

基于改进点线特征与惯性传感单元融合的 SLAM 算法

管玲, 李丹, 李俊祥, 陆钰

(安徽工业大学电气与信息工程学院, 安徽马鞍山 243032)

摘要: 针对传统同步定位与建图 (SLAM) 算法在密集环境中过度提取线段的问题, 提出 1 种基于改进点线特征与惯性传感单元 (IMU) 紧耦合的 SLAM 算法, 即 IPLI-SLAM。在数据预处理阶段采用 Shi-Tomasi 特征进行点特征提取并使用 LK (Lucas-Kanade) 光流进行跟踪和匹配; 引入线特征, 在像素梯度过滤机制的基础上改进直线段检测 (LSD) 提取算法, 筛选过滤线特征密集区域; 将视觉点线信息与 IMU 紧耦合后加入后端, 提高算法精度, 最后将 IPLI-SLAM 算法在数据集 EuRoc 与实际场景下进行测试验证。结果表明: 相较于 LSD 算法, 改进 LSD 算法的提取及匹配时间减少 8.2%; 相较 VINS-mono 与 PL-VINS 算法, IPLI-SLAM 算法定位精度分别提高 50.7%, 13.2%, 在运动速度快、场景模糊的环境中定位精度分别提高 55.6%, 25.1%; 在实际多纹理场景中本文算法定位轨迹起点与终点相差的距离均小于 VINS-mono 算法, 由此证明本文算法在显著提高定位精度的同时, 还具有更高的稳定性和更强的鲁棒性。

关键词: SLAM; IMU; 点线结合; 像素梯度过滤

中图分类号: TP 391 **文献标志码:** A **doi:** 10.12415/j.issn.1671-7872.23165



SLAM Algorithm Based on Fusion of Improved Point and Line Features with Inertial Sensing Units

GUAN Ling, LI Dan, LI Junxiang, LU Yu

(School of Electrical & Information Engineering, Anhui University of Technology, Maanshan 243032, China)

Abstract: In response to the issue of excessive line segment extraction by traditional simultaneous localization and mapping (SLAM) algorithms in dense environments, a SLAM algorithm based on improved point line features and tightly coupled inertial sensing units (IMU) was proposed, namely IPLI-SLAM. During the data preprocessing phase, Shi-Tomasi features were used for point feature extraction, and the Lucas-Kanade (LK) optical flow was used for tracking and matching. Line features were introduced to improve the line segment detector (LSD) extraction algorithm based on pixel gradient filtering mechanism, and to filter out areas with dense line features. By tightly coupling visual point and line information with IMU and adding it to the backend, the algorithm accuracy was improved. Finally, IPLI-SLAM algorithm was tested and validated on the dataset EuRoc and in actual scenes. The results show that compared with the original LSD algorithm, the improved LSD algorithm reduces extraction and matching time by 8.2%. Compared with the VINS-mono and PL-VINS algorithms, the IPLI-SLAM algorithm has

收稿日期: 2023-11-02

基金项目: 安徽省自然科学基金项目 (2108085MF225)

作者简介: 管玲 (1997—), 女, 安徽颍上县人, 硕士生, 主要研究为机器人定位与导航方向。

通信作者: 李丹 (1976—), 女, 安徽马鞍山人, 博士, 教授, 主要研究方向为图像处理与机器人视觉、自主导航等。

引文格式: 管玲, 李丹, 李俊祥, 等. 基于改进点线特征与惯性传感单元融合的 SLAM 算法 [J]. 安徽工业大学学报(自然科学版), 2024, 41(3):305-313.

improved positioning accuracy by 50.7% and 13.2%, respectively. In environments with fast motion speed and blurred scenes, the positioning accuracy is further enhanced by 55.6% and 25.1%, respectively. In actual multi-texture scenes, the distance between the starting and ending points of the localization trajectory using our algorithm is smaller than that of VINS-mono algorithm. This proves that the algorithm proposed in this article not only significantly improves positioning accuracy, but also has higher stability and stronger robustness.

Keywords: SLAM; IMU; dotted line combined; pixel gradient filtering

机器人搭载传感器能在无先验环境信息的情况下实现自主定位与建图,同步定位与建图(simultaneous localization and mapping, SLAM)是机器人导航的重要技术^[1]。根据传感器的不同,SLAM系统可分为激光SLAM和视觉SLAM。视觉SLAM主要以摄像头作为传感器,具有价格低廉、易于安装携带等优势,逐步成为SLAM主流趋势^[2]。根据图像数据的不同处理方法,可将视觉SLAM分为基于直接法与基于特征点法两大类。相较直接法,特征点法对光照变化有更好的跟踪效果^[3]。基于特征点的典型算法有尺度不变特征转换(scale invariant feature transform, SIFT)^[4]、同步跟踪与建图(parallel tracking and mapping for small ar workspaces, PTAM)^[5]、快速稳健特征(speeded-up robust features, SURF)算法^[6]等,但这些算法存在实时性差、容易丢失跟踪等不足。快速特征点提取和描述ORB(oriented fast and rotated brief)-SLAM使特征点SLAM技术达到了顶峰,但存在计算耗时长等不足^[7]。随后又提出ORB-SLAM2^[8]、ORB-SLAM3^[9],但其存在点特征稀疏导致定位精度下降等不足。为此,Li等^[10]提出室内单目零漂移系统,但耗时严重。直线段检测(line segment detector, LSD)-SLAM,通过像素梯度实现半密集跟踪,但对硬件及图像的要求较高^[11];点线结合(point and line, PL)-SLAM可实现实时定位和地图构建,但线特征提取及匹配时间长,且环境线密集区域导致线特征过度提取,造成特征误匹配影响定位精度^[12]。因此,提取稳定、均匀、富含环境信息的点线特征对视觉SLAM定位精度的提高具有重要意义。

视觉SLAM系统易受快速移动等因素的影响,致使点线特征难以建立数据关联,与多种传感器融合成为SLAM技术发展的主要趋势。不同类别的传感器具有各自不同的优点,惯性传感单元(inertial measurement unit, IMU)不依赖外界的感知,能精确记录短时间内的快速运动,IMU静止时视觉图像信息作为IMU的约束能有效减少IMU的漂移和累积误差^[13],使IMU能恢复单目相机的尺度信息且能为

视觉提供快速响应的定位。相机摄像头在低速运动上具有较好的跟踪效果,将相机摄像头与IMU进行融合,能有效提高视觉SLAM的定位精度。Mourikis等^[14]提出多状态约束下的卡尔曼滤波器(multi-state constraint Kalman filter, MSCKF)算法,基于滤波器设计取得了良好的位姿优化效果,但无法用于全局。Leutenegger等^[15]提出基于关键帧的视觉惯性SLAM(open keyframe-based visual-inertial SLAM, OKVIS-SLAM)算法,在滑窗模型中构建非线性优化函数处理视觉的重投影误差和IMU的测量误差,但无法用于重定位。Qin等^[16]提出单目视觉惯性系统(monocular visual inertial navigation system, VINS-mono)联合优化视觉和IMU信息,采用非线性紧耦合方法获得高精度的视觉惯性里程计,但对弱纹理、关键点少的环境鲁棒性和精度差;He等^[17]提出点和线特征的紧耦合单目视觉惯性里程计,通过将视觉信息与IMU紧耦合来提高定位精度,但存在线特征提取匹配时间长的缺点。王立玲等^[18]在ORB-SLAM3基础上增加线特征,采用自适应加权策略提高系统的稳定性,但线特征提取质量较差。鉴于此,提出1种基于改进的点线特征与IMU耦合的SLAM算法(improved point-line feature and IMU coupled SLAM, IPLI-SLAM),采用改进线特征改善点特征稀疏、线特征过度提取问题,同时利用IMU弥补单目视觉的不足,以期提高SLAM算法在密集环境中的定位精度。

1 SLAM的数学描述及改进

改进的点线结合与惯性传感单元(IMU)紧耦合的视觉SLAM算法(IPLI-SLAM)主要分为前端特征提取、IMU数据预处理、初始化和后端滑动窗口联合优化三部分,整体框架如图1。

算法改进主要在前端,对于Shi-Tomasi点特征,基于Lucas-Kanade(LK)光流估计算法进行跟踪和匹配;对于改进的LSD线特征,采用描述符分析(line band descriptor, LBD)^[19]进行匹配。点线重投影残差和IMU数据测量预积分残差及闭环信息部分通过

加入优化窗口进行位姿估计, 筛选关键帧进行闭环优化。系统提取的点线特征质量直接影响后续匹配质量, 提取质量越高匹配质量就越高, 反之提取质量

不高则匹配质量不高。后端线程优化来自前端的点线及 IMU 数据误差, 优化模型为点线及 IMU 的误差函数。

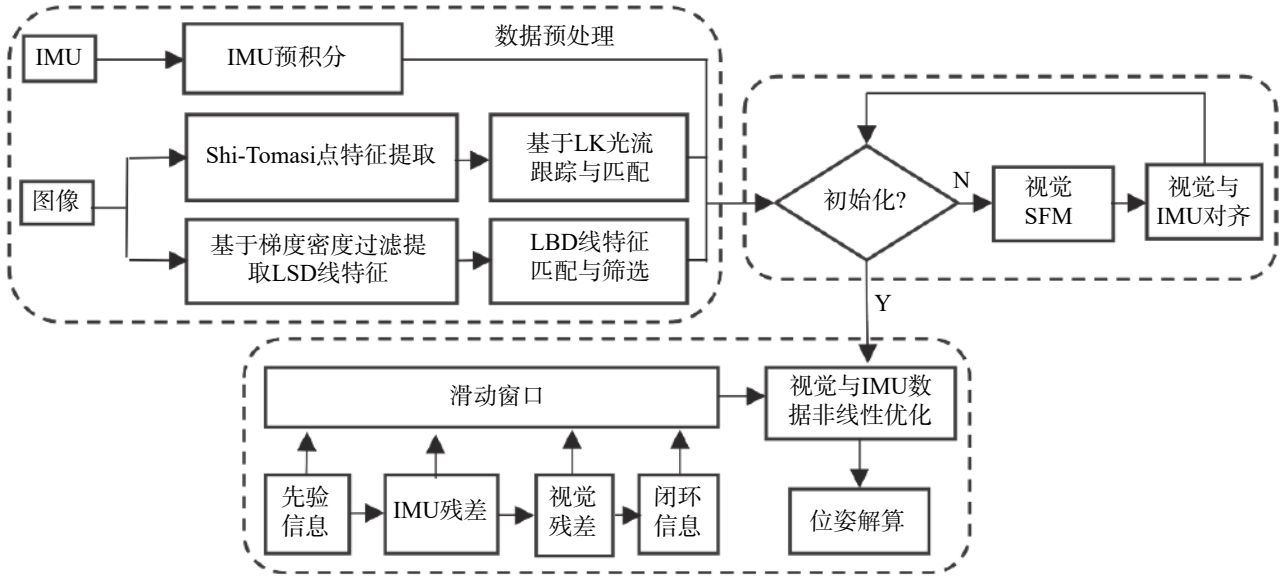


图1 IPLI-SLAM 算法框架

Fig. 1 IPLI-SLAM algorithmic frame

1.1 前端数学描述

1.1.1 点线的表示

根据相机的针孔模型, 世界坐标系某点在相机坐标系下为 p_c , 该点在像素平面上的映射坐标 $p_l = [x, y, 1]^T$ (x, y 分别为点的横、纵坐标), p_c 与 p_l 的关系如下

$$p_c = sK_c p_l \quad (1)$$

其中: s 为相机逆深度; K_c 为相机内参。世界坐标系的 1 条线可用普吕克坐标表示 (包括垂直向量 n , 方向向量 v), 假设从 1 条线上截取 2 个点 p_1 和 p_2 , 则直线可由式 (2) 表达。

$$L = \begin{bmatrix} p_2 \times p_1 \\ p_1 - p_2 \end{bmatrix} = [n, v]^T \quad (2)$$

将世界坐标系下的直线 L_w 通过转换矩阵 H_c^w 转换为相机坐标系下的 L_c , 投影矩阵 K_l 将 L_c 转换到像素平面 L_l , 分别见式 (3), (4)。

$$L_c = [n_c, v_c]^T = H_c^w L_w \quad (3)$$

$$L_l = K_l n_c = [l_1, l_2, l_3]^T \quad (4)$$

1.1.2 点线重投影残差模型

如图 2, 空间中第 k 个点 p_k 在 2 个像素平面的测量值分别为 $z_{p_k}^{c_i} = [u_{p_k}^{c_i}, v_{p_k}^{c_i}, 1]^T$, $z_{p_k}^{c_j} = [u_{p_k}^{c_j}, v_{p_k}^{c_j}, 1]^T$, 像素平面的点经 σ^{-1} 函数转换归一化平面 $\bar{p}_k^{c_i} = [\bar{x}_k^{c_i}, \bar{y}_k^{c_i}, 1]^T$, $\bar{p}_k^{c_j} = [\bar{x}_k^{c_j}, \bar{y}_k^{c_j}, 1]^T$, 见式 (5) 和式 (6)。

$$\bar{p}_k^{c_i} = \sigma^{-1} z_{p_k}^{c_i} \quad (5)$$

$$\bar{p}_k^{c_j} = \sigma^{-1} z_{p_k}^{c_j} \quad (6)$$

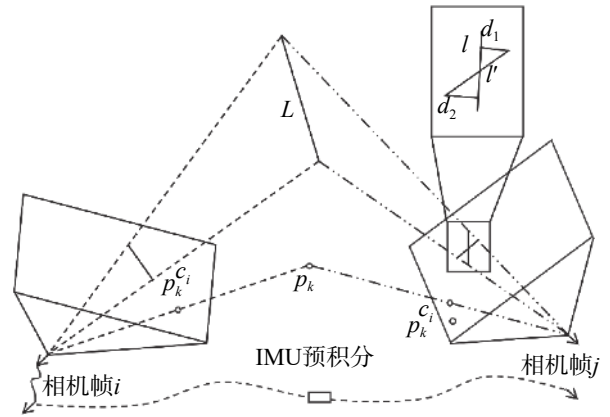


图2 点线重投影误差

Fig. 2 Point line reprojection error

将 $\bar{p}_k^{c_i}$ 转换为 $p_k^{c_j}$, 见式 (7), 其中 ρ 为转换函数。本文在归一化平面表示重投影残差, 见式 (8)。

$$p_k^{c_j} = \rho \bar{p}_k^{c_i} \quad (7)$$

$$r_{p_k} z_{p_k}^{c_j} = \left[\frac{p_k^{c_j}}{\|p_k^{c_j}\|} - \bar{p}_k^{c_i} \right] \quad (8)$$

同理, 直线 L 的投影为 l' , 像素匹配的直线 l 的 2 个端点 p_s 和 p_e , 则线的重投影误差表示为式 (9)。

$$r_l z_{L_l}^{c_j} = [d_1, d_2] \left[\frac{p_s^{c_j T} l'}{\sqrt{l_s^2 + l_2^2}}, \frac{p_e^{c_j T} l'}{\sqrt{l_e^2 + l_2^2}} \right]^T \quad (9)$$

1.1.3 IMU 预积分与残差模型

IMU 是指包含 3 轴加速度计和 3 轴陀螺仪的

6轴传感器。假设已知 t 时刻 IMU 的位置、速度及旋转分别为 \dot{p}_{wb_t} , \dot{v}_{wb_t} , \dot{q}_{wb_t} , 以及 t 时刻到 $t+1$ 时刻 IMU 的加速度 a 和角加速度 ω 的测量值 (\dot{p}_{wb_t} 为 t 时刻的 IMU 本体坐标系相对于世界坐标系的旋转测量值, 其他同理), 则通过连续时间积分可得到 $t+1$ 时刻的位置 $\dot{p}_{wb_{t+1}}$, $\dot{v}_{wb_{t+1}}$, $\dot{q}_{wb_{t+1}}$, 如式 (10)。

$$\begin{aligned}\dot{p}_{wb_{t+1}} &= \dot{p}_{wb_t} + \dot{v}_{wb_t} \Delta t + \iint_t^{t+1} (R_{b_t}^w (a_t - g^w)) dt^2 \\ \dot{v}_{wb_{t+1}} &= \dot{v}_{wb_t} + \int_t^{t+1} (R_{b_t}^w (a_t - g^w)) dt \\ \dot{q}_{wb_{t+1}} &= \dot{q}_{wb_t} \otimes \int_t^{t+1} \left[0, \frac{1}{2} \omega^{b_t} \right]^T dt\end{aligned}\quad (10)$$

其中: $R_{b_t}^w$ 为 t 时刻 IMU 坐标系到世界坐标系的转换矩阵; g^w 为世界坐标系下重力加速度; a_t 为 t 时刻的加速度; ω^{b_t} 为 t 时刻在 IMU 坐标系下的角加速度。

通常 IMU 的帧率高于相机帧率, 相机两帧间 IMU 的相对观测量可用预积分表示, 预积分可使世界坐标系与 IMU 坐标系解耦, 避免 IMU 状态量在优化过程时式 (10) 重复积分, 进而减少计算量。对式 (10) 进行整理, 考虑加速度计的偏置 b_a 、陀螺仪的偏置 b_g ($b_a^{b_t}$ 为 t 时刻 b_a 在 IMU 坐标系下的值, 其他同理), 通过调整偏差的一阶近似值来调整预测积分量, 进一步根据文献 [20–21] 中对 IMU 偏置的处理方法, 则此时广义 IMU 残差模型为:

$$\mathbf{r}_{\text{IMU}} = \begin{bmatrix} r_p r_v r_q r_{b_a} r_{b_g} \\ R_{b_t}^{b_{t+1}} \left(\dot{p}_{wb_{t+1}} - \dot{p}_{wb_t} - \frac{1}{2} g^w \Delta t^2 - v_{wb_t} \Delta t \right) - \tilde{\alpha}_{b_{t+1}}^{b_t} \\ R_{b_t}^{b_{t+1}} \left(\dot{v}_{wb_{t+1}} + g^w \Delta t - \dot{v}_{wb_t} \right) - \tilde{\beta}_{b_{t+1}}^{b_t} \\ 2 \left[\dot{q}_{b_{t+1}}^{w-1} \otimes \dot{q}_{b_{t+1}}^w \otimes \left(\tilde{\gamma}_{b_{t+1}}^{b_t} \right)^{-1} \right]_{xyz} \\ b_a^{b_{t+1}} - b_a^{b_t} \\ b_g^{b_t} - b_g^{b_t} \end{bmatrix} \quad (11)$$

其中: r_p 为位置量; r_v 为速度量; r_q 为旋转量; r_{b_a} 为加速度偏置量; r_{b_g} 为陀螺仪偏置量; $\tilde{\alpha}_{b_{t+1}}^{b_t}$, $\tilde{\beta}_{b_{t+1}}^{b_t}$, $\tilde{\gamma}_{b_{t+1}}^{b_t}$ 为 IMU 预积分项。

1.1.4 线特征提取方法的改进

LSD 算法利用像素信息提取直线, 速度快, 能在短时间内获得亚像素级的高精度信息。但 LSD 算法存在复杂环境下过度提取的问题, 当图像中存在大量密集线条会过度提取, 从而增加特征匹配的难度和时间, 甚至出现误匹配情况, 影响系统定位精度。LSD 算法提取直线时在像素梯度变化率高的地方可能形成直线, 故进行 LSD 检测前, 基于像素梯度设置 1 个滤波器对密集区域进行筛选, 检测灰度图中像素密度高的区域, 并将该区域的像素更改为

统一的像素值。对于更改的区域, 只会检测出更改的边缘框, 不会检测出该区域存在的密集线段。过滤过程如图 3。

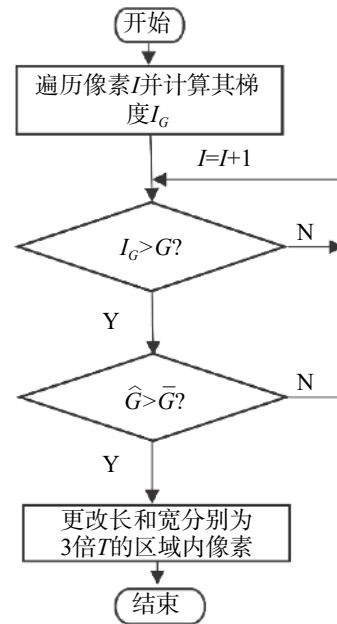


图3 过滤流程图

Fig. 3 Filtration flowchart

过滤过程的主要步骤如下:

- 1) 遍历某灰度图像素并计算其梯度值, 记某坐标点为 (x, y) 的像素点的梯度值为 $G_{(x, y)}$, 设梯度阈值为 G ;
- 2) 定义区域梯度密度数 \hat{G} , 当 $G_{(x, y)} > G$ 时, 以 (x, y) 坐标点为中心, 计算长和宽为 T 的区域内所有像素点的梯度值。区域内像素点的 $G_{(x, y)_T} > G$ 的个数即为区域梯度密度数 \hat{G} ;
- 3) 设置区域梯度密度阈值 \bar{G} , 当 $\hat{G} > \bar{G}$ 时, 更改长和宽为 3 倍 T 的区域内所有的像素值。

T 与 \bar{G} 值的设置直接影响过滤效果, 阈值过大过滤效果不明显; 阈值过小会过滤过度, 导致后续检测的线段减少。综合考虑, 设置 T 值为 5, \bar{G} 值为 18。

LSD 算法会提取出较多的短线段, 但位姿估计不需大量的短线段, 只需提取富含信息的稳定长线段。因此, 需调整 LSD 算法内部的隐参数, 采用金字塔层数为 2、尺度因子为 0.5、最小密度阈值为 0.6、长度阈值为某图长宽中最小值的 0.125。过滤后的灰度图直接进入 LSD 提取线程, 像素变化的区域将被检测为线框。图 4 为 LSD 算法改进前后的线提取结果, 图中百叶窗是线特征密集区域, 属于复杂环境。由图 4 可看出: 经过滤处理的 LSD 算法在百叶窗区域没有大量提取线段, 仅提取边缘线框。由此可见, 提出的滤波处理方法可有效改善密集区域线段过度提取的问题。



(a) 原LSD算法



(b) 改进的LSD算法

图4 LSD 算法改进前后的线特征提取结果

Fig. 4 Line feature extraction results before and after LSD algorithm improvement

1.2 后端联合优化

采用与文献 [16] 相同的方法进行联合初始化, 将视觉信息与 IMU 信息紧耦合, 再利用运动恢复结构 (structure from motion, SFM) 技术估计滑窗内连续帧间的位姿, 最后将视觉与 IMU 融合对齐实现对陀螺仪的偏置校正。初始化后, 通过滑窗视觉约束、IMU 约束及路标点进行非线性优化, 图 5 为三者约束关系。

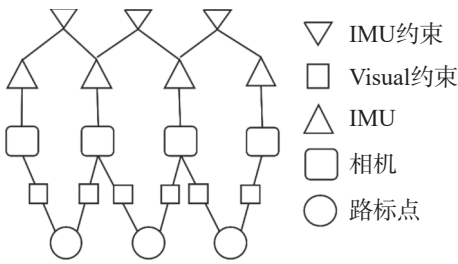


图5 约束关系

Fig. 5 Constraint relationship

定义滑动窗口内的状态变量:

$$\chi = [x_0, x_1, \dots, x_f, \lambda_0, \lambda_1, \dots, \lambda_m, l_0, l_1, \dots, l_o]^T \quad (12)$$

$$\mathbf{x}_u = [p_{wb_t}, v_{wb_{t+1}}, q_{wb_t}, \mathbf{b}_a, \mathbf{b}_g]^T, u \in (0, f) \quad (13)$$

式中: \mathbf{x}_u 为第 u 个 IMU 的状态, 包含加速度计和陀螺仪的偏置以及 IMU 的预积分估计的位姿信息; f 为关键帧的数目; m 为所有点特征的逆深度; o 为线特征的正交坐标参数。利用光束平差法 (bundle adjustment, BA) 进行优化, 最小化所有测量残差的

先验与马氏距离之和, 得到最大后验估计, 获得非线性优化目标函数, 即

$$\mathcal{F} = \min_{\chi} \{ \|r - H\chi\|^2 + \sum \|r_p\|^2 + \sum \|r_l\|^2 + \sum \|r_{imu}\|^2 \} \quad (14)$$

其中 $\{r, H\}$ 表示滑窗边缘化的先验数据。最后使用非线性化求解器 cere-solver 进行运算, 求解非线性问题。

2 实验与结果分析

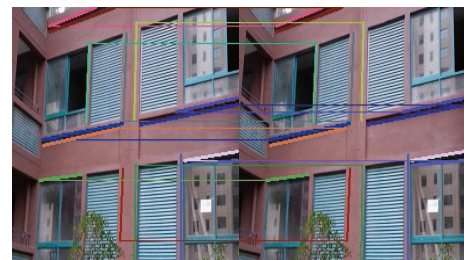
为验证本文算法的有效性, 通过开源数据集 EuRoC^[22] 及实际多纹理环境进行实验。EuRoC 数据集由苏黎世联邦理工学院使用室内无人机收集, 包含不同光线和不同纹理等复杂环境。数据集与实际多纹理环境的测试分别基于 64 位 Ubuntu 18.04 和 Ubuntu 20.04 操作系统。

2.1 改进线特征匹配实验

随机选取具有密集区域的图像进行 LSD 线特征提取且利用 LBD 匹配。原 LSD 及经过滤处理的 LSD 算法线匹配结果如图 6。由图 6 可看出: 采用原 LSD 算法匹配的百叶窗密集区域密集, 匹配结果复杂; 采用经过滤处理 LSD 算法匹配的百叶窗密集区域清晰。



(a) 原LSD算法



(b) 改进的LSD算法

图6 LSD 算法改进前后的线匹配结果

Fig. 6 Line matching results before and after LSD algorithm improvement

复杂的匹配过程会增加系统的处理时间, 过滤处理后的密集区域不再提取大量线段, 可降低匹配的复杂度, 提高系统的实时性, LSD 算法改进前后提取及匹配时间如表 1。由表 1 可知: 改进前后所需时间分别为 0.152 0, 0.139 4 s, 与原 LSD 算法相比,

改进后 LSD 算法的提取及匹配时间大约减少 8.28%, 经过滤处理的 LSD 算法可有效提高提取匹配的实时性, 验证了算法的可行性。

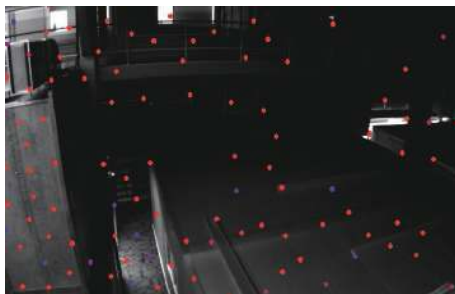
表 1 LSD 算法改进前后的匹配时间

Tab.1 Matching time before and after LSD algorithm improvement

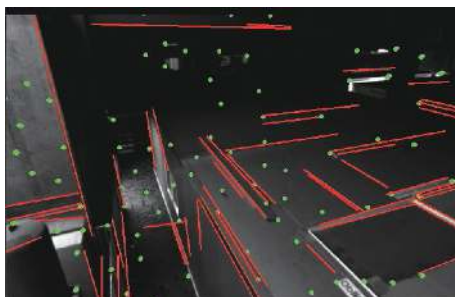
算法	提取及匹配时间/s
原LSD	0.152 0
改进的LSD	0.139 4

2.2 EuRoc 数据集评估实验

为验证本文算法的定位性能, 首先将本文算法与经典算法 VINS-mono 在 MH_04_difficuil 序列下进行实验比较。使用 EVO 工具将时间戳与真实值对齐并计算绝对轨迹误差 (absolute pose error, APE), 使用均方根误差 (root mean square error, RMSE) 对系统性能进行评估。MH_04_difficuil (简称 MH_04) 序列是 1 个复杂的环境, 线条纹理较多, 更能体现系统的精度。在 MH_04 序列本文算法与 VINS-mono 算法的提取效果如图 7。由图 7 可看出: VINS-mono 算法只能提取点特征, 本文算法增加了改进线特征的提取线程, 增加了线特征纹理, 可弥补点特征不足。



(a) VINS-mono 算法



(b) 本文算法

图7 点线特征提取效果

Fig. 7 Extraction effect of point and line feature

图 8 为本文算法与 VINS-mono 算法在 MH_04 下的实验轨迹与真实轨迹之间的 APE 误差热力图。

数据集提供真实运动轨迹以虚线表示, 算法输出的位姿估计用彩色线条表示。线条颜色强度表示该位置与真实轨迹之间的绝对误差, 其变化范围可参见图例的色彩标尺。由图 8 可看出: 由 2 种算法得到的实验轨迹与真实轨迹均非常接近, VINS-mono 的 APE 最大为 1.316 m、最小为 0.021 m, 本文算法的 APE 最大为 0.395 m、最小为 0.035 m; 本文算法的 APE 变化稳定均匀, 而 VINS-mono 的 APE 变化剧烈, 说明本文算法稳定性强, 鲁棒性高。经计算, VINS-mono 与本文算法的 RMSE 分别为 0.398 1, 0.189 3 m, 本文算法的 RMSE 较小, 表明本文算法能有效提高系统的定位精度, 尤其是在线条纹理较多的环境中。

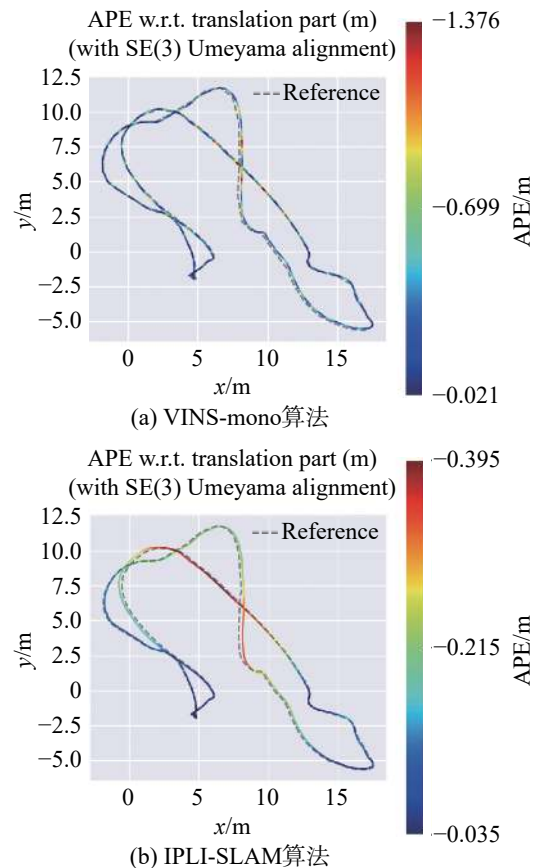


图8 APE 热力图

Fig. 8 APE heat map

图 9 为 VINS-mono 算法和本文算法在 MH_04 下的实验轨迹与真实轨迹在 xyz 轴上随时间的变化情况。由图 9 可看出: 2 种算法在 x 轴上得到的轨迹均接近于真实曲线; VINS-mono 算法在 y 轴和 z 轴上的轨迹曲线参差不齐, 接近锯齿状, 而本文算法的轨迹曲线平滑且整体接近真实曲线。

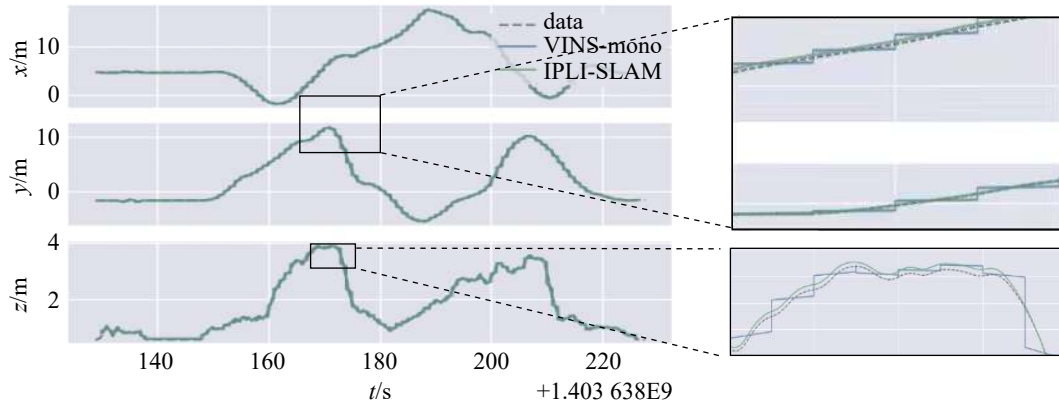


图9 随时间变化的轨迹曲线

Fig. 9 Trajectory curve over time

为进一步验证本文算法鲁棒性,将本文算法与经典算法 VINS-mono 和 PL-VINS^[23] 在 EuRoc 数据集其他序列进行实验比较。表 2 为 VINS-mono, PL-VINS 算法及本文算法在 EuRoc 数据集上测试的 RMSE。使用 MH_01_easy, MH_02_easy, MH_03_medium, MH_04_difficult, MH_05_difficult 共 5 个序列和 VI_01_easy, VI_02_medium, VI_03_difficult 共 3 个序列,分别对应简单、中等和困难 3 个等级,3 种算法均在同一环境中执行,并启用回环检测。

表 2 不同算法测试的 RMSE 结果

Tab. 2 RMSE results of different algorithm tests

数据集序列	RMSE/m		
	VINS-mono	PL-VINS	IPLI-SLAM
MH_01_easy	0.177 0	0.121 8	0.119 0
MH_02_easy	0.201 5	0.118 5	0.107 2
MH_03_medium	0.401 4	0.162 4	0.147 6
MH_04_difficult	0.398 1	0.200 5	0.189 3
MH_05_difficult	0.382 0	0.223 5	0.214 2
VI_01_easy	0.144 3	0.087 8	0.062 1
VI_02_medium	0.310 6	0.113 0	0.083 4
VI_03_difficult	0.330 0	0.260 0	0.208 3

从表 2 可看出:与 VINS-mono 算法相比,本文算法 RMSE 显著降低,平均降低了 50.7%;与 PL-VINS 算法相比,本文算法的 RMSE 平均降低 13.2%。在 MH 序列中,与 VINS-mono 及 PL-VINS 算法相比,本文算法的 RMSE 分别降低 47.8%, 6.1%;在 VI 序列中,与 VINS-mono 及 PL-VINS 算法相比,本文算法的 RMSE 分别降低了 55.6%, 25.1%, VI 序列相比于 MH 序列的 RMSE 下降更明显,即使 VI 序列存在运动速度快、场景模糊及纹理稀疏等不同的问题,但本文算法的 RMSE 仍显著降低,说明本文方法在复杂场景中的适用性强,进一步证明本文算法具有较强的鲁棒性。

2.3 多纹理环境实验

为验证本文算法在密集环境中的定位精度,选取图书馆某一场景进行实验。图书馆内有大量书架,存在大量的线密集区域。实验人员手持笔记本电脑搭载 D435i 相机在室内行走。D435i 相机具有视觉处理器、彩色图像传感器以及内置 6 轴的 IMU 单元,经标定后误差达到要求,可满足实验硬件需要。人员操作存在偶发性和误差、D435i 的标定或硬件存在偶然误差,故只对实验进行定性比较。测试过程中,由实验人员手持设备在图书馆内保持匀速行动,起点出发,经书架中间过道再折进某一书架,转折时光线增强,存在光线变化剧烈情况,最后原路返回到起点,通过比较起点和终点的重合度来衡量算法的精度。图 10 为图书馆内点线特征提取效果。由图 10 可看出:在提取点特征的同时提取线特征,提取线特征能弥补天花板及书架无点特征的不足。

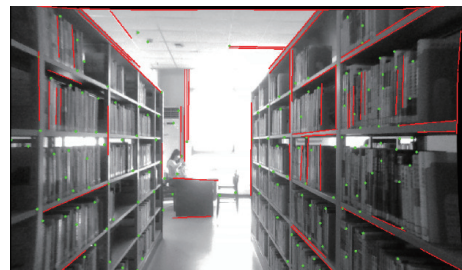


图10 室内点线特征提取效果

Fig. 10 Extraction effect of indoor point and line feature

图 11 为实验运行轨迹在 xy 轴上的平面展示。由图 11 可看出:本文算法的定位轨迹起点与终点的距离相差仅 0.128 m; VINS-mono 算法的定位轨迹偏离真实轨迹,起点与出发点相差 0.371 m,充分说明本文算法的有效性,可有效提高系统的定位精度和鲁棒性。表 3 为本文算法与 VINS-mono 算法在图书馆内不同运行路径下起点与终点相差的距离。由表 3 可看出:路径 1, 2 下本文算法的起点与终点相

差的距离小于 VINS-mono 算法,进一步说明本文算法可提高系统的定位精度且具有更高的鲁棒性。

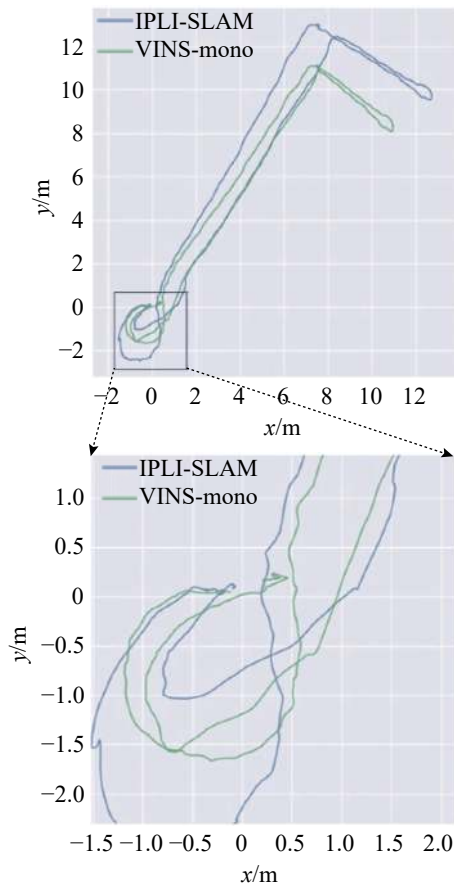


图11 室内定位轨迹对比

Fig. 11 Comparison of indoor positioning trajectory

表 3 不同算法规划路径起点与终点相差的距离

Tab. 3 Distance between the starting and ending points of path planning with different algorithms

路径	相差的距离/m	
	VINS-mono	IPLI-SLAM
路径1	0.371	0.128
路径2	0.238	0.151

3 结论

针对密集环境中 SLAM 算法过度提取线段的不足,提出 1 种改进的点线融合与 IMU 紧耦合的视觉 SLAM 算法。提取 Shi-Tomasi 特征点,并用 LK 光流进行跟踪和匹配;基于像素梯度过滤机制改进 LSD 算法,过滤密集区域以减少多纹理环境中的线段提取,采用 LBD 描述子进行线匹配;再将视觉信息与 IMU 数据添加到后端优化估算位姿,最后将所提算法在开源数据集 EuRoc 和实际场景中进行测试验证。结果表明:本文算法相较 VINS-mono 与

PL-VINS 算法定位精度分别提高 50.7%, 13.2%;在实际多纹理场景中,本文算法定轨迹起点与终点相差的距离均小于 VINS-mono 算法,验证了本文算法可有效改善密集环境中过度提取线段的问题,在提高系统定位精度的同时,还可增强系统的鲁棒性。

参考文献:

- [1] 褚舒春. 动态场景下基于实例分割的视觉SLAM算法研究[D]. 哈尔滨: 哈尔滨理工大学, 2022.
- [2] 王柯赛, 姚锡凡, 黄宇, 等. 动态环境下的视觉SLAM研究评述[J]. 机器人, 2021, 43(6):715-732.
- [3] CADENA C, CARLONE L, CARRILLO H, et al. Past, present, and future of simultaneous localization and map**: toward the robust-perception age[J]. *IEEE Transactions on Robotics*, 2016, 32(6):1309-1332.
- [4] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60:91-110.
- [5] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces[C]//2007 6th IEEE and ACM International Symposium on Mixed And Augmented Reality. Nara, Japan: IEEE, 2007:225-234.
- [6] BAY H, ESS A, TUYTELAARS T, et al. Speeded-up robust features (SURF)[J]. *Computer Vision and Image Understanding*, 2008, 110(3):346-359.
- [7] MUR-ARTAL R, MONTIEL J M M, TARDOS J D. ORB-SLAM: a versatile and accurate monocular SLAM system[J]. *IEEE Transactions on Robotics*, 2015, 31(5):1147-1163.
- [8] 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.
- [9] 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.
- [10] LI Y, BRASCH N, WANG Y, et al. Structure-SLAM: low-drift monocular SLAM in indoor environments[J]. *IEEE Robotics and Automation Letters*, 2020, 5(4):6583-6590.
- [11] ENGEL J, STÜCKLER J, CREMERS D. Large-scale direct SLAM with stereo cameras[C]//2015 IEEE/RSJ International Conference on Intelligent Robots and Systems. Hamburg, Germany: IEEE, 2015:1935-1942.
- [12] GOMEZ-OJEDA R, MORENO F A, ZUNIGA-NOËL D, et al. PL-SLAM: a stereo SLAM system through the combination of points and line segments[J]. *IEEE Transactions on Robotics*, 2019, 35(3):734-746.

- [13] 邹汉达, 袁洪. 视觉辅助下的室内惯导位姿修正[J]. 测绘通报, 2019(5):16–20.
- [14] MOURIKIS A I, ROUMELIOTIS S I. A multi-state constraint Kalman filter for vision-aided inertial navigation[C]//Proceedings 2007 IEEE International Conference on Robotics and Automation. Rome, Italy, IEEE, 2007:3565–3572.
- [15] LEUTENEGGER S, LYNEN S, BOSSE M, et al. Key frame-based visual-inertial odometry using nonlinear optimization[J]. *The International Journal of Robotics Research*, 2015, 34(3):314–334.
- [16] QIN T, LI P, SHEN S. VINS-mono: a robust and versatile monocular visual-inertial state estimator[J]. *IEEE Transactions on Robotics*, 2018, 34(4):1004–1020.
- [17] HE Y, ZHAO J, GUO Y, et al. PL-VIO: tightly-coupled monocular visual-inertial odometry using point and line features[J]. *Sensors*, 2018, 18(4):1159–1183.
- [18] 王立玲, 朱旭阳, 马东, 等. 基于点线特征视觉惯性融合的机器人SLAM算法[J]. 中国惯性技术学报, 2022, 30(6):730–737.
- [19] ZHANG 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.
- [20] FORSTER C, CARLONE L, DELLAERT F, et al. On-manifold preintegration for real-time visual-inertial odometry[J]. *IEEE Transactions on Robotics*, 2016, 33(1):1–21.
- [21] YANG Z, SHEN S. Monocular visual-inertial state estimation with online initialization and camera-IMU extrinsic calibration[J]. *IEEE Transactions on Automation Science and Engineering*, 2016, 14(1):39–51.
- [22] BURRI M, NIKOLIC J, GOHL P, et al. The EuRoC micro aerial vehicle datasets[J]. *The International Journal of Robotics Research*, 2016, 35(10):1157–1163.
- [23] FU Q, WANG J, YU H, et al. PL-VINS:real-time monocular visual-inertial SLAM with point and line features[EB/OL]. (2020–10–28)[2023–11–02]. <http://arxiv.org/pdf/2009.07462v2.pdf>.

责任编辑: 何莉