

一种基于点线融合的室内视觉惯性 SLAM 方法

张善福,胡步发

(福州大学 机械工程及自动化学院,福建 福州 305116)

摘要:以实现室内移动机器人高精度、高鲁棒性的定位与建图为目标,针对单目视觉 SLAM 方法在室内弱纹理环境下由于特征点稀疏、相机移动过快所导致的定位精度下降甚至跟踪失败问题,提出一种结合点线特征的视觉惯性融合 SLAM 方法。利用改进的 LSD 算法提取线特征以提升数据关联的准确性,通过最小化点线及 IMU 等残差构成的目标函数得到准确的位姿估计。根据对比多组公开数据集及实际室内环境中的实验结果表明:该方法提高了室内定位精度,具备较好的鲁棒性。

关键词:同步定位与建图;视觉惯性融合;状态估计;点线特征

中图分类号:TP274 **文献标志码:**B **文章编号:**1671-5276(2023)04-0202-04

An Indoor Visual Inertial SLAM Method Based on Point-line Fusion

ZHANG Shanfu, HU Bufa

(College of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350116, China)

Abstract: In order to realize the high-precision and robust localization and mapping of indoor mobile robot, aiming at the decrease in localization accuracy or even tracking failure due to sparse feature points and too fast camera movement in weak texture environment by monocular vision slam method, a visual inertia fusion slam method combined with point line features is proposed. The improved LSD algorithm is used to extract line features for improving the accuracy of data association. The accurate pose estimation is obtained by minimizing the objective function composed of point line and IMU residuals. The comparison of several groups of public datasets and the experimental results in the actual indoor environment show that the method can improve the indoor positioning accuracy and has good robustness.

Keywords: simultaneous localization and mapping; visual inertial fusion; state estimation; point-line features

0 引言

如今有相当数量的移动机器人在室内运行,而室内环境属于 GPS 不可达的区域,因此如何实现室内移动机器人的自主定位是一个值得研究的问题。SLAM 方法是解决室内环境感知与定位导航问题的有效手段之一,它主要研究搭载各种传感器的主体,在移动过程中感知未知环境,判断自己在环境中的位置并同时对其周围的环境进行建图。

单目视觉 SLAM 是近年来的重要研究分支,这得益于单目摄像头体积小、质量轻、能耗低且应用前景广阔。然而,单目视觉系统存在无法恢复度量尺度且易受外界光线干扰等问题,针对这些问题,研究人员考虑将惯性测量单元(IMU)与相机进行传感器融合。IMU 能够高频记录本体的运动信息,获得相对准确的短期运动姿态估计,而相机能够有效地修正 IMU 的长期漂移,二者优势互补,故可以获得较好的 SLAM 性能。

对于基于优化的 VI-SLAM 方法,研究人员做了许多研究。MUR-ARTAL R 等^[1]提出了 VIORB-SLAM 算法,实现了较高的定位精度并且具有地图重用和回环检测等

功能,但特征点的提取与匹配较为耗时,香港科技大学的 QIN T 等^[2]提出了结合光流法和 IMU 的 VINS-Mono 算法,采用松耦合方式进行视觉惯性联合初始化,使其效果优于其他基于非线性优化的 VI-SLAM 算法,光流法追踪特征点解决了特征点匹配耗时的问题,但由于仅提取环境中的点特征,在弱纹理区域容易出现精度下降问题。

相较于点特征,线特征在室内人造场景中的结构性表达更加丰富。PUMAROL A 等^[3]提出了单目点线结合的 PL-SLAM 算法,HE Y 等^[4]提出了单目点线里程计 PL-VIO 系统,赵良玉等^[5]提出改进的 FLD 线特征提取算法用以替代传统的 LSD 算法,加快线特征提取的运行速度。

因此,针对室内普遍存在的弱纹理环境下基于特征点的 SLAM 系统定位精度下降问题,本文采取单目点线融合的视觉惯性 SLAM 方法,在 VINS-mono 框架的基础上增加线特征的提取与匹配以加强帧间数据关联,通过实验验证,该方法相较于 VINS-mono 提高了精度,且在弱纹理环境中运行依旧有良好的鲁棒性。

1 系统整体框架

视觉惯性融合的 SLAM 算法总体框架如图 1 所示,整

个框架主要分为 4 部分:图像与 IMU 数据预处理、视觉惯性系统联合初始化、基于紧耦合的后端滑动窗口非线性优化与回环检测。

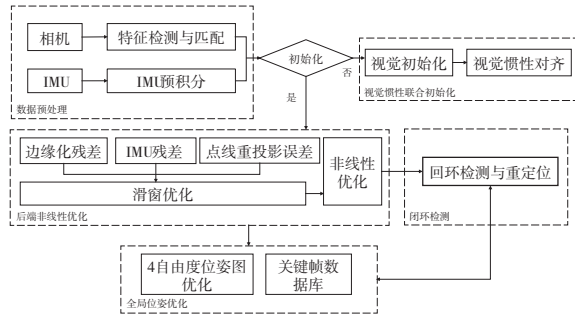


图 1 视觉惯性融合 SLAM 算法框架

首先,提取输入图像中的 Shi-Tomasi 角点特征并在帧间通过光流法追踪匹配,利用改进的 LSD 算法提取线特征,使用 LBD 描述符进行线特征匹配;同时,将两个视觉帧之间 IMU 测量所得数据信息进行预积分处理;之后,采用视觉与 IMU 松耦合的方案进行初始化得到初始位姿及重力方向尺度因子、陀螺仪偏置及每一帧对应的速度;在后端建立一个统一的损失函数,包括视觉信息约束与 IMU 约束,优化求解相机和 IMU 的外参、载体的位姿、路标点的深度;回环检测通过描述符匹配关键帧,将匹配信息输入后端对全局位姿进行优化。

2 线特征提取

2.1 改进的 LSD 线特征提取

LSD^[6]是一种无参数调优的线段检测器,但其设计之初目的是用于描述静态场景的形状,而不是 SLAM 中的位姿估计。为了使其更适应于本文系统,根据文献[7]提出的方法对 LSD 算法部分参数做修改以提升线段提取的实时性。

将修改以下两个参数以提高性能:图像尺度参数 s (scale)用于缩放高斯金字塔每一层的图像;像素密度阈值 d (density)用于剔除候选线段。

使用本文算法在不同参数值下运行数据集 MH-04-Difficult,得到的方均根误差与线特征提取平均时间结果点线图如图 2 所示。其中,图 2 上方正方形标记点线图表示方均根误差变化趋势,下方三角形标记点线图表示特征提取时间变化趋势。根据图 2(a)可以发现,减少 s 的值可以在几乎不影响精度的情况下,大幅度减少线特征的提取时间,因此本文设置参数 s 为 0.5(原参数为 0.8),在此基础上,改变参数 d 得到的点线图如图 2(b),因此本文设置参数 d 为 0.5(原参数为 0.7)。

2.2 长度抑制策略

传统线特征检测算法在提取线段特征时,会检测出较多的短线段特征,这不仅增加了计算成本,也增加了误匹配的概率。

对于点线特征融合的 SLAM 算法,加入线特征是为了提供额外约束以提高位姿估计的精度,因此本文采用线段

长度抑制的方法,保留场景中较长的线段特征用于特征检测,剔除准则为

$$l_i \geq l_{\min} = \eta \times \lceil \min(W_1, H_1) \rceil \quad i \in (1, m) \quad (1)$$

式中: l_i 为单帧图像 m 条线特征中第 i 条线特征的长度; l_{\min} 表示最短线特征长度阈值; W_1 和 H_1 分别为图像的宽度和高度; $\lceil \cdot \rceil$ 表示向上取整; η 为比例系数,取 η 为 0.125。经过长度抑制策略之后,线特征提取效果如图 3 所示。

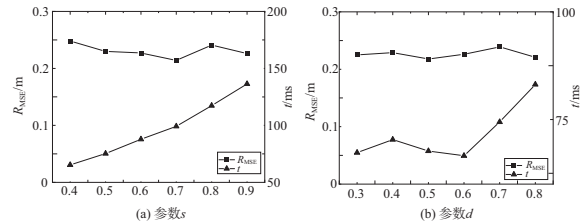


图 2 内部参数 s, d 分析

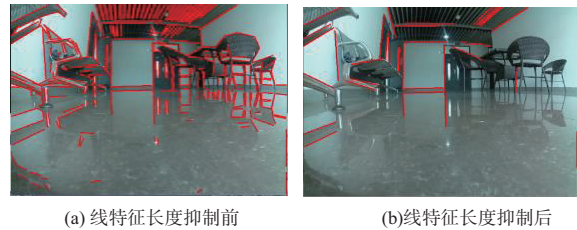


图 3 长度抑制结果对比

3 线特征参数化及重投影误差模型

3.1 线特征的参数化

三维中的刚体有 6 个自由度,而线段是 4 个自由度,如使用三维空间中两点的齐次坐标表示,会产生多余的 4 个参数,使优化问题的约束增加。因此本文使用普吕克坐标^[8]和正交表示法相结合的方式来表示空间直线,其中普吕克直线坐标法可以方便直线参与几何运算,而正交表示法用于直线参与的优化运算。

定义 P_1 和 P_2 为空间中一条直线上两点,则空间直线普吕克坐标如式(2)所示,其中 \mathbf{v} 是直线的方向向量, \mathbf{n} 是由坐标原点与 P_1P_2 构成平面的法向量。

$$\mathbf{L}_w = \begin{bmatrix} P_1 \times P_2 \\ w_1 P_1 - w_2 P_2 \end{bmatrix} = \begin{bmatrix} \mathbf{n} \\ \mathbf{v} \end{bmatrix} \quad (2)$$

正交表示对应两个正交矩阵 $(U, W) \in SO(3) \times SO(2)$,可通过对普吕克坐标进行 QR 分解得到。

3.2 线特征重投影误差模型

线重投影误差与点的重投影残差类似,首先,将直线 \mathbf{L}_w 从世界坐标系转换到相机坐标系,记作 \mathbf{L}_c ,并使用普吕克坐标表示,如式(3)所示,式中 \mathbf{T}_{cw} 为世界坐标到相机坐标的直线变换矩阵, \mathbf{R}_{cw} 为旋转矩阵, \mathbf{t}_{cw} 为位移矩阵, $(\mathbf{t}_{cw})^\wedge$ 表示 \mathbf{t}_{cw} 的反对称矩阵。

$$\mathbf{L}_c = \mathbf{T}_{cw} \mathbf{L}_w = \begin{bmatrix} \mathbf{R}_{cw} & (\mathbf{t}_{cw})^\wedge \mathbf{R}_{cw} \\ 0 & \mathbf{R}_{cw} \end{bmatrix} \begin{bmatrix} \mathbf{n}_w \\ \mathbf{v}_w \end{bmatrix} \quad (3)$$

接着,将 L_c 投影到图像平面得到 l_c ,其中, K_L 表示直线的投影矩阵:

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

最后,参考文献[8]中给出的直线重投影误差的定义方式,采用代数几何距离定义的非线性误差来表示。

如图4所示,匹配所得线段 $x_m x_n$ 与投影所得线段 l_c 之间并不重合,用 d_1 与 d_2 表示二者的误差,即匹配线段端点到投影所得线段的距离,线的重投影误差如下式所示。

$$r_L(z_j^c, \mathcal{X}) = [d_1, d_2]^T = \left[\frac{x_m^T l_c}{\sqrt{l_1^2 + l_2^2}}, \frac{x_n^T l_c}{\sqrt{l_1^2 + l_2^2}} \right]^T \quad (5)$$

式中: L 表示所观察到的线特征的集合; L_j 表示在第 i 帧中观察到的第 j 条线段特征。

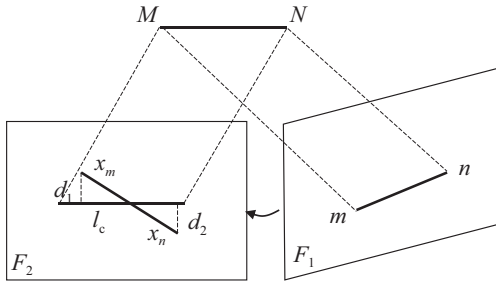


图4 线重投影误差示意图

4 基于点线及惯性信息融合的后端滑动窗口优化

考虑到实时性问题,采用滑动窗口模型进行后端优化。首先,在滑动窗口中定义点、线和 IMU 测量信息相关的全状态向量 \mathcal{X} ,其次,通过最小化滑动窗口中的目标函数来优化窗口中的状态量:

$$\min_{\mathcal{X}} (e_{\text{prior}} + e_{\text{imu}} + e_{\text{point}} + e_{\text{line}} + e_{\text{loop}}) \quad (6)$$

与传统的基于点特征的视觉惯性融合 SLAM 方法相比,本文在目标函数中增加了线特征的重投影残差:

$$e_{\text{line}} = \sum_{(i,j) \in L} (\rho \|r_L(z_j^c, \mathcal{X})\|_{P_{L_j}^c})^2 \quad (7)$$

式中: ρ 为鲁棒核函数,用于抑制异常值; $P_{L_j}^c$ 为线特征噪声项的协方差矩阵。之后采用 LM 算法来进行迭代求解非线性优化问题。

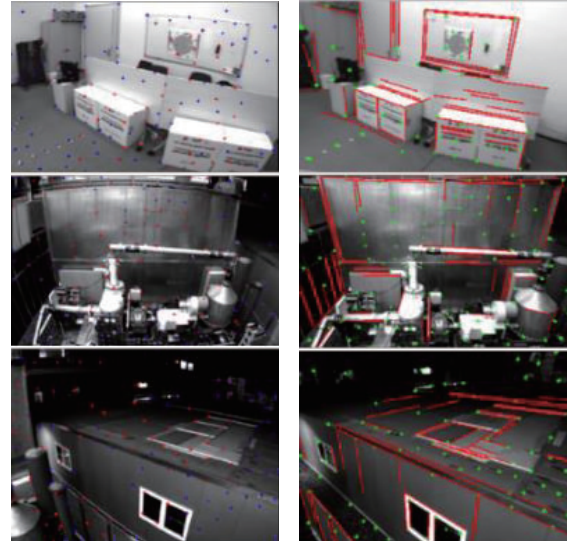
5 实验与分析

为验证点线融合的视觉惯性 SLAM 方法的有效性,将本文方法在公开数据集上进行实验,并在实际环境中验证。实验所用 PC 配置为 Intel (R) Core (TM) i3-4170 CPU @ 3.70 GHz 处理器,4 GB 内存,系统版本为 Ubuntu 18.04 LTS 64 位,ROS Melodic。

5.1 定位精度实验

在公开的 EuRoC 数据集^[9]下对本文的方法进行仿真实验对比,以验证其定位精度。图5为数据集 MH_04_

difficult 序列与 V1_02_medium 序列中部分光照变化及弱纹理场景实验图。可以看到,在图中所示光线变化较大且弱纹理区域,所提取的点特征较为稀少,而室内人造场景中存在较多的物体边缘为线性结构,因此使用线特征提取算法,能够获得丰富的线特征,可以为后续的运动估计提供充足且鲁棒的信息。



(a) 点特征提取

(b) 点线特征提取

图5 部分弱纹理场景特征提取对比

为了进行定量评估,表1给出了以绝对轨迹误差 (ATE) 为指标的二者数据集仿真结果,即计算估计位姿与真实位姿之间欧氏距离的方均根误差 (RMSE)。

表1 EuRoC 数据集方均根误差表

名称	VINS-mono 点+IMU	本文算法点、线+IMU
MH_01_easy	0.181 221	0.149 008
MH_02_easy	0.184 900	0.090 001
MH_03_medium	0.411 024	0.092 633
MH_04_difficult	0.401 103	0.225 026
MH_05_difficult	0.380 402	0.266 130
V1_01_easy	0.142 571	0.058 560
V1_02_medium	0.313 125	0.093 658
V1_03_difficult	0.318 180	0.116 828
V2_01_easy	0.124 754	0.081 699
V2_02_medium	0.305 598	0.097 142
V2_03_difficult	0.320 891	0.131 528
平均值	0.280 343	0.127 474

根据实验结果可以发现,引入线特征后算法准确度较仅有点特征的视觉 SLAM 算法准确度明显提高,与 VINS-mono 相比,平均定位误差下降了 54%。在 V1_02_medium 序列中,尽管场景中存在弱纹理区域,但本文方法依然稳定运行,具备较好的鲁棒性。图6为不同算法在 MH_04_difficult 序列中得到的 xy 平面轨迹与真实轨迹对比图。从图中可以看出,本文的点线融合方法与真实轨迹距离较近,精度较高。

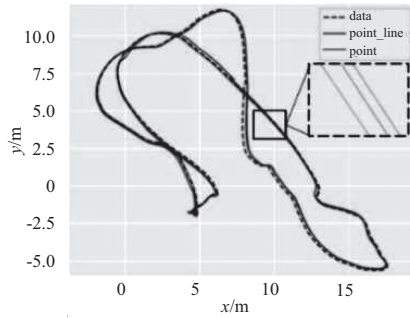


图6 MH_04_difficult 序列轨迹对比图

5.2 实际弱纹理环境测试

为了测试本文算法在实际环境中的应用效果,使用移动小车在室内环境中采集数据进行验证。移动小车配备树莓派 4B 控制机身运动及获取环境数据,视觉数据通过树莓派 500w 广角摄像头获取,惯性数据通过惯性传感器 MPU6050 获取。实验场景为教学楼的室内部分走廊,特点是白墙较多,属于典型的室内弱纹理环境,且光照变化大,对于位姿估计存在一定难度,如图 7 所示。

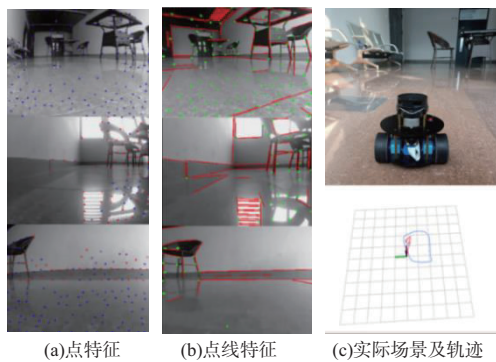


图7 实际环境下实验结果对比

不同算法在实际场景下的运动估计精度依旧采用 ATE 评估,如表 2 所示。可以看到,实际结果与数据集仿真结果基本符合,表明引入线特征对于 SLAM 精度有所提升。

表2 室内实际场景方均根误差表

VINS-mono 点+IMU	本文算法点、线+IMU
0.435 598	0.246 330

6 结语

本文基于 VINS-mono 框架,引入线特征,得到点线融合的视觉惯性 SLAM 方法。通过改进的 LSD 算法提高了特征提取的速度,使用线段长度抑制方法提高了线特征提取的质量。在数据集的仿真实验中表明,点线融合的特征提取与匹配策略可以加强帧间的数据关联,提高定位精度,并验证了该方法在实际室内弱纹理环境中的定位精确性,为后续建立高精度地图及导航等功能实现奠定了基础。

参考文献:

[1] MUR-ARTAL R, TARDÓS J D. Visual-inertial monocular SLAM with map reuse [J]. IEEE Robotics and Automation Letters, 2017, 2(2): 796-803.

[2] 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.

[3] PUMAROLA A, VAKHITOV A, AGUDO A, et al. PL-SLAM: real-time monocular visual SLAM with points and lines [C]// 2017 IEEE International Conference on Robotics and Automation. Singapore: IEEE, 2017: 4503-4508.

[4] HE Y, ZHAO J, GUO Y, et al. PL-VIO: tightly-coupled monocular visual-inertial odometry using point and line features [J]. Sensors (Basel, Switzerland), 2018, 18(4): E1159.

[5] 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.

[6] VON GIOI R G, JAKUBOWICZ J, MOREL J M, et al. LSD: a line segment detector [J]. Image Processing on Line, 2012(2): 35-55.

[7] FU Q, WANG J L, YU H S, et al. PL-VINS: real-time monocular visual-inertial SLAM with point and line [EB/OL]. 2020; arXiv: 2009.07462 [cs. RO]. <https://arxiv.org/abs/2009.07462>.

[8] BARTOLI A, STURM P. The 3D line motion matrix and alignment of line reconstructions [C]// Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001. Kauai, HI, USA: IEEE, 2001: I.

[9] 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.

收稿日期: 2022-02-16