近系统状态进行优化，以找到一个成本函数可以减小视觉重投影的误差，同时还可以使 IMU 测量数据误差和 GPS 数据残差最小化^[21]。该方法利用了 IMU 预积分方法来高效地推导全局位置误差项，通过考虑所有传感器测量之间的相关性来扩展以往的研究，从而得到准确且全球一致的估计结果，而计算成本增加几乎可以忽略不计。Mur-Artal 和 Tardos 提出了 ORB-SLAM 和 ORB-SLAM2 算法^[22-23]，Carlos Campos

等人在其基础上发展了 ORB-SLAM3 算法^[24]。该算法是用于实时双目相机的视觉定位与建图系统，采用 FAST 特征检测器迅速提取图像中的关键点，接着，借助 BRIEF（Binary Robust Independent Elementary Features）描述符对这些关键点进行描述，这些描述符以二进制形式呈现，因而计算速度更快。并且 ORB-SLAM2 具有闭环检测功能，可以检测到相机返回到之前访问过的位置的情况，这有助于纠正累积误差，并提高 SLAM 系统的鲁棒性。Khan 及其团队提出了一项改进的无人机姿态估计方法^[25]，其采用半密集视觉里程计（SDVO）结合单目系统，通过将惯性测量单元（IMU）的姿态信息注入单目深度估计过程，并借助卡尔曼滤波将单目相机数据与 IMU 数据进行融合，从而降低姿态估计误差，提高深度图的鲁棒性。Yanchen Liu 及其团队提出了利用双目视觉里程计和点云数据构建越野场景中的多通行特征图的方法^[26]。Chuyan Wang 等人基于改进后的 MobileNetV3 神经网络模型和 ORB-SLAM2 视觉系统，运用回路检测、定位重定向和地图重复利用，提升了定位的精确性和稳定性^[27]。Liangyu Wang 等人创新性地应用了 ORB-SLAM2 算法，引入了独立的并行目标检测线程。与传统的 ORB-SLAM2 相比，他们成功缩短了地图构建的时间，同时提高了精确度^[28]。Dekai Chen 等学者提出了结合视觉里程计和 IMU 的方法，从而得到高精度和高稳定性的同步定位与建图^[29]。Bo Peng 等人提出了一种基于双目几何相关学习的自监督单目深度估计网络，研究出了一种预测双目相机线索的模块，利用视图间的几何相关性，生成了可以辅助视觉定位的线索，用于单目相机深度估计的自监督学习^[30]。Jiaqi Xu 和其团队提出了一项无人机精准定位算法，结合了 VIO 和 Cartographer 算法，用于在无人机飞行时持续输出准确位置估计信息。该算法旨在解决点云数据中由于特征丢失或坐标系漂移而导致位置估计不准确的问题^[31]。Jinyu Li 等人提出了 RD-VIO 新型视觉惯性里程计设计，采用 IMU-PARSAC 算法进行关键点匹配，并针对纯旋转运动提供额外约束，在动态环境下相比其他方法具有显著优势^[32]。

1.2.3 对国内外研究现状的分析

国内的研究表明，无人机视觉定位在无 GPS 环境下进行，在无 GPS 环境下有着广泛的应用前景。但是由于缺乏 GPS 数据，并且仅使用单目相机和视觉里程计，在复杂环境中可能存在尺度漂移等问题的影响。多摄像头的引入在于整合了多个摄像头的测量数据，提高了无人机在位置复杂环境中的自主导航能力，但是此种方法对硬件要求较高，可能会增加系统成本和复杂性。引入 YOLO 算法增进了视觉惯性里程计的精确性与韧性，但深度学习模型的培训需大量数据，消耗更多计算资源。由于无人机在飞行过程中特征点数目偏少，依靠语义信息剔除动态特征

点的方法在无人机飞行过程中并不适用。点转移校正和点线特征融合方法用以克服点特征退化和卷帘效应的误差问题。然而，在本课题研究中，上述问题并非造成误差的主要原因，因此不适用。在飞行过程中，无人机依赖光流追踪是一种可行方案，通过光流追踪和视觉里程计的融合，可以实时提供更精确的定位。利用 ORB 算法，以及 IMU 和 GPS 相融合的双目视觉里程计构建过程，可以实现更加鲁棒的无人机自主定位系统。

SOFT-SLAM 技术是国外研究中无 GPS 环境下的定位方法，在无 GPS 环境下有着广泛的应用潜力。然而，由于缺乏 GPS 数据，并且仅依赖单目相机和视觉里程计，在复杂环境中可能会受到尺度漂移等问题的影响。利用改进过的 MobileNetV3 神经网络模型，自监督单目深度估计网络可以提高定位的精度和鲁棒性，但训练神经网络模型需要大量数据，计算资源占用量较大。利用回环检测、重定位和地图复用在构建地图的工作中有较大帮助，但在本研究中，构建地图并不是主要工作，因此上述方法不适用于本次研究。VIO 与 Cartographer 算法融合的无人机精准定位算法可以为无人机飞行过程中实时输出高精度的位置信息，但在本次研究中由于无人机载重的问题，不考虑使用相关设备提供点云信息，因此该方法不适用于本次研究。GPS 与 IMU 通过紧耦合的方法，IMU 预积分算法以及 ORB-SLAM 算法等可以帮助无人机在飞行过程中检测特征点，提高定位的精确性和鲁棒性。

根据对国内外研究的综合分析，并结合本次研究中的主要问题，针对户外场地开阔的情况，可供识别和追踪的特征点较少的问题，采用 ORB-SLAM2 算法作为解决该问题参考算法。针对无人机在运动过程中速度较快，定位精度较差的问题，在前人论文提出的多传感器数据融合的方向，本次研究将融合双目视觉里程计、惯性测量仪 IMU 数据与 GPS 数据，利用 GPS 在开阔场景下的稳定性提升无人机在开阔场景下自身定位的精准性和鲁棒性，满足无人机户外开阔场景下电力输电线路巡检的自主定位需求。

1.3 主要研究内容

1.3.1 基于 ORB 算法的特征点检测与跟踪的方法研究

无人机在飞行过程中需要较高的定位精度，所以对于特征点的检测与跟踪直接关系到无人机飞行过程中定位结果的好坏。在本课题的研究中，由于开阔场景下无人机飞行姿态的变化，所以考虑在现有 ORB 算法的基础上，通过引入一些识别和优化算法，以帮助无人机飞行姿态变化中检测和描述在不同角度，不同方向 and 不同尺度上的特征点，同时 ORB 算法本身就对图像噪声和一定程度的遮挡具有

较强的鲁棒性，这也能够帮助空中飞行的无人机对不同视角的物体进行更好的识别和跟踪；此外，由于无人机在工作时飞行速度较快，会导致检测到的特征点出现尺度上的一些变化，描述子出现噪声过大等问题。本课题打算对 ORB 算法使用的 FAST^[33]角点检测算法和 BRIEF^[34]二进制描述子进行一些优化，例如为使 FAST 角点检测具有尺度不变性，可以借助高斯金字塔。在使用 BRIEF 二进制描述子时可以使用高斯卷积核进行滤波降噪，如图 1-2 所示。再加之这些描述子本身计算效率高和空间紧凑的特性，可以极大地减少存储和匹配成本，有利于无人机在飞行过程中出色地计算实时定位，从而得到较高定位精度的视觉里程计。

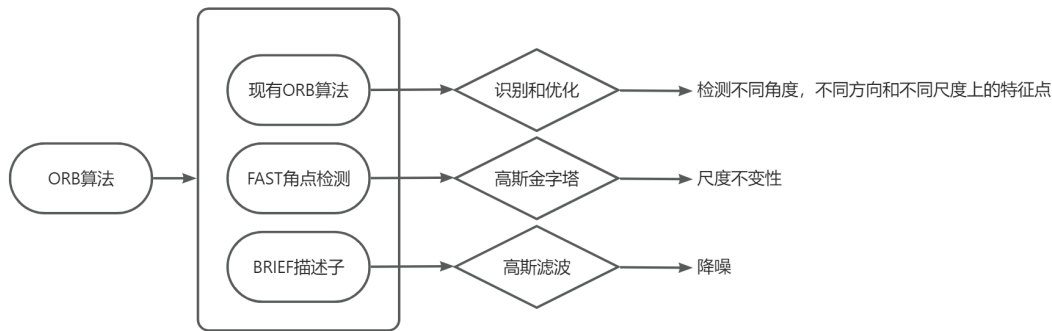


图 1-2 ORB 算法研究内容过程图

1.3.2 基于扩展卡尔曼滤波的双目视觉惯性里程计的实现方法研究

在广阔场景中，无人机飞行时飞行姿态频繁变化，仅依靠双目相机数据不足以准确估计无人机姿态，可能引发定位误差。然而，通过融合无人机上搭载的 IMU 传感器与双目相机获取的数据，并采用基于扩展卡尔曼滤波的方法，可实现更精确的自身位姿状态估计，进而获得更高的定位精度，实现了在无人机上应用双目视觉惯性里程计的目标。

在这个融合过程中，IMU 传感器提供了关于无人机三轴加速度和角速度的信息，视觉里程计则通过 3.1 中的研究已经达到了较高的定位精度。通过将这两种不同类型的数据结合起来，在实际操作过程中则可以通过不断改变扩展卡尔曼滤波算法的参数，使其可以更好地校正由于飞行姿态变化引起的误差，从而提供更准确的位姿估计，帮助无人机更稳定地执行自主导航、避障、目标跟踪等任务，从而更好地完成电力输电巡检作业。

1.3.3 基于双目视觉惯性里程计和 GPS 数据融合的方法研究

获取无人机双目视觉惯性里程计数据后，需考虑研究环境的特殊性，即户外广阔环境。在这种环境下，仅仅依靠 1.3.2 研究中得到的双目视觉惯性里程计可能会带来一些挑战。这是因为无人机在户外环境中的飞行速度可能会非常快，而所

覆盖的空间范围也相对较大，这导致双目相机只能够检测和跟踪到数量相对较少的特征点，从而降低双目视觉惯性里程计的准确性。

因此，在这种情况下，本文可以考虑采用一种更综合的方法，将双目视觉惯性里程计数据与 GPS（全球定位系统）数据相结合，利用扩展卡尔曼滤波技术来融合这两种不同类型的数据源。这种数据融合的方式可以帮助本文实现更高精度的定位结果，特别是在户外环境中，其中 GPS 通常提供了相对准确的位置信息。通过将双目视觉惯性里程计和 GPS 数据结合起来，无人机可以更可靠地确定自身位置，克服了双目相机在特征点有限的情况下定位不精准的限制。此融合方法可增强导航精度，保证无人机完成电路巡检任务，免受位置不确定性干扰。

1.3.4 研究内容过程图

综合以上研究内容，可得到基于 GPS 数据融合的双目视觉里程计研究过程图，如图 1-3 所示：

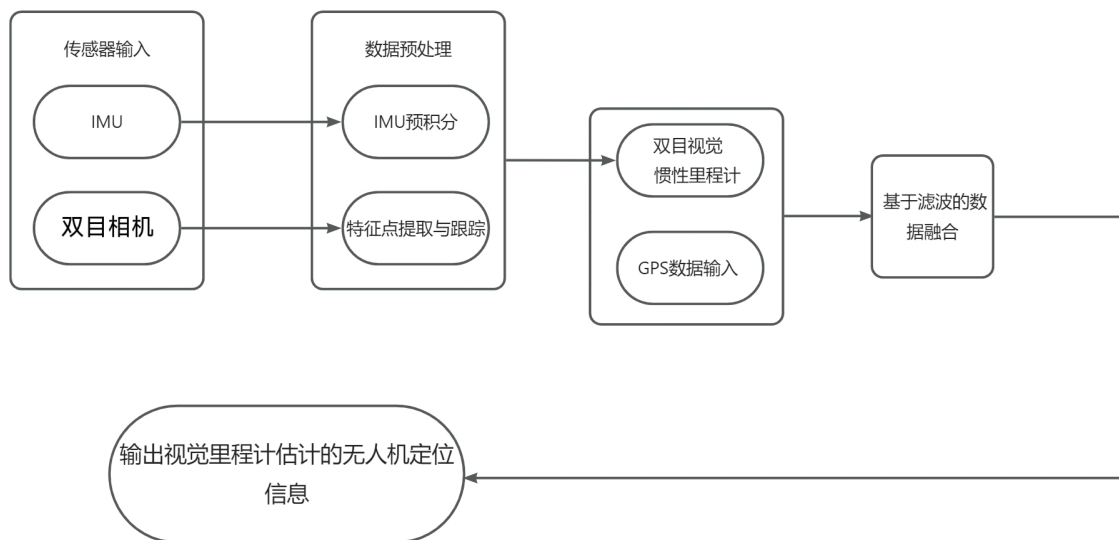


图 1-3 双目视觉里程计研究内容图

第 2 章 基于 ORB 算法的特征点检测与跟踪

本章主要阐述了课题中所涉及的无人机自主定位系统中特征点的检测与跟踪方案，首先讲述了角点检测的方法，然后以检测到的角点为依据进行描述特征点的操作，最后为双目相机左目相机和右目相机中检测到的特征点进行匹配操作。此外，本章节中针对特征点的检测与跟踪设计了一套实验方案，用于检验本章所述方法的有效性。

2.1 角点检测

在该部分，本文主要利用了 FAST 算法和 Harris 角点检测算法相结合的方式来实现精准辨识角点。通过融合这两种算法，本文能够更全面地捕捉图像中重要的角点信息，从而提升检测的准确性和稳定性。FAST 算法以其高效的特点迅速定位出可能的角点位置，而 Harris 角点检测算法则能够对这些候选点进行更精确的评估，进一步过滤出真正的角点。这两种算法的结合能够充分发挥各自的优势，使得角点检测在图像处理和计算机视觉任务中发挥更为重要的作用。

2.1.1 Harris 角点

Harris 角点是一种计算机视觉方法，该方法用于检测图像局部角点^[35]。在图像中，角点定义为灰度值有过大变化的点，或者在两个主方向上的特征点。在图像处理中，这些点通常对应于物体的边缘或角落，具有明显变化的像素梯度。

Harris 角点识别算法的基本思想是使用固定窗口在任何方向上移动图像，并比较滑动前后窗口中像素灰度的变化。如果存在任意方向上的滑动都使得窗口中像素的灰度变化过大，那么这个窗口的中心就被认为是一个角点。在本次研究工作中，Harris 角点检测算法的具体步骤如下：

① 计算图像的梯度：计算 D435i 相机获取到的图像梯度，获取图像中每个像素的梯度幅值和方向。

② 计算自相关矩阵：对于每个像素，计算其周围窗口内的梯度幅值和方向的自相关矩阵。

③ 计算角点响应函数：对于 D435i 相机中获取到的图像中的每个像素，使用自相关矩阵计算它们的角点响应函数，该函数度量了像素周围区域的角点特征，其中角点响应函数公式为：

$$R = \det M - \alpha(\text{trace } M)^2 \quad (2-1)$$

④ 非极大值抑制：比较每个像素的角点响应函数与周围像素的响应函数，保留局部最大值作为角点。**Harris 角点检测**

算法的优点在于其稳定性和对噪声的鲁棒性，能够准确地定位图像中的角点，并且具有较高的计算能力。因此，在本部分研究工作中将其与 FAST 算法结合共同进行角点检测，以提高系统检测的精确性和鲁棒性。

2.1.2 FAST 算法与 Harris 算法结合的角点检测

在研究中，首先应用 FAST 算法对图像进行初步检测，以识别潜在角点作为特征点的候选。然而，仅仅依靠 FAST 算法检测出的特征点可能并不足以满足后续应用的需求。接着使用 Harris 角点的评价方式，评估图像中各像素的角点响应值，以确定其是否为角点。具体来说，Harris 角点响应函数考虑了像素点周围灰度值的改变情况，如果周围灰度值改变过大，则该点具有较高的响应值，表明其为角点的可能性较大。利用 Harris 角点响应函数，计算每个备选特征点的响应值，并依据响应值大小排序，选取前 N 个特征点作为最终结果。通过这种方式，能够确保系统提取出的特征点不仅数量适中，而且具有较高的代表性和稳定性，为后续的特征描述和匹配提供了坚实的基础。

针对 FAST 检测特征点缺乏尺度不变性的难题，本次研究工作引入了尺度空间理论。尺度空间理论描述了各种尺度下的图像特征，其通过对图像进行不同尺度的平滑处理，建立图像的金字塔结构，每层金字塔代表了不同尺度下的图像表达。通过这种方式，本研究可以同时获取图像在不同尺度下的特征信息，从而实现尺度不变性。

在本研究中，首先建立了图像的高斯金字塔^[37]。高斯金字塔是图像处理中至关重要的概念，它以一种有条理的方式协助本研究观察和分析不同尺度下图像的特征。这一过程通过一系列精心设计的步骤对原始图像进行处理而完成，其中包括高斯滤波和下采样。高斯金字塔的基础之一为高斯滤波。该滤波为线性平滑方法，通过在图像中每点，基于周围像素灰度分布进行加权平均，获得新像素值。这种滤波方法可以高效地消除图像中的高频噪声和细节，使图像更平滑。在高斯金字塔的构建中，原始图像会进行多次高斯滤波，每次滤波都会使图像变得更加模糊，从而得到一系列不同模糊程度的图像。紧接着，下采样是构建高斯金字塔的关键步骤之一。下采样是通过减少图像的尺寸来实现的，通常是通过每隔一定数量的像素取一个像素值，或者对一定区域内的像素值进行平均等方式来实现。下采样的结果是图像的尺寸变小，细节信息减少，但主要的结构和特征仍然得以

保留。在高斯金字塔中，每次高斯滤波后的图像都会进行下采样处理，生成一系列尺寸逐渐减小的图像。

通过高斯滤波和下采样这两个步骤的交替进行，最终可以得到一个高斯金字塔，如图 2-1 所示。

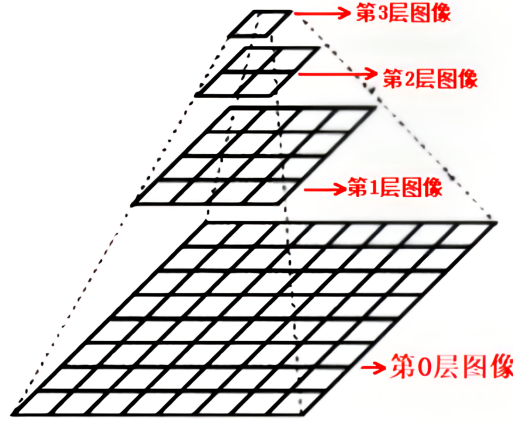


图 2-1 高斯金字塔

金字塔包含多幅图像，每层展示原图在不同尺度上的表示。随着层数增多，图像尺寸逐渐缩小，细节渐显模糊，但核心结构和特征保持。

经过高斯金字塔处理后，可在每个图像层面逐层使用 FAST 和 Harris 角点检测算法进行角点探测。由于高斯金字塔的每一层表示了不同尺度下的图像，因此，在每一层上应用 FAST 算法和 Harris 角点检测算法检测到的角点都具有尺度不变性。因此无论图像的尺寸或分辨率如何变化，都能够在高斯金字塔的相应层上找到稳定的特征点。

除了尺度不变性，FAST 特征点还面临另一个挑战，即缺乏旋转不变性。当图像发生旋转时，FAST 算法可能探测到的特征点发生变化，导致特征匹配失败。为了解决这一问题，原论文中提出了一种利用灰度质心法来确定特征点方向的方法，从而实现了带有方向的 FAST 角点，即 oFAST。

灰度质心法的基本思想是，对于每一个 FAST 特征点，首先要确定一个半径范围，这个范围通常根据图像的具体情况和特征点的性质来设定。然后，在这个半径范围内，计算所有像素的灰度值的加权平均位置，即灰度质心。通过计算得到的灰度质心，实际上代表了该特征点周围像素灰度值的分布中心。因为灰度质心是根据灰度值计算的，对图像亮度变化非常敏感，故灰度质心法可精确反映特征点周围像素的灰度分布特性。对于任意一个特征点 p 来说，定义 p 的邻域像素的矩为：

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

其中 $I(x, y)$ 代表像素灰度值， $\bigcirc_r$ 是半径为 r 离散化的 Bresenham 圆形窗口，零阶矩 m_{00} 即图像领域窗口内每一个像素的灰度总和， m_{10} 和 m_{01} 分别是相对于 x 与 y 的一阶矩，因此可定义图像局部邻域的中心矩或者质心为：

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (2-3)$$

在确定了灰度质心后，继续计算特征点到质心的向量。这个向量不仅描述了它们的位置关系，还隐含了灰度分布的方向。因此，本文可以将这个向量定义为该特征点的主方向。一旦确定了主方向，即使图像发生了旋转，只要特征点的主方向也相应旋转，就可以通过对比它们的主方向来判断特征点之间的对应关系。oFAST 特征点的方向矢量定义为特征点的中心 O 与质心 C 形成的向量 $\overrightarrow{OC}$ 与 X 轴的夹角，即：

$$\theta = \text{atan} 2(m_{01}, m_{10}) \quad (2-4)$$

2.2 特征点描述

BRIEF 特征描述子是利用二进制编码表示和描述已检测到的特征点的方法，与传统的提取特征的方式不同，传统方法是利用图像局部邻域的灰度直方图或梯度直方图。经过上一步的角点检测之后，采用 BRIEF 算法来创建特征描述符。算法具体操作为：在特征点周围的 Patch 内随机选取多对像素点。这些点对的选择并非任意，而是遵循一定的统计规律，以确保它们能够充分反映邻域内的图像纹理和结构信息。随后，算法会比较每一对点的灰度值大小，根据比较结果生成二进制编码。一般而言，如果一对点中第一个点的灰度值大于第二个点，则编码为 1；反之，则编码为 0。灰度测试 τ 定义如下：

$$\tau(p; \mathbf{x}, \mathbf{y}) = \begin{cases} 1 & p(\mathbf{x}) < p(\mathbf{y}) \\ 0 & p(\mathbf{x}) \geq p(\mathbf{y}) \end{cases} \quad (2-5)$$

式中 $p(\mathbf{x})$ 是图像块上点 $X = (u, v)$ 处的灰度值。通过这种方式，BRIEF 算法为每个特征点生成了一个独特的二进制字符串，即特征描述符。这种描述符不仅具有极小的存储开销，而且由于其二进制特性，使得特征匹配过程变得极为高效。匹配时为了衡量特征描述符之间的相似程度，需要计算两个特征点之间的汉明距离。汉明距离是指两个等长二进制字符串在对应位置上不同字符（通常是 0 和 1）的个数，非常适合于快速比较二进制特征描述符。其计算过程简单直接，对于两个等长的二进制字符串，本研究只需逐一比较它们对应位置上的字符。每发现一个不同的字符，就将计数器加 1。最终，计数器的值就是这两个字符串之间的汉明

距离。这种逐位比较的方式不仅易于实现，而且能够确保结果的准确性。利用汉明距离在特征匹配中，系统能够快速找到与给定描述符最接近的其他描述符。通过设定一个阈值，本研究可以筛选出汉明距离小于该阈值的所有候选匹配项，从而进一步缩小匹配范围，提高匹配的准确性和效率。通过比较所有点对，产生了一个长的二进制串：

$$f_n(p) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(p; \mathbf{x}_i, \mathbf{y}_i) \quad (2-6)$$

对于 n 的取值可以设为 128、256 或者 512，在本次研究中经过测试 n 取值为 128、256 和 512 均对实验结果无明显影响。

在实际应用中，BRIEF 特征描述子凭借其高效和紧凑的特性具有明显的速度优势，但正如前文所说，BRIEF 描述子对噪声具有较高的敏感性。噪声是图像处理中常见的问题，它可能来源于图像采集设备的物理限制、环境干扰或者传输过程中的误差。这些噪声可能改变图像中的像素灰度值，进而影响到特征点的检测和描述。对于 BRIEF 来说，噪声的存在可能导致随机选择的点对灰度值比较结果的不稳定，进而影响到最终生成的二进制特征描述符的准确性。针对此难题，本研究采用高斯滤波技术对图像进行预先处理操作。高斯滤波其基本原理是通过卷积操作平滑处理图像。本次研究中对图像的每个像素及其邻域进行高斯核的卷积操作，从而实现将每个像素值替换为其邻域内所有像素的平均值，该平均值根据邻域内每个像素的权值不同进行计算。由于高斯核的形状特性，靠近中心位置的像素会被赋予更高的权重，而距离中心较远的像素权重则逐渐减小，这种加权平均的方式使得图像中的噪声得到平滑处理，同时保留了图像的主要结构和边缘信息。通过以上操作，图像中的每个像素值被替换为其邻域内像素根据权值不同计算出来的平均值，从而有效地平滑图像并减少噪声。这种降噪处理不仅保留了图像的主要结构和特征，还在一定程度上缓解了 BRIEF 特征描述子对噪声的敏感问题，提升了特征点检测与跟踪的准确性和稳定性，为后续的特征匹配和目标识别提供了更坚实的基础。

2.3 双目相机特征点匹配

如图 2-2 所示，首先，图像被双目相机获取，接着系统会进行角点检测和特征描述。通过以上操作，对于双目相机传入的两幅图像，只要角点的特征描述相似，本研究就可以认为它们是匹配的。之后系统处理这些立体特征点。立体特征点由三个坐标 $X_s = (u_L, v_L, u_R)$ 定义，其中 (u_L, v_L) 是左图像上的坐标， u_R 是右图像中的坐标。

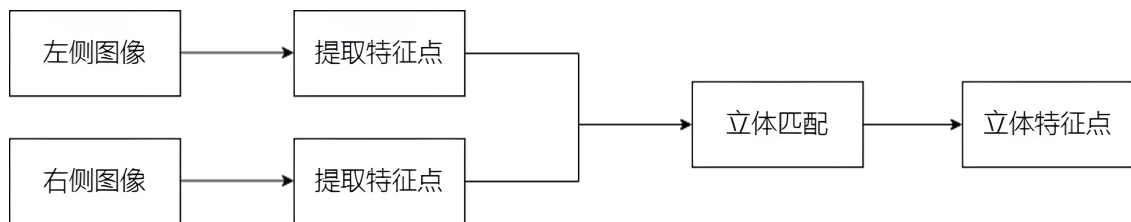


图 2-2 特征点流程图

本研究旨在从两个图像中提取特征点，并在右侧图像中对每个左侧特征点进行匹配搜索。这一过程依赖于特征点描述的相似性，以精确匹配两个图像中的特征点。匹配完成后，通过左侧 ORB 坐标和右侧匹配坐标生成立体特征点，这些点包含了二维图像信息，同时还融合了深度信息，从而为本研究提供了更全面的场景理解。由于双目相机的特殊结构，研究中可以单帧获取深度特征点信息，不必像单目相机那样从运动初始化获取结构信息。系统启动时，需要创建一个关键帧，本次研究中为第一帧图像，同时将原点设置为第一帧图像中的姿势，获得其初始位置后从所有三维特征点搭建地图。

2.4 实验验证

2.4.1 实验设计

本部分首先利用公开的 KITTI 数据集，对所提出的特征点检测与跟踪方法进行了全面评估。本研究利用该数据集进行了检测与跟踪特征点的实验，来评估所提方法的性能和优劣。

其次，本研究利用双目相机 D435i，在不同环境条件下，包括室内和室外环境，对特征点的检测与跟踪进行了进一步的验证。双目相机 D435i 具有较高的分辨率和精度，能够捕捉到更丰富的场景信息，从而更好地评估特征点检测与跟踪算法在实际应用中的表现。室内环境和室外环境的不同光照条件、背景干扰等因素都会对算法的性能产生影响，因此本研究通过在不同环境下进行的实验，以更全面地评估所提出的方法的鲁棒性和适用性。

2.4.2 实验结果及分析

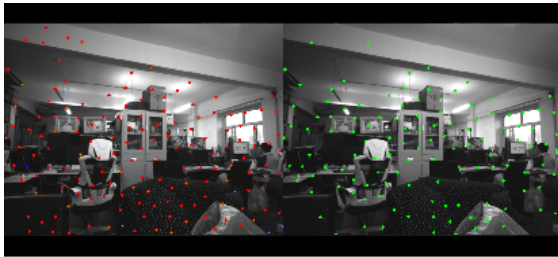
根据上述实验设计进行实验，最终特征点的检测与跟踪效果如图 2-3 和图 2-4 所示。图 2-3 为通过使用 FAST 算法和 Harris 角点检测算法处理 KITTI 数据集中的的一张图像，可以观察到它们对特征点的检测与跟踪效果都很显著。图 2-4(a)展示了在室内环境下双目相机的特征点检测结果，由于室内物体较多，相机能够识别和检测的物体较多，因此特征点的识别和跟踪效果十分出色。图 2-4(b)中为无人机上的双目相机 D435i 特征点检测与跟踪的效果图，由于无人机在空中飞行高度较高，

可供识别与检测的物体较少，因此在无人机上双目相机的朝向为竖直向下，以便双目相机检测和跟踪到更多地特征点。从图中识别效果来看，室外环境特征点检测与跟踪的效果良好，且与室内环境特征点检测与跟踪效果相比，二者检测与跟踪效果几乎相同。

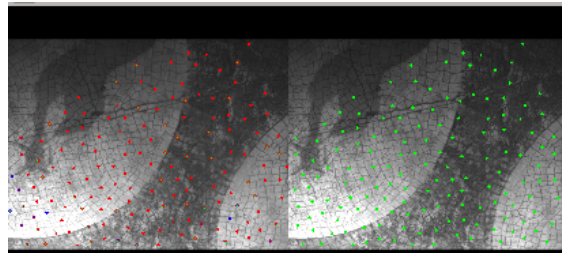
根据上述实验结果和分析，可以得出本章研究的基于 ORB 算法的特征点检测与跟踪方法具备稳定性和有效性，适用于后续无人机飞行中特征点的检测与跟踪，为飞行定位提供重要支持。



图 2-3 KITTI 数据集上检测到的特征点



(a) 室内环境检测到的特征点



(b) 室外环境检测到的特征点

图 2-4 特征点检测

2.5 本章小结

本章针对特征点的检测与跟踪，提出了基于 ORB 算法的检测与跟踪方法，主要包括角点检测和特征点描述两个阶段。本章提出了将 Harris 和 FAST 角点检测算法结合，以提升角点检测的精确性和鲁棒性。提出了构建高斯金字塔的方法使其具备尺度不变性；提出了利用灰度质心法使其具备旋转不变性。在特征点描述中，本章提出了利用高斯滤波的方法进行降噪处理，在一定程度上缓解了 BRIEF 特征描述子对噪声的敏感问题。最后，在双目相机中利用左侧 ORB 的坐标和右侧匹配的坐标来生成立体特征点，从而实现了特征点的检测与跟踪，为后续章节中双目视觉惯性里程计的研究提供了不可或缺的重要数据。

第 3 章 基于扩展卡尔曼滤波的双目视觉惯性里程计

本章主要阐述了双目视觉惯性里程计的构建方法。首先估计无人机飞行状态，然后对待融合的传感器因子进行数学分析，最终利用边缘化和非线性优化处理融合数据，以构建精确且鲁棒的双目视觉惯性里程计。同时，本章节中针对双目视觉惯性里程计设计了一套实验方案，用于检验本章构建出的双目视觉惯性里程计的有效性。

3.1 无人机状态估计

3.1.1 状态定义

首先需估计的状态为无人机中心的三维位置和姿态，相机的三维位置或视觉路标深度，以及 IMU 的变加速度计偏置。所以对视觉和 IMU 进行如下定义：

$$\begin{aligned} \mathbf{X} &= [\mathbf{p}_0, \mathbf{R}_0, \mathbf{p}_1, \mathbf{R}_1, \mathbf{L}, \mathbf{p}_n, \mathbf{R}_n, \mathbf{x}_{cam}, \mathbf{x}_{imu}] \\ \mathbf{x}_{cam} &= [\lambda_0, \lambda_1, \mathbf{L}, \lambda_l] \\ \mathbf{x}_{imu} &= [\mathbf{v}_0, \mathbf{b}_{a_0}, \mathbf{b}_{g_0}, \mathbf{v}_1, \mathbf{b}_{a_1}, \mathbf{b}_{g_1}, \mathbf{L}, \mathbf{v}_n, \mathbf{b}_{a_n}, \mathbf{b}_{g_n}] \end{aligned} \quad (3-1)$$

其中， $\mathbf{p}$ 对应于世界坐标系下的无人机位置， $\mathbf{R}$ 对应于世界坐标系下的机体姿态。 $\mathbf{x}_{cam}$ 为与相机相关的状态， λ 为特征点的深度距离。由加速度计偏置 $\mathbf{b}_a$ ，无人机飞行速度 $\mathbf{v}$ 和 IMU 偏置 $\mathbf{b}_g$ 组成的 $\mathbf{x}_{imu}$ 为与陀螺仪相关的变量，假设通过离线标定得出的 IMU 中心到机体中心的其次变换矩阵一定，为了便于公式表达，之后内容都用陀螺仪中心代替无人机机身中心。

3.1.2 代价函数

状态估计旨在最优地估计机器人的当前状态，包括位置和姿态。这涉及解决一个最大似然估计（MLE）问题^[37]。MLE（maximum likelihood estimation）是一项概率统计方法，利用观测数据对模型参数进行估计，以使观测数据出现的概率最大化。在机器人的状态估计中，MLE 问题通常涉及一段时间内机器人位姿的联合概率分布。这一联合概率分布涵盖了机器人在不同时间点可能的位置和姿态信息，以及它们之间的关联。通过 MLE 方法，本研究可以根据观测数据（如传感器读数、视觉特征点等）来推断机器人的最可能状态序列。在所有测量数据是独立的假设下，问题通常推导为：

$$\mathbf{X}^* = \arg \max \prod_{t=0k \in S}^n p(\mathbf{z}_t^k | \mathbf{X}) \quad (3-2)$$

其中 $\mathbf{S}$ 为一组测量，该资料源自于相机、IMU 以及其他感知器。为了更好地进行数学处理，本研究将测量误差按照高斯分布 $p(\mathbf{z}_t^k, \mathbf{X}) \sim \mathcal{N}(\bar{\mathbf{z}}_t^k, \Omega_t^k)$ 进行处理。因此，式(3-2)的代价函数可以写成：

$$\begin{aligned} \mathbf{X}^* &= \underset{\mathbf{X}}{\operatorname{argmax}} \prod_{t=0}^n \prod_{k \in \mathbf{S}} \exp\left(-\frac{1}{2} \|\mathbf{z}_t^k - h_t^k(\mathbf{X})\|_{\Omega_t^k}^2\right) \\ &= \underset{\mathbf{X}}{\operatorname{argmin}} \sum_{t=0}^n \sum_{k \in \mathbf{S}} \|\mathbf{z}_t^k - h_t^k(\mathbf{X})\|_{\Omega_t^k}^2 \end{aligned} \quad (3-3)$$

因此，状态估计被转化为了非线性最小二乘问题，又称捆绑调整（BA）。捆绑调整（Bundle Adjustment）^[38]是三维重建和计算机视觉中常用的优化技术，旨在优化相机参数（如内外参数）和三维空间点的位置，以降低重投影误差。具体来说，捆绑调整首先将所有的观测数据（包括图像特征点的像素坐标和相机的内外参数）进行整理，形成一个大的优化问题。接着，采用 Levenberg-Marquardt 算法等非线性优化方法，通过迭代求解优化问题，逐步调整相机参数和三维点坐标，直至达到满足条件的收敛状态。在本节双目视觉里程计的研究中，会通过多个视角的图像来确定当前相机所处的位置。这些图像之间会存在大量的同名特征点，即在不同图像中都能识别到的同一物理点。捆绑调整的目标就是对这些特征点的立体空间坐标以及相机的内参和外参进行全局优化，使得这些特征点在立体空间中的投影与它们在图像中的位置尽可能一致。

3.2 传感器因子

3.2.1 相机因子

每个摄像机的内参和不同摄像机之间的外参变换都能通过标定获得，这些参数的准确性直接影响到从图像中重建的场景三维结构的准确性。在图像中，角点特征稳定性高，在不同视角和光照条件下仍能保持相对稳定。

为了跟踪这些特征点在连续帧之间的变化，本研究采用了 KLT 跟踪器^[39]。基于光流法的 KLT 跟踪器，可精准追踪当前帧的特征点，保持高效率 and 精度，尤其在复杂环境下。双目相机系统需要匹配左右图像的特征，确保两侧图像的特征点对应。通过双目匹配可以进一步确定特征点的三维位置，进而获得特征点的轨迹来构建相机因子。具体而言，相机因子实际上涉及到一项重要的工作：重投影。这个过程起始于对某个特征的首次观测。一旦本研究捕获到了这个特征在某一帧中的确切位置，就可以利用已知的相机参数和三维空间模型，尝试将这个特征投影到后续的帧中。这是一个逆向过程，将立体笛卡尔坐标空间的点映射回二维图像平面。在实际运行过程中，由于各种因素（如相机抖动、光照变化、遮挡等）

的影响，实际观测到的特征位置与通过重投影得到的特征位置往往存在偏差。这个偏差就被称为重投影误差。重投影误差的大小直接反映了本研究当前相机参数和三维点位置的准确性。

为了减小这个误差，本次研究中通过迭代的方法对相机参数和立体空间中的位置进行优化。每次迭代中，本研究会根据当前的重投影误差对相机参数和三维点位置进行调整，然后重新计算误差。这个过程会不断重复，直到重投影误差达到规定的阈值范围内，或者优化算法达到收敛状态。通过以上迭代优化过程，本研究可以使得实际观测到的特征位置与通过重投影得到的特征位置更加接近，从而提升三维重建的精度和鲁棒性。

假设图像 i 的初始特征为 l ，其残差值在后续图像 t 中表示为：

$$\begin{aligned} \mathbf{z}_t^l - h_t^l(\mathbf{X}) &= \mathbf{z}_t^l - h_t^l(\mathbf{R}_i, \mathbf{p}_i, \mathbf{R}_t, \mathbf{p}_t, \lambda_l) \\ &= \begin{bmatrix} u_t^l \\ v_t^l \end{bmatrix} - \pi_c(\mathbf{T}_c^{b^{-1}} \mathbf{T}_t^{-1} \mathbf{T}_i \mathbf{T}_c^b \pi_c^{-1}(\lambda_l, \begin{bmatrix} u_i^l \\ v_i^l \end{bmatrix})) \end{aligned} \quad (3-4)$$

其中， $\begin{bmatrix} u_i^l \\ v_i^l \end{bmatrix}$ 为在图像 i 中的初始特征 l 。后续在图像 t 中的残差值为 $\begin{bmatrix} u_t^l \\ v_t^l \end{bmatrix}$ 。 π_c 和 π_c^{-1} 为针孔相机中使用的 Projection-Backprojection Transformation（投影-反投影变换矩阵）。 $\mathbf{T}$ 为 4×4 的齐次变换矩阵： $\begin{bmatrix} \mathbf{R} & \mathbf{p} \\ 0 & 1 \end{bmatrix}$ 。为了简化表示，在这里忽略了一些齐次项。无人机到双目相机的外参变换表示为 $\mathbf{T}_c^b$ ，它通过离线标定获取。在像素坐标系下用 Ω_i^l 表示投影误差，它可以通过相机的内参标定进行获取。

该因子适用于双目相机的左右目，因此可以不受时间何空间的限制，将左图像中的特征映射到任一侧图像中。

3.2.2 IMU 因子

IMU 惯性测量计包含用于测量物体在三维笛卡尔立体空间的角速度和加速度变化的 x ， y ， z 轴的陀螺仪和 x ， y ， z 轴的加速度计，以估计物体在空间中变化的位姿。但是受到噪声的影响，如式(3-5)、(3-6)所示，IMU 惯性测量计的数据偏差较大，通常在真实测量值的基础上，利用添加了高斯白噪声和具有一定偏置和随机噪声的测量数据来标识旋转角速度和加速度的测量值。式(3-5)、(3-6)中 w 定义为角速度的真实值， w_m 定义为 D435i 相机自带 IMU 的旋转角速度， a 定义为加速度的真实值， a_m 定义为 D435i 相机自带 IMU 的加速度， b 定义为陀螺仪的随机噪声， n 定义为陀螺仪的高斯正态分布噪声。

$$w_m = w + b_w + n_w \quad (3-5)$$

$$a_m = a + b_a + n_a \quad (3-6)$$

在立体笛卡尔空间中，刚体运动的位置、速度、旋转等参数随时间的变化而变化，可以用数学语言表示出来，如式(3-7)、(3-8)、(3-9)所示。其中式(3-7)表示在 t 时刻下，无人机机载 IMU 在立体笛卡尔空间中的位置 p_{wb_t} 对时间的导数，式(3-8)表示无人机机载 IMU 载立体笛卡尔空间中的速度 v_t^w 对时间的导数，式(3-9)表示无人机机载 IMU 载立体笛卡尔空间中的四元数 q_{wb_t} 对时间的导数。其中 D435i 相机自带 IMU 加速度用测量真实值 a 来表示，旋转角速度用测量真实值 w 来表示。

$$\dot{p}^{wb_t} = v_t^w \quad (3-7)$$

$$\dot{v}_t^w = a_t^w \quad (3-8)$$

$$\dot{q}_{wb_t} = q_{wb_t} \otimes \begin{bmatrix} 0 \\ \frac{1}{2} w^{b_t} \end{bmatrix} \quad (3-9)$$

其中 $\otimes$ 为克罗内克积，下同。

因此，可以使用拉格朗日中值定理计算任一时间段内无人机机载 IMU 累计变换。如式(3-10)、(3-11)、(3-12)所示， $k+1$ 时刻的状态值可由 k 时刻的状态值、测量值以及 $k+1$ 时刻的测量值求解得到。如式(3-13)、(3-14)所示，D435i 相机自带 IMU 加速度 k 时刻的平均变化值定义为 a ， w 表示 $k+1$ 时刻的 D435i 相机自带 IMU 旋转角速度平均的变化值。

$$p_{wb_{k+1}} = p_{wb_k} + v_t^w \Delta t + \frac{1}{2} a \Delta t^2 \quad (3-10)$$

$$v_{k+1}^w = v_t^w + a \Delta t \quad (3-11)$$

$$q_{wb_{k+1}} = q_{wb_k} \otimes \begin{bmatrix} 0 \\ \frac{1}{2} w \Delta t \end{bmatrix} \quad (3-12)$$

$$a = \frac{1}{2} [q_{wb_k} (a^{b_k} - b_{a_k}) - g^w + q_{wb_{k+1}} (a^{b_{k+1}} - b_{a_k}) - g^w] \quad (3-13)$$

$$w = \frac{1}{2} [(w^{b_k} - b_{w_k}) + (w^{b_{k+1}} - b_{w_{k+1}})] \quad (3-14)$$

D435i 相机惯性测量单元可便捷测量物体运动中的各项数据，然后通过运算得出物体状态的具体变化过程。然而由于 IMU 的测量过程会不可避免地产生一些随机噪声和高斯噪声，随着飞行时间的延长，其计算的输出容易产生漂移过大的问题，因此通常需融合 IMU 传感器测量和其它传感器，以获得更准确的估计。在此

研究中，本研究进行 IMU 预积分操作，即对连续相邻帧图像间的 IMU 积分增量进行计算，对齐视觉数据与 IMU 数据。原因之一是 IMU 传感器相较于图片帧率来说具有较高的测量频率，采用预积分操作可以将较低帧率的视觉图像在每一个时间点上都与一个预积分 IMU 增益相匹配；原因之二在于预积分与初始时刻的位姿毫无关系，因为它只反映了相邻视觉帧之间的 IMU 测量值的变化情况，所以当更新后端非线性优化的状态量时，得益于两个相邻的视觉帧较短时间间隔的特性，预积分增量的值不会发生改变，避免了累计计算过程中误差较大的零点漂移问题。其详细计算如下所示。

$$p_{wb_j} = p_{wb_i} + v_i^w \Delta t - \frac{1}{2} g^w \Delta t^2 + q_{wb_i} \iint_{t \in [i, j]} (q_{b_i b_t} a^{b_t}) \delta t^2 \quad (3-15)$$

$$v_j^w = v_i^w - g^w \Delta t + q_{wb_i} \int_{t \in [i, j]} (q_{b_i b_t} a^{b_t}) \delta t \quad (3-16)$$

$$q_{wb_j} = q_{wb_i} \int_{t \in [i, j]} q_{b_i b_t} \otimes \begin{bmatrix} 0 \\ \frac{1}{2} w^{b_t} \end{bmatrix} \delta t \quad (3-17)$$

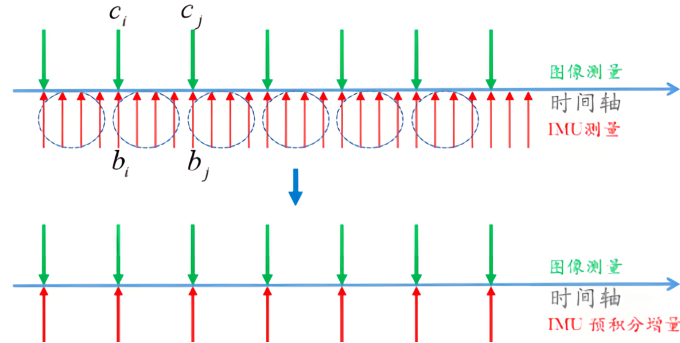


图 3-1 IMU 预积分增量大致过程图

如公式(3-15)、(3-16)、(3-17)和图 3-1 所示，两个相邻的视觉帧 c_i ， c_j 中间包含了 $b_i \dots b_j$ 多个 IMU 测量数据帧， p_{wb_i} 表示 i 时刻下 IMU 在世界坐标参考系中的定位矩阵、 v_i^w 表示在 i 时刻 IMU 在世界坐标系中的速度、 q_{wb_i} 表示 i 时刻下 IMU 在世界坐标系中的四元数，公式(3-15)表示参考 i 时刻 IMU 的位置求解出 j 时刻 D434i 相机自带的 IMU 在世界坐标参考系中的位置 p_{wb_j} ，(3-16)表示参考 i 时刻 IMU 的速度求解出 j 时刻 D434i 相机自带的 IMU 在世界坐标参考系下的速度 v_j^w ，(3-17)表示参考 i 时刻 IMU 的四元数为参考求解出 j 时刻 D434i 相机自带的 IMU 在世界坐标参考系下的四元数 q_{wb_j} 。其中 i ， j 时刻的预积分增益分别用 $\alpha_{b_i b_j}$ 、 $\beta_{b_i b_j}$ 、 $\gamma_{b_i b_j}$ 表示，详细公式如(3-18)、(3-19)、(3-20)所示。

$$\alpha_{b_i b_j} = \iint_{t \in [i, j]} (q_{b_i b_t} a^{b_t}) \delta t^2 \quad (3-18)$$

$$\beta_{b_i b_j} = \int_{t \in [i, j]} (q_{b_i b_t} a^{b_t}) \delta t \quad (3-19)$$

$$\gamma_{b_i b_j} = \int_{t \in [i, j]} q_{b_i b_t} \otimes \begin{bmatrix} 0 \\ \frac{1}{2} w^{b_t} \end{bmatrix} \delta t \quad (3-20)$$

为了便于数学运算，用离散空间内的状态量近似化简上述公式，则连续 D435i 相机自带的 IMU 旋转可表示为式(3-22)，速度表示为式(3-25)，位置表示为式(3-24)。公式(3-21)、(3-22)中 a 表示以 i 时刻作为起始时刻，从 k 时刻到 $k+1$ 时刻的相机自带 IMU 加速度的测量均值， w 表示以 i 时刻作为起始时刻，从 k 时刻和 $k+1$ 时刻的相机自带 IMU 旋转角速度的测量均值。

$$w = \frac{1}{2} [(w^{b_k} - b_{w_k}) + (w^{b_{k+1}} - b_{w_{k+1}})] \quad (3-21)$$

$$\gamma_{b_i b_{k+1}} = \gamma_{b_i b_k} \otimes \begin{bmatrix} 1 \\ \frac{1}{2} w \delta t \end{bmatrix} \quad (3-22)$$

$$a = \frac{1}{2} [q_{b_i b_k} (a^{b_k} - b_{a_k}) + q_{b_i b_{k+1}} (a^{b_{k+1}} - b_{a_k})] \quad (3-23)$$

$$\alpha_{b_i b_{k+1}} = \alpha_{b_i b_k} + \beta_{b_i b_k} \delta t + \frac{1}{2} \alpha \delta t^2 \quad (3-24)$$

$$\beta_{b_i b_{k+1}} = \beta_{b_i b_k} + a \delta t \quad (3-25)$$

3.3 边缘化

处理连续帧图像数据时，随着时间推移，需要追踪的状态数量（比如相机姿态、三维点位置等）会逐渐增多。这导致优化问题规模逐步扩大，从而使得计算复杂度爆炸式增长。这种增长趋势不仅加大了计算资源的消耗，还可能影响到实时处理的性能，在无人机飞行过程中会对无人机的定位产生较大干扰。为了降低计算复杂度，同时又不损失含有价值的信息，本次研究工作中引入了边缘化技术。边缘化是优化问题中处理信息冗余、降低计算量的有效手段。在三维重建和计算机视觉中，边缘化过程实际上是将先前的测量数据转化为一个先验项，这个先验项包含了过去帧中的有用信息，但不再作为独立的变量参与优化。具体来说，边缘化过程涉及到将先前帧中的状态变量从优化问题中移除，同时保留这些状态变量与当前帧中状态变量之间的约束关系。这些约束关系被转化为一个先验项，以

矩阵的形式嵌入到新的优化问题中。通过这种方式，本研究能够在不丢失有用信息的前提下，显著减小优化问题的规模，从而降低计算复杂度。

在本次优化过程中，定义 $\mathbf{X}_m$ 为需要边缘化处理的无人机因子， $\mathbf{X}_r$ 表示为剩余的因子。通过累加所有的边缘化因子，得到一个新的 $\mathbf{H}$ 和 $\mathbf{b}$ 。在重新调整这些因子的顺序之后，获得如下关系：

$$\begin{bmatrix} \mathbf{H}_{mm} & \mathbf{H}_{mr} \\ \mathbf{H}_{rm} & \mathbf{H}_{rr} \end{bmatrix} \begin{bmatrix} \delta \mathbf{X}_m \\ \delta \mathbf{X}_r \end{bmatrix} = \begin{bmatrix} \mathbf{b}_m \\ \mathbf{b}_r \end{bmatrix} \quad (3-26)$$

使用舒尔补^[40]来执行边缘化，如下：

$$\underbrace{(\mathbf{H}_{mm} - \mathbf{H}_{mr} \mathbf{H}_{rr}^{-1} \mathbf{H}_{rm})}_{\mathbf{H}_p} \delta \mathbf{X}_r = \underbrace{(\mathbf{b}_m - \mathbf{H}_{mr} \mathbf{H}_{rr}^{-1} \mathbf{b}_r)}_{\mathbf{b}_p} \quad (3-27)$$

来获取剩余状态新的先验 $\mathbf{H}_p$ 、 $\mathbf{b}_p$ 。在不损失有用信息的前提下，将包含边缘化状态的元素转化为先验项。详细来说，就是在空间坐标系中存留无人机定位系统的十个相机帧。当传入新的关键帧时，边缘化与第一帧的状态有关的惯性因子和视觉因子。在使用贝叶斯法则^[41]获得有关当前状态的先验信息之后就能够计算后验，即：似然与先验的乘积： $p(\mathbf{X} | \mathbf{z}) \propto p(\mathbf{z} | \mathbf{X})p(\mathbf{X})$ ，将其转化为一个最大后验（MAP）问题^[42]。具体操作为边缘处理在时刻 m 之前的状态，将其转化为一个先验项。因此，最大后验问题改写成如下表达式：

$$\begin{aligned} \mathbf{X}_{m:n}^* &= \underset{\mathbf{X}_{m:n}}{\operatorname{argmax}} \prod_{t=m}^n \prod_{k \in \mathbf{S}} p(\mathbf{z}_t^k | \mathbf{X}_{m:n}) p(\mathbf{X}_{m:n}) \\ &= \underset{\mathbf{X}_{m:n}}{\operatorname{argmin}} \sum_{t=m}^n \sum_{k \in \mathbf{S}} \|\mathbf{z}_t^k - h_t^k(\mathbf{X}_{m:n})\|_{\Omega_t^k}^2 + (\mathbf{H}_p \delta \mathbf{X}_{m:n} - \mathbf{b}_p) \end{aligned} \quad (3-28)$$

3.4 非线性优化

在前面研究的视觉惯性距离测量仪中，使用 IMU 点获取 D435i IMU 传感器的帧图像和位置，并通过空间定位技术计算特征点在空间笛卡尔坐标系中的位置。但是，陀螺仪的运动模型和相机投影模型会受到误差的影响而无法正确调整，因此需要排除噪声误差的影响，利用非线性优化的方法将正确的模型重新构建出来。如式(3-29)所示，非线性优化的无人机因子包含参考路标的深度信息 λ_i 、任意一帧图像匹配上的无人机机体的状态参数 $[\mathbf{p}_{wb_k}, \mathbf{v}_k^w, \mathbf{q}_{wb_k}, \mathbf{b}_k^a, \mathbf{b}_k^g]$ 、以 IMU 坐标系为参考，相机的外参矩阵 $\mathbf{X}_{bc}$ 。将所有的优化变量放在一个集合 $\mathbf{X}$ 中，有：

$$\begin{aligned}
 \mathbf{X} &= [\mathbf{x}_0, \mathbf{x}_1 \dots \mathbf{x}_n, \mathbf{x}_{bc}, \lambda_0, \lambda_1, \dots \lambda_m] \\
 \mathbf{x}_k &= [\mathbf{p}_{wb_k}, \mathbf{v}_k^w, \mathbf{q}_{wb_k}, \mathbf{b}_k^a, \mathbf{b}_k^g], k \in [0, n] \\
 \mathbf{x}_{bc} &= [\mathbf{p}_{bc}, \mathbf{q}_{bc}]
 \end{aligned} \tag{3-29}$$

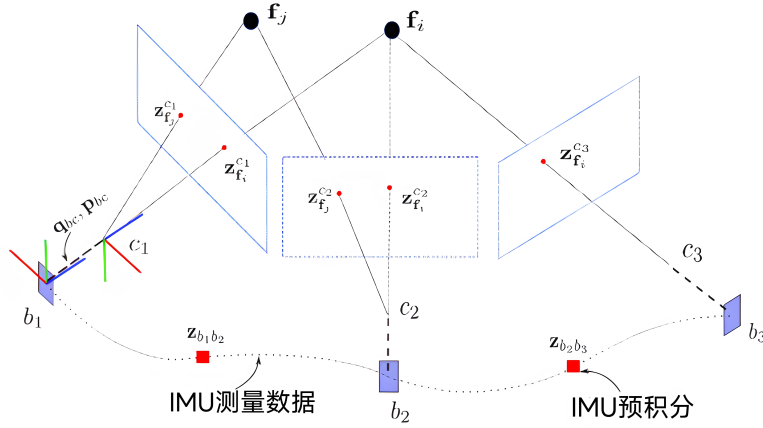


图 3-2 视觉里程计与 IMU 融合的位姿图

图 3-2 所示为视觉里程计与机载 IMU 紧耦合的非线性优化大致过程图，连续两帧图像之间有多帧 IMU 测量，每一帧图像都将双目相机位置 c_i 与无人机机体位置 b_i 进行匹配，对相邻帧图像间的陀螺仪数据进行预积分处理得预积分增益 $\mathbf{z}_{b_i b_j}$ ，并且同一个参考点能够在连续相邻的相机帧中被检测到，比如图中参考点 f_i 在自身的参考系中被标记为像素点 $\mathbf{z}_{f_i}^{c_1}$ 、 $\mathbf{z}_{f_i}^{c_2}$ 、 $\mathbf{z}_{f_i}^{c_3}$ 就是因为相机帧中被 c_1 、 c_2 、 c_3 检测到。上图中的节点是需要通过优化算法来计算的参数，边界是由求解问题的限制条件所产生的约束因子，可通过这些节点和边界建立 IMU 积分计算过程中的误差和视觉重投影过程的误差。使用高斯-牛顿迭代法^[43]、莱文伯格-马夸特迭代^[44]等方法进行优化，使得误差取得最小值并获取局部最优解。以下将说明如何构建 IMU 预积分误差和视觉重投影误差，同时也将说明如何利用高斯-牛顿法解决非线性优化问题。

视觉再投影误差是指在图像中观测到一个特征点 c_i ，在后续图像 c_j 中也被跟踪的情况下，由 c_i 再投影到标准化 c_j 平面，如式 3-30 所示，并将二者的值相减得到再投影偏差值，其计算公式如式(3-31)所示。

$$\begin{bmatrix} x_{c_j} \\ y_{c_j} \\ z_{c_j} \\ 1 \end{bmatrix} = \mathbf{T}_{bc}^{-1} \mathbf{T}_{wb_j}^{-1} \mathbf{T}_{wb_i} \mathbf{T}_{bc} \begin{bmatrix} \frac{1}{\lambda} u_{c_i} \\ \frac{1}{\lambda} v_{c_i} \\ \frac{1}{\lambda} \\ 1 \end{bmatrix} \quad (3-30)$$

$$\mathbf{r}_c = \begin{bmatrix} \frac{x_{c_j}}{z_{c_j}} - u_{c_j} \\ \frac{y_{c_j}}{z_{c_j}} - v_{c_j} \end{bmatrix} \quad (3-31)$$

在公式(3-30)、(3-31)中, (u_{c_i}, v_{c_i}) 表示标准相机平面下特征点的位置, $\lambda = 1/z_{c_i}$ 为标准相机平面下特征点的深度的逆距离, $\mathbf{T}_{bc}$ 表示以 IMU 为参考坐标系将相机坐标系转换到 IMU 坐标系的齐次矩阵, $\mathbf{T}_{wb_i}$ 、 $\mathbf{T}_{wb_j}$ 变换矩阵表示 c_i 、 c_j 在全局参考坐标系下的运动状态, 将特征点映射到普通摄像平面上, 则 $(x_{c_j}, y_{c_j}, z_{c_j})$ 为特征点的定位值, 再将特征点映射到标准摄像平面上, 则 (u_{c_j}, v_{c_j}) 为特征点的定位值, 将再投影偏差 r 的计算同时为二者的坐标值相减。对再投影偏差进行归一化处理, 化简得到标准成像平面 c_j 下的坐标 $(\frac{x_{c_j}}{z_{c_j}}, \frac{y_{c_j}}{z_{c_j}})$, 最后分解其次变换矩阵 $\mathbf{T} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$,

则公式(3-30)化简得到公式(3-32)。

$$\mathbf{f}_{c_j} = \begin{bmatrix} x_{c_j} \\ y_{c_j} \\ z_{c_j} \end{bmatrix} = \mathbf{R}_{bc}^T \mathbf{R}_{wb_j}^T \mathbf{R}_{wb_i} \mathbf{R}_{bc} \frac{1}{\lambda} \begin{bmatrix} u_{c_i} \\ v_{c_i} \\ 1 \end{bmatrix} + \mathbf{R}_{bc}^T (\mathbf{R}_{wb_j}^T ((\mathbf{R}_{wb_i} \mathbf{p}_{bc} + \mathbf{p}_{wb_i}) - \mathbf{p}_{wb_j}) - \mathbf{p}_{bc}) \quad (3-32)$$

然后计算待调整参数的再投影偏差的微分 $\mathbf{J}_c$, 如下所示:

$$\mathbf{J}_c = \begin{bmatrix} \frac{\partial \mathbf{r}_c}{\partial [\mathbf{p}_{wb_i}, \boldsymbol{\theta}_{wb_i}]} & \frac{\partial \mathbf{r}_c}{\partial [\mathbf{p}_{wb_j}, \boldsymbol{\theta}_{wb_j}]} & \frac{\partial \mathbf{r}_c}{\partial [\mathbf{p}_{bc}, \boldsymbol{\theta}_{bc}]} & \frac{\partial \mathbf{r}_c}{\partial \lambda} \end{bmatrix} \quad (3-33)$$

本研究的估计值是通过相邻两帧图像所经历的时间进行积分从而计算出陀螺仪在这段时间内的变化累计量, 测量值则是通过陀螺仪在实际运行中直接获取, 将二者的值作差, 即可得到陀螺仪的预积分误差, 如等式 (3-34) 所示。

$$\mathbf{r}_B = \begin{bmatrix} \mathbf{r}_p \\ \mathbf{r}_q \\ \mathbf{r}_v \\ \mathbf{r}_{ba} \\ \mathbf{r}_{bg} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{wb_i}^T (\mathbf{p}_{wb_j} - \mathbf{p}_{wb_i} - \mathbf{v}_i^w \Delta t + \frac{1}{2} \mathbf{g}^w \Delta t^2) - \mathbf{a}_{b_i b_j} \\ 2[\gamma_{b_i b_j}^{-1} \otimes (\mathbf{q}_{wb_i}^{-1} \otimes \mathbf{q}_{wb_j}^{-1})]_{xyz} \\ \mathbf{R}_{wb_i}^T (\mathbf{v}_j^w - \mathbf{v}_i^w + \mathbf{g}^w \Delta t) - \mathbf{b}_{b_i b_j} \\ \mathbf{b}_j^a - \mathbf{b}_i^a \\ \mathbf{b}_j^g - \mathbf{b}_i^g \end{bmatrix} \quad (3-34)$$

IMU 预积分误差可以用具有 15 个维度的向量 $\mathbf{r}_B = (\mathbf{r}_p \ \mathbf{r}_q \ \mathbf{r}_v \ \mathbf{r}_{ba} \ \mathbf{r}_{bg})^T$ 来表示，其中每个向量都对无人机机体位置有一定的干扰作用，为了便于分析，本研究将其进行分类， i 时刻下有两组，分别是 IMU 运行的速度矢量和偏置 $[\mathbf{v}_i^w \ \mathbf{b}_i^a \ \mathbf{b}_i^g]$ 以及该时刻下的位置和姿态 $[\mathbf{p}_{wb_i} \ \mathbf{q}_{wb_i}]$ ， j 时刻下也有两组，分别是 IMU 运行的速度矢量和偏置 $[\mathbf{v}_j^w \ \mathbf{b}_j^a \ \mathbf{b}_j^g]$ 以及该时刻下的位置和姿态 $[\mathbf{p}_{wb_j} \ \mathbf{q}_{wb_j}]$ 。于是将以上 15 个维度的向量通过求导的方式进行化简，用 Jacobi 矩阵 $\mathbf{J}_B$ 表示，如下式(3-35)所示。

$$\mathbf{J}_B = \begin{bmatrix} \frac{\partial \mathbf{r}_B}{\partial [\mathbf{p}_{wb_i}, \mathbf{q}_{wb_i}]} & \frac{\partial \mathbf{r}_B}{\partial [\mathbf{v}_i^w, \mathbf{b}_i^a, \mathbf{b}_i^g]} & \frac{\partial \mathbf{r}_B}{\partial [\mathbf{p}_{wb_j}, \mathbf{q}_{wb_j}]} & \frac{\partial \mathbf{r}_B}{\partial [\mathbf{v}_j^w, \mathbf{b}_j^a, \mathbf{b}_j^g]} \end{bmatrix} \quad (3-35)$$

随着系统运行时间的不断增加，双目相机所捕获的图像会不断增加，其按照时间顺序有序排列，由于双目相机捕获图像的增加，导致检测与识别到的特征点数量也随之增加，因此，在陀螺仪的约束方程中，系统有了足够多数量的特征点来计算预积分增益和误差，通过匹配相邻图像间的特征点，可以计算出两者图像间的预积分增益 $\mathbf{z}_{b_i b_j}$ ，再与这两者图像中无人机的机体状态量进行匹配，可以得到陀螺仪的预积分误差向量 $\mathbf{r}_B(\mathbf{z}_{b_i b_j}, \mathbf{x}_i, \mathbf{x}_j)$ ，同时由于每一帧图像检测识别到的特征点数量足够多，因此对于每一帧图像都可以利用误差公式构建一个误差 $\mathbf{r}_c(f_i, c_j)$ 。由于运行时间的增加，窗口也会随着时间的推移向后逐级检测双目相机所捕获的图像帧，并设置优化变量以获取初始信息和生成系统残差损失函数 $F(\chi)$ ，如表达式 (3-36) 所示。对当前滑动窗口中的所有项目进行加权，从而生成系统函数的最佳损失值，即 χ^* 为该系统的最优解。

$$\min_{\chi} \left\{ \|\mathbf{r}_p - \mathbf{J}_p \chi\| + \sum_{(l,j) \in C} \rho(\|\mathbf{r}_C(\mathbf{z}_{f_i}^{c_j} - \chi)\|_{P_{f_i}^{c_j}}^2) + \sum_{(i,j) \in B} \|\mathbf{r}_B(\mathbf{z}_{b_i b_j} - \chi)\|_{P_{b_j}^{b_i}}^2 \right\} \quad (3-36)$$

在优化的求解计算过程中，本研究采用一个非线性最小二乘求解问题来解决视觉里程计状态量的优化问题。如式(3-37)所示，需要找到一个使得损失函数 $F(\mathbf{x})$

具有局部最小值的 n 维向量 $\mathbf{x}^* \in \mathbf{R}^n$ ，其中 f_i 为残差函数，那么对于任意 $\|\mathbf{x} - \mathbf{x}^*\| < \delta$ 有 $F(\mathbf{x}^*) \leq F(\mathbf{x})$ 。

$$F(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^m (f_i(\mathbf{x}))^2 \quad (3-37)$$

令 $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}) \cdots f_m(\mathbf{x})]^T$ ，则损失函数可以用向量表示，如式(3-38)所示，每个损失函数相对于优化后的向量取微分，如式(3-39)所示。

$$F(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^m (f_i(\mathbf{x}))^2 = \frac{1}{2} \mathbf{f}^T(\mathbf{x}) \mathbf{f}(\mathbf{x}) \quad (3-38)$$

$$\mathbf{J} = \frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}} = \left[\frac{\partial f_1(\mathbf{x})}{\partial \mathbf{x}}, \frac{\partial f_2(\mathbf{x})}{\partial \mathbf{x}} \cdots \frac{\partial f_m(\mathbf{x})}{\partial \mathbf{x}} \right] \quad (3-39)$$

在式(3-40)中，对残差函数 $\mathbf{f}(\mathbf{x})$ 用一阶泰勒展开，并带回式(3-38)中，可得近似值如下所示：

$$\mathbf{f}(\mathbf{x} + \Delta \mathbf{x}) \approx \mathbf{f}(\mathbf{x}) + \mathbf{J} \Delta \mathbf{x} \quad (3-40)$$

$$\begin{aligned} F(\mathbf{x} + \Delta \mathbf{x}) &= \frac{1}{2} \mathbf{f}^T(\mathbf{x} + \Delta \mathbf{x}) \mathbf{f}(\mathbf{x} + \Delta \mathbf{x}) \\ &\approx \frac{1}{2} \mathbf{f}^T \mathbf{f} + \Delta \mathbf{x}^T \mathbf{J}^T \mathbf{f} + \frac{1}{2} \Delta \mathbf{x}^T \mathbf{J}^T \mathbf{J} \Delta \mathbf{x} \\ &= F(\mathbf{x}) + \Delta \mathbf{x}^T \mathbf{J}^T \mathbf{f} + \frac{1}{2} \Delta \mathbf{x}^T \mathbf{J}^T \mathbf{J} \Delta \mathbf{x} \end{aligned} \quad (3-41)$$

在等式（3-41）中，损失函数是与 $\Delta \mathbf{x}$ 相关的二次函数，如果雅可比矩阵是满秩的， $\mathbf{J}^T \mathbf{J}$ 则是正定阵，在这种情况下，函数具有在整个定义域内具有下边界，即从函数的最低点两端分别取极限逼近可以得到同一个值。因此，取方程（3-41）从两端逼近得到的下极限值，得到方程（3-42）所示的高斯-牛顿法迭代公式：

$$\mathbf{J}^T \mathbf{J} \Delta \mathbf{x} = -\mathbf{J}^T \mathbf{f} \quad (3-42)$$

有了 Gauss-Newton 迭代公式后，就可以采用数值分析的办法，循环往复计算 $\Delta \chi$ ，并利用公式 $\chi \leftarrow \chi + \Delta \chi$ 将求解出的新的状态值赋值给原先的状态值，在实际程序中通过 ceres 求解即可以求出相对应的最优解 χ^* 。

3.5 实验验证

3.5.1 实验设计

在本部分的实验中，本研究使用的设备为双目相机 D435i 及其集成的惯性测量单元 (IMU)。首先，本研究在室内环境中进行了测试。通过操作双目相机 D435i，本研究使其在室内空间中环绕一圈，并保证其起始位置和终点位置相同。在此过

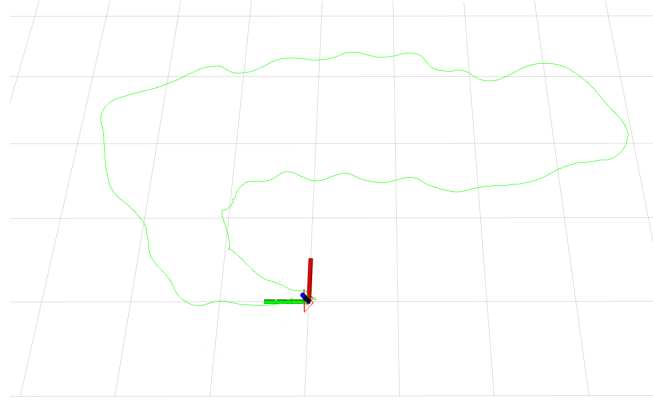
程中，相机不断捕捉周围环境的图像，并通过本章节所述的算法处理图像，从而实时估计出相机的位置和姿态。为了更直观地验证定位的准确性，本研究将实验结果导入到 Rviz（Robot Visualization Tool）这一强大的机器人可视化工具中。在 Rviz 中，本研究可以清晰地看到相机在三维空间中的移动轨迹。通过比较起点位置与终点位置相差的距离，本研究能够评估出定位系统的性能。如果偏差在可接受的范围内，那么说明本研究的双目相机和 IMU 组合在室内环境中能够实现准确的定位。

在户外条件下，本研究运用一架搭载双目相机 D435i 的无人机。该无人机具备更大的灵活性和移动速度，使本研究能够在更宽广的空间进行试验。本研究操纵无人机在天空中盘旋，确保其起点与终点位置一致。同时让双目相机不断捕捉周围的图像信息。同样地，本研究将室外环境下的实验结果也导入到 Rviz 中进行可视化。通过比较无人机起点与终点的位置偏差，本研究可以评估出定位系统在室外环境中的性能。如果偏差在可接受的范围内，那么说明本研究的双目相机和 IMU 组合在室外环境中也能够实现准确的定位。最终，本研究对比了室外环境和室内环境下的定位效果，以评估本研究章节中所探讨的算法对于无人机在广阔环境下的定位有效性。

3.5.2 实验结果及分析

D435i 相机在室内环境环绕一圈所得到的实验结果如图 3-3(a)所示。其中绿线为相机的运动轨迹图，从图中可以看出，相机环绕一圈后的末位置与初位置的距离差在误差范围允许内（两者基本重合），此双目视觉惯性里程计精确度和稳定性满足实验要求。图 3-3（b）所示为载有 D435i 相机的无人机在室外飞行一圈利用双目视觉惯性里程计定位轨迹图，从图中可以看出，在室外环境下，由于场地开阔导致可供识别的特征点数目减少，同时由于无人机飞行速度过快使其定位精度下降，最终无人机的起始位置和终点位置之间存在一定偏差。与室内环境相比，在室外环境中仅靠双目视觉惯性里程计无法为无人机提供高精度和稳定的定位。

综合以上结果和分析，可以得出在室外开阔场景下，本章节所研究的双目视觉惯性里程计仍具备较为准确的自主定位功能，但与室内环境相比定位精度和稳定性都有所下降，一定程度上说明了本章节所研究的双目视觉惯性里程计方法的有效性。



(a) 室内环境定位轨迹



(b) 室外环境定位轨迹

图 3-3 双目视觉惯性里程计

3.6 本章小结

本章针对基于滤波的双目视觉惯性里程计的研究，主要包含双目视觉里程计的构建以及双目视觉里程计与 IMU 数据进行融合这两个部分。在双目视觉里程计的构建中，本章利用 KLT 跟踪器和相机因子进行构建。在双目视觉里程计与 IMU 数据融合中，本章采用 IMU 预积分操作将视觉里程计数据和 IMU 数据进行对齐，采用 Gauss-Newton 法进行非线性优化，采用舒尔补进行边缘化，从而使得双目视觉惯性里程计具有较高的精确性和稳定性。通过实验验证，说明了在室外环境中仅靠双目视觉惯性里程计无法为无人机飞行提供精准、稳定的定位数据，因此在下一章中将 GPS 数据与双目视觉惯性里程计进行融合，以获得在开阔场景下精确性更高，鲁棒性更好的定位数据。

第 4 章 双目视觉惯性里程计与 GPS 的数据融合

由于仅依靠双目视觉惯性里程计无法为无人机飞行提供精准、稳定的定位数据，因此本章主要阐述了双目视觉惯性里程计与 GPS 数据融合的构建方法。首先对 GPS 定位的原理进行了简单阐述，其次将 GPS 数据和双目视觉惯性里程计的数据转换到同一坐标系下，然后通过 EKF 的方式进行数据融合，最后利用非线性最小二乘方法对融合后的数据进行优化。同时，本章节中针对双目视觉惯性里程计和 GPS 数据融合后的系统设计了一套实验方案，用于检验本章所提出融合方法的有效性。

4.1 GPS 定位原理

全球定位系统（GPS）是由地面站、接收机以及一系列分布在地球轨道上的卫星组成。这些卫星以地球为中心飞行，每颗卫星定期向地面发送包括卫星的准确位置和时间信息在内的信号。这些信号覆盖了地球上的几乎所有区域，从而使得无论在地球的哪个角落，用户都能够接收到至少四颗以上的卫星信号^[45]，如图 4-1 所示。通过接收来自不同卫星的信号，并结合卫星的位置信息和时间数据，GPS 接收机可以计算出用户的准确位置、速度和时间。

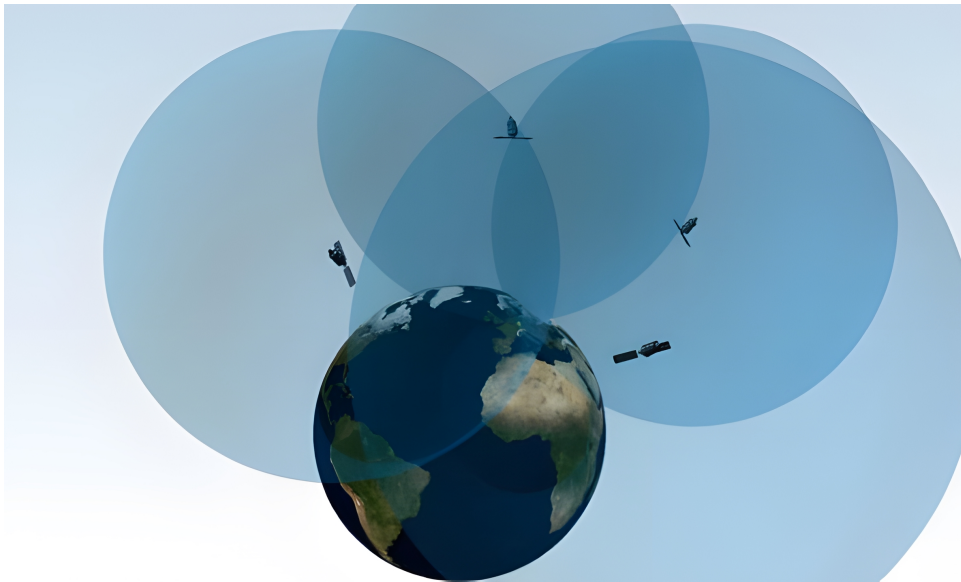


图 4-1 GPS 定位

在本研究中，GPS 数据可用于支持定位，通过无人机在开放空间的自主操作和良好的 GPS 信号，提高定位精度和稳定性。在三维空间中对无人机进行精确定位，需要依赖于至少四颗卫星的信号接收以确保位置信息的准确性和可靠性。这

一过程不仅要求接收机能够连续捕获并稳定跟踪每颗卫星的信号，还需要对接收到的信号进行精确处理和解码，以提取出用于定位的关键信息。其中在无人机的 GPS 定位技术中起着至关重要的作用的是伪距测量。伪距是指通过发送卫星信号的延迟计算出的接收器与卫星之间的近似距离。这一测量过程涉及复杂的数学模型和编码机制，在本次无人机的 GPS 定位研究中，设 t^j 表示卫星的钟面时刻，即真实发射信号时刻， t_{GPS}^j 表示第 j 颗卫星发出信号时的 GPS 标准时间， δt^j 表示卫星的时钟差， t^i 表示接收机的钟面时刻，即接收机真实接收信号时刻，对应卫星的接收信号时刻的 GPS 标准时间用 t_{GPS}^i 表示，接收机的时钟差用 δt^i 表示，因此可以表示卫星时间和接收时间：

$$\begin{aligned} t^j &= t_{GPS}^j + \delta t^j \\ t^i &= t_{GPS}^i + \delta t^i \end{aligned} \quad (4-1)$$

根据距离等于速度乘以时间的基本物理学公式，伪距的计算公式如下：

$$\tilde{p}_i^j(t) = C * (t^i - t^j) \quad (4-2)$$

设 p_i^j 表示卫星 j 在空间笛卡尔坐标系中的坐标与接收天线在空间笛卡尔坐标系中的坐标连线的距离，有：

$$p_i^j(t) = C * (t_{GPS}^i - t_{GPS}^j) \quad (4-3)$$

将(4-1)和(4-2)代入(4-3)，得到新的伪距表达式：

$$\tilde{p}_i^j(t) = p_i^j(t) + C * (\delta t^i - \delta t^j) \quad (4-4)$$

实际上，由于大气中的环境因素电离层干扰 δI_i^j 、对流层误差 δT_i^j 以及其它误差对于距离的计算有较大影响，所以在实际运算过程中需要加入一些系数进行修正处理。因为电离层、对流层都有相应的误差模型，可以计算具体的误差值，因此得到完整的伪距表达式：

$$\tilde{p}_i^j(t) = p_i^j(t) + C * (\delta t^i - \delta t^j) + \delta I_i^j + \delta T_i^j \quad (4-5)$$

如上所述，无人机定位以地球中心为基准点，假设地球中心是空间笛卡尔坐标系的原点，卫星和接收天线都位于这个巨大的空间笛卡尔坐标系中，在这种情况下，基于空间几何原理，可以根据卫星的坐标 $P^j(t)(X^j(t), Y^j(t), Z^j(t))$ 和无人机上 GPS 接收机的坐标 $P_i(X_i, Y_i, Z_i)$ 计算出卫星到无人机上接收机的空间几何距离。因此 p_i^j 也可以表达为：

$$p_i^j(t) = C * (t_{GPS}^i - t_{GPS}^j) = \sqrt{(X^j(t) - X_i)^2 + (Y^j(t) - Y_i)^2 + (Z^j(t) - Z_i)^2} \quad (4-6)$$

将(4-5)伪距表达式代入式(4-6)后，化简得到以下测码伪距观测方程：

$$\tilde{p}_i^j(t) = \sqrt{(X^j(t) - X_i)^2 + (Y^j(t) - Y_i)^2 + (Z^j(t) - Z_i)^2} + C * (\delta t^i - \delta t^j) + \delta I_i^j + \delta T_i^j \quad (4-7)$$

其中 X_i 、 Y_i 、 Z_i 、 δt^i 即使本研究需要求解的参数。假设其中的 X_i 、 Y_i 、 Z_i 都是线性方程，即：

$$\begin{aligned} X_i &= X_i^0 + \delta X_i \\ Y_i &= Y_i^0 + \delta Y_i \\ Z_i &= Z_i^0 + \delta Z_i \end{aligned} \quad (4-8)$$

其中，点 $P_i(X_i, Y_i, Z_i)$ 是无人机 GPS 接收器在笛卡尔空间坐标系中的真实坐标，点 (X_i^0, Y_i^0, Z_i^0) 表示无人机接收器在笛卡尔空间坐标系中位置的估算值，两者坐标分别对应相减即可得到误差值 $(\delta X_i, \delta Y_i, \delta Z_i)$ ，使用一阶泰勒展开，即线性展开后的 $p_i^j(t)$ 的公式如下：

$$\begin{aligned} p_i^j(t) &= f(X_i, Y_i, Z_i) \\ &= f(X_i^0, Y_i^0, Z_i^0) + f'_{X_i}(X_i^0, Y_i^0, Z_i^0)(X_i - X_i^0) \\ &\quad + f'_{Y_i}(X_i^0, Y_i^0, Z_i^0)(Y_i - Y_i^0) + f'_{Z_i}(X_i^0, Y_i^0, Z_i^0)(Z_i - Z_i^0) \end{aligned} \quad (4-9)$$

其中各导数的公式如下：

$$\begin{aligned} f'_{X_i}(X_i^0, Y_i^0, Z_i^0) &= \left(\frac{\partial p_i^j(t)}{\partial X_i} \right)_0 = -\frac{1}{(\rho_i^j(t))_0} (X^j(t) - X_i^0) = -l(t)^j \\ f'_{Y_i}(X_i^0, Y_i^0, Z_i^0) &= \left(\frac{\partial p_i^j(t)}{\partial Y_i} \right)_0 = -\frac{1}{(\rho_i^j(t))_0} (Y^j(t) - Y_i^0) = -m(t)^j \\ f'_{Z_i}(X_i^0, Y_i^0, Z_i^0) &= \left(\frac{\partial p_i^j(t)}{\partial Z_i} \right)_0 = -\frac{1}{(\rho_i^j(t))_0} (Z^j(t) - Z_i^0) = -n(t)^j \end{aligned} \quad (4-10)$$

将式(4-10)代入(4-9)化简，有：

$$p_i^j(t) = (p_i^j(t))_0 - l(t)^j \delta X_i - m(t)^j \delta Y_i - n(t)^j \delta Z_i \quad (4-11)$$

其中， $(p_i^j(t))_0$ 表示表示卫星 j 到点的几何距离，按照最小二乘法的原理，得到卫星 j 到点的几何距离误差方程如式(4-12)所示：

$$\begin{aligned} V_i^j(t) &= (p_i^j(t))_0 - l(t)^j \delta X_i - m(t)^j \delta Y_i - n(t)^j \delta Z_i \\ &\quad + C \delta t_i - C \delta t^j + \delta I_i^j(t) + \delta T_i^j(t) - \tilde{p}_i^j(t) \end{aligned} \quad (4-12)$$

其中 $(p_i^j(t))_0 - C\delta t^j + \delta I_i^j(t) + \delta T_i^j(t)$ 是通过现有参数计算出来的伪距值，用 $\tilde{p}_i^j(t)$ 是接收机真实获取的伪距值，令参数 $\tilde{\lambda} = (p_i^j(t))_0 - C\delta t^j + \delta I_i^j(t) + \delta T_i^j(t) - \tilde{p}_i^j(t)$ ，代入式(4-12)，有：

$$V_i^j(t) = -l(t)^j \delta X_i - m(t)^j \delta Y_i - n(t)^j \delta Z_i + C\delta t_i - \tilde{\lambda}_i^j(t) \quad (4-13)$$

因为无人机想要获取 GPS 数据至少需要接收 4 颗卫星数据，所以 $j \geq 4$ ，令 $n(n \geq 4)$ ，则有以下误差方程组：

$$\begin{aligned} V_i^1(t) &= -l(t)^1 \delta X_i - m(t)^1 \delta Y_i - n(t)^1 \delta Z_i + C\delta t_i - \tilde{\lambda}_i^1(t) \\ V_i^2(t) &= -l(t)^2 \delta X_i - m(t)^2 \delta Y_i - n(t)^2 \delta Z_i + C\delta t_i - \tilde{\lambda}_i^2(t) \\ &\dots \end{aligned} \quad (4-14)$$

$$V_i^n(t) = -l(t)^n \delta X_i - m(t)^n \delta Y_i - n(t)^n \delta Z_i + C\delta t_i - \tilde{\lambda}_i^n(t)$$

为了便于编程运算，将其转化为矩阵表达，如式(4-15)所示：

$$\begin{aligned} V &= Q \times \delta X - A \\ &= \begin{bmatrix} V_i^1 \\ V_i^2 \\ \vdots \\ V_i^n \end{bmatrix} = \begin{bmatrix} -l(t)^1 & -m(t)^1 & -n(t)^1 & C \\ -l(t)^2 & -m(t)^2 & -n(t)^2 & C \\ \vdots & & \dots & \vdots \\ -l(t)^n & -m(t)^n & -n(t)^n & C \end{bmatrix} \times \begin{bmatrix} \delta X_i \\ \delta Y_i \\ \delta Z_i \\ \delta t_i \end{bmatrix} - \begin{bmatrix} \tilde{\lambda}_i^1(t) \\ \tilde{\lambda}_i^2(t) \\ \vdots \\ \tilde{\lambda}_i^n(t) \end{bmatrix} \end{aligned} \quad (4-15)$$

设误差的平方和函数用 J 函数表示，根据最小二乘法计算出：

$$\begin{aligned} J &= V_i^1 * V_i^1 + V_i^2 * V_i^2 + \dots + V_i^n * V_i^n = [V_i^1 V_i^2 \dots V_i^n] * \begin{bmatrix} V_i^1 \\ V_i^2 \\ \vdots \\ V_i^n \end{bmatrix} = V^T * V \\ &= (Q * \delta X - A)^T * (Q * \delta X - A) = [(Q * \delta X)^T - A^T] * (Q * \delta X - A) \\ &= (\delta X^T * Q^T - A^T) * (Q * \delta X - A) \end{aligned} \quad (4-16)$$

求出 J 函数对 δX 偏导函数的零点，该零点即使误差函数取得下极限值的解，如式(4-17)所示：

$$\delta X = \frac{(Q^T A)}{(Q^T Q)} = (Q^T Q)^{-1} (Q^T A) \quad (4-17)$$

将 (X_i^0, Y_i^0, Z_i^0) 初始值设为 $(0, 0, 0)$ ，根据式(4-8)，多次迭代计算获取无人机位置的 GPS 定位数据。

本次研究中所用的 GPS 接收机为 pixhawk4 GPS MODULE，如图 4-2 所示。

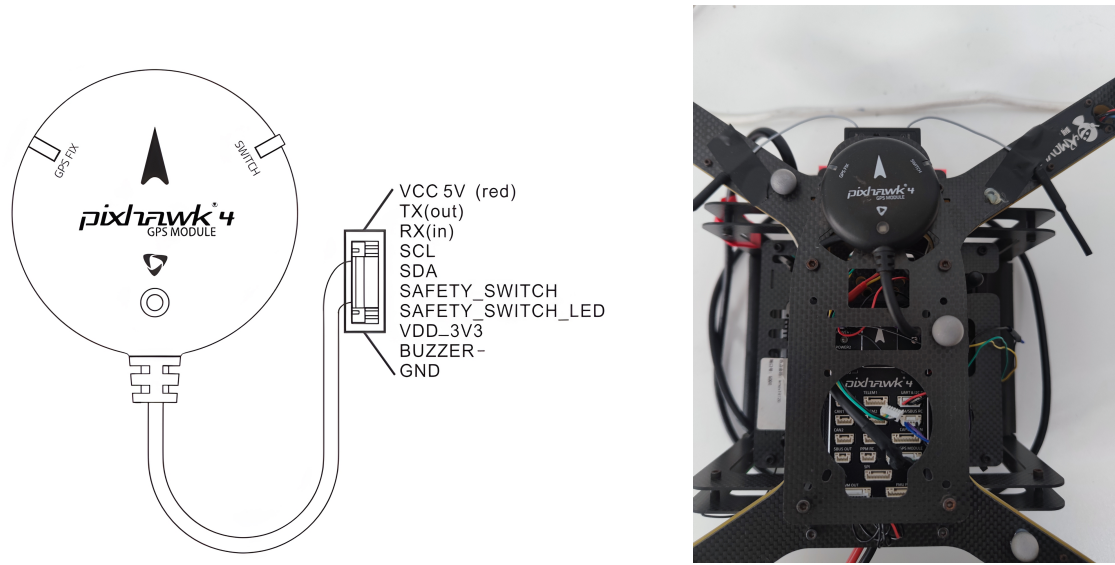


图 4-2 无人机 GPS 模块

4.2 双目视觉惯性里程计与 GPS 数据融合

4.2.1 坐标转换

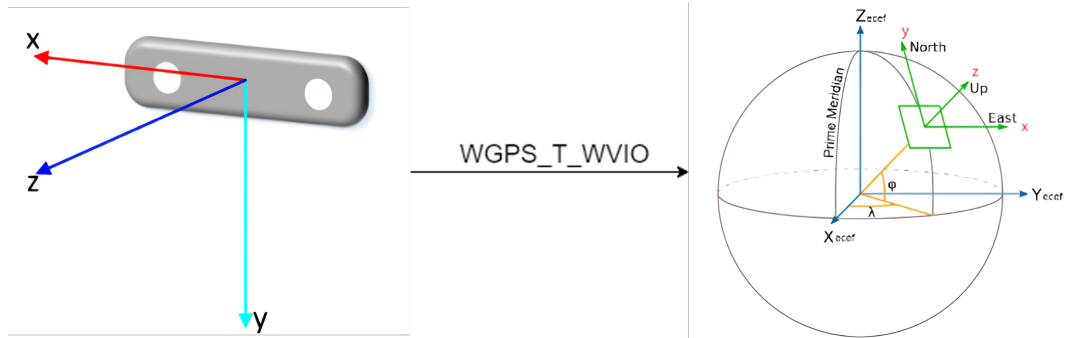


图 4-3 IMU 坐标系转换到 GPS 坐标系下

在之前章节的研究中，如图 4-3 所示，取得了双目视觉惯性里程计的数据和 GPS 数据，这两类定位数据分别以各自的坐标系作为参考，如上图 4-3 所示，其中左图为双目相机惯性里程计的 IMU 坐标系，右图为 GPS 的 ENU (Earth-North-Up) 坐标系。为了实现双目视觉惯性里程计和 GPS 数据的融合，本研究需要将这两种数据进行坐标转换，本节研究中采用的是将双目视觉惯性里程计变换到 ENU 坐标系下，并定义双目相机 IMU 坐标系为无人机的机体坐标系，而 ENU 坐标系则为全局定位和导航的参考坐标系。

在研究中，本研究首先将原始 GPS 数据变换为 ENU 坐标系下的坐标。这些原始 GPS 数据提供了无人机的经纬度和高度信息。接着，本研究以第一帧 GPS 数据作为坐标系的原点，通过一系列坐标转换操作，将通过双目视觉里程计获取到的

无人机位置转换到 GPS 坐标系下。在这个过程中，一个重要的转换矩阵起到了关键作用，即 $W_{GPS_T_WVIO}$ 矩阵。这个矩阵的作用是将双目视觉里程计系统获取到的无人机位置从 IMU 坐标系转换到 GPS 坐标系下。该坐标转换过程可以有效地结合双目视觉惯性里程计的数据和 GPS 数据，实现无人机位置的准确定位和导航。

4.2.2 基于扩展卡尔曼滤波的数据融合

本节数据融合的方式采用扩展卡尔曼滤波进行数据融合的过程中，以双目视觉惯性里程计的位置预测无人机在空中飞行时的状态，为了便于后续使用扩展卡尔曼滤波进行融合，所以在无人机机体坐标系下，本研究将其状态量定义为：

$$\mathbf{X} = (\mathbf{p}, \delta\boldsymbol{\theta}, \mathbf{v}_b, \delta\mathbf{b}_g, \delta\mathbf{b}_a)^T \quad (4-18)$$

在式(4-18)中，各物理量的含义表示如下，IMU 坐标系下无人机的飞行速度 $\mathbf{v}_b$ ，无人机飞行过程中由于转动产生的误差向量 $\delta\boldsymbol{\theta}$ ，全局坐标系下的机体坐标 $\mathbf{p}$ 以及 IMU 的偏移误差 $\delta\mathbf{b}_g$ 、 $\delta\mathbf{b}_a$ 。把式(4-18)对时间求导，得 $\mathbf{X}'$ ：

$$\begin{aligned} \mathbf{p}' &= \mathbf{R} \exp([\delta\boldsymbol{\theta}]^\wedge) \mathbf{v}_b \\ \delta\boldsymbol{\theta}' &= \exp([\delta\boldsymbol{\theta}]^\wedge) (\mathbf{w} - \delta\mathbf{b}_g - \mathbf{n}_g) \\ \mathbf{v}_b' &= -[(\mathbf{w} - \delta\mathbf{b}_g - \mathbf{n}_g)^\wedge \mathbf{v}_b + (\mathbf{R} \exp([\delta\boldsymbol{\theta}]^\wedge))^T \mathbf{g} + \mathbf{a} - \delta\mathbf{b}_a - \mathbf{n}_a] \\ \delta\mathbf{b}_g' &= \mathbf{r}_g \\ \delta\mathbf{b}_a' &= \mathbf{r}_a \end{aligned} \quad (4-19)$$

在式(4-19)中， $\mathbf{w}$ 表示机载相机 D435i 的 IMU 传感器的加速度， $\mathbf{a}$ 表示机载相机 D435i 的 IMU 传感器的旋转角速度测量值， $[\delta\boldsymbol{\theta}]^\wedge$ 为反对称矩阵，表示旋转向量的偏差， $\mathbf{n}_a$ 、 $\mathbf{n}_g$ 表示服从高斯分布的 IMU 测量噪声值， $\mathbf{b}_a$ 、 $\mathbf{b}_g$ 为 IMU 受到随机干扰影响的测量值，则上述无人机系统状态量在扩展卡尔曼滤波中的输入噪声可用如下公式表示：

$$\mathbf{W} = (\mathbf{n}_g, \mathbf{n}_a, \mathbf{b}_g, \mathbf{b}_a) \quad (4-20)$$

根据公式(4-19)和(4-20)，上述状态量的雅克比矩阵如式(4-21)所示：

$$\begin{aligned} \mathbf{J}_X &= \frac{\partial \mathbf{X}'}{\partial \mathbf{X}} dt + I_{15 \times 15} \\ \mathbf{J}_W &= \frac{\partial \mathbf{X}'}{\partial \mathbf{W}} dt \end{aligned} \quad (4-21)$$

本研究所研究的扩展卡尔曼滤波系统包括双目视觉惯性里程计和 GPS 数据两种信息。利用双目视觉惯性里程计进行预测，利用 GPS 数据更新状态，GPS 提供无人机在 ENU 坐标系的位置信息，双目视觉惯性里程计提供无人机姿态信息。双

目视觉惯性距离测量仪的发射频率为 30Hz, GPS 数据仅为 10Hz, 因此可以利用先前的观测数据和预测状态, 参考 stochastic cloning 方法^[46], 增加了延迟位姿 $\mathbf{p}_l$, $\delta\mathbf{\theta}_l$, 如式(4-22)所示。具体而言, 它将观测数据与预测状态对齐, 考虑它们之间的时间差, 然后在状态更新时进行修正, 以更准确地反映系统的当前状态。

$$\tilde{\mathbf{X}} = (\mathbf{X}^T, \mathbf{p}_l, \delta\mathbf{\theta}_l)^T \quad (4-22)$$

上式(4-22)为一个由系数矩阵和常数向量合并而成的表示无人机状态量的矩阵, 对 $\mathbf{p}_l$, $\delta\mathbf{\theta}_l$ 求导函数并解得该函数的零点值, 求得 $\tilde{\mathbf{X}}$ 的雅克比矩阵为:

$$\tilde{\mathbf{J}}_X = \begin{pmatrix} \mathbf{J}_X & \mathbf{0}_{15 \times 6} \\ \mathbf{0}_{6 \times 15} & \mathbf{I}_{6 \times 6} \end{pmatrix} \quad (4-23)$$

$$\tilde{\mathbf{J}}_W = \begin{pmatrix} \mathbf{J}_W \\ \mathbf{0}_{6 \times 12} \end{pmatrix} \quad (4-24)$$

最终扩展卡尔曼滤波系统的在预测阶段的协方差传递函数可以根据无人机状态量的增广矩阵表示为:

$$\tilde{\mathbf{P}}(k+\Phi \ k) = \tilde{\mathbf{J}}_X \tilde{\mathbf{P}}(k) \tilde{\mathbf{J}}_X^T + \tilde{\mathbf{J}}_W \tilde{\mathbf{Q}}(k) \tilde{\mathbf{J}}_W^T \quad (4-25)$$

在使用扩展卡尔曼滤波更新无人机空中飞行状态的过程中, 首先考虑结合视觉惯性里程计的六自由度相对位置。通过观察双目相机所捕获的视觉图像, 根据空间笛卡尔坐标系的相关理论, 利用相邻两个图像帧同一特征点的坐标变换情况可以计算相邻两帧图像中同一特征点的移动的距离 $\Delta\mathbf{p}_{vo}$ 和转动的角度大小 $\Delta\mathbf{\theta}_{vo}$ 。

同时, 将双目相机捕获到的图像与当前无人机的状态结合进行预测, 可以计算出相邻两个图像帧的距离变化和角度变化的预测增量:

$$\begin{aligned} \Delta\mathbf{p} &= \exp(-[\partial\mathbf{\theta}_l]^\wedge) \mathbf{R}_l^T (\mathbf{p} - \mathbf{p}_l) \\ \Delta\mathbf{\theta} &= -[\partial\mathbf{\theta}_l]^\wedge \mathbf{R}_l^T \mathbf{R} \exp(-[\partial\mathbf{\theta}]^\wedge) \end{aligned} \quad (4-26)$$

双目视觉惯性里程计的实际测量值代表了无人机飞行过程中真实测量的定位, 而预测增量值代表了通过扩展卡尔曼滤波计算出来的无人机飞行过程中位置的变化, 将二者作差, 同时用矩阵的形式表达出来, 有:

$$\mathbf{r} = \begin{pmatrix} \Delta\mathbf{p}_{vo} - \Delta\mathbf{p} \\ \Delta\mathbf{\theta}_{vo} - \Delta\mathbf{\theta} \end{pmatrix} \quad (4-27)$$

$$\mathbf{H}_{vo} = \begin{pmatrix} \mathbf{R}_l^T & \mathbf{0} & -\mathbf{R}_l^T & [\mathbf{R}_l^T (\mathbf{p} - \mathbf{p}_l)]^\wedge \\ \mathbf{0} & \mathbf{R}_l^T \mathbf{R} & \mathbf{0} & -\mathbf{I} \end{pmatrix} \quad (4-28)$$

如下式(4-29)所示, 其中 $\mathbf{K}$ 表示扩展卡尔曼滤波的增益 (Kalman Gain), $\tilde{\mathbf{X}}(k+1)$ 表示状态量更新, $\mathbf{P}(k+1)$ 表示协方差更新, 有:

$$\begin{aligned}
 \mathbf{K} &= \mathbf{P}(k+\Phi \ k) \mathbf{H}_{vo}^T ((\mathbf{P}(k+\Phi \ k) \mathbf{H}_{vo}^T + \mathbf{R})^{-1}) \\
 \mathbf{X}(k+1) &= \mathbf{X}(k) + \mathbf{K} \mathbf{r} \\
 \mathbf{P}(k+1) &= (\mathbf{I} - \mathbf{K} \mathbf{H}_{vo}) \mathbf{P}(k+\Phi \ k) (\mathbf{I} - \mathbf{K} \mathbf{H}_{vo})^T + \mathbf{K} \mathbf{R} \mathbf{K}^T
 \end{aligned} \tag{4-29}$$

在本文描述的无人飞机定位系统中，一旦机载相机的 IMU 基于可视里程表预测的条件变量被更新，就需要使用 GPS 信息来更新条件变量。GPS 读数可以观察无人机在 ENU 坐标系中的位置和速度，从而校正状态变量的整体一致性。GPS 测量的观测值如公式(4-30)所示，观测方程的雅可比矩阵如公式(4-31)所示。

$$\mathbf{Z}_{gps} = \begin{pmatrix} \mathbf{p} + \mathbf{R} \exp([\delta \boldsymbol{\theta}]^\wedge) \mathbf{t}_{bg} \\ \mathbf{R} \exp([\delta \boldsymbol{\theta}]^\wedge) (\mathbf{v}_b + [\mathbf{w} - \delta \mathbf{b}_g]^\wedge \mathbf{t}_{bg}) \end{pmatrix} \tag{4-30}$$

$$\mathbf{H}_{gps} = \begin{pmatrix} \mathbf{I} & -\mathbf{R}[\mathbf{t}_{bg}]^\wedge & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{R}[\mathbf{v}_b]^\wedge & \mathbf{R} & [\mathbf{t}_{bg}]^\wedge & \mathbf{0} \end{pmatrix} \tag{4-31}$$

式(4-32)表示 GPS 状态更新的公式，其中 GPS 更新的卡尔曼增益（Kalman Gain）用 $\mathbf{K}$ 表示，更新后的状态量用 $\mathbf{X}(k+1)$ 表示，GPS 观测数据与当前预测的状态量的误差用 $\mathbf{y}$ 表示，更新状态量之后的协方差用 $\mathbf{P}(k+1)$ 表示。

$$\begin{aligned}
 \mathbf{y} &= \mathbf{Z}_{gps} - \mathbf{H}_{gps} \mathbf{X} \\
 \mathbf{K} &= \mathbf{P}(k+1|k) \mathbf{H}_{gps}^T ((\mathbf{P}(k+1|k) \mathbf{H}_{gps}^T + \mathbf{R})^{-1}) \\
 \mathbf{X}(k+1) &= \mathbf{X}(k) + \mathbf{K} \mathbf{y} \\
 \mathbf{P}(k+1) &= (\mathbf{I} - \mathbf{K} \mathbf{H}_{gps}) \mathbf{P}(k+\Phi \ k) (\mathbf{I} - \mathbf{K} \mathbf{H}_{gps})^T + \mathbf{K} \mathbf{R} \mathbf{K}^T
 \end{aligned} \tag{4-32}$$

4.2.3 后端非线性优化

将 GPS 数据和双目视觉惯性里程计的数据通过扩展卡尔曼滤波的方式融合后，为了提高整个视觉里程计的定位精度与鲁棒性，本次研究还继续采用了非线性最小二乘的方式对 GPS 和双目视觉惯性里程计融合后的数据进行优化处理，定义 GPS 和双目 VIO 融合后的位置为 p_{gv} ，原有的 GPS 数据观测值为 p_g ，并采用欧氏距离作为残差函数，则二者之间的残差可以表示为：

$$r(x) = p_{gv} - p_g \tag{4-33}$$

之后，将残差函数代入非线性最小二乘优化中，利用公式(4-17)求出残差平方和的最小值，其中 N 为观测数据的数量， $r_i(x)$ 是第 i 个观测数据对应的残差。

$$\min_x \sum_{i=1}^N \|r_i(x)\|^2 \tag{4-34}$$

接下来，本研究需要计算残差函数关于参数 x 的雅可比矩阵 J_r ，以便进行迭代更新，如式(4-18)所示。

$$J_r = \begin{bmatrix} \frac{\partial r_1}{\partial x_1} & \frac{\partial r_1}{\partial x_2} & \cdots & \frac{\partial r_1}{\partial x_m} \\ \frac{\partial r_2}{\partial x_1} & \frac{\partial r_2}{\partial x_2} & \cdots & \frac{\partial r_2}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial r_N}{\partial x_1} & \frac{\partial r_N}{\partial x_2} & \cdots & \frac{\partial r_N}{\partial x_m} \end{bmatrix} \quad (4-35)$$

最后，采用 Gauss-Newton 法来更新参数 x ，以最小化残差的平方和。

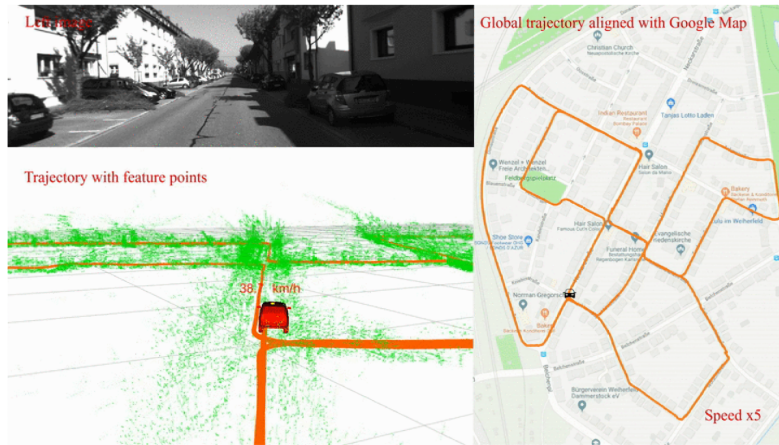
4.3 实验验证

4.3.1 实验设计

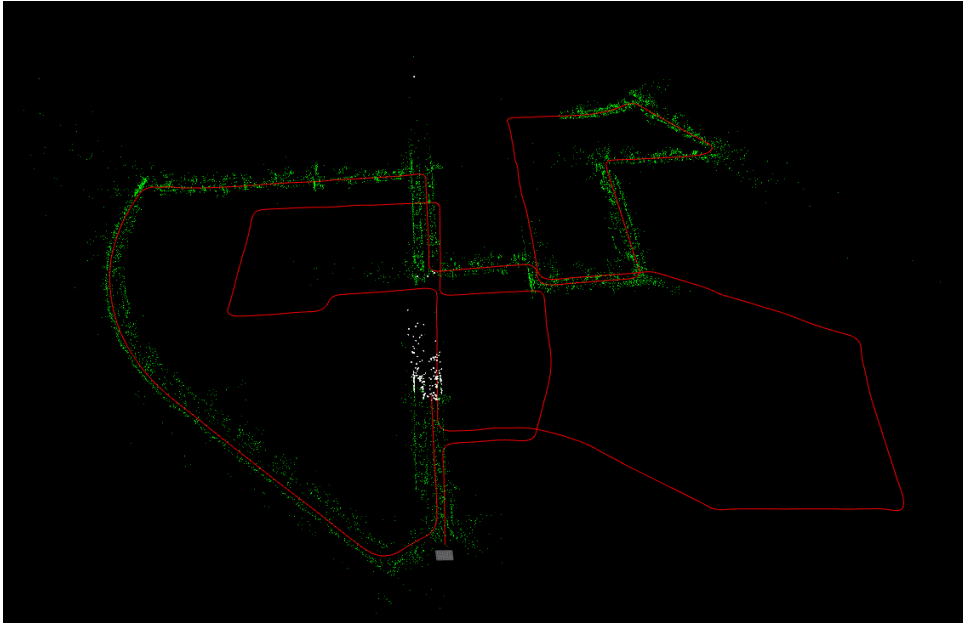
本部分的实验分为两部分进行。第一部分，利用 KITTI 数据集验证本文所研究定位算法的有效性。第二部分，设计在室外环境中，操控一架配备有双目相机 D435i 的无人机在空中飞行一圈，并保证其起始位置和终点位置相同，并将其运动轨迹用 Rviz 可视化工具显示出来。本次实验中将进行两组对照实验：1、比较无人机仅在双目视觉惯性里程计下定位的轨迹路线图和无人机在双目视觉惯性里程计和 GPS 数据融合后的定位轨迹路线图。2、比较无人机在双目视觉惯性里程计和 GPS 数据融合下飞行一圈后起始位置和终点位置的偏差。在本次实验过程中所用设备的参数为：双目视觉里程计输出帧率为 10Hz，GPS 数据接收频率为 10Hz，而 IMU 惯性测量计的输出频率为 200Hz。

4.3.2 实验结果及分析

KITTI 数据集的效果如图 4-4 所示，其中红线表示双目视觉惯性里程计融合 GPS 数据后的轨迹，绿线表示双目视觉惯性里程计的轨迹。



(a) KITTI 数据集原本实验效果



(b) KITTI 数据集在本文算法下的效果

图 4-4 KITTI 数据集实验效果

可以看出，本文所研究的无人机自主定位算法在公开数据集上的定位效果较好，且误差在可接受范围内，证明了本文所研究算法的有效性。

在室外开阔环境下无人机飞行一圈的效果如图 4-5 所示。其中绿线表示双目视觉惯性里程计在自身相机坐标系下的轨迹，蓝线表示双目视觉惯性里程计在 GPS 坐标系下的轨迹，黄线表示 GPS 数据在 GPS 坐标系下的轨迹，红线表示双目视觉惯性里程计与 GPS 融合后的位置在 GPS 坐标系下的轨迹。

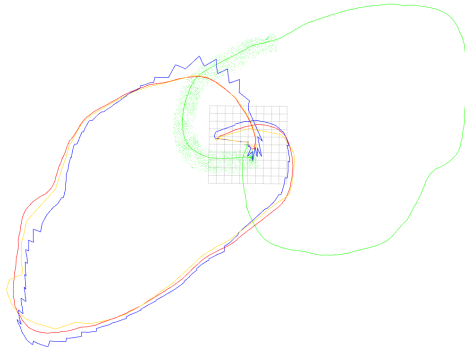


图 4-5 双目视觉惯性里程计与 GPS 数据融合效果图

首先对比蓝线和红线，可以看出，在一开始无人机起飞的过程中，仅靠双目视觉惯性里程计定位的轨迹有较大偏差，而融合了双目视觉惯性里程计和 GPS 数据的定位轨迹则具有较好的稳定性，在之后无人机空中飞行和降落的过程中，从图中可以看出，尽管 GPS 数据下的定位轨迹和双目视觉惯性里程计下的定位轨迹有较多波动，但融合了二者数据的定位轨迹则一直保持相对稳定的平滑曲线。其

次对比无人机在双目视觉惯性里程计和 GPS 数据融合下飞行一圈后的起始位置和终点位置的偏差，对图 4-4 中的起点和终点附近的图像放大，如图 4-6 所示：



图 4-6 起点和终点位置附近图像

可以看出仅依靠双目视觉惯性里程计（绿线）的定位轨迹，其起始位置和终点位置偏差较大，而在融合了 GPS 数据之后的定位轨迹（红线），其起始位置和终点位置则几乎重合，因此可以证明本章所研究的双目视觉惯性里程计和 GPS 数据融合算法的有效性，该算法在开阔环境下可以提升无人机定位的精确性和鲁棒性，能够满足无人机在开阔场景下自主作业的需求。

4.4 本章小结

本章研究主要介绍了 GPS 定位原理和将 GPS 数据与前面章节所研究的视觉里程计进行融合这两个方面。在 GPS 定位原理中，本章介绍了在无人机上获取 GPS 数据的方式；在研究双目视觉惯性里程计融合 GPS 数据时，本章采用扩展卡尔曼滤波进行数据融合，采用 Newton-Gauss 和非线性最小二乘法进行优化。通过以上操作，设计出了一套无人机在开阔场景下的自主定位系统，提升了无人机自主定位精度，满足了无人机户外作业需求。

结 论

本文研究了基于 GPS 数据融合的开阔场景下双目视觉里程计，梳理了机器人领域视觉里程计的国内外研究现状，深入探讨了无人机在开阔场景下自主定位的难点和技术，针对于无人机在开阔场景下较高要求的自主定位精度和鲁棒性问题，本文提出了一种在双目视觉里程计的基础上，融合 IMU 和 GPS 数据的无人机自主定位算法。同时，搭建了挂载双目立体相机的阿木无人机，在室外开阔环境中进行了大量的实验。实验结果有力地证明了本课题研究所设计的无人机自主定位算法的精确性和稳定性。在特征点的检测与跟踪章节中，本文提出了基于 ORB 算法的特征点检测与跟踪的算法，主要包含角点检测和特征点描述这两个部分。在角点检测中，本文提出了将哈里斯角点检测算法和快速角点检测算法结合共同进行角点检测，来增强角点检测的精确性和鲁棒性；提出了搭建 Gauss 金字塔的方法使其具备尺度不变性；提出了利用灰度质心法使其具备旋转不变形。在特征点描述中，本文提出了利用高斯滤波的方法进行降噪处理，在一定程度上缓解了 BRIEF 特征描述子对噪声的敏感问题。最后，在双目相机中利用左侧 ORB 的坐标和右侧匹配的坐标来生成立体特征点，从而实现了特征点的检测与跟踪。

在双目视觉里程计与 IMU 数据相融合得到双目视觉惯性里程计的章节中，主要包含双目视觉里程计的构建以及双目视觉里程计与 IMU 数据进行融合这两个部分。在双目视觉里程计的构建中，本文详细阐述了利用 KLT 跟踪器和相机因子进行构建的过程。在双目视觉里程计与 IMU 数据融合中，本章阐述了利用 IMU 预积分操作在时间上对齐双目视觉里程计数据和 IMU 数据，并且阐述了利用高斯-牛顿法进行非线性优化，采用舒尔补进行边缘化等操作。在双目视觉惯性里程计和 GPS 数据融合部分，本文阐述了双目视觉里程计与 GPS 数据和 IMU 数据通过扩展卡尔曼滤波的方式进行融合的无人机自主定位的算法。

在实验验证部分，本文在阿木无人机平台上设计了操作可行的实验，验证了双目视觉惯性里程计和 GPS 数据融合的无人机自主定位算法的精度和稳定性。最后在户外开阔环境中进行了无人机飞行过程中的自主定位实验，取得了显著的实验效果，充分验证了本课题研究的无人机自主定位算法的有效性。

参考文献

- [1] He, M., Zhu, C., Huang, Q. *et al.* A review of monocular visual odometry. *Vis Comput* **36**, 1053–1065 (2020).
- [2] 慈文彦, 黄影平, 胡兴. 视觉里程计算法研究综述[J]. 计算机应用研究, 2019, 36(09): 2561-2568.
- [3] X. Yan, Z. Shi and Y. Zhong, "Vision-based Global Localization of Unmanned Aerial Vehicles with Street View Images," 2018 37th Chinese Control Conference (CCC), Wuhan, China, 2018, 4672-4678.
- [4] Shaowu Yang, Sebastian A. Scherer, Xiaodong Yi, Andreas Zell, Multi-camera visual SLAM for autonomous navigation of micro aerial vehicles, *Robotics and Autonomous Systems*, 116-134, ISSN 0921-8890.
- [5] K. Wang, S. Ma, J. Chen, F. Ren and J. Lu, "Approaches, Challenges, and Applications for Deep Visual Odometry: Toward Complicated and Emerging Areas," in *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 1, 35-49, March 2022.
- [6] 职恒辉, 尹晨阳, 李慧斌. 基于深度学习的视觉里程计方法综述[J]. 计算机工程与应用, 2022, 58(20): 1-15.
- [7] 王文森, 黄凤荣, 王旭等. 基于深度学习的视觉惯性里程计技术综述[J]. 计算机科学与探索, 2023, 17(03): 549-560.
- [8] 刘海涛, 戴娟, 朱胜涛, 等. 基于深度学习的移动机器人视觉里程计[J/OL]. 控制工程: 1-8[2024-05-10]. <https://doi.org/10.14107/j.cnki.kzgc.20220902>.
- [9] 崔立志, 杨啸乾, 杨艺. 基于深度学习的动态环境视觉里程计研究[J]. 空间控制技术与应用, 2023, 49(06): 58-67.
- [10] 李晋, 陈威, 刘羽鹤, 等. 融合视觉里程计和 BP 神经网络的自适应行人航迹推算方法[J]. 中国现代教育装备, 2024(07): 27-32.
- [11] 李伟东, 朱旭浩. 一种结合光流追踪的双目视觉里程计算法[J]. 计算机测量与控制, 2024, 32(02): 174-180. DOI: 10.16526/j.cnki.11-4762/tp.2024.02.026.
- [12] 廖士楠, 王伟, 王世勇. 基于多旋翼无人机视觉惯性里程计的设计与实现[J]. 兵工自动化, 2024, 43(02): 89-92.
- [13] 闫河, 黄奎霖, 王潇棠, 等. 结合语义信息的视觉里程计动态特征点剔除方法[J]. 信息与控制, 2023, 52(05): 616-625. DOI: 10.13976/j.cnki.xk.2023.2383.
- [14] 张梦龙, 张凯杰, 刘昌林, 等. 基于点转移矫正的视觉惯性里程计[J]. 中国惯性技术学报, 2024, 32(02): 146-152. DOI: 10.13695/j.cnki.12-1222/o3.2024.02.006.

- [15] 田应仲,刘伊铭,杨晓东,等.基于点线特征融合的视觉惯性里程计方法研究[J].计量与测试技术,2024,51(03):45-48.DOI:10.15988/j.cnki.1004-6941.2024.3.013.
- [16] 钟家跃. 未知环境下无人机的双目视觉 SLAM 方法研究[D].南京航空航天大学,2020.
- [17] Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* 60, 91–110 (2004).
- [18] E. Rublee, V. Rabaud, K. Konolige and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," 2011 International Conference on Computer Vision, Barcelona, Spain, 2011, 2564-2571
- [19] 吕华龙. 基于双目视觉的无人机自主感知与导航系统设计与实现[D].天津大学,2022.
- [20] Cvišić, I, Ćesić, J, Marković, I, Petrović, I. SOFT-SLAM: Computationally efficient stereo visual simultaneous localization and mapping for autonomous unmanned aerial vehicles. *J Field Robotics*. 2018; 35: 578–595.
- [21] G. Cioffi and D. Scaramuzza, "Tightly-coupled Fusion of Global Positional Measurements in Optimization-based Visual-Inertial Odometry," 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 2020, 5089-5095.
- [22] R. Mur-Artal, J. M. M. Montiel and J. D. Tardós, "ORB-SLAM: A Versatile and Accurate Monocular SLAM System," in *IEEE Transactions on Robotics*, vol. 31, no. 5, 1147-1163, Oct. 2015.
- [23] R. Mur-Artal and J. D. Tardós, "ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras," in *IEEE Transactions on Robotics*, 1255-1262, Oct. 2017.
- [24] C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. M. Montiel and J. D. Tardós, "ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial, and Multimap SLAM," in *IEEE Transactions on Robotics*, 1874-1890, Dec. 2021.
- [25] Khan and M. Hebert, "Multi modal pose fusion for monocular flight with unmanned aerial vehicles," 2018 IEEE Aerospace Conference, Big Sky, MT, USA, 2018, 1-7, doi: 10.1109/AERO.2018.8396806.
- [26] Liu Y, Zhao J, Zhang J, et al. Construction of Terrain Multidimensional Traversability Feature Map for Off-Road Scenarios Based on Binocular Vision[R]. SAE Technical Paper, 2024.
- [27] C. Wang *et al.*, "Research on Intelligent Inspection Robot Based on Neural Network and Visual SLAM Technology," 2023 *IEEE International Conference on Electrical, Automation and Computer Engineering (ICEACE)*, Changchun, China, 2023, 1452-1457, doi: 10.1109/ICEACE60673.2023.10442067.

-
- [28] L. Wang and Z. Xu, "Improved Mapping Technique Based on ORB-SLAM2 in Dynamic Scenes," *2023 2nd International Conference on Automation, Robotics and Computer Engineering (ICARCE)*, Wuhan, China, 2023, 1-5, doi: 10.1109/ICARCE59252.2024.10492577.
 - [29] D. Chen, K. Xu and W. Ma, "Binocular vision localization based on vision SLAM system with multi-sensor fusion," *2023 4th International Conference on Computer Vision, Image and Deep Learning (CVIDL)*, Zhuhai, China, 2023, 94-97, doi: 10.1109/CVIDL58838.2023.10166820.
 - [30] Peng B, Sun L, Lei J, et al. Self-Supervised Monocular Depth Estimation via Binocular Geometric Correlation Learning[J]. *ACM Transactions on Multimedia Computing, Communications and Applications*, 2024.
 - [31] J. Xu, Z. Chen, J. Chen, J. Zhou, and X. Du, "A precise localization algorithm for unmanned aerial vehicles integrating visual-internal odometry and cartographer," *Journal of Measurements in Engineering*, Mar. 2024, doi: 10.21595/jme.2024.23726.
 - [32] J. Li *et al.*, "RD-VIO: Robust Visual-Inertial Odometry for Mobile Augmented Reality in Dynamic Environments," in *IEEE Transactions on Visualization and Computer Graphics*, doi: 10.1109/TVCG.2024.3353263.
 - [33] Viswanathan D G. Features from accelerated segment test (fast)[C]//*Proceedings of the 10th workshop on image analysis for multimedia interactive services*, London, UK. 2009: 6-8.
 - [34] Calonder M, Lepetit V, Strecha C, et al. Brief: Binary robust independent elementary features[C]//*Computer Vision—ECCV 2010: 11th European Conference on Computer Vision*, Heraklion, Crete, Greece, September 5-11, 2010, *Proceedings, Part IV* 11. Springer Berlin Heidelberg, 2010: 778-792.
 - [35] Harris C, Stephens M. A combined corner and edge detector[C]//*Alvey vision conference*. 1988, 15(50): 10-5244.
 - [36] Babaud J, Witkin A P, Baudin M, et al. Uniqueness of the Gaussian kernel for scale-space filtering[J]. *IEEE transactions on pattern analysis and machine intelligence*, 1986 (1): 26-33.
 - [37] Fisher R A. On the mathematical foundations of theoretical statistics[J]. *Philosophical transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character*, 1922, 222(594-604): 309-368.
 - [38] Triggs B, McLauchlan P F, Hartley R I, et al. Bundle adjustment—a modern synthesis[C]//*Vision Algorithms: Theory and Practice: International Workshop on Vision Algorithms Corfu, Greece, September 21–22, 1999 Proceedings*. Springer Berlin Heidelberg, 2000: 298-372.

- [39] Lucas B D, Kanade T. An iterative image registration technique with an application to stereo vision[C]//IJCAI'81: 7th international joint conference on Artificial intelligence. 1981, 2: 674-679.
- [40] The Schur complement and its applications[M]. Springer Science & Business Media, 2006.
- [41] Joyce, James, "Bayes' Theorem", *The Stanford Encyclopedia of Philosophy* <https://plato.stanford.edu/archives/fall2021/entries/bayes-theorem>.
- [42] Murphy K P. Machine learning: a probabilistic perspective[M]. MIT press, 2012.
- [43] Björck Å. Numerical methods for least squares problems[M]. Society for Industrial and Applied Mathematics, 1996.
- [44] Moré J J. The Levenberg-Marquardt algorithm: implementation and theory[C]//Numerical analysis: proceedings of the biennial Conference held at Dundee, June 28–July 1, 1977. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006: 105-116.
- [45] 王永祥. 基于视觉 GPS 融合二维占据网格地图构建[D]. 太原理工大学, 2022. DOI:10.27352/d.cnki.gylgu.2021.000644.
- [46] Mourikis A I, Roumeliotis S I, Burdick J W. Sc-kf mobile robot localization: A stochastic cloning kalman filter for processing relative-state measurements[J]. IEEE Transactions on Robotics, 2007, 23(4): 717-730.

哈尔滨工业大学本科毕业论文(设计)

原创性声明和使用权限

本科毕业论文(设计)原创性声明

本人郑重声明:此处所提交的本科毕业论文(设计)《基于GPS数据融合的开闾场景下双目视觉里程计研究》,是本人在导师指导下,在哈尔滨工业大学攻读学士学位期间独立进行研究工作所取得的成果,且毕业论文(设计)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。对本毕业论文(设计)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:李明谦

日期:2024年5月28日

本科毕业论文(设计)使用权限

本科毕业论文(设计)是本科生在哈尔滨工业大学攻读学士学位期间完成的成果,知识产权归属哈尔滨工业大学。本科毕业论文(设计)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本科生上交的毕业论文(设计),并向有关部门报送本科毕业论文(设计);
- (2)根据需要,学校可以将本科毕业论文(设计)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)本科生毕业后发表与此毕业论文(设计)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。

本人知悉本科毕业论文(设计)的使用权限,并将遵守有关规定。

作者签名:李明谦

日期:2024年5月28日

导师签名:佟明基

日期:2024年5月28日

致 谢

衷心感谢李湛老师和佟明斯老师对本人毕业论文的精心指导。二位老师的言传身教将使我终身受益。从开始选题到之后写论文、做实验的过程中，二位老师都以严谨认真负责的态度向我提供科学合理的建议，让我得以解决实验过程中遇到的困难，顺利完成实验。

感谢吴桐师兄，对我的论文修改以及实验验证都给予了充分的指导。感谢实验室的各位学长学姐，在我做实验时给予了我充分的帮助。

光阴似箭，白驹过隙，本科四年的学习生涯转瞬即逝，感谢赵德宇同学和杨壬辰同学在四年的学习期间陪我一起打球，一起放松，一起成长，感谢我的家人在这四年里给予我的支持和陪伴，感谢四年生活里所有相识的朋友，感谢你们陪我一路前行。

最后，由于我的学术水平有限，所写论文难免有不足之处，恳请各位老师和同学提出批评和指正。