引文格式: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),2025,50(10):2048-2063.DOI:10.13203/j.whugis20230347

Citation: ZHANG Chenyang, YANG Jian. A Visual SLAM Method Coupled with Adaptive Point-Line Features and IMU[J]. Geomatics and Information Science of Wuhan University, 2025, 50(10):2048-2063.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[1,2] 杨 健[1,3,4]

1 南京理工大学计算机科学与工程学院,江苏 南京,210094

2 常州工学院土木建筑工程学院,江苏 常州,213032

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

4 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,视觉惯导融合的即时定位与建图(simultaneous localization and mapping, 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惯性测量单元(inertial measurement unit, IMU)耦合的视觉SLAM方法。首先设计一种自适应的快速角点特征检测算法,以增强图像特征点检测的鲁棒性。另外,快速线特征检测算法易检测短线、断线,且图像因光照变化易导致线特征的"过提取"或"错提取"。因此,利用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,并对比几种经典SLAM方法。实验结果表明,所构建的SLAM方法平均定位精度至少提高12%,同时具有较强的鲁棒性。

关键词:SLAM;自适应点线特征;IMU;松耦合;紧耦合

中图分类号:P237 **文献标识码**:A **收稿日期**:2024-06-20

DOI:10.13203/j.whugis20230347 **文章编号**:1671-8860(2025)10-2048-16

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

ZHANG Chenyang[1,2] *YANG Jian*[1,3,4]

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

2 School of Civil Engineering and Architecture, Changzhou Institute of Technology, Changzhou 213032, China

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

4 Jiangsu Key Laboratory of Image and Video Understanding for Social Safety, Nanjing University of Science and Technology, Nanjing 210094, China

Abstract: **Objectives**: Visual-inertial simultaneous localization and mapping (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 (IMU) is often underutilized, resulting in reduced pose estimation accuracy and limited robustness. **Methods**: An adaptive fast corner feature detection approach has been developed to enhance the robustness of feature

基金项目:中国博士后科学基金面上项目(2023M741702);国家自然科学基金(42401533,62361166670);江苏省高等学校基础科学研究面上项目(24KJB420002)。

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,教授。csjyang@njust.edu.cn

point detection in images. Additionally, the fast line segment detector algorithm makes it easy to detect short lines 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:** The proposed 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 the proposed algorithm improves average localization accuracy by at least 12% and displays significant robustness.

Key words: SLAM; adaptive point-line features; IMU; 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这一概念自1986年在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，被认为是实现机器人自主性的关键。SLAM作为机器人实际生产和生活应用（如定位导航、自动驾驶、仓储物流等）中的核心基础[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，SLAM的发展受社会信息化和智能化浪潮的推动，也极大地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，同时实现场景地图的构建[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早期算法基本是以视觉图像作为数据源输入逐渐发展起来的，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[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]、香港科技大学机器人研究所基于事件相机的ES-VO[14]等相关的经典（state of the art, 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。

惯性测量单元（inertial measurement unit, IMU）不受上述外界因素的影响，其与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主要为紧耦合和松耦合两类。松耦合是指IMU和相机分别估计自身运动状态，再对两者的位姿估计结果融合，该方法算法复杂度较低；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]。早期紧耦合多是通过滤波来实现的，比如采用扩展的卡尔曼滤波MSCKF算法[16]，该算法在标准扩展的卡尔曼滤波框架下采用滑动窗口来进行姿态估计。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，并用历史图像帧来进行优化求解，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，容易导致运动轨迹的误差累积；后续推出的OKVIS2具有回环检测功能，并在开源数据基准EuRoC上测试验证，结果表明其精度和稳

健性相比OKVIS均有显著性提升。

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,文献[18]推导了流形空间下IMU预积分的数学表达式,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],以及香港科技大学的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,该团队又发布了支持多源传感器的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,许多研究工作通过改进前端位姿估计以提高整个VINS的精度。例如,文献[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。文献[24]在对比分析了视觉里程计特征点法和直接法的优缺点后,将惯性测量单元与这两种方法进行融合,提出了一种半直接法的视觉惯导里程计,以提高相机定位的精度和鲁棒性。此外,文献[25]建立图像金字塔结合LK光流法来改进VINS-Mono的前端性能,并将陀螺仪的旋转量和里程计的平移量作为先验条件,同时融合车轮里程计信息,将里程计约束应用于初始化和SLAM后端非线性优化中,实现了视觉、IMU和车轮里程计的高效融合。

除了点特征,线特征也引入了VINS算法的研究中。文献[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对线特征参数化表示,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,文献[27]通过对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。文献[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。文献[29]和文献[30]使用IMU测量数据为视觉定位算法提供先验和尺度信息,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。此外,文献[31]通过优化加速度计偏置的初始化策略与点线特征数据的选择策略,获取了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的研究现状分析,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,由于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算法采用快速线特征提取(line segment detector,LSD)算法[32]。然而,LSD算法对图像上的噪声和光照敏感,算法的性能会因此受影响,导致部分检测的线特征并非真正来源于场景边缘,从而对算法的定位建图是不利的。此外,现有视觉惯性算法在初始化及前端位姿估计阶段,往往未能充分挖掘IMU预积分所提供的先验信息。针对上述问题,本文首先设计了一种自适应阈值角点特征检测算法,并结合Kd-tree均匀化图像上的特征点;其次,基于Canny算子边缘检测生成的二值图像,设计一种自适应线特征提取算法,并利用消影点特性对线特征进行筛选与聚类,尽可能确保所提取的线特征真实反映场景中的物理边缘。然后,前端基于点线视觉特征和IMU预积分观测量,以松耦合方式获取稳健的初始化参数,而后端则通过紧耦合优化框架,联合优化点线重投影误差与IMU预积分量,最终搭建室内场景下的自适应点线特征与IMU耦合的视觉SLAM方法。

1 方法原理

图1展示了本文SLAM方法的整体流程。算法采用多线程并行架构,从左至右依次运行3个线程:前端位姿估计初始化和特征追踪、后端局部光束法平差(bundle adjustment,BA)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物理边缘结构信息；(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的坐标；(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，便执行全局光束法平差，以此提升视觉惯导SLAM算法定位结果的准确性和鲁棒性。

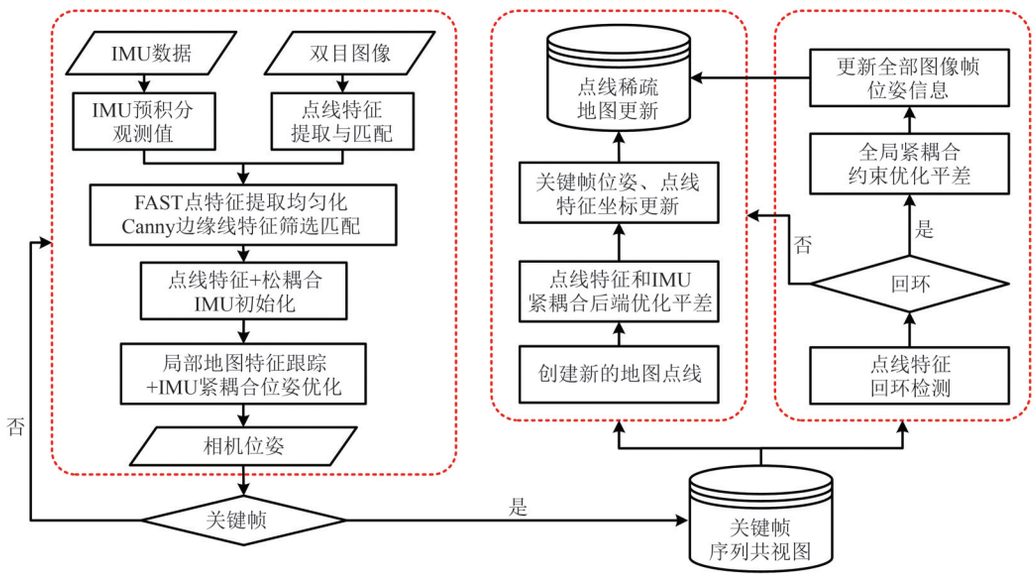


图1　方法流程图

Fig. 1　The Flow Chart of Method

1.1　自适应点特征提取

快速角点特征检测(features from accelerated segment test，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，进而确定图像中关键点的位置。该方法能够实现高效的角点特征检测，并对图像噪声具备一定的鲁棒性，已成为图像处理领域中广泛应用的特征检测算法之一[33]。然而，FAST算法倾向于在纹理丰富区域提取大量关键点，而在纹理稀疏区域则提取较少，导致角点响应值偏高、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提取算法，并引入Kd-tree结构以实现图像角点的均匀化分布。算法执行过程见图2，具体分为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，考虑计算成本和存储开销，金字塔层数设置为8层，逐层计算待提取特征点数量；(3)计算划分为若干单元格，并分配每个单元格中需检测的特征点数目；(4)在各单元格内进行FAST角点检测，并利用Kd-tree算法均匀化单元格内角点特征；(5)计算角点特征总数，判断特征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，响应阈值降低1，并返回步骤4；(6)若角点响应阈值已降至下限，或角点特征总数已达到预设值的90%，则终止当前算法。

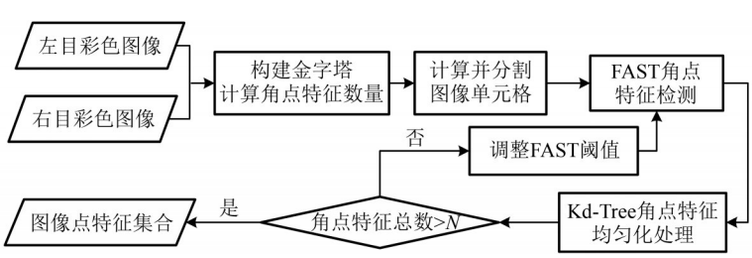
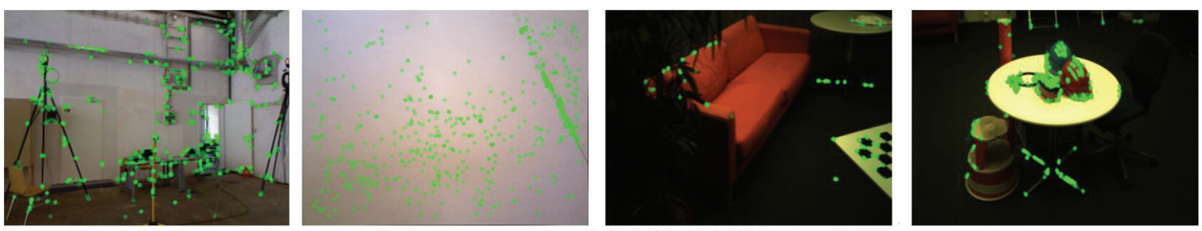


图2　自适应角点特征提取流程

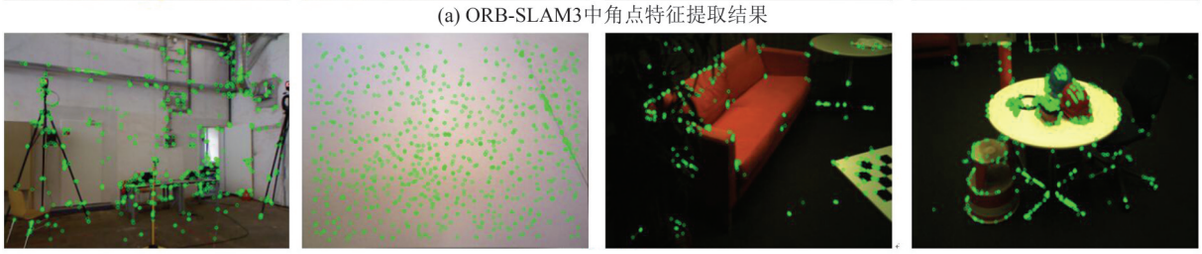
Fig. 2　Flowchart of Adaptive Corner Feature Extraction

为检验上述自适应角点特征检测算法，这里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角点特征提取算法与

ORB-SLAM3中的角点特征提取算法进行对比实验。定性对比实验中图像特征点数量设置为1 000,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仅需循环迭代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,从定性对比结果中不难看出,本文构建的自适应角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角点特征提取结果。



(a) ORB-SLAM3中角点特征提取结果

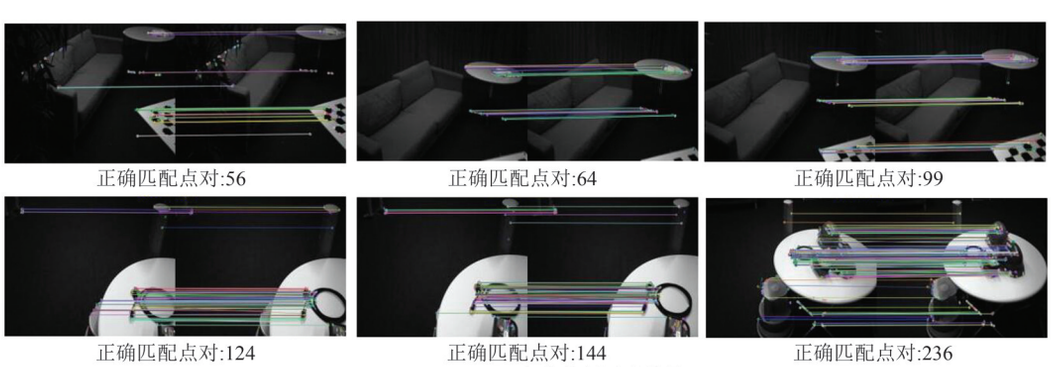


(b) 自适应角点特征提取算法结果

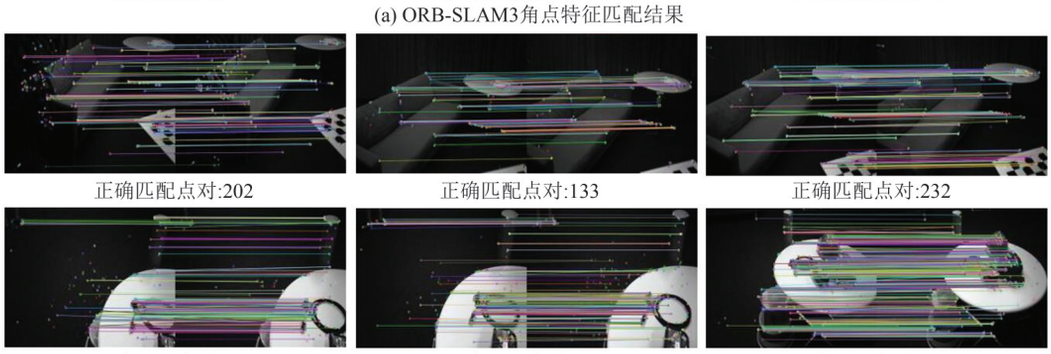
图3 不同场景下两种角点特征提取算法定性对比结果

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

图4展示了本文自适应角点特征提取和ORB-SLAM3的角点特征提取结果的特征匹配阶段的定性定量实验结果。这两种方法都基于角点特征描述子完成匹配。从特征点数量和整体分布来看,本文方法提取的角点特征数量更多,并借助Kd-tree实现了特征点在图像中的均匀分布。相比ORB-SLAM3,更丰富的特征点有助于提高潜在匹配对的数量。根据图4中的实验对比结果,基于自适应角点特征算法的匹配结果比ORB-SLAM3中的角点特征匹配结果能得到更



(a) ORB-SLAM3角点特征匹配结果



(b) 自适应角点特征匹配结果

图4 两种角点特征提取算法匹配结果定性对比

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

多特征匹配对,验证了其在提升匹配性能方面的有效性。

1.2　图像边缘信息自适应线特征提取

目前,图像线特征的检测与提取常采用LSD算法。该算法基于边缘像素的直方图统计实现直线段检测,具有计算速度快、检测精度高等优点,并可通过参数调整适应不同场景下的线特征提取需求[32]。然而,LSD算法也存在一定的局限性,容易产生过多短线段或断裂线段,同时对光照变化较为敏感,可能导致提取的线特征与真实物体边缘结构不一致。为解决上述问题,本文提出一种基于图像边缘信息的自适应线特征提取算法,其流程如图5所示。首先对输入图像进行高斯模糊预处理,并计算整幅图像灰度值的中位数;随后,基于该中位数自适应设定高低阈值,调用OpenCV中的Canny边缘检测函数获取边缘二值图像。接下来,算法遍历边缘图像中的每一个非零像素点,将其作为种子点,沿8个邻域方向进行搜索,统计连续非零像素的数量,并判断是否满足预设的线特征长度阈值。对符合长度条件的像素集合,利用最小二乘法进行线段拟合,从而提取出候选线特征。之后,对空间距离较近的线段进行合并处理,以优化线特征的结构完整性。为提升算法可用性与可复用性,本文将该线特征提取方法嵌入到OpenCV开源库中,封装为新的线特征提取接口,并扩展了高斯金字塔构建与多尺度线特征检测功能,最终实现了自适应边缘线特征提取函数的集成与调用。

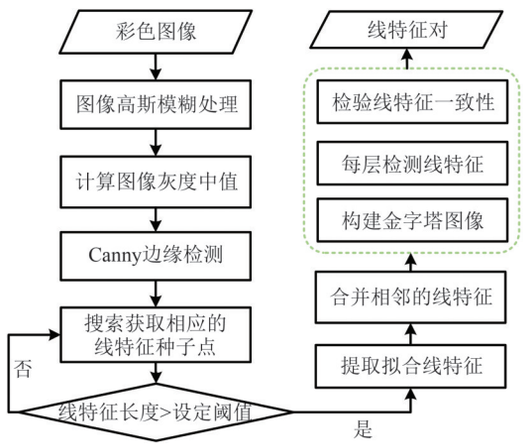
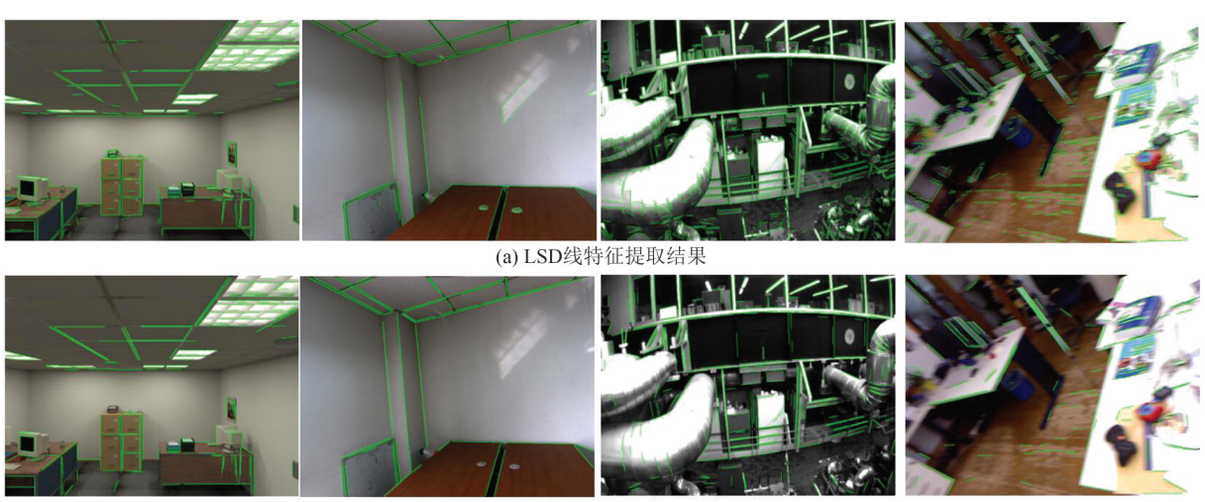


图5　基于图像边缘信息的自适应线特征提取流程

Fig. 5　Flowhart of Adaptive Line Feature Extraction Based on Image Edge Information

为验证上述基于图像边缘特征的自适应线特征提取算法,本文开展了多组定性对比实验。将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,结果见图6。



(a) LSD线特征提取结果

(b) 自适应线特征提取结果

——— 提取的线特征

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

实验结果表明,在存在噪声或者光照变化时,LSD算法容易受干扰,导致检测结果稳定性差。LSD算法提取的线特征往往较为碎片化,出现大量短线和断裂线段,且在个别光照条件复杂的图像中易发生误检。相比之下,本文构建的基于图像边缘信息的自适应线特征提取算法结果不存在短线和断裂线段,基本未出现因光照因素导致的错误线段提取。与LSD算法相比,本文设计的提取检测算法在线特征的完整性、稳定性和抗干扰能力方面均表现出明显优势。

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阈值为20、30、40像素下测试对比,结果见图7。实验结果表明,当长度阈值为20像素时,线特征检测数量相对较多,但结果存在大量短线片段;当长度预定阈值为提升为30像素时,线特征提取效果有明显改善;当阈值进一步增大至40像素时,场景中部分边缘线特征存在损失。综合考虑点线特征联合位姿估计的准确性与算法效率,本文最终将线特征检测的长度阈值设定为30像素,即图像上检测提取的线段最短长度为30像素。

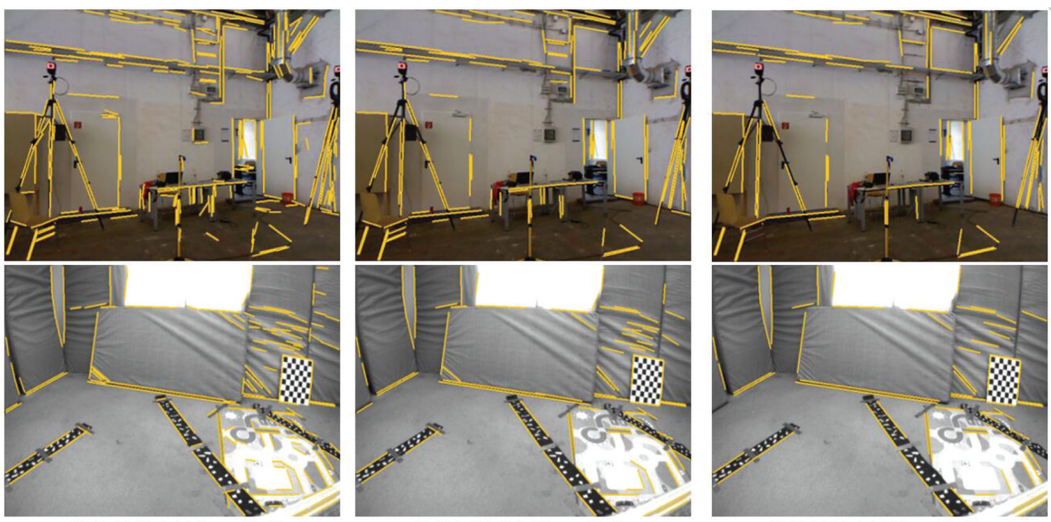


图7 不同长度阈值条件下的线特征提取结果

Fig. 7 Line Feature Extraction Results Under Different Length Threshold 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。基于消影点这一特性,对线特征作进一步筛选,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,首先采用双向匹配来获取粗糙的线特征匹配对集合;其次,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。以上几何约束的计算公式为:

$$\begin{cases} d_1 \times d_2 < T_1 \\ \alpha < T_2 \end{cases} \tag{1}$$

式中,线特征匹配对阈值T_1为30×30像素,角度阈值T_2为$\pi/8$。相机输出频率高,相邻图像帧旋转运动量较小,所以T_2设为$\pi/8$。本文线特征提取算法最短线段长度为30像素,式(1)中d_1和d_2应取$\min(d_1,30)$和$\min(d_2,30)$,所以T_1设为30×30像素。

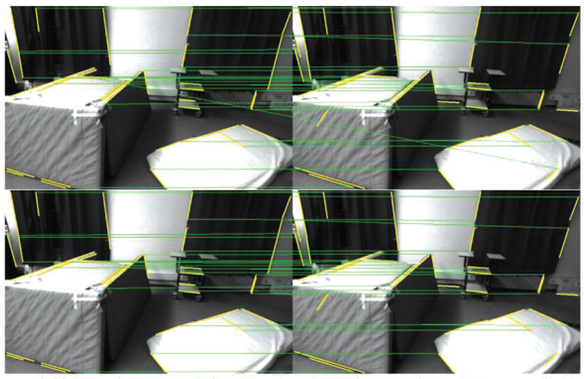
本文选取EuRoc数据集上V202和MH05两组不同序列的图像开展线特征匹配实验,线特征匹配对比结果见图8。首先,经过消影点线特征的筛选,本文自适应线特征的检测提取结果基本符合场景内的物体边缘信息;其次,根据图8特征匹配策略结果对比,增加几何约束能有效剔除错误的线匹配对,提高线特征匹配精度。

1.4 联合松耦合和紧耦合的视觉惯导优化

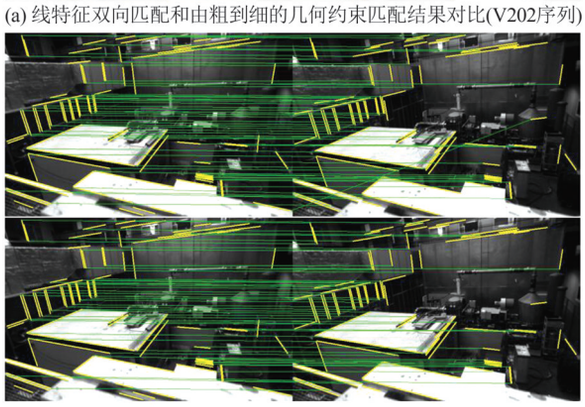
为提升算法整体定位的鲁棒性和精度,本文结合松耦合和紧耦合的优势,充分融合IMU的测量信息,联合点线特征重投影误差模型和IMU预积分姿态信息,在前端和后端协调同步完成相机位姿优化。具体而言,前端特征追踪主要通过点线视觉特征重投影误差和IMU先验位姿信息,通过松耦合得到系统稳健的初始化结果;后端则是通过点线视觉特征重投影误差和IMU预积分观测量,通过紧耦合建立误差方程,进而实现局部地图优化。

1.4.1 视觉惯导前端松耦合IMU初始化

IMU惯性测量单元数据的输出频率较高,而



(a) 线特征双向匹配和由粗到细的几何约束匹配结果对比(V202序列)



(b) 线特征双向匹配和由粗到细的几何约束匹配结果对比(MH05序列)

提取的线特征　线特征匹配对连接线

图 8　由粗到细的匹配策略的线特征匹配结果对比

Fig. 8　Comparison of Matching Result of Line Feature Matching Based on Matching Strategy from Coarse to Fine

相机图像的输出频率较低，为了实现两类传感器数据时间戳能够对齐，本文首先要对IMU测量值进行预积分处理，IMU的全部待优化变量定义为：

$$\boldsymbol{X}=\begin{bmatrix}\boldsymbol{p} & \boldsymbol{q} & \boldsymbol{v} & \boldsymbol{b}_g & \boldsymbol{b}_a\end{bmatrix} \tag{2}$$

式中，$\boldsymbol{q}$、$\boldsymbol{p}$、$\boldsymbol{v}$分别表示刚体旋转的四元数、刚体位移和速度变量；$\boldsymbol{b}_g$和$\boldsymbol{b}_a$分别表示三轴陀螺仪和三轴加速度计的偏置。$\boldsymbol{q}$为：

$$\boldsymbol{q}=\begin{bmatrix}q_x & q_y & q_z & q_w\end{bmatrix} \tag{3}$$

式中，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$\boldsymbol{X}$中的变量依次求导可得：

$$\dot{\boldsymbol{p}}_{bi}^{w}=\boldsymbol{v}_{bi}^{w},\dot{\boldsymbol{v}}_{bi}^{w}=\boldsymbol{a}_{bi}^{w},\dot{\boldsymbol{q}}_{bi}^{w}=\boldsymbol{q}_{bi}^{w}\otimes\begin{bmatrix}0\\\frac{1}{2}\boldsymbol{\omega}_{bi}\end{bmatrix} \tag{4}$$

式中，$\boldsymbol{p}_{bi}^{w}$、$\boldsymbol{v}_{bi}^{w}$、$\boldsymbol{q}_{bi}^{w}$依次表示惯性测量单元在i时刻世界坐标系中的位置、速度、旋转量；$\otimes$表示四元数之间的乘法操作；$\boldsymbol{\omega}_{bi}$为陀螺仪在i时刻的角速度变量。对上述变量在i、j时刻Δ_t时间段内的预积分观测量表达形式对时间小量δ_t的积分得：

$$\boldsymbol{p}_{bj}^{w}=\boldsymbol{p}_{bi}^{w}+\boldsymbol{v}_{i}^{w}\Delta t+\iint\boldsymbol{R}_{bt}^{w}(\boldsymbol{a}_{bt}-\boldsymbol{g}^{w})\delta t^{2} \tag{5}$$

$$\boldsymbol{v}_{bj}^{w}=\boldsymbol{v}_{bi}^{w}+\int\boldsymbol{R}_{bt}^{w}(\boldsymbol{a}_{bt}-\boldsymbol{g}^{w})\delta t \tag{6}$$

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

式中，$\boldsymbol{p}_{b*}^{w}$、$\boldsymbol{v}_{b*}^{w}$、$\boldsymbol{q}_{b*}^{w}$分别是不同时刻对应的位移、速度和旋转；$\boldsymbol{R}_{b*}^{w}$表示相对世界坐标系下对应时刻间的IMU本体坐标系旋转量。

不难看出，在求解过程中每一次迭代都要进行相应的积分操作，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：

$$\boldsymbol{\alpha}_{bibj}=\iint\boldsymbol{R}_{bibt}(\boldsymbol{a}_{bt}-\boldsymbol{g}^{w})\delta t^{2} \tag{8}$$

$$\boldsymbol{\beta}_{bibj}=\int\boldsymbol{R}_{bibt}(\boldsymbol{a}_{bt}-\boldsymbol{g}^{w})\delta t \tag{9}$$

$$\boldsymbol{q}_{bibj}=\int\boldsymbol{q}_{bibt}\otimes\begin{bmatrix}0\\\frac{1}{2}\boldsymbol{\omega}_{bt}\end{bmatrix}\delta t \tag{10}$$

式中，$\boldsymbol{\alpha}_{bibj}$、$\boldsymbol{\beta}_{bibj}$和$\boldsymbol{q}_{bibj}$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。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，首先恢复IMU本体坐标在世界坐标系下的旋转量$\boldsymbol{R}$和平移向量$\boldsymbol{t}$，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：

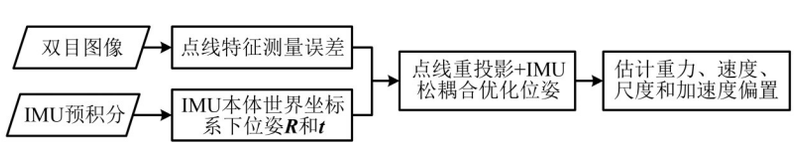


图 9　基于点线特征测量误差和IMU松耦合的初始化步骤

Fig. 9　Initialization Step Based on Measurement Error of Point-Line Feature and IMU Loose Coupling

$$\chi^{*}=\arg\min\Big(\rho_{p}\sum_{p=1}^{n}\boldsymbol{e}_{p}^{\mathrm{T}}\boldsymbol{\Sigma}_{p}\boldsymbol{e}_{p}+\rho_{l}\sum_{l=1}^{m}\boldsymbol{e}_{l}^{\mathrm{T}}\boldsymbol{\Sigma}_{l}\boldsymbol{e}_{l}+\lambda\left\|T-T_{\mathrm{IMU}}\right\|\Big)\tag{11}$$

式中，χ^{*}为全部待优化变量$\boldsymbol{X}$；λ表示松耦合正则化比例因子；ρ_{p}和ρ_{l}分别表示点线优化因子对应的权函数；$\boldsymbol{e}_{p}$和$\boldsymbol{e}_{l}$分别表示点线误差向量（具体见§1.4.2内容）。此外，本文采用Huber核函数，$\boldsymbol{\Sigma}_{*}$表示点线特征测量值对应的信息矩阵，T、T_{IMU}分别表示待优化的4×4位姿转移矩阵和IMU积分后的旋转量$\boldsymbol{R}$和平移量$\boldsymbol{t}$组建的4×4转移矩阵。

1.4.2 点线视觉特征误差方程

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氏距离误差建立点特征的视觉测量误差模型：

$$\boldsymbol{e}_{p}=\boldsymbol{u}-\frac{1}{z}\boldsymbol{K}^{-1}\left[\boldsymbol{R}_{cb}\left(\boldsymbol{R}_{bw}\boldsymbol{P}+\boldsymbol{t}_{bw}\right)+\boldsymbol{t}_{cb}\right]\tag{12}$$

式中，$\boldsymbol{u}$是图像上特征点二维坐标；$\boldsymbol{K}$为相机的参数矩阵；$\boldsymbol{P}$为世界坐标系下的三维特征点坐标；$\boldsymbol{R}_{bw}$和$\boldsymbol{t}_{bw}$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；$\boldsymbol{R}_{cb}$和$\boldsymbol{t}_{cb}$分别表示从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（3自由度旋转量、3自由度平移量）求解，则误差函数$\boldsymbol{e}_{p}$对$\boldsymbol{R}_{bw}$和$\boldsymbol{t}_{bw}$中的6自由度位姿变量$\boldsymbol{\xi}$求导为：

$$\frac{\partial\boldsymbol{e}_{p}}{\partial\boldsymbol{\xi}}=\frac{\partial\boldsymbol{e}_{p}}{\partial\boldsymbol{P}_{c}}\frac{\partial\boldsymbol{P}_{c}}{\partial\boldsymbol{P}_{b}}\frac{\partial\boldsymbol{P}_{b}}{\partial\boldsymbol{\xi}}\tag{13}$$

式中，$\boldsymbol{P}_{c}$和$\boldsymbol{P}_{b}$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向量；$\frac{\partial\boldsymbol{e}_{p}}{\partial\boldsymbol{P}_{c}}$、$\frac{\partial\boldsymbol{P}_{b}}{\partial\boldsymbol{\xi}}$可以参考文献[36]具体推导过程；$\frac{\partial\boldsymbol{P}_{c}}{\partial\boldsymbol{P}_{b}}=\boldsymbol{R}_{cb}$。

本文建立的线特征误差模型见图10，图10中PQ表示世界坐标系下的三维线段，p_{*}和q_{*}是相对应的图像上检测的线特征端点，p^{*}和q^{*}是PQ投影到图像帧上对应的线特征端点，O_{1}、O_{2}、O_{3}表示相机摄影中心。

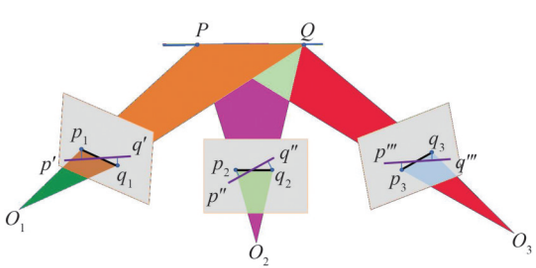


图10 线特征重投影误差方程示意图

Fig. 10 Diagram of Re-projection Error Equation of Line Feature

对于任意一对匹配的线特征，基于空间三维线段重投影至图像平面后所得二维线段的两端点，以图像中实际提取的对应二维线特征之间的距离作为约束条件，建立线特征的视觉测量约束方程。如图10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的二维表达式$ax+by+c=0$，以此建立线特征重投影误差方程：

$$\boldsymbol{e}_{l}=\begin{cases}d_{s}=\dfrac{\left|au_{s}+bv_{s}+c\right|}{\sqrt{a^{2}+b^{2}}}\\[2ex] d_{e}=\dfrac{\left|au_{e}+bv_{e}+c\right|}{\sqrt{a^{2}+b^{2}}}\end{cases}\tag{14}$$

式中，u_{*}、v_{*}分别表示线特征两个端点二维坐标分量。

同理，对于位姿求解关于$\boldsymbol{R}_{bw}$和$\boldsymbol{t}_{bw}$的6自由度变量$\boldsymbol{\xi}$求导建立误差方程：

$$\frac{\partial\boldsymbol{e}_{l}}{\partial\boldsymbol{\xi}}=\begin{cases}a\dfrac{\partial\boldsymbol{u}_{s}}{\partial\boldsymbol{p}_{cs}}\dfrac{\partial\boldsymbol{p}_{cs}}{\partial\boldsymbol{p}_{bs}}\dfrac{\partial\boldsymbol{p}_{bs}}{\partial\boldsymbol{\xi}}+b\dfrac{\partial\boldsymbol{v}_{s}}{\partial\boldsymbol{p}_{cs}}\dfrac{\partial\boldsymbol{p}_{cs}}{\partial\boldsymbol{p}_{bs}}\dfrac{\partial\boldsymbol{p}_{bs}}{\partial\boldsymbol{\xi}}\\[2ex] a\dfrac{\partial\boldsymbol{u}_{e}}{\partial\boldsymbol{p}_{ce}}\dfrac{\partial\boldsymbol{p}_{ce}}{\partial\boldsymbol{p}_{be}}\dfrac{\partial\boldsymbol{p}_{be}}{\partial\boldsymbol{\xi}}+b\dfrac{\partial\boldsymbol{v}_{e}}{\partial\boldsymbol{p}_{ce}}\dfrac{\partial\boldsymbol{p}_{ce}}{\partial\boldsymbol{p}_{be}}\dfrac{\partial\boldsymbol{p}_{be}}{\partial\boldsymbol{\xi}}\end{cases}\tag{15}$$

式中，$\boldsymbol{p}_{c*}$和$\boldsymbol{p}_{b*}$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向量；$\boldsymbol{u}_{*}$、$\boldsymbol{v}_{*}$分别表示线特征两个端点二维坐标向量；$\frac{\partial\boldsymbol{u}_{*}}{\partial\boldsymbol{p}_{c*}}$、$\frac{\partial\boldsymbol{v}_{*}}{\partial\boldsymbol{p}_{c*}}$参考文献[36]中推导过程。

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，系统基于点线特征和IMU预积分量，分别与前端的位姿跟踪和后端的局部地图做耦合优化。前端位姿跟踪模块负责估计并优化系统的整体位姿，但不更新世界坐标系中点、线特征的三维坐标。同时，系统根据当前图像帧的位姿信息判断其是否作为关键帧。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方法中后端没有基于滑动窗口优化的算法，而是借鉴ORB-SLAM3框架构建点线特征共视图，并基于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，还有点线特征三维坐标，以及上述相应的IMU待优化变量$\boldsymbol{X}$，运用IMU预积分结果和点线特征测量信息构建统一的优化代价函数：

$$\chi^{*}=\arg\min\left\{\rho_{p}\sum_{p=1}^{n}\boldsymbol{e}_{p}^{\mathrm{T}}\boldsymbol{\Sigma}_{p}\boldsymbol{e}_{p}+\rho_{l}\sum_{l=1}^{m}\boldsymbol{e}_{l}^{\mathrm{T}}\boldsymbol{\Sigma}_{l}\boldsymbol{e}_{l}+\rho_{i}\sum_{i=1}^{k}\boldsymbol{e}_{i}^{\mathrm{T}}\boldsymbol{\Sigma}_{i}\boldsymbol{e}_{i}\right\} \tag{16}$$

式中，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；$\boldsymbol{\Sigma}_{*}$表示点线视觉特征和IMU预积分量的信息矩阵；$\boldsymbol{e}_{*}$分别表示构建的点线特征和IMU预积分量的误差向量。

2　可行性实验与分析

为了验证本文自适应点线特征和IMU耦合的视觉惯导SLAM方法，本文开展了多项定性和定量实验研究分析，重点评估其在双目惯性配置下的位姿估计精度，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算法运行平台是Ubuntu 16.04，并基于OpenCV3.2.0、Eigen和g2o等C++开源程序库编程实现。位姿精度评价采用绝对运动轨迹的均方根误差（absolute trajectory error root mean square error，ATE-RMSE）为指标，其能直观地反映相机全局位姿的一致性，是衡量SLAM系统性能的重要依据。假定给定n幅图像帧的估计位姿信息$\hat{\boldsymbol{X}}_{i}$和相应的真实运动轨迹信息$\boldsymbol{X}_{i}$，则ATE-RMSE E_{t}的计算公式为：

$$E_{t}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left\| T(\hat{\boldsymbol{X}}_{i})-T(\boldsymbol{X}_{i})\right\|_{2}^{2}} \tag{17}$$

式中，n为数据序列中图像帧个数；$T(\ )$表示位姿信息中的平移部分。

本文采用开源SLAM算法评估工具EVO对不同算法在多个数据序列上的表现进行定量分析。绘制相机运动轨迹误差对比图、误差分布箱线图以及各序列ATE-RMSE随时间戳的变化曲线，进一步验证所提算法的有效性与鲁棒性。

2.1　消融实验定性定量验证

首先，对前端联合视觉特征与IMU松耦合初始化的视觉SLAM方法的有效性与鲁棒性进行了验证。基于本文提出的自适应点线特征检测算法，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。实验结果如表1所示，评价指标基于式（17）中的E_{t}，Trans.和Rot.分别表示ATE-RMSE中平移与旋转部分误差分量，空白则表示相应算法在该序列中未能完成测试或跟踪失败。

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

Table 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＋自适应线＋紧耦合）		本文方法（自适应点＋自适应线＋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IMU松耦合的方式来估计先验位姿，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）下连续时序下图像帧绝对位姿估计误差随时间变化的对比图。在平移量变化波动较大对应的时间戳下，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

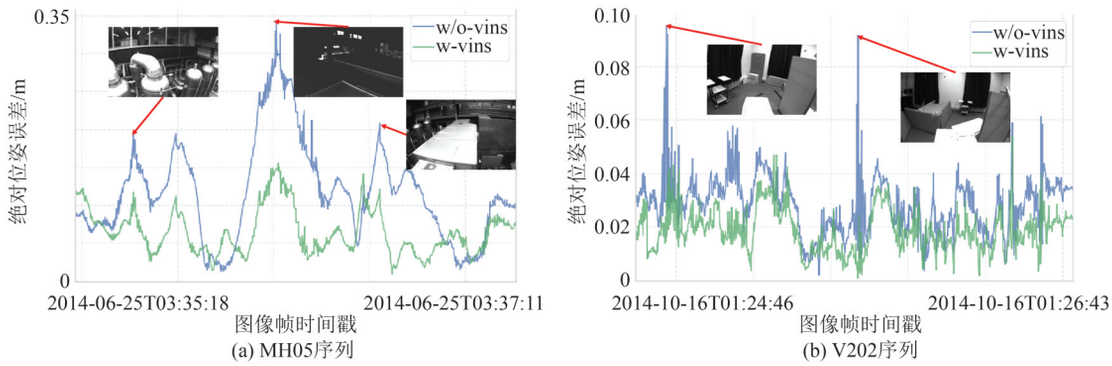


图11 MH05和V202序列下图像帧绝对位姿误差随时间变化对比

Fig. 11 Comparison of Absolute Pose Errors of Image Frames Over Time Under MH05 and V202 Sequences

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。

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,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组消融对比实验,实验具体结果见表2。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

Table 2 Comparison of Ablation Experimental Results Based on ORB Point Feature, LSD Line Feature, and Adaptive Point-Line Feature

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	0.110	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

为全面评估本文算法的性能,本文选取了3种当前主流的视觉惯性SLAM方法做了对比实验,包括VINS-Fusion、PL-VINS和ORB-SLAM3。VINS-Fusion是由香港科技大学提出的双目视觉和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3种算法位姿估计精度高且鲁棒性较强。表3列出了本文方法以及上述3种方法在数据集EuRoC上的测试结果。在评价指标方面，仍采用ATE-RMSE中的平移和旋转误差量来综合对比，以综合衡量各方法的定位精度，↑表示本文算法相对3种算法在平移和旋转量精度有提升，括号内百分数表示结果提升的百分比。

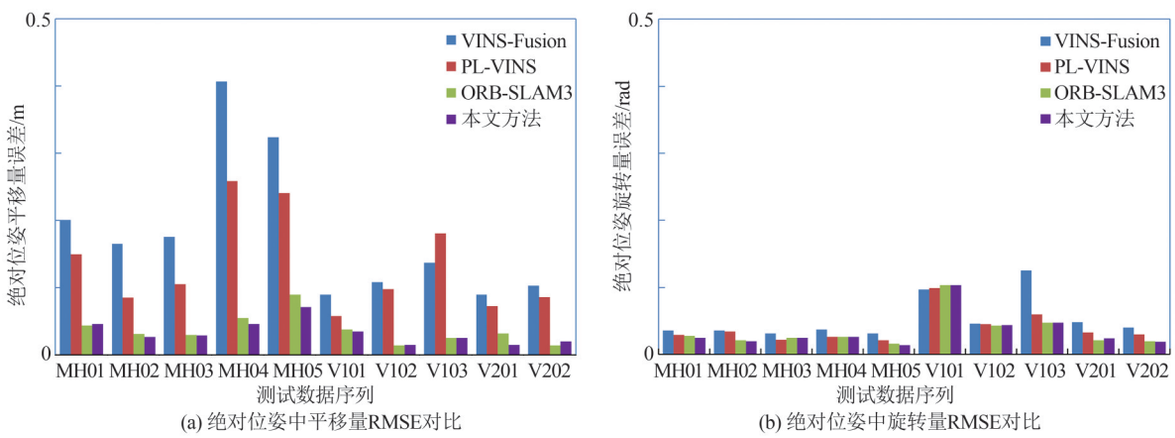
表3　EuRoc数据集下4种视觉惯导SLAM方法位姿实验结果对比

Table 3　Comparison of Pose Experimental Results of 4 Visual Inertial SLAM Methods 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
MH01	0.201	0.036	0.150	0.029	0.044	0.028	0.046	0.025
MH02	0.165	0.036	0.085	0.034	0.031	0.021	0.027	0.020
MH03	0.176	0.031	0.105	0.022	0.030	0.025	0.029	0.025
MH04	0.407	0.037	0.259	0.026	0.055	0.026	0.046	0.026
MH05	0.324	0.031	0.241	0.021	0.090	0.016	0.071	0.014
V101	0.090	0.097	0.058	0.099	0.038	0.103	0.035	0.103
V102	0.108	0.046	0.098	0.045	0.014	0.043	0.015	0.044
V103	0.137	0.125	0.181	0.060	0.025	0.047	0.025	0.047
V201	0.090	0.048	0.073	0.033	0.032	0.021	0.015	0.024
V202	0.103	0.040	0.086	0.030	0.014	0.020	0.020	0.019
平均	0.180（82%）↑	0.053（34%）↑	0.130（75%）↑	0.040（13%）↑	0.037（12%）↑	0.035（0.8%）↑	0.033	0.0347

从表3的实验结果可以看出，本文的视觉惯导SLAM方法在8组数据序列上整体表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，本文方法的实验结果相比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RMSE。



(a) 绝对位姿中平移量RMSE对比　　(b) 绝对位姿中旋转量RMSE对比

图12　4种SLAM方法在EuRoC数据集下的绝对位姿平移、旋转量RMSE误差柱状图

Fig. 12　Bar Charts of RMSE of Absolute Pose Translation and Rotation from 4 SLAM Methods with EuRoC Dataset

从图12中的位姿平移量和旋转RMSE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RMSE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3种SLAM方法位姿估计结果,本文方法在整体位姿估计精度方面展现出较强的竞争力与鲁棒性。

最后,为定性对比研究4种SLAM方法的数据集实验测试结果,图13和14分别给出了上述4种SLAM方法在MH04、MH05和V201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,红色线段则代表估计值与真实值之间的偏差。结合图14的箱线图分析可知,本文SLAM方法相较于VINS-Fusion与PL-VINS所得轨迹误差的中位数明显偏低,表明其定位结果更为精确,且整体稳定性更高,波动范围较小。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。

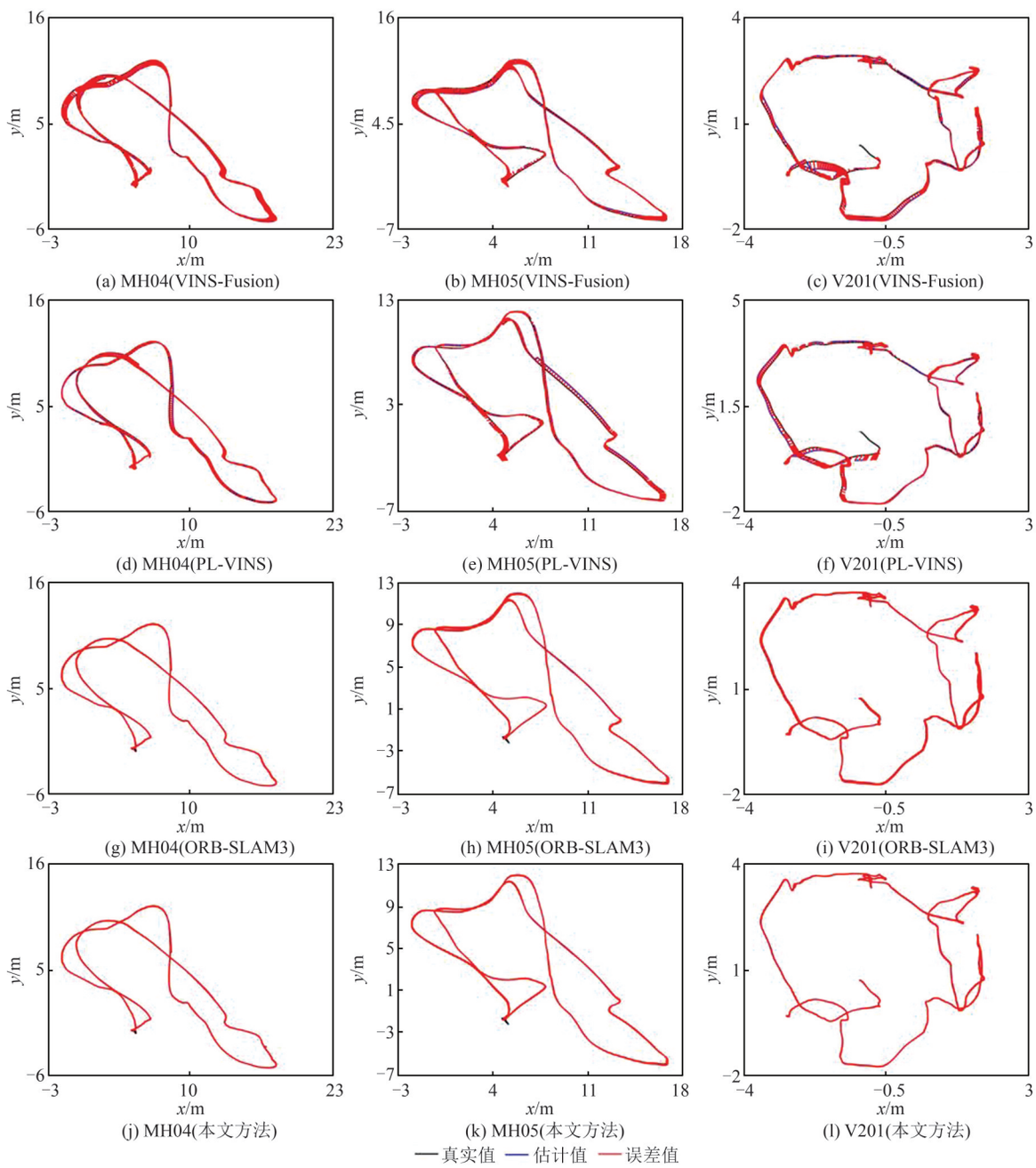


图13 4种SLAM方法在不同数据序列运动轨迹结果对比

Fig. 13 Comparison of Motion Trajectory Results from 4 SLAM Methods in Different Data Sequences

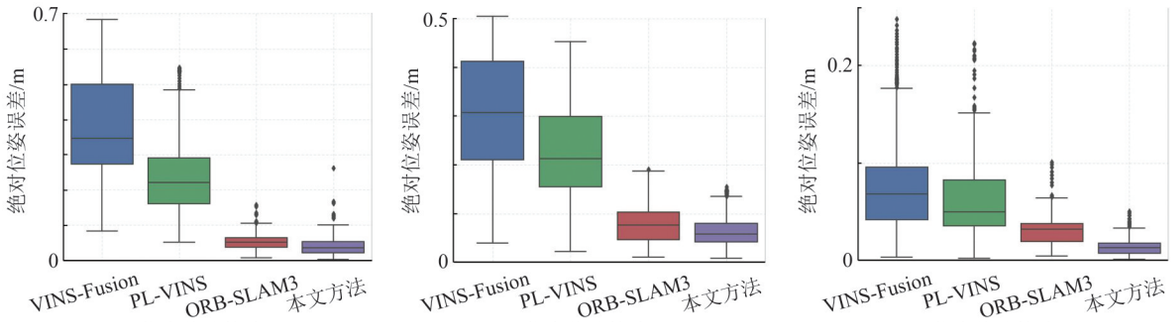


图 14　不同数据序列下各个SLAM方法位姿估计箱线图对比结果

Fig. 14　Comparison Results of Box-Plots for Pose Estimation Using Different SLAM Methods Under Different Data Sequences

2.3　时间效率分析

本文最后对所提SLAM系统中各模块的运行效率进行了分析，具体如表4所示。在双目图像上提取点线特征大约需要0.05 s，平均每幅图像需要耗费0.025 s左右。前端点线视觉特征和IMU预积分测量数据构建图模型需0.001 s，而前端优化迭代需要0.018 s。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0.070 s。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观测量，所以耗时较长，点线特征位姿更新大约需要0.022 s。整个系统运行于一台搭载Intel i5处理器、未使用GPU加速的普通PC机上。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。

表4　各模块运行的时间/s

Table 4　Execution Time of Each Module/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时间	图像线特征提取平均时间	跟踪处理每帧图像平均时间	前端点线-IMU模型构建平均时间	前端优化迭代过程平均时间	后端优化模型构建平均时间	后端图优化迭代过程时间	点线特征位姿更新时间
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，以及IMU测量信息利用不充分等问题，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3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先验观测量，通过前端松耦合的方式，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稳健的初值。(3)后端基于紧耦合的方式，联合点线特征信息构造非线性优化平差模型，实现自适应点线特征与IMU数据的高效融合，从而构建了自适应点线特征与IMU耦合的视觉SLAM。

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。实验结果表明，相比其他视觉惯导SLAM方案，本文构建的SLAM方法的定位精度相比ORB-SLAM3、PL-VINS和VINS-Fusion算法至少提升了12%，充分证明了自适应点线特征与IMU松-紧耦合策略在提升视觉SLAM精度与鲁棒性方面的有效性。

参　考　文　献

[1]　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.
SHI Junyi，ZHA Fusheng，SUN Lining，et al. A Survey of Visual-Inertial SLAM for Mobile Robots [J]. *Robot*，2020，42(6)：734-748.

[2]　廖健驰，李星星，冯绍权. GVIL：基于图优化的GNSS PPP/视觉/惯性/激光雷达紧组合算法[J].

武汉大学学报(信息科学版), 2023, 48(7): 1204-1215.
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.

[3] 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.
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.

[4] 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.
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.

[5] 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.
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.

[6] MUR-ARTAL R, MONTIEL J M M, TARDÓS J D. ORB-SLAM: A Versatile and Accurate Monocular SLAM System[J]. *IEEE Transactions on Robotics*, 2015, 31(5): 1147-1163.

[7] MUR-ARTAL R, TARDÓS J D. ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras[J]. *IEEE Transactions on Robotics*, 2017, 33(5): 1255-1262.

[8] ENGEL J, KOLTUN V, CREMERS D. Direct Sparse Odometry[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 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. Large-Scale Direct SLAM with Stereo Cameras[C]//2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 2015.

[11] LIU H M, ZHANG G F, BAO H J. Robust Keyframe-Based Monocular SLAM for Augmented Reality[C]//IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct), Merida, Mexico, 2016.

[12] ZHOU H Z, ZOU D P, PEI L, et al. StructSLAM: Visual SLAM with Building Structure Lines[J]. *IEEE Transactions on Vehicular Technology*, 2015, 64(4): 1364- 1375.

[13] CHEN L, SUN L B, YANG T, et al. RGB-T SLAM: A Flexible SLAM Framework by Combining Appearance and Thermal Information[C]//IEEE International Conference on Robotics and Automation (ICRA), Singapore, 2017.

[14] ZHOU Y, GALLEGO G, SHEN S J. Event-Based Stereo Visual Odometry[J]. *IEEE Transactions on Robotics*, 2021, 37(5): 1433-1450.

[15] 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(6): 1007-1020.
ZENG Qinghua, LUO Yixue, SUN Kecheng, et al. Review on SLAM Technology Development for Vision and Its Fusion of Inertial Information[J]. *Journal of Nanjing University of Aeronautics and Astronautics*, 2022, 54(6): 1007-1020.

[16] MOURIKIS A I, ROUMELIOTIS S I. A Multi-state Constraint Kalman Filter for Vision-Aided Inertial Navigation[C]//IEEE International Conference on Robotics and Automation, Rome, Italy, 2007.

[17] LEUTENEGGER S, LYNEN S, BOSSE M, et al. Keyframe-Based Visual-Inertial Odometry Using Nonlinear Optimization[J]. *The International Journal of Robotics Research*, 2015, 34(3): 314-334.

[18] FORSTER C, CARLONE L, DELLAERT F, et al. On-Manifold Preintegration for Real-Time Visual: Inertial Odometry[J]. *IEEE Transactions on Robotics*, 2017, 33(1): 1-21.

[19] MUR-ARTAL R, TARDÓS J D. Visual-Inertial Monocular SLAM with Map Reuse[J]. *IEEE Robotics and Automation Letters*, 2017, 2(2): 796-803.

[20] CAMPOS C, ELVIRA R, RODRÍGUEZ J J G, et al. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial, and Multimap SLAM[J]. *IEEE Transactions on Robotics*, 2021, 37(6): 1874-1890.

[21] QIN T, LI P L, SHEN S J. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State

Estimator [J]. *IEEE Transactions on Robotics*, 2018, 34(4): 1004-1020.

[22] QIN T, CAO S Z, PAN J, et al. A General Optimization-Based Framework for Global Pose Estimation with Multiple Sensors[EB/OL]. (2019-1-11)[2022-7-10]. https://arxiv.org/pdf/1901.03642v1.

[23] 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(5)：595-605.
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.

[24] 许智宾. 基于双目视觉与惯导融合的移动机器人室内定位技术研究[D]. 郑州：郑州大学，2021.
XU Zhibin. Research on Indoor Positioning Technology of Mobile Robot Based on Fusion of Binocular Vision and Inertial Navigation[D]. Zhengzhou: Zhengzhou University, 2021.

[25] 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.
ZHANG Fubin, ZHANG Bingshuo, YANG Yushuai. SLAM Algorithm Based on Monocular/IMU/Odometer Fusion[J]. *Acta Armamentarii*, 2022, 43(11): 2810-2818.

[26] HE Y J, ZHAO J, GUO Y, et al. PL-VIO: Tightly-Coupled Monocular Visual-Inertial Odometry Using Point and Line Features[J]. *Sensors*, 2018, 18(4): 1159.

[27] FU Q, WANG J L, YU H S, et al. PL-VINS: Real-Time Monocular Visual-Inertial SLAM with Point and Line Features[EB/OL]. (2020-9-16)[2022-7-15]. https://arxiv.org/abs/2009.07462.

[28] LEE J, PARK S Y. 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] 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)：363-377.
ZHAO Liangyu, JIN Rui, ZHU Yeqing, et al. Stereo Visual-Inertial SLAM Algorithm Based on Merge of Point and Line Features[J]. *Acta Aeronautica et Astronautica Sinica*, 2022, 43(3): 363-377.

[30] 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(6)：106-112.
YING Wenjian, PAN Linhao, SHE Bo, et al. Stereo Visual-Inertial SLAM Algorithm Based on Point and Line Features[J]. *Journal of Naval University of Engineering*, 2021, 33(6): 106-112.

[31] 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(4)：20-27.
WEI Shuangfeng, SHI Xianjie, LIU Zhenbin, et al. Point-and-Line Joint Optimization Visual Inertial Odometer[J]. *Science of Surveying and Mapping*, 2021, 46(4): 20-27.

[32] GIOI R G, JAKUBOWICZ J, MOREL J M, 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] 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(5)：677-699.
JIA Di, ZHU Ningdan, YANG Ninghua, et al. Image Matching Methods [J]. *Journal of Image and Graphics*, 2019, 24(5): 677-699.

[34] LU X H, YAOY J, LI H A, et al. 2-Line Exhaustive Searching for Real-Time Vanishing Point Estimation in Manhattan World[C]//IEEE Winter Conference on Applications of Computer Vision (WACV), Santa Rosa, CA, USA, 2017.

[35] ZHANG L L, KOCH R. 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] 高翔，张涛，刘毅，等. 视觉SLAM十四讲：从理论到实践[M]. 北京：电子工业出版社，2017.
GAO Xiang, ZHANG Tao, LIU Yi, et al. Fourteen Lectures on Visual SLAM: From Theory to Practice[M]. Beijing: Publishing House of Electronics Industry, 2017.