

# 一种单目 VIO 定位精度与跟踪稳定性优化方法

谢波<sup>1,2</sup> 张国良<sup>1,2</sup> 李歆<sup>1,2</sup> 张自杰<sup>1,2</sup> 汪坤<sup>1,2</sup>

(1. 四川轻化工大学自动化与信息工程学院 宜宾 644000; 2. 人工智能四川省重点实验室 宜宾 644000)

**摘要:**为了提高视觉惯性同时定位与建图(visual-inertial simultaneous localization and mapping, VISLAM)系统的系统性能,提出了一种单目视觉惯性里程计(visual-inertial odometry, VIO)定位精度与跟踪稳定性优化方法。在相机位姿优化阶段,通过多残差项对相机位姿进行优化,提高系统的定位精度。在特征跟踪丢失时,通过惯性测量单元(inertial measurement unit, IMU)积分信息与特征点深度增强方法对系统进行重定位,提高系统的跟踪稳定性。针对所提方法,基于 ORB-SLAM3 代码框架进行改进,并在开源数据集与真实环境中验证方法的可行性。实验结果表明,所提方法能够有效提高视觉惯性里程计的定位精度与跟踪稳定性。

**关键词:**同时定位与建图;视觉惯性里程计;多残差项;深度增强;ORB-SLAM3

**中图分类号:** TP242 **文献标识码:** A **国家标准学科分类代码:** 520.604

## Optimization method for positioning accuracy and tracking stability of monocular VIO

Xie Bo<sup>1,2</sup> Zhang Guoliang<sup>1,2</sup> Li Xin<sup>1,2</sup> Zhang Zijie<sup>1,2</sup> Wang Kun<sup>1,2</sup>

(1. School of Automation and Information Engineering, Sichuan University of Science & Engineering, Yibin 644000, China; 2. Key Laboratory of Artificial Intelligence in Sichuan Province, Yibin 644000, China)

**Abstract:** In order to improve the system performance of visual-inertial simultaneous localization and mapping (VISLAM) system, an optimization method for positioning accuracy and tracking stability of monocular visual-inertial odometry (VIO) is proposed. In the camera pose optimization stage, in order to improve the positioning accuracy of the system, the camera pose is optimized through multiple residual items. When feature tracking is lost, in order to improve the tracking stability of the system, the integral information of inertial measurement unit (IMU) and the depth enhancement method of feature point are used to relocate the system. The proposed method is improved based on the ORB-SLAM3 code framework, and the feasibility of the method is verified in open source data sets and real environments. The experimental results show that the proposed method can effectively improve the positioning accuracy and tracking stability of the visual-inertial odometry.

**Keywords:** simultaneous localization and mapping; visual-inertial odometry; multiple residual items; depth enhancement; ORB-SLAM3

## 0 引言

视觉惯性即时定位与地图构建(simultaneous localization and mapping, SLAM)技术<sup>[1-2]</sup>通过处理相机与惯性测量单元(inertial measurement unit, IMU)获取的信息,可在陌生的环境中完成实时定位与周围环境地图构建任务,目前已经广泛应用于自动驾驶<sup>[3]</sup>、智能机器人<sup>[4]</sup>、无人机<sup>[5]</sup>等众多领域。

视觉惯性 SLAM 按照处理视觉信息与 IMU 信息的方法不同可分为两大类,松耦合法和紧耦合法。松耦合法分别处理视觉与 IMU 信息完成系统状态估计,然后将二者的估计结果进行融合,而紧耦合法同时处理视觉信息与 IMU 信息直接完成系统状态估计。基于松耦合法,Weiss 等<sup>[6]</sup>提出了一种基于微型飞行器的视觉惯性导航系统,该系统通过扩展卡尔曼滤波器将视觉 SLAM 模块与 IMU 估计的状态量进行融合,实现了微型飞行器的实时导航。

Shen 等<sup>[7]</sup>提出了一种模块化、可扩展的实时定位系统,该系统将松耦合的视觉惯性里程计作为系统中的一部分,然后通过无迹卡尔曼滤波器融合所有传感器的信息来完成定位任务。Cvišić 等<sup>[8]</sup>提出了 SOFT-SLAM 系统,该系统分线程完成定位与建图任务,有效提高了系统运行效率。松耦合法的灵活性、扩展性好,但是定位精度低,相比于松耦合法,紧耦合法虽然灵活性、扩展性偏低,但是定位精度高、系统鲁棒性好,因此紧耦合法应用的更加广泛。基于紧耦合法,Leutenegger 等<sup>[9]</sup>提出了 OKVIS 系统,该系统提出了一种相机外参在线标定方法,并采用滑窗对系统状态变量进行优化,能够在存在外点的情况下稳定运行,有效提高了特征跟踪稳定性。Bloesch 等<sup>[10]</sup>提出了 ROVIO 系统,不同于前面算法基于图像帧间特征点灰度残差构建优化问题,该系统通过图像帧间特征点周围图像块的灰度残差构建优化问题,最终提高了系统的定位精度与特征跟踪稳定性,但是降低了系统的运行效率。Mur-Artal 等<sup>[11-12]</sup>提出了 VIORB 系统,该系统基于 ORB-SLAM 系统,能够实现地图重用并在完成建图的区域实现零偏移定位,有效改善了 ORB-SLAM 系统的性能。Qin 等<sup>[13]</sup>提出 VINS-Mono 系统,该系统用光流法追踪特征点,并结合滑窗约束与回环约束对系统状态变量进行优化,拥有较高的定位精度,但特征跟踪不够稳定,很难实现系统重定位。Campos 等<sup>[14]</sup>在 ORB-SLAM2 系统<sup>[15]</sup>的基础上提出 ORB-SLAM3 系统,该系统的定位精度与特征跟踪稳定性优于前面的算法,但该系统在特征跟踪丢失后会重新初始化新的地图,这导致如果系统检测不到回环就无法获得完整的位姿轨迹。

上述文献在改善 SLAM 系统的系统性能上都做出了很大贡献,但在系统的定位精度与特征跟踪稳定性方面仍有很大的提升空间。为此本文提出了一种单目视觉惯性里程计(visual-inertial odometry, VIO)定位精度与跟踪稳定性优化方法,进一步提高 SLAM 系统的系统性能。在相机位姿优化阶段,同时最小化重投影像素点坐标、灰度残差与 IMU 预积分残差,提高相机位姿估计精度。在特征跟踪丢失时,通过 IMU 恢复相机位姿,并用深度增强的方法恢复特征点空间三维坐标,提高特征跟踪稳定性。

## 1 系统框架

本文通过处理相机与 IMU 数据来构建视觉惯性里程计系统,并对相机位姿优化方法与特征点跟踪丢失后系统重定位方法进行改进,以提高视觉惯性里程计系统的定位精度与特征跟踪稳定性,系统流程如图 1 所示。

如图 1 所示,当系统利用图像数据与 IMU 数据完成系统初始化后,首先会提取图像帧中的 ORB 特征点<sup>[16]</sup>,并对所提取特征点进行图像帧间特征跟踪。然后判断特征跟踪是否成功,如果特征跟踪成功则在图像帧间构建优化问题,并利用列文伯格-马夸尔特算法对相机位姿进行优化,否则利用 IMU 信息与特征点深度增强的方法对系

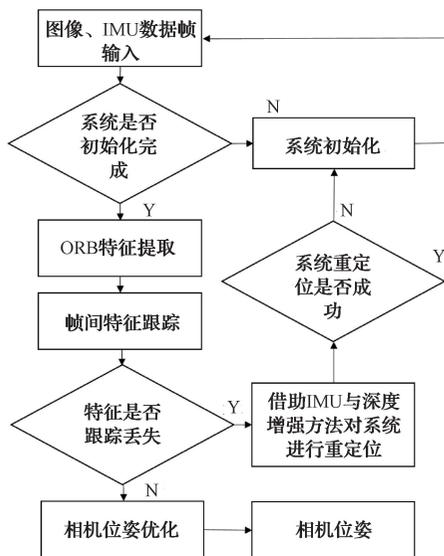


图 1 系统流程

统进行重定位,如果重定位失败就重新初始化系统。

## 2 系统流程

### 2.1 坐标系定义

视觉惯性里程计主要涉及的坐标系有世界坐标系,记作  $w$  系;相机坐标系,记作  $c$  系;IMU 坐标系,记作  $b$  系;坐标系定义示意图如图 2 所示。

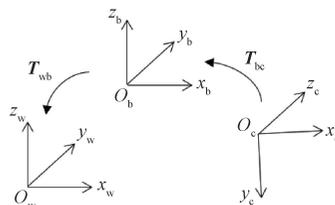


图 2 坐标系示意图

图 2 中各坐标系均符合右手准则,  $T_{wb}$  是 IMU 坐标系到相机坐标系的变换矩阵,  $T_{bc}$  是相机坐标系到 IMU 坐标系的变换矩阵。

### 2.2 系统初始化

系统初始化的流程与系统初始化过后的流程都大致相同,都会通过提取并追踪图像帧中的 ORB 特征来恢复相机位姿。不同之处在于系统初始化时会借助 IMU 信息来恢复场景尺度<sup>[17]</sup>,并且如果特征跟踪丢失后并不会进行重定位,而会直接抛弃特征跟踪丢失前的图像帧,并将当前帧作为原始第 1 帧,追踪当前帧的特征点。另外,初始化时根据对极约束来恢复相机位姿,并通过三角化的方法恢复空间路标点的坐标。

### 2.3 特征提取与特征跟踪

在特征提取阶段,首先构建图像金字塔,然后分层提取图像中的 ORB 特征点。在提取完图像特征后通过帧间

特征匹配方法进行特征跟踪,常用的特征匹配方法有词袋向量法与区域搜索法。词袋向量法速度快,但是在一些相似场景中,由于各个特征点的描述子差异不大,可能会造成许多误匹配。相比之下,区域搜索法虽然速度稍慢,但是由于加上了范围的限制,可有效抑制误匹配点的存在。考虑上述因素,选用区域搜索法进行帧间特征匹配,基于区域搜索法的帧间特征匹配原理如图3所示。

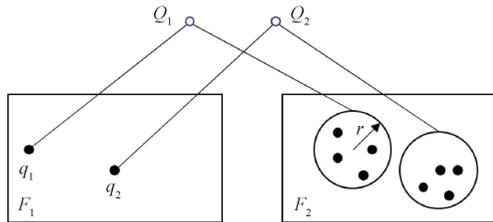


图3 帧间特征匹配原理

图3中  $F_1$  为参考帧,  $F_2$  为当前帧,  $q_1, q_2$  为参考帧中的特征点,  $Q_1, Q_2$  为  $q_1, q_2$  对应的空间路标点。因为  $F_1$  在世界坐标系下的相机位姿  $T_{w1}$  在前面跟踪过程中已经求得,所以根据  $F_1 \sim F_2$  间IMU积分所得的相对位姿  $T_{21}$  可计算出  $F_2$  在世界坐标系下的相机位姿初值  $\tilde{T}_{w2}$ :

$$\tilde{T}_{w2} = T_{w1} T_{21}^{-1} \quad (1)$$

利用  $\tilde{T}_{w2}$  将  $Q_1, Q_2$  投影到  $F_2$  中,在投影点处半径为  $r$  的图像区域内搜索与  $q_1, q_2$  匹配度最高的特征点,完成帧间特征匹配。搜索半径  $r$  的大小由式(2)决定。

$$r = n\lambda \times 7 \quad (2)$$

式中:  $n$  为特征点所在图像金字塔的层级 ( $1 \leq n \leq 8$ );  $\lambda = 1.2$  为图像金字塔的缩放因子。

#### 2.4 基于多残差项相结合的相机位姿优化

帧间特征匹配成功后通过多残差项对相机位姿进行优化,得到最优相机位姿,相机位姿优化原理如图4所示。

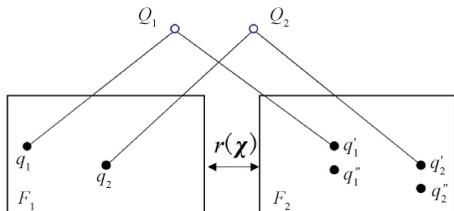


图4 相机位姿优化原理

图4中参考帧  $F_1$  中的特征点  $q_1$  与  $q_2$  在当前帧  $F_2$  中的匹配点为  $q'_1$  与  $q'_2$ ;  $r(\chi)$  为  $F_1$  与  $F_2$  间的IMU预积分残差,  $\chi = [R_{wb} \ t_{wb} \ v_w \ b_a \ b_g]$  为IMU的状态向量,其中  $R_{wb}$  与  $t_{wb}$  分别为IMU在世界坐标系下的旋转量与平移量,  $v_w$  为IMU在世界坐标系下的速度,  $b_a$  与  $b_g$  分别为IMU加速度计与陀螺仪的零偏;  $q''_1, q''_2$  为  $Q_1, Q_2$  根据  $F_2$  的相机位姿初值  $\tilde{T}_{w2}$ , 在  $F_2$  中的投影点。

因为  $\tilde{T}_{w2}$  精度不高,导致  $q''_1, q''_2$  与  $q'_1, q'_2$  间存在重

投影像素点坐标与灰度残差。为了提高  $\tilde{T}_{w2}$  的精度,通过结合特征点匹配对间的重投影像素点坐标、灰度残差与帧间IMU预积分残差,对IMU的状态向量  $\chi$  进行优化,从而间接优化  $\tilde{T}_{w2}$ 。重投影像素点坐标残差为:

$$c(\chi, Q_i^w, u_i) = u_i - K(R_{cb}(R_{wb}^{-1}Q_i^w - R_{wb}^{-1}t_{wb}) + t_{cb}) \quad (3)$$

重投影像素点灰度残差为:

$$o(\chi, Q_i^w, u_i) = I(u_i) - I(K(R_{cb}(R_{wb}^{-1}Q_i^w - R_{wb}^{-1}t_{wb}) + t_{cb})) \quad (4)$$

式中:  $u$  为在  $F_2$  中匹配到的特征点的像素坐标;  $K$  为相机内参矩阵;  $R_{cb}$  与  $t_{cb}$  为相机与IMU间的外参;  $Q_i^w$  为特征点匹配对对应的空间路标点的坐标;  $I(\cdot)$  为像素点坐标处的像素值;下标  $i$  为特征点匹配对的序号。

根据式(3)、(4)与预积分残差项  $r(\chi)$  可构建最终的优化方程:

$$\chi^* = \underset{\chi}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^n \| c(\chi, Q_i^w, u_i) \|_2^2 + \frac{1}{2} \sum_{i=1}^n \| o(\chi, Q_i^w, u_i) \|_2^2 + r(\chi) \right\} \quad (5)$$

利用列文伯格-马夸尔特算法对式(5)进行优化求解后,可得到IMU的最优状态  $\chi^*$ 。求解得到  $\chi^*$  后,一方面更新当前时刻IMU的状态,以消除IMU的累计误差;另一方面通过优化后的  $R_{wb}$  与  $t_{wb}$  得到最优当前帧相机位姿  $T_{w2}^*$  为:

$$T_{w2}^* = T_{wb} T_{cb}^{-1} \quad (6)$$

式中:  $T_{wb} = [R_{wb} \ | \ t_{wb}]$ ,  $T_{cb} = [R_{cb} \ | \ t_{cb}]$ 。

#### 2.5 特征点跟踪丢失后系统的重定位

特征点跟踪丢失后对系统的重定位过程就是恢复当前帧相机位姿与当前帧特征点在三维空间中的坐标。一方面,因为在特征点跟踪丢失前的每一次相机位姿优化过程中都会消除IMU的累计误差并更新IMU的零偏  $b_a$  与  $b_g$ ,所以IMU在参考帧与当前帧间的积分结果是比较准确的,可通过IMU的积分信息恢复当前帧的相机位姿。另一方面,可利用系统在局部地图<sup>[12]</sup>中的空间路标点,对当前帧的特征点进行深度增强,恢复当前帧特征点的空间三维坐标,特征点深度增强原理如图5所示。

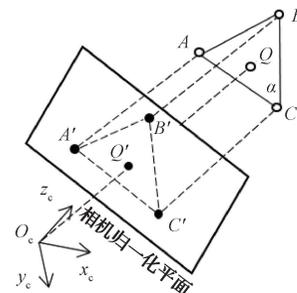


图5 特征点深度增强原理

如图 5 所示,当前帧中的特征点  $q$  在相机归一化平面上的点为  $Q'$ , 利用 IMU 积分恢复的当前帧相机位姿可将局部地图中在当前帧视野范围内的空间路标点转换到当前帧相机归一化平面上,然后通过 KT-tree 可搜索到局部地图中离  $Q'$  最近的 3 个归一化平面上的点  $A'$ 、 $B'$ 、 $C'$ , 它们分别对应当前帧坐标系下的空间路标点  $A$ 、 $B$ 、 $C$ 。通过  $A$ 、 $B$ 、 $C$  三点可构成平面  $\alpha$ , 那么相机光心  $O_c$  与  $Q'$  形成的射线与平面  $\alpha$  的交点  $Q$  就为  $q$  对应的空间路标点。设  $O_c$ 、 $Q'$ 、 $Q$  的空间三维坐标分别为  $O_c = [O_1, O_2, O_3]^T$ 、 $Q' = [X', Y', Z']^T$ 、 $Q = [X, Y, Z]^T$ , 那么可得:

$$[X \ Y \ Z] = [O_1 + X's \ O_2 + Y's \ O_3 + Z's] \quad (7)$$

其中,只有  $s$  为未知数,只要求出  $s$  就能求出  $Q$ 。设空间路标点  $A$ 、 $B$ 、 $C$  的空间三维坐标分别为  $A = [a_1, a_2, a_3]^T$ 、 $B = [b_1, b_2, b_3]^T$ 、 $C = [c_1, c_2, c_3]^T$ , 那么可求出  $\alpha$  平面的法向量  $N$  为:

$$N = (A - B) \times (B - C) \quad (8)$$

设  $N = [N_1, N_2, N_3]^T$ , 并选取  $\alpha$  平面上的空间路标点  $A$  可得到:

$$N_1(X - a_1) + N_2(Y - a_2) + N_3(Z - a_3) = 0 \quad (9)$$

因为相机光心  $O_c$  的坐标值为 0, 所以联合式(7)与(9)可求解出  $s$  的值:

$$s = \frac{N_1 a_1 + N_2 a_2 + N_3 a_3}{N_1 X' + N_2 Y' + N_3 Z'} \quad (10)$$

求解出的  $s$  实际上就是当前帧坐标系下特征点  $q$  的深度值, 根据  $s$  可求解出  $Q$ :

$$Q = sQ' \quad (11)$$

为了保障恢复出的  $Q$  更接近于真实值, 当  $Q'$  与  $A'$ 、 $B'$ 、 $C'$  任意一点的距离大于阈值 10 cm 或  $A$ 、 $B$ 、 $C$  三点间任一点对的距离大于 10 cm 时放弃恢复  $Q$  的值。如果最终通过特征点深度增强的方法恢复出来的空间路标点数量

不足, 则说明重定位失败, 此时需从当前帧开始, 重新进行系统初始化。

### 3 实验与分析

本文在开源数据集与实际环境中验证所提方法的可行性, 实验设备如图 6 所示。



图 6 实验设备

图 6 中 PC 的型号为 Lenovo XiaoXin I2000, 处理器为 Intel i7-5557U, 主频为 3.10 GHz, 运行内存为 8.00 GB, 系统为 Ubuntu 16.04; 相机的型号为 Intel RealSense D435i, 该相机带有一对双目相机、一个 RGB 相机, 并且内置 IMU, 相机的分辨率设置为  $640 \times 480$ , 频率设置为 30 Hz, IMU 的频率设置为 200 Hz; PC 与相机间的接口为 USB 3.0 接口; 移动平台为 Smart 智能移动平台; 轨迹评估工具为 EVO。

#### 3.1 数据集系统性能实验分析

在数据集实验中, 选取了 EuRoC 数据集和 TUM-VI 数据集里一共 9 个不同的图像序列在单目模式下与 ORB-SLAM3 算法进行轨迹绝对误差对比实验。表 1 与图 7 分别为本文所提方法、ORB-SLAM3 算法在 EuRoC 数据集上估计轨迹与真实轨迹间的绝对误差对比结果与轨迹对比。

表 1 EuRoC 数据集上轨迹的绝对误差比较

(m)

图像序列	算法	最大误差	平均误差	最小误差	均方根误差
MH02	ORB-SLAM3	0.242	0.114	0.010 8	0.127
	本文	0.106	0.036 8	0.003 00	0.042 8
MH03	ORB-SLAM3	0.110	0.053 6	0.009 64	0.057 4
	本文	0.100	0.042 9	0.007 79	0.046 9
V102	ORB-SLAM3	0.039 8	0.018 7	0.002 62	0.020 7
	本文	0.034 6	0.015 4	0.002 19	0.016 7
V103	ORB-SLAM3	0.082 9	0.032 3	0.003 25	0.038 4
	本文	0.059 4	0.024 2	0.002 54	0.028 3

表 2 与图 8 分别为本文所提方法、ORB-SLAM3 算法在 TUM-VI 数据集上估计轨迹与真实轨迹间的绝对误差对比结果与轨迹对比。

EuRoC 数据集里的图像序列是通过无人机采集的, 多为室内场景; 而 TUM-VI 数据集里的图像序列是通过手持设备采集的, 有室内与室外场景。图 7 中的所有图像

序列与图 8(a)~(d) 图像序列都为比较常见的室内或室外场景, 在这些场景中一般不会出现特征点跟踪丢失的情况。而图 8(e) 图像序列为极具挑战性的场景, 在这种场景下很有可能会出现特征点跟踪丢失的情况。如图 7、8 所示, 不管是在什么场景下, 本文所提方法都能完整估计整个数据集的轨迹, 这说明本文所提方法具有很好的特征跟

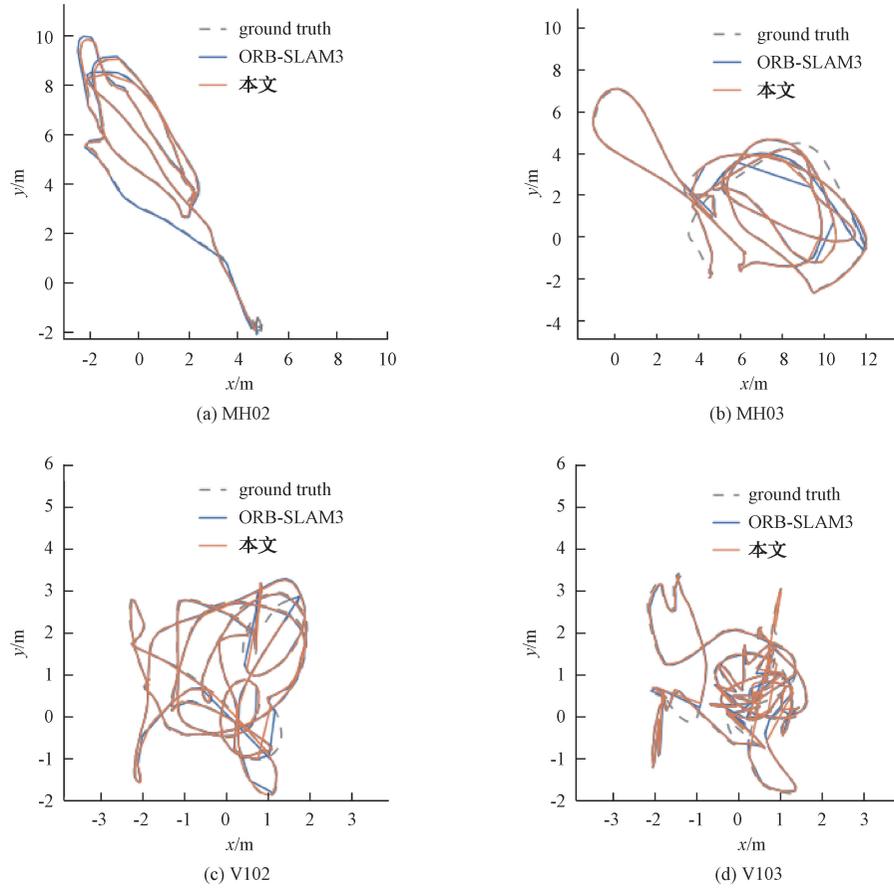


图7 EuRoC数据集上轨迹的对比

表2 TUM-VI数据集上轨迹的绝对误差比较

(m)

图像序列	算法	最大误差	平均误差	最小误差	均方根误差
Corridor2	ORB-SLAM3	0.096 4	0.065 8	0.040 8	0.067 4
	本文	0.052 6	0.028 4	0.014 1	0.030 0
Magistrale2	ORB-SLAM3	0.405	0.329	0.274	0.333
	本文	0.233	0.172	0.114	0.175
Room2	ORB-SLAM3	0.018 0	0.009 06	0.002 76	0.009 45
	本文	0.013 7	0.006 14	0.002 56	0.006 72
Outdoor2	ORB-SLAM3	16.818	11.163	6.990	11.561
	本文	11.128	8.240	5.474	8.318
Slide2	ORB-SLAM3	1.554	1.298	1.120	1.303
	本文	0.996	0.751	0.536	0.761

踪稳定性。并且如表1、2所示,在这些场景中本文所提方法估计轨迹与真实轨迹间的最大误差、平均误差、最小误差、均方根误差都明显低于ORB-SLAM3算法。相比于ORB-SLAM3算法,本文所提方法在9个数据集中轨迹的最大误差降低了32.0%,平均误差降低了37.3%,最小误差降低了37.2%,均方根误差降低了36.9%,这说明本文所提方法能有效提高SLAM系统的定位精度。

### 3.2 实际环境系统性能实验分析

在实际环境实验中,选取室内场景、走廊场景与室外场景在单目模式下验证所提方法的实用性,图9所示为实际环境场景。

图9中,在室内、走廊与室外场景的定位建图轨迹大约为27、225、343 m,整个定位建图过程从起点开始,最终又回到起点位置。图10所示为在各个场景中的定位建图过程,图11所示为在各个场景中建立的点云地图与相机轨迹。

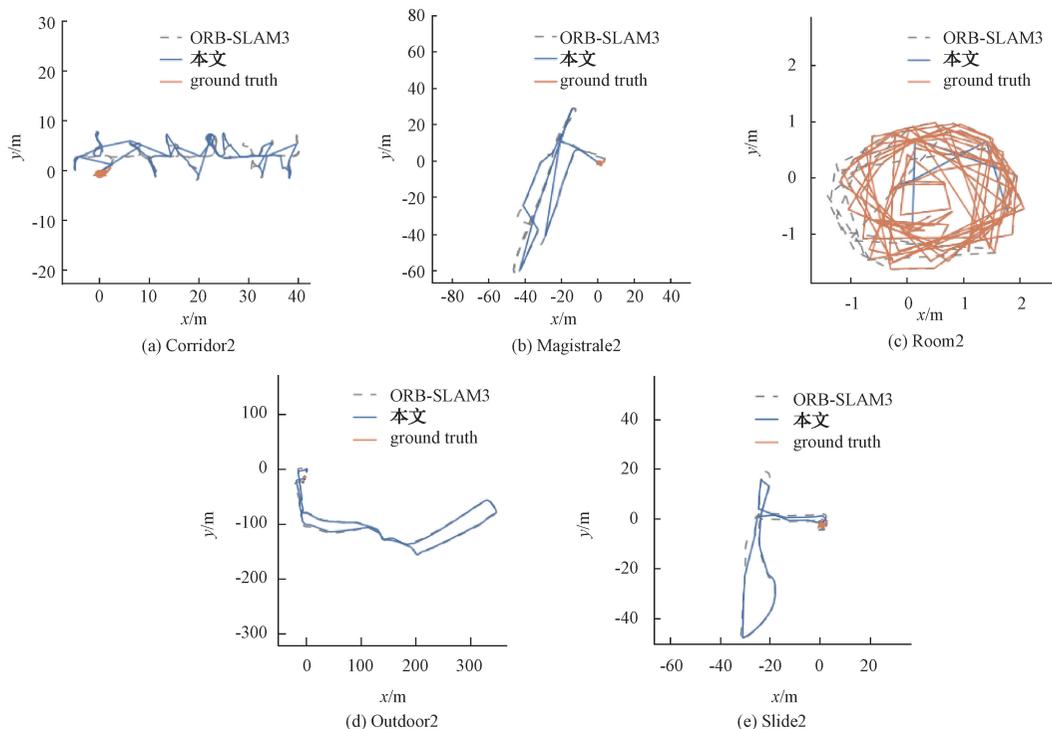


图8 TUM-VI数据集上轨迹的对比



图9 实际环境场景

图10中用红色方框框起来的特征跟踪丢失的图像帧所示,即使在跟踪过程中出现特征跟丢的情况,在经过本文所提方法对系统进行重定位后也能重新恢复特征跟踪,估计出完整的相机轨迹,这说明本文所提方法能够提高SLAM系统的特征跟踪稳定性。另外,图11中用红色虚线椭圆圈起来的相机轨迹为室内场景、走廊场景与室外场景中相机的起点轨迹与终点轨迹。根据实验数据,相机在室内场景、走廊场景与室外场景中的起点轨迹与终点轨迹间的偏差分别不超过0.08、0.25、0.3 m,说明本文所提方法在实际环境中能保持较高的定位精度,有很好的实用性。

#### 4 结论

本文为了提高VISLAM系统的系统性能,提出了一种单目VIO定位精度与跟踪稳定性优化方法。当特征跟

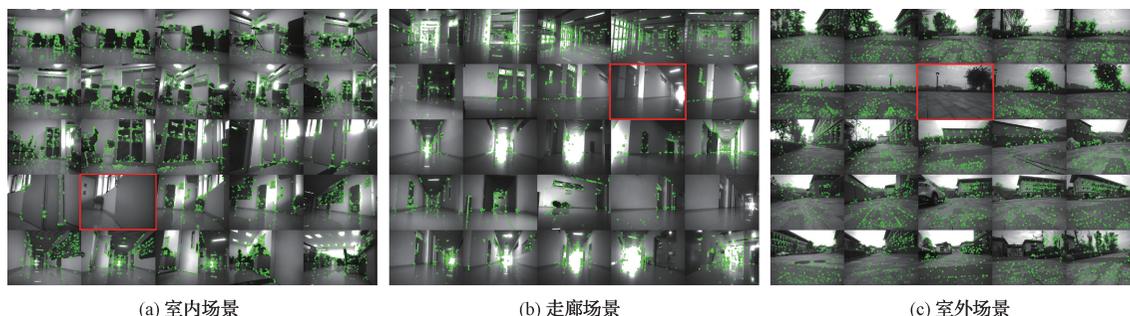


图10 定位建图过程

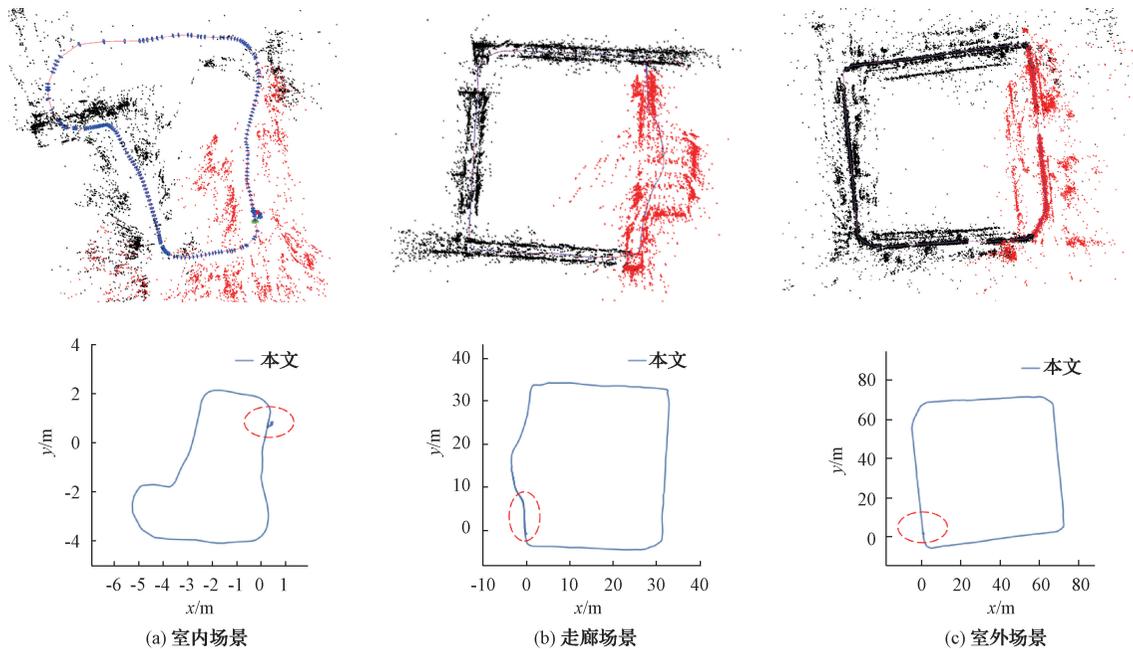


图 11 点云地图与相机轨迹示意图

踪丢失时,一方面用 IMU 的积分信息恢复相机位姿,另一方面用特征点深度增强的方法恢复特征点的空间三维坐标,实现系统的重定位。在对相机位姿进行优化时,结合特征点匹配对间的最小化重投影像素点坐标残差、灰度残差与 IMU 的预积分残差对相机位姿进行优化,从而得到精度较高的相机位姿。在开源数据集与实际环境中的实验证明,本文所提方法能够有效提高 VISLAM 系统的定位精度与特征跟踪稳定性。

#### 参 考 文 献

- [1] 刘建军, 卢大威, 胡雪花, 等. 基于点线特征快速单目惯性 SLAM 算法[J]. 国外电子测量技术, 2022, 41(3): 14-19.
- [2] 余洪山, 郭丰, 郭林峰, 等. 融合改进 SuperPoint 网络的鲁棒单目视觉惯性 SLAM[J]. 仪器仪表学报, 2021, 42(1): 116-126.
- [3] SINGANDHUPE A, LA H M. A review of SLAM techniques and security in autonomous driving[C]. 2019 3rd IEEE International Conference on Robotic Computing (IRC). IEEE, 2019: 602-607.
- [4] 吴玉香, 王超, 洗颖宪. 基于稀疏直接法和图优化的移动机器人 SLAM[J]. 仪器仪表学报, 2018, 39(4): 257-263.
- [5] 付林, 郑佳楠, 何洪磊, 等. 面向大型室内场景的无人机三维激光雷达解耦 SLAM 方法[J]. 电子测量技术, 2022, 45(13): 96-103.
- [6] WEISS S, ACHELNIK M W, LYNEN S, et al. Real-time onboard visual-inertial state estimation and self-calibration of MAVs in unknown environments[C]. 2012 IEEE International Conference on Robotics and Automation. IEEE, 2012: 957-964.
- [7] SHEN S, MULGAONKAR Y, MICHAEL N, et al. Multi-sensor fusion for robust autonomous flight in indoor and outdoor environments with a rotorcraft MAV[C]. 2014 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2014: 4974-4981.
- [8] CVIŠIĆ I, ĆESIĆ J, MARKOVIĆ I, et al. SOFT-SLAM: Computationally efficient stereo visual simultaneous localization and mapping for autonomous unmanned aerial vehicles[J]. Journal of Field Robotics, 2018, 35(4): 578-595.
- [9] 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.
- [10] BLOESCH M, OMARI S, HUTTER M, et al. Robust visual inertial odometry using a direct EKF-based approach[C]. 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2015: 298-304.
- [11] MUR-ARTAL R, TARDÓS J D. Visual-inertial monocular SLAM with map reuse[J]. IEEE Robotics and Automation Letters, 2017, 2(2): 796-803.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] RUBLEE E, RABAUD V, KONOLIGE K, et al. ORB: An efficient alternative to SIFT or SURF[C]. 2011 International Conference on Computer Vision. IEEE, 2011: 2564-2571.
- [17] CAMPOS C, MONTIEL J M M, TARDÓS J D. Inertial-only optimization for visual-inertial initialization[C]. 2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2020: 51-57.

## 作者简介

谢波, 硕士研究生, 主要研究方向为视觉惯性 SLAM 技术。

E-mail: 1342407974@qq.com

张国良, 博士, 教授, 主要研究方向为先进控制理论、机器人技术、组合导航。

E-mail: zhgl@sohu.com

李歆, 硕士研究生, 主要研究方向为机器人实时定位与建图技术。

E-mail: 576145181@qq.com