



武汉大学学报(信息科学版)

Geomatics and Information Science of Wuhan University

ISSN 1671-8860, CN 42-1676/TN

《武汉大学学报(信息科学版)》网络首发论文

题目：一种自适应点线特征和 IMU 耦合的视觉 SLAM 方法
作者：张晨阳, 杨健
DOI: 10.13203/j.whugis20230347
收稿日期: 2024-06-28
网络首发日期: 2024-07-17
引用格式: 张晨阳, 杨健. 一种自适应点线特征和 IMU 耦合的视觉 SLAM 方法[J/OL]. 武汉大学学报(信息科学版). <https://doi.org/10.13203/j.whugis20230347>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

DOI:10.13203/j.whugis20230347

引用格式：

张晨阳, 杨健. 一种自适应点线特征和 IMU 耦合的视觉 SLAM 方法[J]. 武汉大学学报(信息科学版), 2024, DOI: 10.13203/j.whugis20230347 (ZHANG Chenyang, YANG Jian. A Visual SLAM Method Coupled with Adaptive Point-line Features and IMU[J]. Geomatics and Information Science of Wuhan University, 2024, DOI: 10.13203/j.whugis20230347)

一种自适应点线特征和 IMU 耦合的视觉 SLAM 方法

张晨阳¹, 杨健^{1,2,3}

1 南京理工大学, 计算机科学与工程学院, 江苏 南京, 210094

2 南京理工大学高维信息智能感知与系统教育部重点实验, 江苏 南京, 210094

3 南京理工大学江苏省社会安全图像与视频理解重点实验室, 江苏 南京, 210094

摘要: 室内低/弱纹理、光照不足的场景下, 视觉惯导融合的 SLAM 定位精度明显优于纯视觉 SLAM 方法。然而, 当前基于点特征的视觉惯导 SLAM 方法通常难以检测并追踪足够的特征, 同时惯性测量单元的先验测量信息亦未充分利用, 导致 SLAM 整体定位精度低、鲁棒性弱。针对这些问题, 构建一种自适应点线特征和 IMU 耦合的视觉 SLAM 方法。首先设计一种自适应的 FAST 阈值角点算法, 以增强图像特征点检测的鲁棒性。另外, LSD 线特征算法易检测短线、断线, 且图像因光照变化易导致线特征的“过提取”或“错提取”。为此, 基于边缘检测二值图像构造自适应线特征提取算法, 并借助消影点的特性筛选聚类线特征。然后, 由点线特征重投影误差和 IMU 先验预积分位姿估计量, 通过松耦合为 SLAM 前端位姿估计和算法提供稳健的初始化结果。随后, 后端利用紧耦合建立视觉和 IMU 观测量的统一非线性最小化残差函数, 并优化得到准确的图像帧位姿。最后, 在开源数据集上测试验证, 并对比几种 SOTA (State Of The Art) SLAM 方法。实验结果表明, 本文构建的 SLAM 方法平均定位精度至少提高 12%, 同时兼顾较强的鲁棒性。

关键词: SLAM; 自适应点线特征; 惯性测量单元; 松耦合; 紧耦合

A Visual SLAM Method Coupled with Adaptive Point-line Features and IMU

ZHANG Chenyang¹, YANG Jian^{1,2,3}

1 School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

2 Key Lab of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education, Nanjing University of Science and Technology, Nanjing 210094, China

3 Jiangsu Key Lab of Image and Video Understanding for Social Security, Nanjing University of Science and Technology, Nanjing 210094, China

Abstract: Objectives: Visual-inertial SLAM typically outperforms pure visual SLAM in indoor scenes characterized by low or sparse textures and varying lighting conditions. Nonetheless, most existing visual-inertial SLAM encounter challenges in detecting and tracking sufficient feature points. Moreover, the prior pose measurement information from the Inertial Measurement Unit is often underutilized, resulting in reduced pose estimation accuracy and limited robustness. **Methods:** An adaptive point detection approach has been developed to enhance the robustness of feature point detection in images. Additionally, the LSD line feature algorithm makes it easy to detect short lines

收稿日期: 2024-06-28

基金项目: 中国博士面上项目 (2023M741702), 国家自然科学基金 (62361166670)。

第一作者: 张晨阳, 博士, 博士后, 研究方向为视觉 SLAM。zcynj@hhu.edu.cn

通信作者: 杨健, 博士, 博士生导师, 教授。csjyang@njust.edu.cn

and broken line features, and the performance of the algorithm is affected by the change of illumination, resulting in "over-extraction" or "wrong-extraction" of line features. Accordingly, an adaptive algorithm for extracting line features was introduced, utilizing edge-detected binary images and incorporating the removal of erroneous lines based on the geometry characteristics of the vanishing point. Following this, the algorithm integrates the visual measurements from point-line features with the pre-integration measurement of IMU to yield reliable outcomes for front-end pose estimation and initialization parameters in a loosely coupled manner. In the back-end section of our proposed SLAM method, a unified nonlinear minimization residual function is established for visual and IMU measurements through tight coupling, optimizing for obtaining precise pose of the image or camera. **Results:** Our SLAM method has been validated and tested on publicly available benchmarks, showcasing its performance through ablation experiments and qualitative as well as quantitative comparative analyses against several state-of-the-art visual-inertial SLAM algorithms. **Conclusions:** The results indicate that our algorithm improves average localization accuracy by at least 12% and displays significant robustness.

Key words: SLAM, Adaptive point-line features, Inertial Measurement Unit, Loose coupling, Tight coupling

当前即时定位与建图 (Simultaneous Localization and Mapping, SLAM) 被认为是解决未知环境下机器人感知和定位问题的有效手段之一^[1]。SLAM 概念自 1986 年美国旧金山举行的机器人与自动化会议上提出以来, 由于其重要的理论和应用价值, 许多学者认为其是实现机器人自主性的关键。即时定位与建图作为机器人实际生产和生活应用中的核心基础 (如定位导航、自动驾驶、仓储物流等)^[2-3], 对其进行深入研究具有重要意义。此外, 作为近景摄影测量学的重要研究领域, 加之 SLAM 的发展受社会信息化和智能化浪潮的推动, 也极大促进了测绘科学技术向智能化方向的转型升级^[4]。

SLAM 是指机器人搭载特定传感器, 在外界环境先验信息未知的情况下, 通过构建环境感知测量模型并估计自身位姿来实现^[5]。相机具有成本低、体积小以及能实时捕捉场景中结构纹理信息等优势, 因此 SLAM 算法最初研究基本是以视觉图像作为数据源输入逐渐发展起来, 并在国内外取得显著进展。典型的代表算法如西班牙萨拉戈萨大学的 ORB-SLAM 算法^[6-7], 苏黎世联邦理工学院的 DSO^[8], 慕尼黑工业大学的 DVO^[9]和 LSD-SLAM^[10]等; 国内前沿研究有浙江大学计算机视觉组提出的 RK-SLAM^[11], 上海交通大学感知导航研究所结合场景结构信息提出的 Struct-SLAM^[12], 中山大学数据科学与计算机学院的 RGB-T SLAM^[13], 港科大机器人研究所基于事件相机的 ESVO^[14]等相关的 SOTA SLAM 算法。然而, 这些视觉 SLAM 方案多数受限于实验室场景。当相机置身于低/弱纹理、光照变化的场景中, 或者受图像质量、平台快速旋转移移动等因素影响时, 视觉特征的关联准确性和鲁棒性会下降, 这直接导致视觉 SLAM 算法定位精度的退化。

IMU 不受上述外界因素的影响, IMU 和视觉相机的数据互补性显现出显著的优势。近年来, 通过耦合 IMU 和相机的视觉惯导定位导航算法 (Visual Inertial Navigation System, VINS) 得到广泛的关注, 并发展成为多传感器耦合 SLAM 领域研究的重点和热点。视觉惯导耦合的方式分为紧耦合和松耦合, 松耦合指 IMU 和相机分别估计自身运动状态, 随后将两者的位姿估计结果融合, 这种方式算法的复杂度较低; 紧耦合则是基于 IMU 和相机测量数据共同构建运动和观测方程, 并联合完成状态估计; 与松耦合相比, 紧耦合的计算复杂度较高, 在定位精度和鲁棒性方面具有明显的优势^[15]。紧耦合早期是通过滤波来实现的, 比如基于扩展的卡尔曼滤波 MSCKF 算法^[16], 该算法采用滑动窗口基于标准 EKF 框架来实现。基于优化的紧耦合算法是将传感器的测量数据建模为一个非线性问题, 并将历史图像帧来做优化求解, 是视觉惯导最主要的耦合方式。OKVIS^[17]是最早一个基于非线性优化的视觉惯导 SLAM, 算法构建非线性函数通过滑动窗口优化相机位姿, 但缺少闭环检测, 这容易导致运动轨迹误差累积; 后续作者推出具有闭环检测的 OKVIS2, 并在开源数据基准 EuRoC 上测

试验证, 相比 OKVIS 的精度和稳健性有明显提升。

IMU 的采样频率通常高于相机的输出频率, 为构建 IMU 测量观测量, 并与图像帧时间戳对齐, Foster 等人^[18]推导了流形空间下 IMU 预积分数学表达式, 在后续的视觉惯导 SLAM 研究中得到广泛应用, 如西班牙萨拉戈萨大学的 VI-SLAM^[19]和 ORB-SLAM3^[20], 以及港科大的单目视觉惯导 SLAM 算法 Vins-Mono^[21]等。Vins-Mono 前端用光流法跟踪特征, 紧耦合 IMU 预积分量与视觉特征, 通过非线性优化来实现定位和建图。在 Vins-Mono 的基础上, 港科大研究团队又发布了多源传感器耦合的 SLAM 算法 Vins-Fusion^[22]。考虑到 VINS 算法前端姿态估计精度与系统后端非优化之间存在较大的关联, 研究人员通过改进前端位姿估计初值来提高整个 VINS 的精度。例如, 龚赵慧等^[23]将 IMU 测量数据与图像结合, 利用 IMU 加权的运动先验模型, 通过 IMU 预积分获取状态估计, 并使用 IMU 先验信息的加权值为前端提供精确的初值, 进而提出一种半直接法的视觉惯导里程计。在对比分析了视觉里程计特征点法和直接法的优缺点后, 许宾智等^[24]将惯性测量单元与两种方法进行融合, 提出了一种半直接法的视觉惯导里程计, 以提高相机定位精度和鲁棒性。此外, 张福斌等^[25]建立图像金字塔结合 LK 光流法来改进 Vins-Mono 的前端, 并将陀螺仪的旋转量和里程计的平移量作为先验条件, 同时加入车轮里程计测量量, 将里程计约束应用于初始化和后端非线性优化中, 充分融合了视觉、IMU 和车轮里程计测量信息。

除了点特征, 线特征也引入了 VINS 算法研究。He 等人^[26]将 IMU 测量数据和点线特征信息融合, 提出基于点线特征的单目视觉里程计—PL-VIO。该算法用普吕克坐标来表示线特征, 并通过滑动窗口联合点线特征来优化相机位姿。结合 Vins-Mono 和 PL-VIO, Fu 等^[27]通过对 LSD 线特征检测结果做阈值筛选处理, 并用线特征中点到对应匹配线段的距离作为几何约束条件, 在 PL-VIO 的基础上增加了回环检测模块, 来构建融合 IMU 和点线特征的单目视觉惯导 SLAM 算法—PL-VINS。Lee^[28]结合点线特征融合了平行三维线, 并基于消影点聚类线特征, 在多视图线特征聚类过程中, 利用一致性检校方法剔除并行三维线的外点特征, 最后构造对应的三维线残差函数, 建立点线和并行线的单目视觉惯导 SLAM。赵良玉^[29]和应文健等^[30]使用 IMU 测量数据为视觉定位算法提供先验和尺度信息, 他们对提取的线特征做合并、修补或匹配, 并通过联合最小化残差函数提出了基于点线特征并融合 IMU 的双目视觉惯导 SLAM 算法, 该算法在保证定位精度的同时能够用多种形式表达环境地图。危双丰^[31]等人通过优化加速度计偏置初始化与点线特征数据选择策略, 获取更稳健的初始化结果, 从而增强算法鲁棒性和精度。

综合上述对视觉惯导 SLAM 研究现状分析, 不难发现合理有效地处理前端视觉图像特征是提高前端位姿估计稳健性和整体 SLAM 算法精度的关键^[15]。对于点特征而言, 现有算法通常采用特定的特征检测算子来直接提取点特征。然而, 室内场景下图像纹理信息或光照变化因素的存在, 直接使用特征算子提取点特征易导致前端特征跟踪数量较少且特征分布不均匀。对于线特征, 大多数基于点线特征的视觉惯导 SLAM 算法采用 LSD^[32]线特征提取算法。然而, LSD 算法对图像上噪声和光照敏感, 算法的性能会因此受到影响, 导致部分检测的线特征并非真正来源于场景边缘, 这对于算法的定位建图是不利的。此外, 在现有多数视觉惯导算法中, IMU 预积分的先验信息在初始化和前端位姿估计阶段也未能充分利用。针对上述问题, 本文首先设计了一种自适应阈值特征检测算法, 并结合 Kd-tree 均匀化图像上的特征点; 然后, 基于 Canny 算子边缘检测生成的二值图像, 设计自适应线特征算法来检测和提取线特征, 同时利用消影点特性筛选聚类线特征, 以尽可能保证图像线特征符合场景物体边缘。再后, 前端基于点线视觉特征和 IMU 预积分观测量, 以松耦合方式来获取稳健的初始化参数结果, 而后端则采用点线特征重投影误差与 IMU 预积分观测量进行紧耦合优化平差, 最终搭建室内场景下自适应点线特征与 IMU 耦合的视觉 SLAM 方法。

1 自适应点线特征和 IMU 耦合的视觉 SLAM

图 1 给出了本文 SLAM 方法的流程图，算法基于多线程并行运算从左到右依次分三个线程：前端位姿估计初始化和特征追踪，后端局部 BA（Bundle Adjustment）和回环检测。算法大致流程为：1）基于双目相机视觉图像，利用多线程并行来自适应提取图像点特征和线特征；2）利用消影点特性聚类线特征，以便筛选的线特征符合场景边缘结构信息；3）基于点线特征和 IMU 先验测量信息建立前端位姿松耦合优化模型，并完成 SLAM 系统的初始化；4）基于 IMU 预积分和点线特征重投影误差方程构建紧耦合优化方程，并更新点线特征以及图像帧在世界坐标系下坐标信息；5）回环检测采用点线词袋模型来检测回环，一旦检测到回环，便执行全局 BA，以此提升视觉惯导 SLAM 算法定位结果的准确性和鲁棒性。

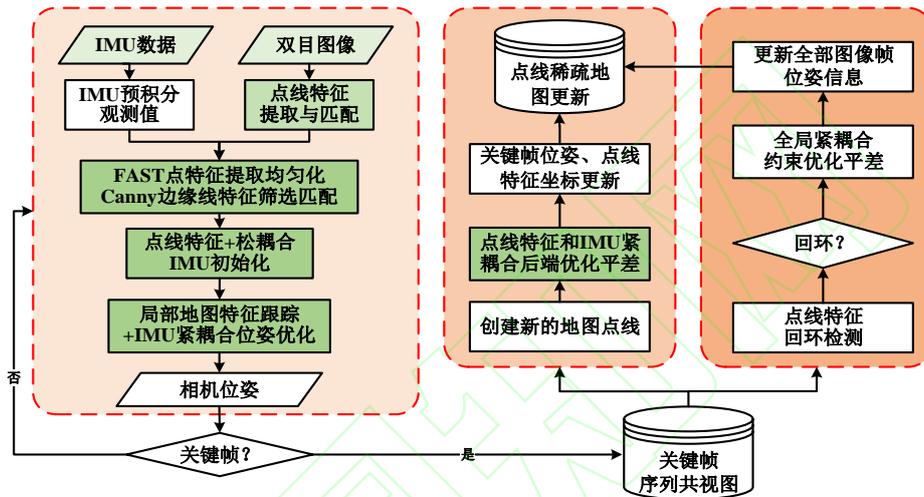


图 1 本文方法流程图

Fig. 1 The Flow Chart of Proposed Method

1.1 自适应点特征提取

FAST 算法主要基于邻域像素灰度值的变化来寻找具有显著特征的像素点，从而确定图像关键点的位置，能够快速实现特征点检测，对图像噪声有一定的鲁棒性，已成为图像处理领域比较流行的特征检测算法^[33]。FAST 算法倾向于在纹理丰富区域生成大量的关键点，在纹理稀疏区域则较少，且角点检测响应值高，造成关键点分布不均匀。针对这些问题，本文构建一种自适应的 FAST 角点特征提取算法，同时运用 Kd-tree 算法来均匀化图像角点特征。算法执行过程主要为 6 个步骤：1）对双目视觉图像作畸变校正以及均衡化预处理；2）调用 OpenCV 高斯金字塔接口函数来构建图像金字塔，考虑到计算成本和存储实际需求，算法高斯金字塔层数设置为 8 层，依次计算每层金字塔图像特征点待提取数量；3）计算划分每层图像单元格，并计算各单元格需检测的点特征数量；4）FAST 角点检测，并利用 Kd-tree 算法均匀化单元格内点特征；5）计算特征点总数，判断特征点总数是否达到预先设定的阈值，如果满足则输出对应特征点集，否则，角点响应阈值减去 1，并返回步骤 4；6）若阈值调整至下限值或者点总数超过预设总数的 90%，则跳出当前图像点特征检测过程。

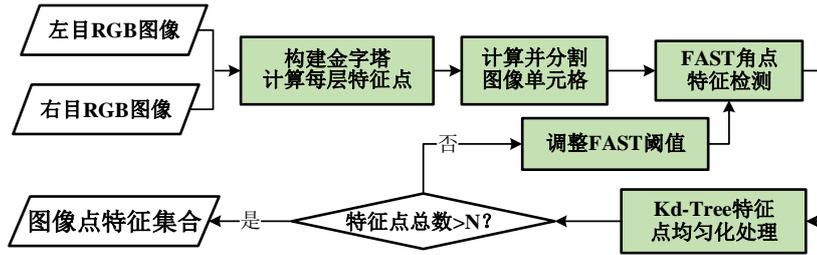


图2 自适应角点特征提取流程
Fig. 2 The Flow Chart of Adaptive Corner Feature Extraction and Detection

本文中自适应角点特征提取算法伪代码如下所示：

算法输入 (input)：双目 RGB 图像 $C_{左}$ 、 $C_{右}$ ，图像特征点总的数量 N ，点特征阈值 T

算法输出 (output)：鲁棒性强，特征显著的 FAST 点特征集合

Step1：对图像 $C_{左}$ 、 $C_{右}$ 畸变校正预处理，图像质量增强并直方图均衡化->ImageEnhance();

Step2：构建图像金字塔->ComputePyramid(), 计算每一层金字塔特征点个数->ComputeLevel();

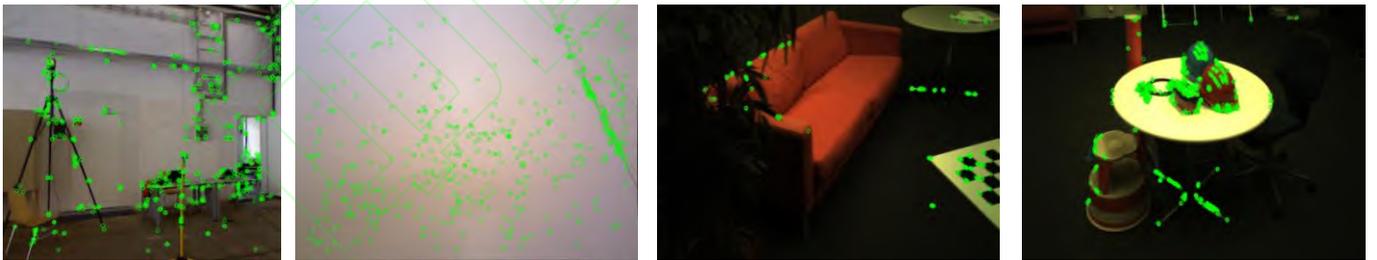
Step3：分割图像单元格，并计算每一图像单元格中点需检测的特征数量->DividingCells();

Step4：检测图像特征点->Fast(), 基于 Kd-tree 特征点均匀化->Kd-tree(), 并判断特征点总数 N 是否满足要求;

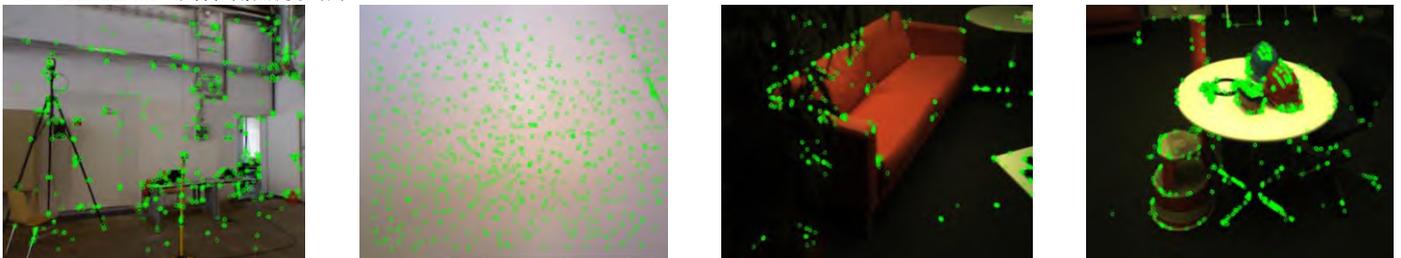
Step5：如果满足预设总数要求，输出特征点集合，计算并判断特征点总数是否大于 N ，如何不满足继续返回 Step 4, $T--$;

Step6：判断 T 是否超过预先阈值，或者特征点总数达到总数的 0.9, Exit。

为检验上述自适应点特征检测算法，本文列出部分典型场景图像定性实验结果，分别选取几组室内常规纹理场景图像和低\弱纹理图像，结合本文设计的自适应特征角点提取算法与 ORB-SLAM3 中的特征点提取算法结果开展对比验证实验研究。本文定性对比实验中图像特征点数量设置为 1000，角点特征检测最低阈值为 9，图像金字塔层数设置为 8 层。实验中常规纹理下算法提取只需循环迭代 3 次左右，低纹理场景下需循环迭代 10 次左右。特征点检测结果见图 3，图中绿色小圆圈均表示检测的特征点。从图 3 中定性的对比结果中不难看出，本文构建的自适应点特征检测算法相比 ORB-SLAM3 算法中的角点特征提取结果，在室内低\弱纹理、常规纹理以及光照不足的图像上呈现相对较好且稳健的特征提取结果。



(a) ORB-SLAM3 中特征角点提取结果



(b) 本文构建的自适应角点特征提取算法结果

图3 不同场景下两种特征角点提取算法定性对比结果

Fig.3 Qualitative Comparison Results of Two Feature Corner Extraction Algorithms in Several Scenarios

图 4 展示了本文自适应角点特征提取和 ORB-SLAM3 的特征提取结果的特征匹配部分的定性定量实验结果。这两种方法都基于点特征描述子完成特征匹配。从特征点数量和整体

分布的角度来看, 本文的自适应角点特征算法结果中具有更多角点特征, 而且这些特征点基于 Kd-tree 均匀化地分布在图像上。相较 ORB-SLAM3 中的点特征检测结果, 图像上特征点数量多有利于获取更多可能的特征匹配对。根据图 4 中的实验对比结果不难发现, 基于自适应点特征算法的匹配结果比 ORB-SLAM3 中的点特征匹配结果能得到更多的特征匹配对。

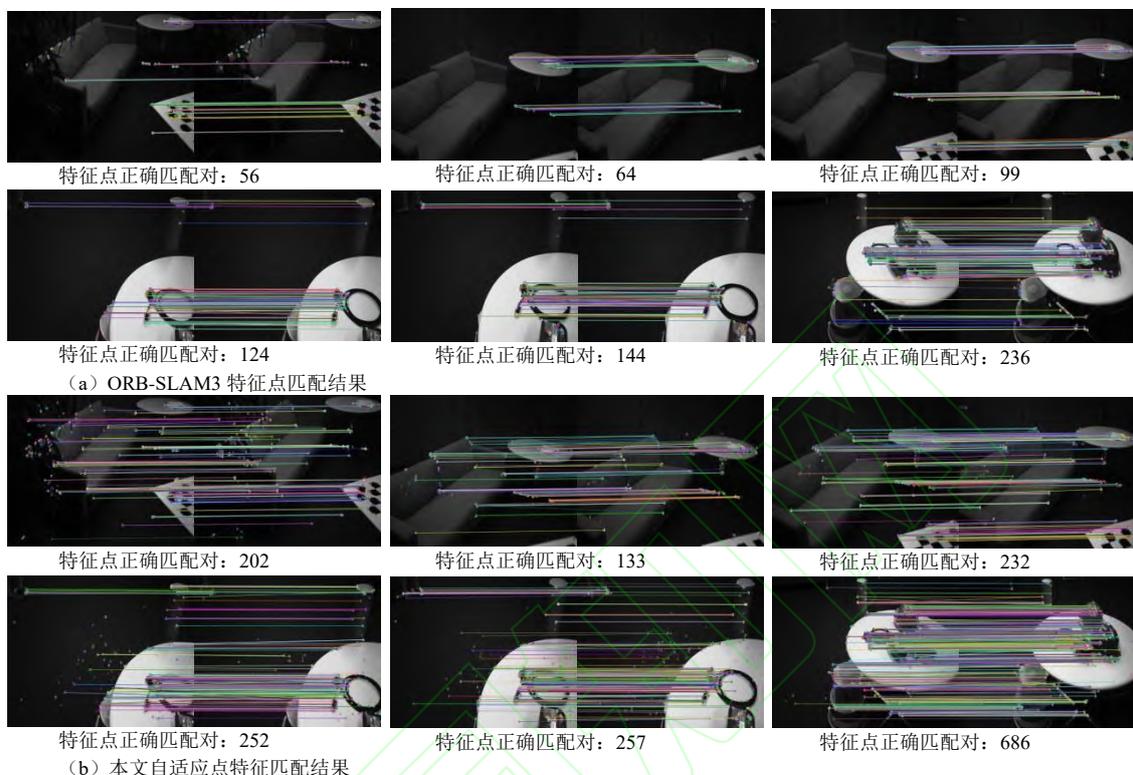


图 4 两种特征角点提取算法匹配结果定性对比

Fig. 4 Qualitative Comparison Matching Results using Two Feature Corner Extraction Algorithms

1.2 自适应边缘线特征提取

当前图像线特征常用 LSD 算法检测和提取, LSD 是基于边缘像素的直方图统计来检测图像中的直线段, 算法计算速度快、检测结果准确, 能通过调整参数适应不同场景线特征检测需求^[32]。然而, LSD 算法容易出现短线或断线问题, 并且容易受外部光照的干扰, 导致提取的线特征与实际物体边缘结构信息不一致。为解决这一问题, 本文基于边缘检测设计一种自适应二值图像的线特征提取算法, 其具体实现流程如图 5 所示。算法首先对图像做高斯模糊预处理, 并计算整幅图像灰度值的中位数, 然后基于灰度值中位数, 调用 OpenCV 的 Canny 边缘检测函数接口, 并基于灰度值中位数设置自适应阈值进行图像边缘检测。接下来, 算法遍历边缘二值图像上的每一个非零像素点, 并将其作为种子点向八个相邻方向搜索, 逐个计算每个非零相邻像素点的数量, 以确定是否满足线特征的长度阈值条件。一旦满足预定的线特征长度阈值条件, 便利用最小二乘原理对像素点集合做线段拟合处理, 即而获得图像上的线特征。随后, 又将距离较近的线段合并处理, 得到图像线特征检测结果。最后, 本文将该算法嵌入并封装为 OpenCV 开源程序库线特征提取接口, 增加图像高斯金字塔构建和线特征检测模块, 最终实现自适应边缘线特征提取函数接口的封装调用。

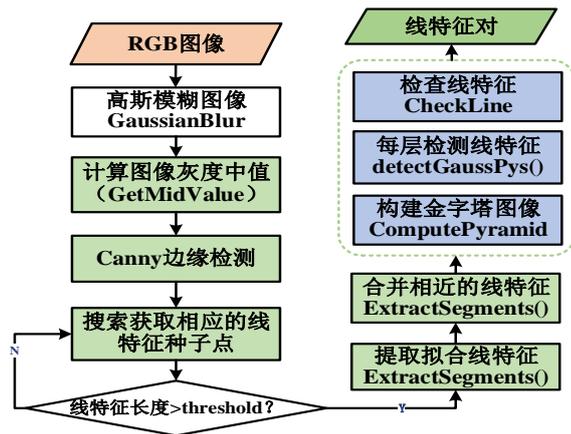


图5 自适应线特征提取检测流程
Fig.5 The Flow Chart of Adaptive Line Feature Extraction and Detection

为验证上述基于自适应二值图像边缘特征的自适应线特征提取算法，本文开展了几组定性对比实验。实验分别在不同场景下的图像线特征检测结果与 LSD 算法结果进行对比，实验结果见图 6。从图 6 结果中对比可知，在图像中存在噪声或者光照变化时，LSD 算法容易受到干扰，导致检测结果不够稳定。LSD 算法提取的图像线特征结果相对比较碎，图像上线特征短线、断线较多，且个别图像因为光照因素影响产生线特征的误提取，本文构建的自适应二值图像线特征提取结果不存在短线和断线，几乎不存在因图像光照因素导致的线段错提取，相较 LSD 算法线特征检测结果显现出显著的优势。

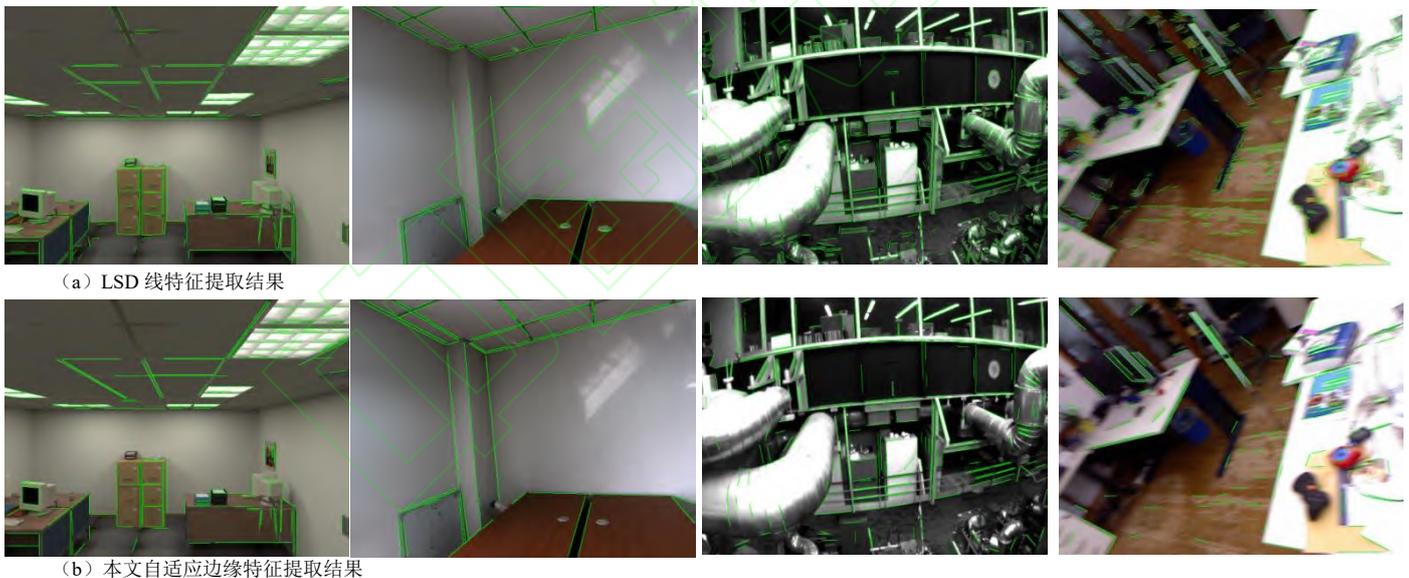


图6 不同场景下两种线特征算法定性对比结果
Fig. 6 Qualitative Comparison Results of Two Line Feature Algorithms in Several Scenarios

为确定算法线特征提取设置的最佳长度阈值，本文给出了室内场景线特征检测部分消融实验结果。算法分别在线特征长度为 20、30、40 像素下测试对比，实验对比结果见图 7。根据图 7 实验结果，长度阈值为 20 时，线特征检测数量相对较多，但线特征检测结果存在短线；当长度预定阈值为 30 时，线特征提取效果有明显改善；长度阈值为 40 时，场景中部分边缘线特征存在损失。因此，结合本文点线视觉特征位姿估计效率的实际考虑，线特征检测算法像素长度阈值设置为 30 像素。



Length:20 pixel Length:30 pixel Length:40 pixel
 图 7 不同长度阈值条件下的线特征提取结果
 Fig.7 The Line feature Extraction Results under Different Length Thresholds Conditions

1.3 基于消影点的线特征筛选与匹配

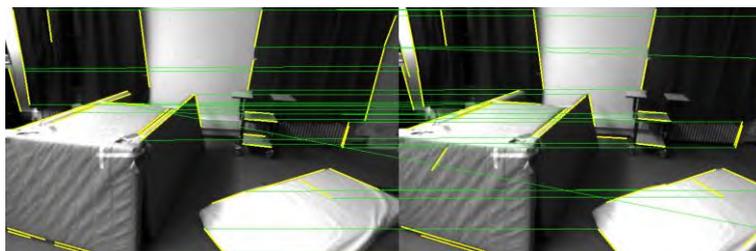
为了提升本文构建的双目视觉惯导 SLAM 算法的鲁棒性和精度，基于线特征匹配对建立测量误差方程之前，需要对检测的线特征做预筛选处理。由透视几何的特性可知，二维图像上平行线特征将会相交于一点，该点为图像的消影点（Vanishing Point）。基于消影点这一特性，我们对线特征做进一步筛选，从而消除那些与场景三维物体边缘结构不相符的线特征，从而提高整个 SLAM 位姿估计的精度和稀疏地图构建的有效性，本文基于消影点筛选线特征方法具体参考文献[34]。

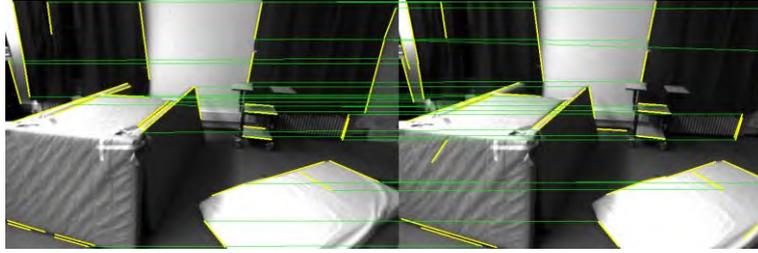
本文线特征匹配策略首先基于 LBD 线特征描述子^[35]，结合消影点的特性，本文首先采用双向匹配来获取粗糙的线特征匹配对集合；其次，这里线特征匹配策略通过增加一定的角度和距离约束来精细化线特征匹配对。角度和距离约束需要满足的两个约束条件：（1）相邻线特征匹配对之间的夹角应小于一定的阈值；（2）线特征匹配对应的中点距离也要小于预先设置的阈值，以上几何约束由公式（1）表示为：

$$\begin{cases} d_1 \times d_2 < threshold_1 \\ \alpha < threshold_2 \end{cases} \quad (1)$$

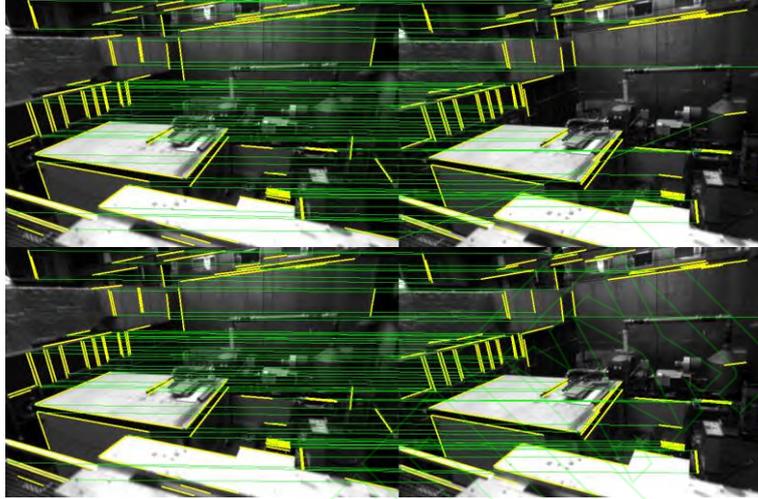
公式（1）中线特征匹配对阈值 $threshold_1$ 为 30×30 ， $threshold_2$ 为 $\pi/8$ ，由于相机输出频率高，相邻图像帧旋转运动量较小， $threshold_2$ 这里阈值定为 $\pi/8$ ，本文线特征提取算法最短线段长度为 30 像素，结合本文线特征匹配实践经验以及角度阈值 $threshold_2$ ， $threshold_1$ 设定为 30×30 。

本文选取 EuRoc 数据集上 V202 和 MH05 两组不同序列的图像开展线特征匹配验证实验，线特征匹配对比结果见下图所示。从图 8 中不难发现，经过消影点线特征的筛选，本文自适应线特征的检测提取结果基本符合场景内的物体边缘信息；其次，根据图 8 特征匹配策略结果对比，增加几何约束能有效剔除错误的线匹配对，提高线特征匹配精度。





(a) 线特征双向匹配结果和“由粗到细”的几何约束匹配结果对比（V202 序列）



(b) 线特征双向匹配结果和“由粗到细”的几何约束匹配结果对比（MH05 序列）

图 8 本文构建的由粗到细的匹配策略的线特征匹配结果对比

Fig.8 Matching Result of Line Feature Matching Based on the Matching Strategy From Coarse to Fine

1.4 联合松耦合和紧耦合的视觉惯导优化

为提升算法整体定位的鲁棒性和精度，算法分别结合松耦合和紧耦合的优点充分融合 IMU 的测量信息，联合点线特征重投影误差模型和 IMU 预积分姿态信息分别在前端和后端共同完成相机位姿的优化。本文算法前端特征追踪主要通过点线视觉特征重投影误差和 IMU 先验位姿信息，通过松耦合得到系统稳健的初始化结果；后端则是通过点线视觉特征重投影误差和 IMU 预积分观测量，通过紧耦合建立误差方程，实现后端的局部地图优化。

1.4.1 视觉惯导前端松耦合 IMU 初始化

IMU 惯性测量单元数据输出频率较高，相机图像输出频率低于 IMU 惯性测量单元，为了使两类传感器数据时间戳能够对齐，本文首先要对 IMU 测量值预先做预积分处理，IMU 全部待优化变量定义为：

$$\mathbf{X} = [\mathbf{p} \quad \mathbf{q} \quad \mathbf{v} \quad \mathbf{b}_g \quad \mathbf{b}_a] \quad (2)$$

$$\mathbf{q} = [q_x \quad q_y \quad q_z \quad q_w] \quad (3)$$

公式 (2) 中 $\mathbf{q}, \mathbf{p}, \mathbf{v}$ 分别表示刚体旋转的四元数，刚体位移和速度变量， \mathbf{b}_g 和 \mathbf{b}_a 分别表示三轴陀螺仪和三轴加速度计的偏置；公式 (3) 中 q_x, q_y, q_z 是旋转四元数对应的虚部， q_w 为四元数的实部。对上述 \mathbf{X} 中的变量依次求导可得公式 (4)：

$$\dot{\mathbf{p}}_b^w = \mathbf{v}_t^w, \dot{\mathbf{v}}_t^w = \mathbf{a}_t^w, \dot{\mathbf{q}}_b^w = \mathbf{q}_b^w \otimes \begin{bmatrix} 0 \\ \frac{1}{2} \boldsymbol{\omega}_{bt} \end{bmatrix} \quad (4)$$

其中，字母上标依次表示在世界坐标系下的位置，速度，旋转量； \otimes 表示四元数之间的乘法， $\boldsymbol{\omega}$ 为陀螺仪角速度变量。对上述变量间的预积分观测量表达形式对时间的积分得式 (5)-(7)：

$$\mathbf{p}_{b_j}^w = \mathbf{p}_{b_i}^w + \mathbf{v}_i^w \Delta t + \iint \mathbf{R}_{bt}^w (\mathbf{a}_{bt} - \mathbf{g}^w) \delta t^2 \quad (5)$$

$$\mathbf{v}_j^w = \mathbf{v}_i^w + \int \mathbf{R}_{bt}^w (\mathbf{a}_{bt} - \mathbf{g}^w) \delta t \quad (6)$$

$$\mathbf{q}_{bj}^w = \int \mathbf{q}_{bt}^w \otimes \begin{bmatrix} 0 \\ \frac{1}{2} \boldsymbol{\omega}_{bt} \end{bmatrix} \delta t \quad (7)$$

上式中 $\mathbf{p}_{b_*}^w$, \mathbf{v}_*^w , $\mathbf{q}_{b_*}^w$ 分别是不同时刻对应的位移、速度和旋转, $\mathbf{R}_{b_*}^w$ 表示相对世界坐标系下对应时刻的 IMU 本体坐标系旋转量。从上面公式中不难看出, 在求解过程中每一次迭代, 都要做相应的积分操作, 计算量相对较大, 因此为避免大量的重复积分运算, 对上述公式做如下变形:

$$\boldsymbol{\alpha}_{b_i b_j} = \iint \mathbf{R}_{b_i b_t} (\mathbf{a}_{bt} - \mathbf{g}^w) \delta t^2 \quad (8)$$

$$\boldsymbol{\beta}_{b_i b_j} = \int \mathbf{R}_{b_i b_t} (\mathbf{a}_{bt} - \mathbf{g}^w) \delta t \quad (9)$$

$$\mathbf{q}_{b_i b_j} = \int \mathbf{q}_{b_i b_t} \otimes \begin{bmatrix} 0 \\ \frac{1}{2} \boldsymbol{\omega}_{bt} \end{bmatrix} \delta t \quad (10)$$

$\boldsymbol{\alpha}_{b_i b_j}$ 、 $\boldsymbol{\beta}_{b_i b_j}$ 和 $\mathbf{q}_{b_i b_j}$ 便是相邻图像帧间的 IMU 预积分观测量。

前端视觉图像匹配以及前端位姿估计精度的稳健性是 IMU 获取准确的初始化结果的关键。考虑到实际运动中视觉图像质量会影响点线特征的提取与匹配, 最终影响前端位姿优化收敛的速度和精度。本文结合预积分观测量, 恢复 IMU 本体坐标在世界坐标系下的旋转 \mathbf{R} 和平移 \mathbf{t} , 然后联合点线投影误差建立视觉和 IMU 的松耦合相机位姿估计模型, 优化并获取稳健的位姿结果, 从而完成 IMU 的初始化, 整个初始化流程见图 9。

在系统前端初始化过程中, 本文运用 IMU 预积分量和点线视觉特征构建统一的松耦合位姿优化代价函数:

$$\boldsymbol{\chi}^* = \arg \min \left\{ \rho_p \sum_{i=1}^n \mathbf{e}_p^T \boldsymbol{\Sigma}_p \mathbf{e}_p + \rho_l \sum_{i=1}^m \mathbf{e}_l^T \boldsymbol{\Sigma}_l \mathbf{e}_l + \lambda \|\mathbf{T} - \mathbf{T}_{imu}\| \right\} \quad (11)$$

公式 (11) 中, λ 表示松耦合正则化比例因子, ρ_* 表示点线优化因子对应的权函数, 这里的权函数采用 Huber 核函数, $\boldsymbol{\Sigma}_*$ 表示点线特征测量值对应的信息矩阵, \mathbf{T} , \mathbf{T}_{imu} 分别表示待优化的 4×4 位姿转移矩阵和 IMU 的旋转 \mathbf{R} 和平移 \mathbf{t} 组建的 4×4 转移矩阵。

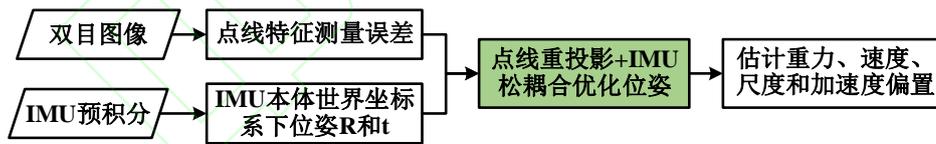


图 9 基于点线特征测量误差和 IMU 松耦合的初始化步骤
Fig. 9 The Initialization Step based on Measurement Error of Point Line feature and IMU Loose Coupling

1.4.2 点线视觉特征误差方程

本文首先对于点特征采用二进制描述子完成点特征的匹配, 通过场景中的三维点特征的重投影与图像对应特征点欧式距离误差, 建立点特征的视觉测量误差模型, 见公式 (12):

$$\mathbf{e}_p = \mathbf{u} - \frac{1}{z} \mathbf{K}^{-1} [\mathbf{R}_{cb} (\mathbf{R}_{bw} \mathbf{P} + \mathbf{t}_{bw}) + \mathbf{t}_{cb}] \quad (12)$$

\mathbf{u} 是图像特征点提取坐标位置, \mathbf{K} 为相机的参数矩阵, \mathbf{P} 为世界坐标系下的三维特征点坐标, \mathbf{R}_{bw} 和 \mathbf{t}_{bw} 表示从世界坐标系转换到惯导本体坐标系下旋转量和平移量, \mathbf{R}_{cb} 和 \mathbf{t}_{cb} 表示从 IMU 本体坐标系转换到像平面坐标系下旋转量和平移量。分别对位姿求解关于 \mathbf{R}_{bw} 和 \mathbf{t}_{bw} 的 6 自由度变量 $\boldsymbol{\zeta}$ 求导建立误差方程 (13):

$$\frac{\partial e_p}{\partial \zeta} = \frac{\partial e_p}{\partial \mathbf{p}_c} \frac{\partial \mathbf{p}_c}{\partial \mathbf{P}_b} \frac{\partial \mathbf{P}_b}{\partial \zeta} \quad (13)$$

其中 \mathbf{P}_c 和 \mathbf{P}_b 分别是场景点特征在相机坐标系和本体坐标系下三维坐标, $\frac{\partial e_p}{\partial \mathbf{p}_c}$ 、 $\frac{\partial \mathbf{P}_b}{\partial \zeta}$ 可以参考文献[36]具体推导过程, $\frac{\partial \mathbf{P}_c}{\partial \mathbf{P}_b} = \mathbf{R}_{cb}$ 。

本文建立的线特征误差模型见图 10, 图中 PQ 表示世界坐标系下的三维线段, p_*q_* 是相对应的图像上检测的二维线特征, pq^* 是 PQ 投影到图像上对应的二维线特征。

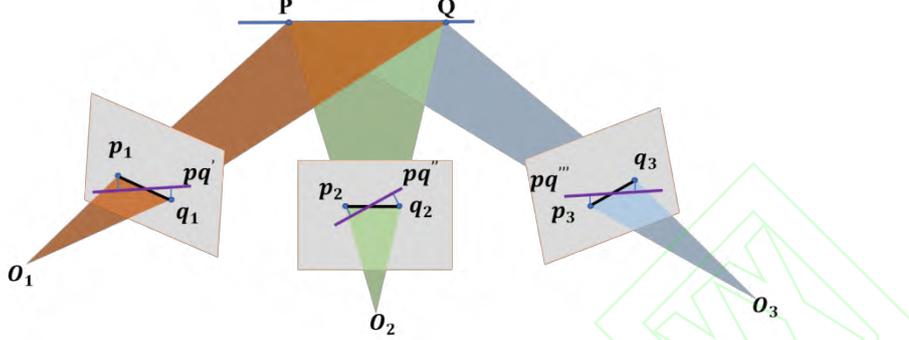


图 10 线特征重投影误差方程示意图
Fig. 10 Schematic Diagram of the Re-projection Error Equation of Line Feature

对于任意线特征匹配对, 基于空间三维线特征重投影到图像上的两个端点到图像对应的二维线特征的距离为约束, 从而建立线特征的视觉测量约束方程。如图 11 所示, 已知线特征端点为 PQ, 线特征重投影到二维图像上为 pq, 由先前图像检测获得的线特征端点二维坐标可以计算线特征直角坐标系下二维表达式 $ax+by+c=0$, 以此建立线特征重投影误差方程:

$$e_l = \begin{cases} d_s = \frac{|au_s + bv_s + c|}{\sqrt{a^2 + b^2}} \\ d_e = \frac{|au_e + bv_e + c|}{\sqrt{a^2 + b^2}} \end{cases} \quad (14)$$

同理, 对于位姿求解关于 \mathbf{R}_{bw} 和 \mathbf{t}_{bw} 的 6 自由度变量 ζ 求导建立误差方程 (15):

$$\frac{\partial e_l}{\partial \zeta} = \begin{cases} a \frac{\partial u_s}{\partial \mathbf{p}_{cs}} \frac{\partial \mathbf{p}_{cs}}{\partial \mathbf{p}_{bs}} \frac{\partial \mathbf{p}_{bs}}{\partial \zeta} + b \frac{\partial v_s}{\partial \mathbf{p}_{cs}} \frac{\partial \mathbf{p}_{cs}}{\partial \mathbf{p}_{bs}} \frac{\partial \mathbf{p}_{bs}}{\partial \zeta} \\ a \frac{\partial u_e}{\partial \mathbf{p}_{ce}} \frac{\partial \mathbf{p}_{ce}}{\partial \mathbf{p}_{be}} \frac{\partial \mathbf{p}_{be}}{\partial \zeta} + b \frac{\partial v_e}{\partial \mathbf{p}_{ce}} \frac{\partial \mathbf{p}_{ce}}{\partial \mathbf{p}_{be}} \frac{\partial \mathbf{p}_{be}}{\partial \zeta} \end{cases} \quad (15)$$

其中 \mathbf{P}_{c*} 和 \mathbf{P}_{b*} 分别是线段端点在相机坐标系和 IMU 本体坐标系下三维坐标, \mathbf{u}_* 、 \mathbf{v}_* 分别表示线特征两个端点二维坐标, $\frac{\partial u_*}{\partial \mathbf{P}_c}$ 、 $\frac{\partial v_*}{\partial \mathbf{P}_c}$ 参考文献[36]中推导过程, 其余内容与 (12) 一致。

1.4.3 视觉惯导后端紧耦合优化模型

在本文构建的视觉与 IMU 融合的 SLAM 方法中, 为进一步提高 SLAM 定位精度和前端位姿跟踪的鲁棒性, 我们的方法基于点线特征和 IMU 预积分量分别与前端的位姿跟踪和后端的局部地图做紧耦合。本文前端位姿跟踪并优化整体的位姿信息, 没有更新世界坐标系下的点线特征的三维坐标信息, 并结合图像帧位姿信息来判定是否为关键帧。一旦确定当前帧为关键帧, 算法进入后端局部地图执行模块, 本文 SLAM 方法后端没有基于滑动窗口优化的算法, 而是基于 ORB-SLAM3 通过构建点线特征共视图, 根据关键帧来实现后端优化。本文方法的待优化变量 χ^* 除了对包括位姿的 6 个变量求导外, 还有点线视觉特征三维坐标, 以及上述相应的 IMU 待优化变量 \mathbf{X} , 运用 IMU 预积分结果和多元视觉特征测量信息值构建

统一的优化代价函数:

$$\mathbf{x}^* = \arg \min \left\{ \rho_p \sum_{i=1}^n \mathbf{e}_p^T \Sigma_p \mathbf{e}_p + \rho_l \sum_{i=1}^m \mathbf{e}_l^T \Sigma_l \mathbf{e}_l + \rho_i \sum_{i=1}^k \mathbf{e}_i^T \Sigma_i \mathbf{e}_i \right\} \quad (16)$$

公式 (16) 中, ρ_* 表示点线优化因子的相应的权函数, 权函数采用 Huber 核函数, Σ_* 表示点线视觉特征和 IMU 预积分量的信息矩阵, \mathbf{e}_* 分别表示构建的点线特征和 IMU 预积分量的误差方程。

2 可行性实验与分析

为了验证自适应点线特征和 IMU 耦合的视觉惯导 SLAM 方法, 我们开展了多项定性和定量实验研究分析, 重点测试本文双目惯性视觉 SLAM 位姿估计精度, 并通过消融实验并与当前 SOTA 算法对比, 从而验证本文提出的自适应点线特征和 IMU 融合的视觉 SLAM 方法的有效性。为进一步验证定位的精度和鲁棒性, 本文在开源数据基准 EuRoC 上开展了一系列定性和定量测试验证实验, 并与 VINS-Fusion、ORB-SLAM3 和 PL-VINS 等视觉惯导 SLAM 方法进行对比分析。本文 SLAM 方法运行平台是 Ubuntu 16.04, 基于 OpenCV3.2.0、Eigen 和 g2o 等 C++ 开源程序库实现。相机位姿精度评价采用绝对运动轨迹的均方根误差为指标 (ATE_RMSE), 该指标能直观地反映相机全局位姿的一致性, 也能准确地反映 SLAM 的性能。假定给定图像帧的估计位姿信息 $\hat{\mathbf{X}}_i$ 和相应的真实运动轨迹信息 \mathbf{X}_i , 则 ATE-RMSE 的计算公式如 (17) 所示, 其中 n 为数据序列中图像帧个数。

$$ATE_RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left\| \text{trans}(\hat{\mathbf{X}}_i) - \text{trans}(\mathbf{X}_i) \right\|^2} \quad (17)$$

最后, 我们利用开源 SLAM 算法评估工具 EVO (<https://github.com/MichaelGrupp/evo>) 来给出不同算法部分数据序列下的相机运动轨迹误差对比图和箱线图, 以及算法在不同序列下绝对位姿均方根误差随时间戳变化的对比图, 从而进一步验证算法的有效性。

2.1 消融实验定性定量验证

实验首先验证了本文方法前端的联合视觉特征和 IMU 松耦合初始化的视觉 SLAM 方法的有效性和鲁棒性。基于本文提出的自适应点线特征检测方法, 在数据集上进行了自适应点线视觉惯导 SLAM (w/o-vins) 和联合 IMU 松耦合初始化的点线视觉惯导 SLAM 方法 (w/vins) 的消融实验。实验结果见表 1, 实验评价指标采用绝对位姿的均方根误差, 表中 Trans. 和 Rot. 分别表示相机运动轨迹平移和旋转的均方根误差, “—” 表示在该数据序列下算法未能完成测试与验证。

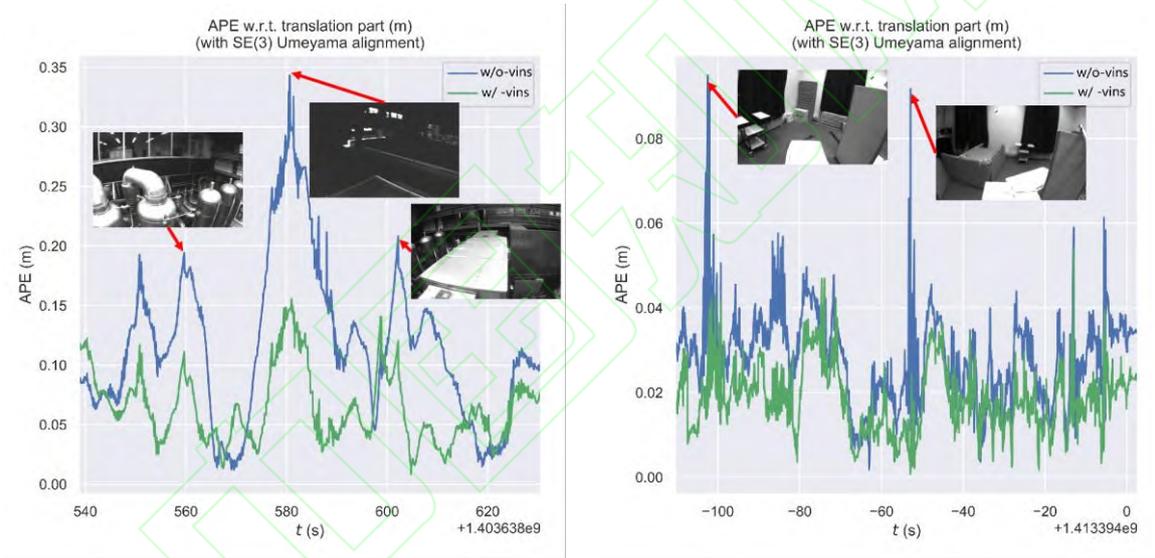
表 1 联合松耦合和紧耦合的视觉惯导 SLAM 消融实验结果对比
Tab. 1 Comparison of Ablation Experimental Results of Visual Inertial SLAM Based on Combination of Loose and Tight Coupling

数据序列	本文方法 (自适应点+Canny 线+紧耦合)		本文方法 (自适应点+Canny 线+松耦合+紧耦合)	
	Trans.(m)	Rot.(rad)	Trans.(m)	Rot.(rad)
MH01	0.052	0.026	0.046	0.025
MH02	0.029	0.022	0.027	0.020
MH03	0.027	0.025	0.029	0.025
MH04	0.060	0.026	0.046	0.026
MH05	0.133	0.017	0.071	0.014
V101	0.036	0.104	0.035	0.103
V102	0.016	0.043	0.015	0.044
V103	0.027	0.050	0.025	0.047

V201	—	—	0.017	0.024
V202	0.028	0.020	0.020	0.019

从表 1 结果中不难发现，前端位姿优化模块通过松耦合 IMU 先验位姿估计，系统能够获取精确稳健的初始化参数，且完成上述所有数据序列下的测试实验，算法定位精度和鲁棒性相比前端未联合 IMU 松耦合的定位精度要高。

图 11 展示了数据集中两组场景挑战性较高的序列（MH05 和 V202）下相机绝对位姿估计量随时间变化的对比图。由图可以明显看出，在相机位姿平移量变化波动较大对应的的时间戳下，视觉图像的光照变化大、纹理稀疏或图像模糊，这对图像特征稳健提取与匹配构成挑战，进而影响整体初始化状态估计。此外，这类环境下视觉和 IMU 初始化结果参数容易存在明显偏差，这些因素会累加影响相机运动轨迹的估计精度。因此，前者（w/o-vins）初始化全部依赖于视觉特征信息，初始化结果稳健性不强，后续相机位姿局部平差过程容易出现较大的偏差。而相比之下，后者（w/-vins）将 IMU 的预积分位姿信息与视觉特征进行松耦合约束，能够有效地增强初始化结果参数的鲁棒性，从而保证后端在局部平差过程中相机运动位姿估计精度。因此，后者添加上 IMU 先验位姿估计数据后相较前者而言，具有较高的整体定位精度和鲁棒性。



(a) MH05 序列下消融实验相机轨迹误差随时间变化

(b) V202 序列消融实验相机位姿误差随时间变化

图 11 MH05 和 V202 数据序列下相机运动的平移分量均方根误差随时间变化示意图

Fig. 11 Schematic Diagram of Root Mean Square Error of Translation Component of Camera Motion over time under MH05 and V202

其次，为验证本文构建的自适应点线特征对 SLAM 整体位姿估计和定位精度的有效性，实验分别基于 ORB 点特征、LSD 线特征以及本文自适应点线特征进行了多项消融对比实验，实验具体结果见表 2。由表中对比结果知，基于自适应点线特征的算法整体位姿估计精度和鲁棒性优于基于 ORB 和 LSD 点线视觉特征下的定位精度。这主要是由于 ORB 点特征中 FAST 角点在检测提取点特征时，通过判断像素的灰度值是否明显高于或低于周围像素，阈值角点的响应阈值较高，导致纹理稀疏或光照不足的图像区域提取的特征数量较少，且关键点分布不均匀。此外，当图像存在光照不均匀或强烈光照变化，LSD 算法性能受到影响，易产生短线、断线，同时检测的部分线段可能与场景物体的边缘结构不符。这类因素会影响线特征的匹配以及误差方程优化过程的收敛速度，从而综合影响整体的定位精度。

表 2 基于 ORB 点特征和 LSD 线特征、自适应点线特征的消融实验结果对比

Tab. 2 Comparison of Ablation Experimental Results based on ORB point features, LSD line features, and Adaptive Point Line Features

数据 序列	本文方法 (ORB+LSD)		本文方法 (自适应点+LSD)		本文方法 (自适应点+自适应线)	
	Trans.(m)	Rot.(rad)	Trans.(m)	Rot.(rad)	Trans.(m)	Rot.(rad)
MH01	0.069	0.027	0.060	0.025	0.046	0.025
MH02	0.028	0.021	0.030	0.023	0.027	0.020
MH03	0.026	0.025	0.025	0.025	0.029	0.025
MH04	0.059	0.027	0.052	0.026	0.046	0.026
MH05	0.080	0.015	0.11	0.014	0.071	0.014
V101	0.038	0.104	0.037	0.103	0.035	0.103
V102	0.017	0.043	0.016	0.043	0.015	0.044
V103	0.029	0.044	0.025	0.048	0.025	0.047
V201	0.037	0.022	0.018	0.024	0.017	0.024
V202	—	—	0.103	0.040	0.020	0.019

2.2 对比实验定性定量验证

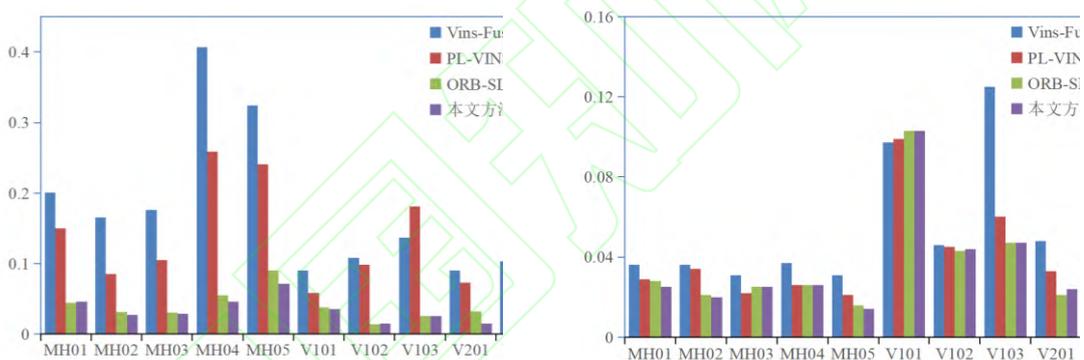
我们又开展与现存的3种常见的经典视觉惯导SLAM方法的对比实验,包括Vins-Fusion、PL-VINS和ORB-SLAM3。Vins-Fusion是由港科大提出的双目视觉和IMU融合的开源视觉惯导SLAM; PL-VINS是湖南大学推出的单目点线视觉惯导SLAM; ORB-SLAM3是西班牙萨拉戈萨大学的视觉和IMU融合的SOTA视觉惯导SLAM算法,上述三种算法位姿估计精度高且鲁棒性较强。表3列出了本文算法以及上述三种算法在数据集上的测试结果。在评价指标方面,实验仍采用相机绝对位姿的平移和旋转量的均方根误差来综合对比。

表3 EuRoc数据集下四种视觉惯导SLAM方法位姿实验结果对比
Tab. 3 Comparison of Pose Experimental Results of Four Visual Inertial SLAM Algorithms on the EuRoc dataset

数 据 序 列	VINS-Fusion (双目+点+IMU)		PL-VINS (单目+点+线+IMU)		ORB-SLAM3 (双目+点+IMU)		本文方法 (自适应点线+IMU)	
	Trans.(m)	Rot.(rad)	Trans.(m)	Rot.(rad)	Trans.(m)	Rot.(rad)	Trans.(m)	Rot.(rad)
M H01	0.201	0.036	0.150	0.029	0.044	0.028	0.046	0.025
M H02	0.165	0.036	0.085	0.034	0.031	0.021	0.027	0.020
M H03	0.176	0.031	0.105	0.022	0.030	0.025	0.029	0.025
M H04	0.407	0.037	0.259	0.026	0.055	0.026	0.046	0.026
M H05	0.324	0.031	0.241	0.021	0.090	0.016	0.071	0.014
V1 01	0.090	0.097	0.058	0.099	0.038	0.103	0.035	0.103
V1 02	0.108	0.046	0.098	0.045	0.014	0.043	0.015	0.044
V1 03	0.137	0.125	0.181	0.060	0.025	0.047	0.025	0.047
V2 01	0.090	0.048	0.073	0.033	0.032	0.021	0.015	0.024
V2 02	0.103	0.040	0.086	0.030	0.014	0.020	0.020	0.019
平	0.18 ↑	0.053 ↑	0.13 ↑	0.040 ↑	0.037 ↑	0.035 ↑	0.033 ↑	0.0347 ↑

从表 3 的实验结果可以明显看出，本文的视觉惯导 SLAM 方法在 8 组数据序列上展现了相对较高的定位精度。相机绝对位姿的平移量和旋转量在大多数数据序列优于上述 3 种视觉惯导 SLAM 算法。尽管在 MH01、V102 和 V202 数据序列中，本文方法虽未能达到最佳结果，但基本上能够获得所有测试实验结果中的次优结果，并且与 ORB-SLAM3 算法的精度差距较小（MH01 和 V102）。在数据序列 V202 中，我们的实验结果相比 ORB-SLAM3 算法定位精度下降 42%，这是因为数据序列 V202 有大量的快速旋转和平移，这对于算法线特征的匹配和追踪是不利的，从而影响位姿估计算法的收敛速度和鲁棒性，如何保证在图像质量下降下的线特征匹配稳健是本文后续需继续研究方向之一。整体综合表 3 的实验对比结果来看，本文方法在平移量和旋转量，相对上述 3 种算法有 0.8%~82%不同程度的提升。

图 12 列出了表 3 中每种 SLAM 方法在不同数据序列旋转和平移量的均方根误差。从图 12 中的位姿和旋转均方根误差柱状图不难看出，VINS-Fusion 在平移和旋转量偏差较大，PL-VINS 增加了线特征，在绝对位姿的平移和旋转量均方根误差有所下降，ORB-SLAM3 相比较 VINS-Fusion 和 PL-VINS 精度有明显提升。综合柱状图，本文的 SLAM 方法相比 ORB-SLAM3 虽在个别数据序列上（V202）位姿估计精度未能达到最优，但能获得次优结果。综合对比以上三种 SLAM 方法位姿估计结果，我们的位姿估计结果具有较强的竞争力。



(a) 算法绝对位姿平移量 (RMSE-Trans.) 均方根误差对比 (b) 算法绝对位姿旋转量 (RMSE-Rot.) 均方根误差对比
图 12 四种 SLAM 方法绝对位姿平移、旋转量均方根误差误差柱状图
Fig. 12 Bar Charts of Root Mean Square Error of Absolute Pose Translation and Rotation from Four SLAM Algorithms

最后，为定性对比研究四种 SLAM 方法的数据集实验测试结果，图 13 和 14 给出了上述四种 SLAM 方法在 MH04、MH05 和 V201 三组数据序列下的相机绝对运动轨迹误差示意图和箱线图。其中图 13 中蓝色线段表示每一时刻算法估计的相机绝对位姿信息，黑色线段是相机的真实运动轨迹，红色线段表示相机估计的运动轨迹信息和实际运动轨迹的差。图中 (a) - (1) 分别表示 VINS-Fusion、PL-VINS、ORB-SLAM3 和本文 SLAM 方法在上述数据序列下的相机运动轨迹结果。从图 13 以及图 14 的箱线图中可以看出，我们的 SLAM 方法得到的相机运动轨迹精度中位数明显小于 Vins-Fusion 和 PL-VINS 两种算法，定位结果整体比较稳定，结果波动幅度较小。同样，对比 ORB-SLAM3 相机运动轨迹，本文方法在 MH05 和 V201 序列下的相机运动轨迹定位精度明显优于前者。

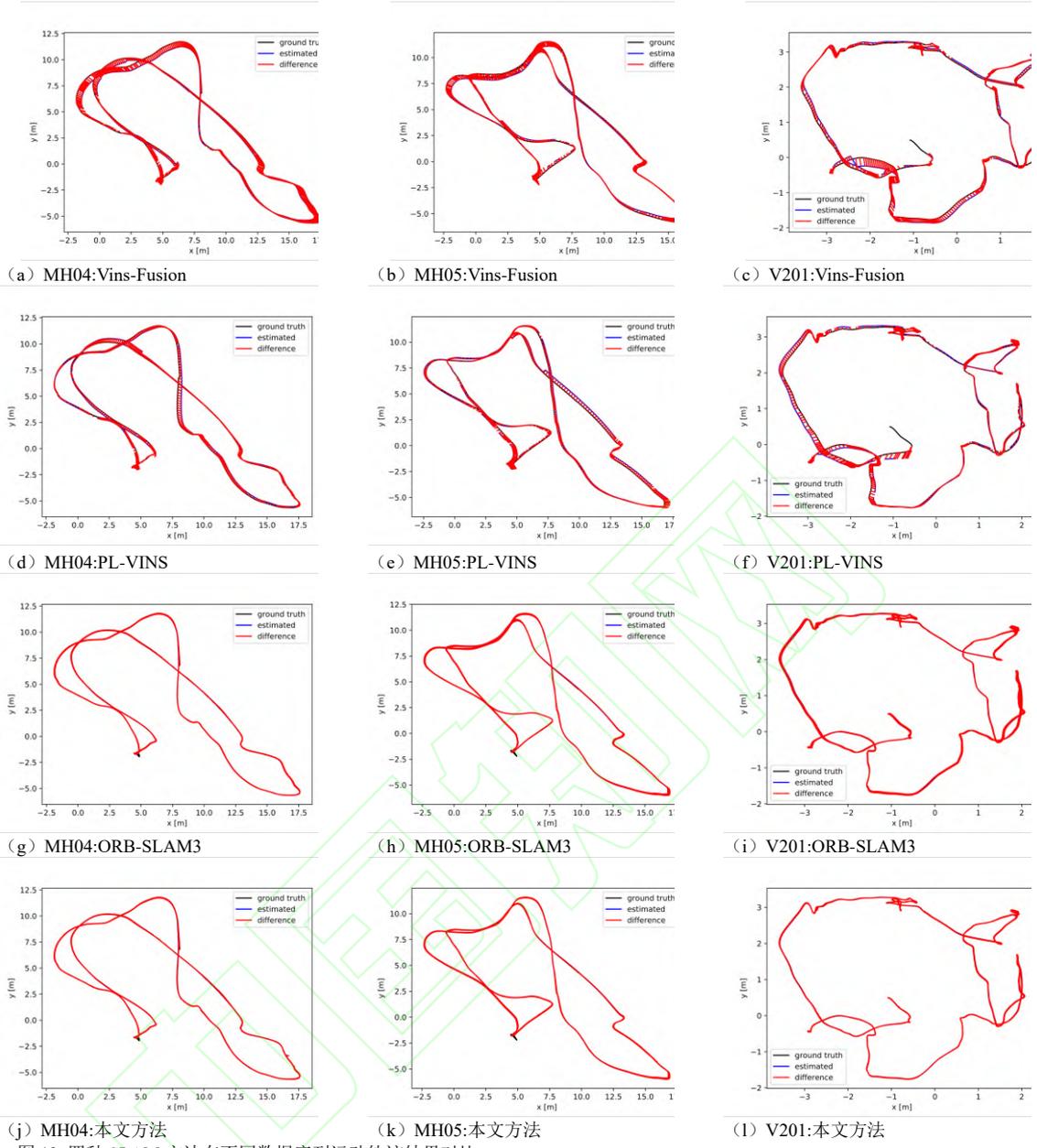


图 13 四种 SLAM 方法在不同数据序列运动轨迹结果对比
 Fig.13 Comparison of Motion Trajectory Results from Four SLAM Algorithms in Different Data Sequences

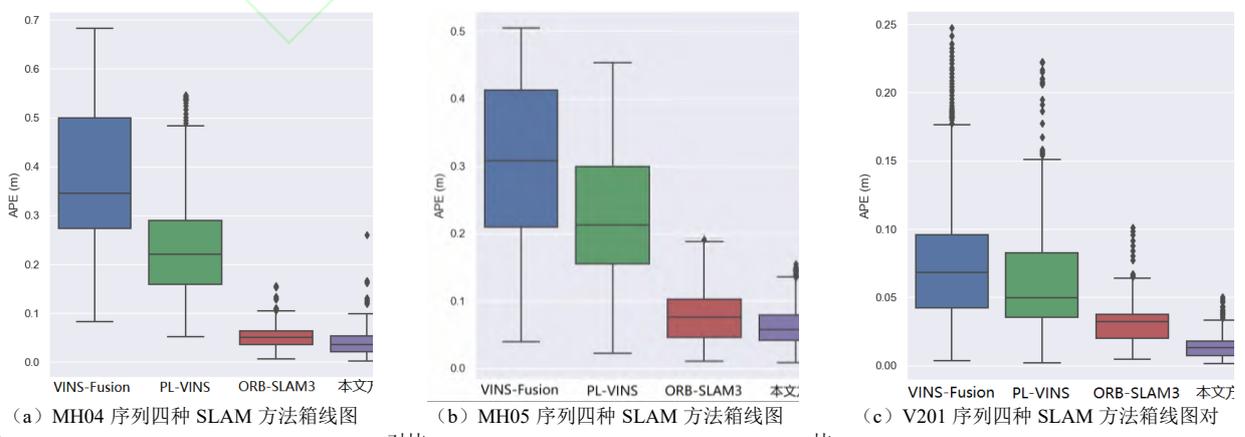


图 14 不同数据序列下各个 SLAM 方法位姿估计箱线图对比结果
 Fig. 14 Comparison Results of Box-plots for Pose Estimation using Different SLAM Algorithms under Different Data Sequences

2.3 时间效率分析

本文最后给出了本文 SLAM 中各个模块的时间运行效率,具体见表 4。从表 4 的统计结果可以看出,在双目图像上提取点线特征大约需要 0.05 秒左右,平均每幅图像需要耗费 0.025 秒左右。前 endpoint 视觉特征和 IMU 预积分测量数据构建图模型需 0.001 秒,而前端优化迭代需要 0.018 秒。此外,考虑到视觉特征跟踪处理和匹配等因素,平均跟踪每帧图像耗费时间大约为 0.07 秒。在后端优化中,视觉特征优化迭代过程需要同时处理点线视觉特征和 IMU 观测数据,所以耗时较长,点线特征位姿更新大约需要 0.022 秒。整个 SLAM 的运行平台是一台常规的 PC 机器,搭载 i5CPU 处理器,没有 GPU 加速。本文 SLAM 方法在同时处理点线特征和 IMU 数据的情况下,能够每秒处理大约 14 帧,基本满足实时性要求。然而,对于进一步提升方法的运行效率,我们将继续在下一步研究工作中优化代码,从而提高程序执行效率。

表 4 各模块运行的时间效率(单位: s)

Tab. 4 The Execution efficiency of Each Module Involved in Our Method (unit: s)

特征提取		前端特征跟踪处理与图优化			后端图优化与特征位姿更新		
图像点特征提取平均效率	图像线特征提取平均效率	跟踪处理每帧图像平均时间效率	前 endpoint 视觉特征和 IMU 预积分测量数据构建图模型平均时间效率	前端优化迭代过程平均时间效率	后端优化模型构建平均效率	后端图优化迭代过程时间效率	点线特征位姿更新时间效率
0.022	0.028	0.070	0.001	0.009	0.018	0.40	0.022

3 结语

针对室内低/弱纹理、光照不足的场景下,视觉惯导 SLAM 算法难以检测并追踪足够的视觉特征,并且未能充分利用 IMU 测量信息,从而导致算法定位精度低、鲁棒性弱的问题,本文提出了一种基于自适应点线特征与 IMU 融合的视觉 SLAM 方法。本文方法的创新点主要有三个方面:(1) 构建了自适应的点线特征检测提取方法,改善了在不同场景图像中点线特征的提取效果,并增强了特征提取和匹配结果的鲁棒性;(2) 充分利用 IMU 先验观测数据,基于前端松耦合的方式,在初始化阶段融合自适应点线特征,提供了稳健的初始化结果,为后续的特征跟踪和后端优化提供了较好的初值。(3) 后端基于紧耦合的方式,融合点线视觉特征信息构造后端非线性优化平差模型,从而构建了自适应点线特征与 IMU 耦合的视觉 SLAM。

最后,我们在开源数据基准上开展测试验证实验,并与几种典型的视觉惯导 SLAM 方法做了定性和定量对比分析。实验结果显示,相比其他视觉惯导 SLAM 方案,本文构建的 SLAM 方法的定位精度至少提升了 12%,有效证明了基于自适应点线特征并联合 IMU 松耦合与紧耦合对于提高视觉 SLAM 整体的定位精度和鲁棒性的有效性。

参考文献

- [1] Shi Junyi, Zha Fusheng, Sun Lining, et al. A Survey of Visual-Inertial SLAM for Mobile Robots[J]. *Robot*, 2020, 42(6): 734-748.(施俊屹,查富生,孙立宁,等.移动机器人视觉惯性 SLAM 研究进展[J].*机器人*,2020,42(6): 734-748.)
- [2] Liao Jianchi, Li Xingxing, Feng Shaoquan. GVIL: Tightly-Coupled GNSS PPP/Visual/INS/LiDAR SLAM Based on Graph Optimization [J]. *Geomatics and Information Science of Wuhan University*, 2023, 48(7): 1204-1215.(廖健驰,李星星,冯绍权. GVIL: 基于图优化的 GNSS PPP/视觉/惯性/激光雷达紧耦合算法[J].*武汉大学学报(信息科学版)*, 2023, 48(7):1204-1215.)
- [3] Lou Yidong, Wang Yusheng, Tu Zhiyong, et al. Real Time Localization and Mapping Integrating Multiple Prism LiDARs/IMU/RTK on Railway Locomotive[J].*Geomatics and Information Science of Wuhan University*, 2021, 46(12): 1802-1807.(楼益栋,王昱升,涂智勇,等.融合多棱镜式雷达/IMU/RTK 的轨道车辆高精度实时定位与建图[J].*武汉大学学报(信息科学版)*, 2021, 46(12):1802-1807.)
- [4] Gong Jianya. Chances and Challenges for Development of Surveying and Remote Sensing in the Age of Artificial Intelligence[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(12): 1788-1796.(龚健雅.人工智能时代测绘遥感技术的发展机遇与挑战[J].*武汉大学学报(信息科学版)*, 2018, 43(12): 1788-1796.)
- [5] Di Kaichang, Wan Wenhui, Zhao Hongying, et al. Progress and Applications of Visual SLAM [J]. *Acta Geodaetica et Cartographica Sinica*, 2018, 47(6): 770-779.(邸凯昌,万文辉,赵红颖,等.视觉 SLAM 技术的进展与应用[J].*测绘学报*, 2018, 47(6): 770-779.)

- [6] R.Mur-Artal, J.M.M.Montiel, J.D.Tardos. ORB-SLAM: A versatile and accurate monocular SLAM system[J]. *IEEE Transactions on Robotics*,2015,31(5):1147-1163.
- [7] R.Mur-Artal, J.D.Tardós. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras[J]. *IEEE Transactions on Robotics*,2016,33(5):1255-1262.
- [8] J.Engel, V.Koltun, D.Cremers. Direct Sparse Odometry[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017,40(3):611-625.
- [9] Kerl C, Sturm J, Cremers D. Dense visual SLAM for RGB-D cameras[C]//2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. Tokyo, Japan, 2013.
- [10] Engel J,Schöps T, Cremers D. LSD-SLAM: Large-scale direct monocular SLAM[C]// European conference on computer vision. Springer, Cham,2014.
- [11] H. Liu, G. Zhang and H. Bao. Robust Keyframe-Based Monocular SLAM for Augmented Reality[C]//2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct), Merida, Mexico, 2016
- [12] H. Zhou, D. Zou, L. Pei, R. Ying, P. Liu and W. Yu, StructSLAM: Visual SLAM With Building Structure Lines[J]. *IEEE Transactions on Vehicular Technology*,2015,64(4),1364-1375.
- [13] L. Chen, L. Sun, T. Yang, et al. RGB-T SLAM: A flexible SLAM framework by combining appearance and thermal information[C]// 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore,2017.
- [14] Y. Zhou, G. Gallego and S. Shen.Event-Based Stereo Visual Odometry[J]. *IEEE Transactions on Robotics*,2021,37(5),1433-1450.
- [15] Zeng Qinghua, Luo Yixue, Sun Kecheng, et al. Overview of the Development of SLAM Technology for Vision and Its Fusion Inertia [J]. *Journal of Nanjing University of Aeronautics and Astronautics*, 2022,54 (06):1007-1020.(曾庆化,罗怡雪,孙克诚,等.视觉及其融合惯性的SLAM技术发展综述[J].南京航空航天大学学报,2022,54(06):1007-1020.)
- [16] A.Mourikis, S.Roumeliotis. A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation[C]//Proceedings 2007 IEEE International Conference on Robotics and Automation, Rome, Italy, 2007.
- [17] S.Leutenegger,Furgalep,V.Rabaud,et al. Keyframe based visual inertial SLAM using nonlinear optimization[J].*The international Journal of Robotics and Research*, 2015,34(3):314-334.
- [18] C.Forster, L.Carlone, F.Dellaer, et al. On-Manifold Pre-integration for Real-Time Visual Inertial Odometry[J]. *IEEE Transactions on Robotics*,2017,33(1):1-21.
- [19] R.Mur-Artal, J.D.Tardos. Visual-inertial monocular SLAM with map reuse[J]. *IEEE Robotics and Automation Letters*, 2017, 2 (2): 796-803.
- [20] C.Campos,R.Elvira, J.J.G.Rodríguez, J.M.M.Montiel, J.D.Tardós. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial, and Multi-map SLAM[J]. *IEEE Transactions on Robotics*, 2021,37(6),1874-1890.
- [21]T.Qin,P.Li,S.Shen.VINS-Mono:a robust and versatile monocular visual inertial state estimator [J].*IEEE Transactions on Robotics*, 2018,34(4):1004-1020.
- [22] T.Qin, S.Cao, J.Pan, et al. A General Optimization-based Framework for Global Pose Estimation with Multiple Sensors[J]. ArXiv. /abs/1901.03642.
- [23] Gong Zhaohui, Zhang Xiaoli, Peng Xiafu, et al. Semi-Direct Monocular Visual Odometry Based on Visual-Inertial Fusion[J]. *Robot*, 2020,42(5): 595-605. (龚赵慧,张霄力,彭侠夫,等.基于视觉惯性融合的半直接单目视觉里程计[J].机器人, 2020, 42(05): 595-605.)
- [24] Xu Zhibin, Li Hongwei, Zhang Bin, et al. Localization method of mobile robot based on binocular vision and inertial navigation[J]. *Acta Geodaetica et Cartographica Sinica*,2021,50(11):1512- 1521.(许智宾,李宏伟,张斌,等.双目视觉与惯导融合的移动机器人定位方法[J].测绘学报,2021, 50(11): 1512-1521.)
- [25] Zhang Fubing, Zhang Bingshuo, Yang Yushuai. SLAM Algorithm Based on Monocular IMU Odometer Fusion[J]. *Acta Armamentarii*, 2022,43(11):2810-2817.(张福斌,张炳烁,杨玉帅.基于单目/IMU/里程计融合的SLAM算法[J].兵工学报,2022,43 (11): 2810-2818.)
- [26] Y.He, J.Zhao, Y.Guo, et al. PL-VIO: Tightly-Coupled Monocular Visual-Inertial Odometry Using Point and Line Features[J]. *Sensors*,2018,18(4):1159-1184.
- [27] Q.Fu, J.Wang, H.Yu, et al. PL-VINS: Real-Time Monocular Visual-Inertial SLAM with Point and Line [J]. arXiv preprint arXiv: 2009.07462, 2020.
- [28] J. Lee, S.-Y. Park. PLF-VINS: Real-Time Monocular Visual-Inertial SLAM With Point-Line Fusion and Parallel-Line Fusion[J]. *IEEE Robotics and Automation Letters*,2021,6(4):7033-7040.
- [29] Zhao Liangyu, Jin Rui, Zhu Yeqing, et al. Stereo visual-inertial SLAM algorithm based on merge of point and line features[J]. *acta aeronauticaet astronautica sinica*,2022,43(3): 325117-325117.(赵良玉,金瑞,朱叶青,等.基于点线特征融合的双目惯性SLAM算法[J].航空学报,2022,43 (3):325117-325117.)
- [30] Ying Wenjian, Pan Linhao, She Bo, et al. Binocular Vision Inertial Navigation SLAM Algorithm Based on Fusion of Point Line Features [J]. *Journal of Naval Engineering University*, 2021,33(06): 106-112.(应文健,潘林豪,佘博,等.融合点线特征的双目视觉-惯导SLAM算法[J].海军工程大学学报, 2021,33(06):106-112.)
- [31] Wei Shuangfeng, Shi Xianjie, Liu Zhenbin,et al. Optimization of Visual Inertial Odometer with Point Line Joint [J]. *Surveying and Mapping Science*, 2021,46 (04): 20-27+36.(危双丰,师现杰,刘振彬,等.点线联合的优化视觉惯性里程计[J].测绘科学,2021,46(04):20- 27+36.)
- [32] R. Grompone von Gioi, J. Jakubowicz, J. Morel, et al. LSD: A fast line segment detector with a false detection control[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(4): 722-732.
- [33] Jia D, Zhu Ningdan, Yang Ning Hua, et al. Image matching methods[J]. *Journal of Image and Graphics*,2019,24(05):0677-0699. (贾迪,朱宁丹,杨宁华,等.图像匹配方法研究综述[J].中国图象图形学报,2019,24(05):677-699.)
- [34] X. Lu, J. Yaoy, H. Li. 2-Line Exhaustive Searching for Real-Time Vanishing Point Estimation in Manhattan World[C]//2017 IEEE Winter Conference on Applications of Computer Vision (WACV), Santa Rosa, CA, USA, 2017,.
- [35] L. Zhang, R.Koch. An efficient and robust line segment matching approach based on LBD descriptor and pairwise geometric consistency[J].*Journal of Visual Communication and Image Representation*, 2013, 24(7):794-805.
- [36] Gao Xiang, Zhang Tao, Yan Qinrui,et al. Visual SLAM Lecture 14: From Theory to Practice[M]. Beijing: Electronic Industry Press, 2017.(高翔,张涛,颜沁睿,等.视觉SLAM十四讲:从理论到实践[M].北京:电子工业出版社, 2017.)

网络首发:

标题: 一种自适应点线特征和 IMU 耦合的视觉 SLAM 方法

作者: 张晨阳, 杨健

收稿日期: 2024-06-28

DOI:10.13203/j.whugis20230347

引用格式:

张晨阳, 杨健. 一种自适应点线特征和 IMU 耦合的视觉 SLAM 方法[J]. 武汉大学学报(信息科学版), 2024, DOI: 10.13203/j.whugis20230347 (ZHANG Chenyang, YANG Jian. A Visual SLAM Method Coupled with Adaptive Point-line Features and IMU[J]. Geomatics and Information Science of Wuhan University, 2024, DOI: 10.13203/j.whugis20230347)

