

# 改进视觉惯性里程计融合 GPS 的 无人机定位方法研究\*

周恩达 胡敏森 唐嘉宁 宋一鸣 陈伟 和雪梅  
(云南民族大学电气信息工程学院 昆明 650504)

**摘要:** 为提升无人机大范围弱纹理场景下的状态估计,提出一种改进视觉惯性里程计融合 GPS 的定位方法。首先,通过在视觉惯性里程计中加入线特征来表示环境的几何结构信息,提升位姿估计的准确性;其次,通过引入长度阈值筛选,剔除对位姿估计贡献不大的短线段,改善特征追踪的鲁棒性;最后,使用非线性优化的方式,将 GPS 测量信息和改进的视觉惯性里程计融合,校正视觉惯性里程计的累积误差。基于 EuRoC 数据集仿真实验以及应用于无人机的真实场景实验表明,相较于原算法,加入线特征算法的定位误差在仿真实验中降低了 39.14%,室内场景降低了 23.48%,室外场景降低了 33.58%。融合了 GPS 的点线特征算法相较于原算法,定位误差降低了 53.99%。

**关键词:** 同时定位与地图构建;线特征;长度阈值筛选;多传感器融合

**中图分类号:** TP391 **文献标识码:** A **国家标准学科分类代码:** 510.4050

## Research on UAV localization method based on improved visual inertial odometry and GPS

Zhou Sida Hu Minsen Tang Jianing Song Yiming Chen Wei He Xuemei  
(School of Electrical Information Engineering, Yunnan Minzu University, Kunming 650504, China)

**Abstract:** In order to improve the state estimation of UAV in a large range of weak texture scenes, an improved visual inertial odometer combined with GPS positioning method is proposed. Firstly, the geometric structure information of the environment was represented by adding line features into the visual inertial odometer to improve the accuracy of pose estimation. Secondly, by introducing length threshold screening, the short line segments that do not contribute much to pose estimation are eliminated to improve the robustness of feature tracking. Finally, the GPS measurement information is fused with the improved visual inertial odometer in a nonlinear optimization way to correct the cumulative error of the visual inertial odometer. The simulation experiment based on EuRoC dataset and the real scene experiment applied to UAV show that, compared with the original algorithm, the positioning error of the line feature algorithm is reduced by 39.14% in the simulation experiment, 23.48% in the indoor scene and 33.58% in the outdoor scene. The point and line feature algorithm integrated with GPS. The positioning error was reduced by 53.99%.

**Keywords:** simultaneous localization and mapping; line features; length threshold screening; multi-sensor fusion

### 0 引言

随着小型无人机的智能化发展,智能型无人机在抢险救灾、电力巡检、农林植保等领域发挥着越来越重要的作用。因此,无人机拥有准确的位置信息以及能够感知周围可靠的环境信息对其进行路径决策时变得尤为重要。全球定位系统(global positioning system,GPS)可以提供绝对定位信息,但是其定位精度存在一定误差,尤其在室内、楼宇

间、山林或峡谷等场景中,GPS 信号不充分导致其定位准确度下降。因此,无人机在实际应用中需要更精确的测量方式来补充定位方案。惯性测量单元(inertial measurement unit,IMU)通过解算线加速度和角速度可以得到载体的位置、姿态等信息,但其存在累积误差。同时,高精度 IMU 通常需要较高的成本。激光雷达传感器虽然获取的地图可以直接用来定位和导航,但其测量范围较小,成本高昂,且主要应用于室内。相机传感器成本低廉、小型

收稿日期:2022-10-04

\* 基金项目:国家自然科学基金(61963038)项目资助

轻便,其获取的图像含有丰富的环境信息,此外,它可以通过利用连续帧之间的一致性从图像序列中恢复相机位移和旋转。因此,越来越多学者聚焦于通过结合视觉的多传感器融合来构建同步定位与建图(simultaneous localization and mapping,SLAM)系统。

传统视觉SLAM通常以特征点法处理图像。然而,在遇到玻璃、白墙或天空等弱纹理场景中,或在相机快速运动导致图像模糊的情况下,图像中可检测的特征点数量显著降低,算法难以正常工作。为提升视觉SLAM在复杂环境下的准确性和鲁棒性,有效解决方案之一是将视觉与惯性传感器进行结合,视觉惯性SLAM数据根据融合的方法不同分为松耦合和紧耦合。当前主流的方案是紧耦合,其将视觉和IMU的原始参数进行联合优化,从而生成全局一致性的位姿估计。典型方案有MSCKF<sup>[1]</sup>、VINS-Mono<sup>[2]</sup>、ORB-SLAM3<sup>[3]</sup>等。尽管视觉惯性传感器在小范围内能够实现局部精确的位姿估计,但由于传感器噪声、传感器标定误差、设计算法近似解等因素,在大范围环境下长期运行累计漂移问题较为显著。GPS可以提供基于地球坐标起点的绝对测量值,可有效避免累计漂移的问题。然而,GPS数据通常不平滑,含有噪声,且输出频率低,难以实现高精度的六自由度估计。因此,将局部传感器与全局传感器进行优势互补是实现精确状态估计的有效方案。Angelino等<sup>[4]</sup>和Lynen等<sup>[5]</sup>提出了一种基于扩展卡尔曼滤波(extended Kalman filter, EKF)的视觉惯性融合GPS的状态估计算法,Shen等<sup>[6]</sup>提出了一种无迹卡尔曼滤波算法(unscented Kalman filter, UKF),该算法融合了视觉惯性、激光雷达和GPS等多种传感器,能够在不同环境中实现平滑一致的位姿估计。然而,基于滤波的方法对时间同步比较敏感,并且在GPS信号退化时,会影响到卡尔曼滤波对于测量噪声的不相关性假设,从而影响系统的性能。Qin等<sup>[7]</sup>采用非线性优化的方法将视觉惯性里程计的结果与GPS的结果融合,该方法比基于滤波的方法更具有优势。Cao等<sup>[8]</sup>提出了一种基于非线性优化的方法,将GPS的伪距和多普勒频移测量与视觉惯性里程计融合,其使用的原始GPS数据,是当前较为先进的方法。然而,Qin等<sup>[7]</sup>和

Cao等<sup>[8]</sup>皆没有考虑到弱纹理场景中的特征跟踪问题。因此,解决弱纹理环境中的特征跟踪问题的方案之一是在SLAM的前端添加更多的几何结构信息——线特征。线特征具有良好的光照和旋转不变性,在一些特征点稀疏的室内人造场景,往往具有丰富的线性结构。Lu等<sup>[9]</sup>从RGB-D图像数据中提取点和线特征,分析它们的测量不确定性,并使用极大似然估计相机运动,实验证明,点线特征比点特征具有更小的运动不确定性。Vakhitov等<sup>[10]</sup>提出了一种将3D点到2D线对应关系引入多点视角(perspective-n-point, PnP)公式的新方法,允许同时处理点和线,实验表明,与仅基于点或仅基于线的解决方案相比,所提出的方法有着显著的性能提升。然而传统的线特征是专为场景结构设计,不适用于位姿估计问题。在严纪强等<sup>[11]</sup>描述中,基于点线的SLAM系统中存在以下问题,1)线特征提取存在场景边缘长线提取后断裂成数根短线特征的情况,这些短线特征容易在下一帧中变化甚至消失,导致SLAM在线特征匹配时容易出现误匹配问题,降低系统精度;2)在设计长度阈值筛选策略时,根据经验确定长度因子常数,不适用于特征线段数量变化大的情况,导致计算量增加。

针对上述问题,本文在VINS-Mono的基础上,对应用于视觉惯性SLAM系统的线段特征提取(line segment detector, LSD)<sup>[12]</sup>算法进行改进,提出基于锚点的主动分组策略,使弱梯度区域的特征线段相对完整,不易断裂,从而减少线特征的误匹配率,提高系统精度;在此基础上,提出一种改进的长度阈值筛选策略,筛选出有利于位姿估计的长线段,剔除对位姿估计贡献不大的不稳定短线段。从而在降低线特征计算量大的同时,提高SLAM系统特征跟踪的鲁棒性。在此基础上,将改进后的视觉惯性里程计结果与GPS的测量结果融合,实现局部精确,全局无漂移的无人机六自由度位姿估计。

## 1 系统框架

本文是在VINS-Mono的基础上进行改进,设计的系统整体框架如图1所示,系统采用相机和IMU紧耦合的策

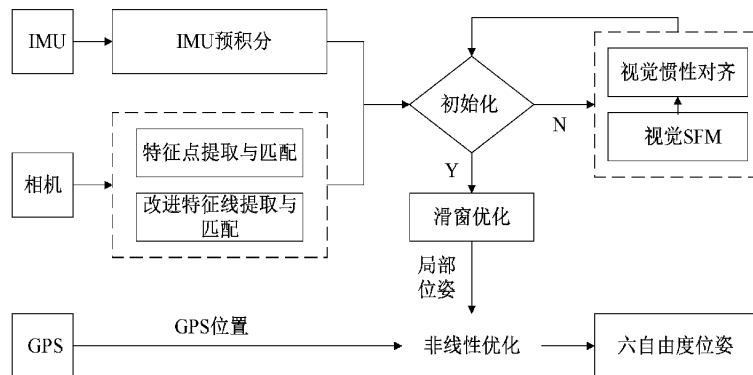


图1 系统整体框架

略以非线性优化的方式联合估计相机的自我运动,本系统的研究重点是在系统中添加了改进的线特征提取与匹配算法,并且以基于非线性优化的方式将改进后的视觉惯性里程计结果与 GPS 进行融合。

首先,系统输入的是 IMU 的加速度和角速度信息,以及相机的单目图像信息。由于 IMU 的频率高于图像的频率,因此需要对两帧图像间的 IMU 信息进行预积分。图像的特征点提取采用了 Harris 等<sup>[13]</sup>的角点算法,通过机器人操作系统(the robot operating system,ROS)的回调函数,对相邻图像帧使用金字塔光流法(Kanade Lucas Tomasi,KLT)跟踪。添加的线特征提取算法,采用了改进的 LSD 算法进行特征线提取并使用 LBD<sup>[14]</sup>描述符进行匹配,并采用普吕克坐标表示。接着,将纯视觉运动恢复结构(structure from motion,SFM)估计的相机位姿和特征点的深度信息与 IMU 积分进行联合校准,对齐。最后,通过最小化目标函数(先验残差、IMU 的测量残差、点和线的重投影误差)优化出最终的相机位姿。最后,将改进后的视觉惯性里程计输出的局部位姿和 GPS 观测数据融合,得到局部一致,全局无漂移的六自由度位姿。

## 2 改进点线特征 SLAM 算法

当前基于点线特征的 SLAM 系统中常用的线特征提取算法一般是调用的是 opencv 中的传统 LSD 算法,如 He 等<sup>[15]</sup>和 Wen 等<sup>[16]</sup>。传统 LSD 算法能够在线性时间内得到亚像素极准确度,但是它为场景结构而设计的,不适用于位姿估计问题,其在图像中会检测出大量的短线特征,

然而有些短线是由于不稳定像素点断开的。针对断线问题,赵良玉等<sup>[17]</sup>对传统算法进行改进,其根据相邻线段的角度、端点距离进行分组,然后在比较短线间的端点距离和线段长度,以此来判断短线间是否需要合并。但是占用的计算资源较多。

### 2.1 改进的 LSD 特征提取算法

针对线特征提取存在场景边缘长线提取后断裂成数根短线特征导致线特征误匹配的问题,本文受 Zhang 等<sup>[18]</sup>启发,对线特征的分组策略进行了针对性改进。首先对输入图像进行灰度化,并计算梯度幅度和梯度方向,计算过程与 LSD 算法相同,其次,提出一种锚点概念,锚点解释为提取线段通过的概率高的像素,得到的锚点图作用是通过丢弃大部分无用像素同时确认线段提取中的重要像素来简化图像,然后根据线段的几何形状和梯度方向对齐,将锚点主动分组为线段。最后根据梯度幅度分布和线几何形状对线段进行验证和细化。

如图 2 所示,线段提取图由锚点和线段组成,线段是可能会断裂的边缘主轴,线段由锚点拟合,小箭头的方向和长度表示每个像素的方向和大小,在弱梯度区域不稳定的像素可能会导致锚点不连续,从而会发生主轴断裂。在本文应用于 SLAM 的线特征算法中,采取主动分组策略,直接以主轴和上一个已经分组的锚点(设为 PA)计算下一个待分组的候选锚点(设为 NA)。添加新锚点后会更新主轴,主轴如式(1)表示。在主轴的引导下,可以扩展候选者 NA,从而可以跳过不稳定的像素,图 2(b)~(c)所示。当分组过程完成时,线段由主轴和锚点的末端拟合。

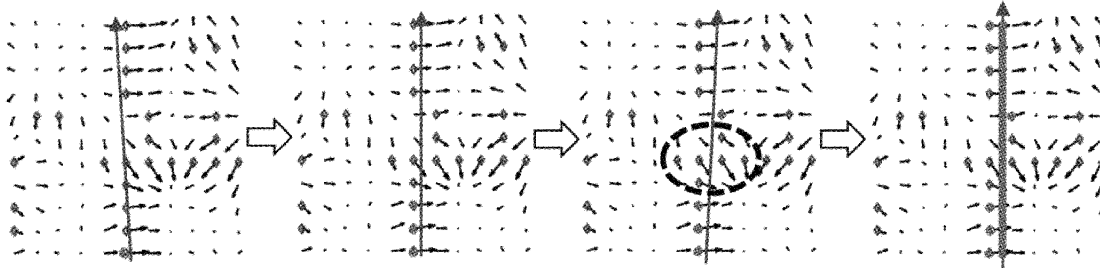


图 2 线段提取区域锚点分组图

主轴的表达式如下:

$$l = \{(x, y) \mid y = \tan(\theta_l)(x - x_c) + y_c\} \quad (1)$$

其中,  $c = (x_c, y_c)$  表示主轴中心,  $x_c = \frac{1}{n} \sum_{i=1}^n x_i$ ,

$y_c = \frac{1}{n} \sum_{i=1}^n y_i$ ,  $(x_i, y_i)$  是每个锚点坐标,  $\theta_l$  表示组角度,

$\theta_l$  可以由惯性矩阵  $M$  的特征值确定,表达式如下:

$$\theta_l = \begin{cases} \arctan(\lambda - M_{11})/M_{12}, & M_{11} > M_{22} \\ \arctan(M_{12}/(\lambda - M_{22})), & M_{11} \leq M_{22} \end{cases} \quad (2)$$

其中,  $M_{11} = \sum_{i=1}^n (x_i - x_c)^2$ ,  $M_{22} = \sum_{i=1}^n (y_i - y_c)^2$ ,

$M_{12} = M_{21} = \sum_{i=1}^n (x_i - x_c)(y_i - y_c)$  以及最小特征值:

$$\lambda = \frac{1}{2} \{M_{11} + M_{22} - \sqrt{(M_{11} - M_{22})^2 + 4M_{12}M_{21}}\} \quad (3)$$

在定义了主轴和锚点的概念后,进行主动分组以拟合线段,算法步骤如下:

步骤 1) 定义主轴的方向  $dir_{PA} = (\cos(\theta_l), \sin(\theta_l))$ , 已分组的锚点 PA 的坐标定义为  $X_{PA} = (x_{PA}, y_{PA})$ , 计算最近的候选锚点 NA 坐标,  $X_{NA} = \text{round}(X_{PA} + dir_{PA})$ 。

步骤 2) 验证候选锚点。首先候选锚点的方向与已分组锚点群的平均方向之间的差异应该小于  $\pi/8$ 。其次,候

选锚点与主轴的正交距离应在 1 pixel 以内。如果一个候选锚点满足这两个约束条件,则算法进入步骤 3)。

步骤 3) 展开候选锚点。为了减少不稳定像素引起的断裂,当最近的 3 个候选锚点不能满足步骤 2) 中的验证时,候选锚点通过在线几何约束下的跳跃进行扩展,这是通过在方程中添加一个比率来实现的,公式如下:

$$X_{NA} = \text{round}(X_{PA} + j \times \text{dir}_{PA}), j \in N^*, j \leq t \quad (4)$$

$j$  从 2 增加到  $t$ , 直到一个候选锚点 NA 满足步骤 2)

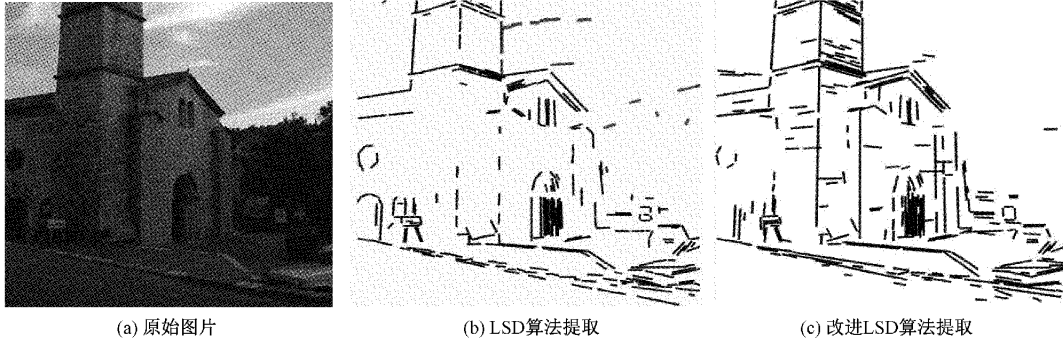


图3 算法改进测试效果对比

### 2.2 长度阈值筛选策略

LSD 内部参数如金字塔数、比例因子等使用 LSD 算法调整的隐藏参数,以加快提取的速度。为了提高线特征提取质量,提出一种长度阈值筛选策略,剔除对位姿估计不稳定的短线。公式如下:

$$\left\{ \begin{array}{l} \text{len}_{\min} \geq \alpha \times [\max(W_l, H_l)] \\ \alpha = 1 - e^{-0.001k} \end{array} \right\}, i \in \{1, 2, \dots, k\} \quad (5)$$

式中:  $\text{len}_{\min}$  是最短线段长度,  $\alpha$  表示长度因子,  $W_l$  是图像帧宽度,  $H_l$  是图像帧高度,  $k$  表示当前帧提取线特征的个数。

Fu 等<sup>[19]</sup> 使用类似的策略来提高线段的质量,但仅凭经验选择固定的长度因子常数,不适用于线段数量变化很大的情况。如果长度因子选择的值较大,则线段数较少时有效跟踪线段数不足。如果选取的值较小,当线段较多时,无法有效消除多余的短线段。在本文提出的长度阈值筛选公式中,线段的数量通过公式  $\alpha$  代替长度因子常数。当线段数量较少时,增加长度阈值以获得更多的线段用于帧间跟踪。当线段数量较多时,降低长度阈值可以减少短线段对位姿估计稳定性的影响,从而提高长度阈值筛选策略在不同线段数量下的适应性,使 SLAM 系统减少计算量的同时,提高系统的鲁棒性。

### 3 视觉惯性里程计融合 GPS

为消除视觉惯性里程计长期运行下的累积误差,将视觉惯性里程计的结果与 GPS 进行融合。首先将 GPS 经纬度坐标系转化笛卡尔坐标系,其起点为第一帧获取的 GPS 坐标,视觉惯性里程计和 GPS 坐标转换公式如下:

中的验证。候选锚点 NA 是显著边缘像素,  $t$  设置为 10, 这是一个宽松的阈值,可以跳过大多数噪声干扰。

步骤 4) 更新式(1) 主轴,并将加入的锚点设置为已分组锚点 PA,然后算法重新进入步骤 1)。

如图 3 所示,图 3(b) 是传统 LSD 算法的提取效果,传统算法在弱梯度边缘区域提取后存在很多断线。图 3(c) 是改进后的 LSD 算法,其在墙面,马路等弱梯度区域能够提取到相对完整的长线。与传统 LSD 算法相比,改进的算法能够提取更多有利于位姿估计的长线段。

$$\min F(R_{R^v}, T_{R^v}) = \sum_i^N P_i^s - (R_{R^v} \cdot P_i^v + T_{R^v}) \quad (6)$$

其中,  $R_{R^v}$  和  $T_{R^v}$  是两个待优化的参数,分别是视觉惯性里程计坐标系到 GPS 笛卡尔坐标系的旋转和平移矩阵,  $N$  是可以用来融合的 GPS 坐标数量,  $P_i^s$  是无人机在  $i$  时刻的 GPS 坐标,  $P_i^v$  是无人机在  $i$  时刻的视觉惯性里程计坐标。

构造非线性优化目标函数如下所示:

$$G(\chi) = \chi - \sum_{i=0}^N \sum_{j=0}^M E^v(i, j, k_i, k_j, t^v, q^v) - w^s \sum_{i=0}^N E^s(i) \quad (7)$$

式中:  $\chi = [\xi_1, \dots, \xi_N]$ ,  $\xi_i$  是无人机在  $i$  时刻优化后的位姿,是公式的优化变量,  $M$  是视觉惯性里程计的位姿数量。因为视觉惯性里程计存在累积误差,但短时间内精度较高,所以视觉惯性里程计观测值只信任  $i \sim j$  时刻的位姿变化量,  $E^v(i, j, k_i, k_j)$  是由视觉惯性里程计得到的  $i \sim j$  时刻的位姿观测变化量与无人机真实位姿变化量的残差。  $k_i$  和  $k_j$  分别是视觉惯性里程计在  $i$  和  $j$  时刻得到的位姿,  $t^v$  和  $q^v$  分别是位移和旋转之间的权重分配比例,根据经验  $t^v$  设置成 0.1,  $q^v$  设置成 0.01。

$$w^s = \frac{1}{\kappa} \cdot \frac{M}{N} \quad (8)$$

其中,  $w^s$  是 GPS 融合的权重,  $\kappa$  是 GPS 定位的几何精度因子(geometric dilution of precision, GDOP),是衡量定位精度的重要参数,其值越大,定位精度越差,通常由 GPS 信号接收器给出。  $E^s(i)$  是 GPS 在  $i$  时刻观测的位置与无人机真实位姿的残差。没有 GPS 信号时,  $E^s(i) = 0$ 。

### 4 实验结果与分析

#### 4.1 仿真实验环境和数据集介绍

本文实验的设备是主频 1.99 Hz,i7-8550U CPU、内存 8 GB 的笔记本电脑,实验平台是 Ubuntu16.04 操作系统,ROS 机器人操作系统的版本是 kinetic。实验的数据集是苏黎世联邦理工学院录制的 EuRoC<sup>[20]</sup> 数据集,场景在室内,并且具有从低速运动到高速运动情况以及重复纹理的人造复杂场景。另外,EuRoC 数据集提供了地面真值轨迹。本文选用 EuRoC 数据集里的 8 个场景,其中,MH\_01、MH\_02 和 V1\_01 是简单环境下的场景,MH\_03 和 V2\_02 是中等环境下的场景,MH\_04、MH\_05 和 V1\_03 是复杂环境下的场景。

对于算法的轨迹精度,采用 SLAM 轨迹评估 (evaluation of odometry and SLAM,EVO)精度评估工具进行评估。EVO 可以通过对真值与估计位姿进行时间戳对齐来计算估计轨迹和真值轨迹的绝对位姿误差 (absolute pose error,APE),通常由均方根误差 (root mean square error,RMSE)来统计绝对位姿误差,如式(9)所示。

$$APE_{\text{trans}} = \sqrt{\frac{1}{N} \sum_{i=1}^N \| \text{trans}(T_{k,i}^{-1} T_{e,i}) \|^2} \quad (9)$$

其中,trans 表示平移误差,  $T_{k,i}$  表示真实位姿,  $T_{e,i}$  表示估计位姿。

#### 4.2 数据集场景仿真实验结果与分析

表 1 比较了本文改进的算法、原算法 VINS-Mono 以及基于点线特征同类型的算法 PL-VIO<sup>[15]</sup> 和 PL-VINS<sup>[19]</sup> 在 EuRoC 数据集上的均方根误差,以上算法皆是在没有闭环检测的情况下进行评测。从表 1 可以看出,与 VINS-Mono 基于点特征的 SLAM 算法相比,本文基于点线特征的改进 SLAM 较优,定位误差与 VINS-Mono 相比平均降低了 39.14%,与同类型的基于点线特征的 PL-VIO 相比平均降低了 32.31%,与 PL-VINS 相比平均降低了 4.54%。实验证明了相较于传统基于点特征的视觉惯性 SLAM 算法,加入线特征后系统精度具有不错的定位精度提升,此外,相较于同类型其他算法,本文算法在没有回环检测的情况下依然一定的精度提升。VINS-loop 是在原算法中加入了回环检测进行的误差实验,可以看出,本文的改进算法相较于有回环检测的原算法,也有一定的精度优势。

表 1 本文算法与其他算法的均方根误差

数据集序列	VINS-Mono	VINS-loop	PL-VIO	PL-VINS	本文算法
MH_01_easy	0.213 600	0.132 763	0.165 691	0.144 752	0.120 021
MH_02_easy	0.233 520	0.166 334	0.193 785	0.130 378	0.129 738
MH_03_medium	0.265 210	0.142 485	0.325 684	0.165 453	0.147 514
MH_04_difficult	0.382 511	0.221 846	0.391 475	0.183 220	0.185 918
MH_05_difficult	0.341 728	0.184 529	0.301 265	0.169 120	0.167 432
V1_01_easy	0.152 721	0.117 976	0.143 288	0.137 701	0.132 029
V2_02_medium	0.173 921	0.131 578	0.162 479	0.133 472	0.130 251
V1_03_difficult	0.248 369	0.152 482	0.184 526	0.152 821	0.150 041

本文算法轨迹相对于真值轨迹的误差可视化比较如

图 4 所示,展示了本文算法在 MH\_01\_easy 场景中的可视

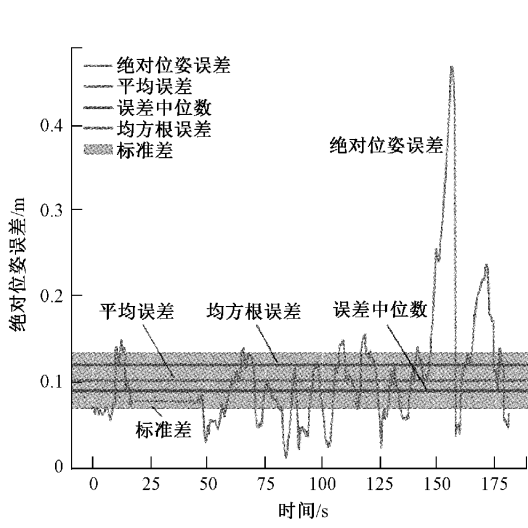
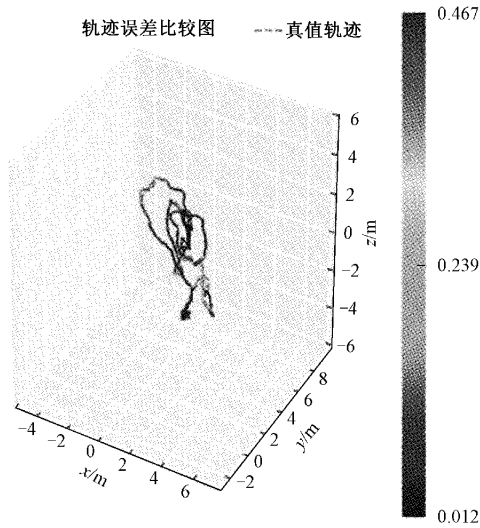


图 4 本文算法轨迹相对于真值轨迹的误差比较图

化轨迹和误差图,图4(a)展示了本文算法轨迹相对于数据集真值轨迹的误差,对齐方式采用的是 Sim(3) Umeyama<sup>[21]</sup>对齐,其可以同时处理旋转平移和尺度。颜色由红到蓝,误差逐渐减小。由图4(a)可以看出,本文算法在相机发生旋转时也能保持较好的精度。图4(b)展示了本文算法的均方根误差、平均误差、误差中位数、标准差等可视化图,横坐标为算法运行的时间,可以看出,在整个算法运行的时间内,均方根误差总体保持在0.1附近。

### 4.3 真实物理实验结果与分析

#### 1) 室内场景实验结果与分析

本文的实验硬件平台是一架自主组装的小型四旋翼无人机,实验硬件平台如图5所示,其由一个 Intel RealSense D435i 双目深度相机,一个 pixhawk4 飞控以及一台大疆妙算 manifold-2c 机载计算机等硬件组成,本文的算法在机载计算机中实时运行,算法估计出的飞行轨迹和位姿数据通过远程方式可视化显示在地面端电脑。



图5 实验硬件平台图

本文的室内实验场景如图6所示,飞行轨迹和起点已标注,飞行总长度约为14.3 m。实验场景中布置了由16个摄像头组成的 OptiTrack 室内定位系统,其可以精确捕捉无人机的飞行轨迹,采集无人机的六自由度位姿数据,误差可以达到毫米级精度,如图7所示。因此,本文将 OptiTrack 室内定位系统采集的位姿当作地面真值(Ground\_truth),以此来定量评估本文算法的精度。

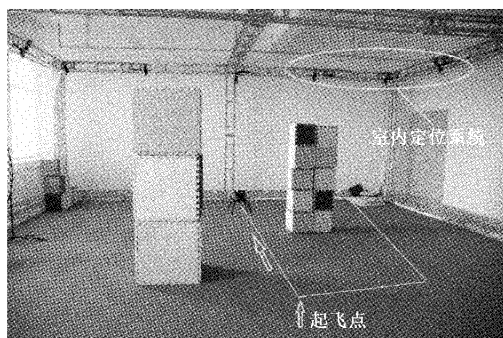


图6 室内实验场景图

原算法与本文改进算法的估计轨迹以及其与地面真值轨迹可视化比较如图8所示。从图8可以初步看出,原

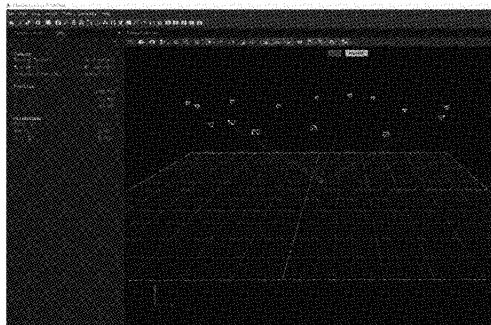


图7 室内定位系统软件示意图

算法的估计轨迹在相机姿态发生变化时容易发生漂移,本文算法的估计轨迹在相机姿态变化时能更好的保持定位精度,更接近真实轨迹。表2是原算法 VINS-Mono 与本文改进的算法的不同误差。从表2可以得出,在室内环境下,与基于点特征的原算法相比,本文加入线特征的算法均方根误差降低了23.48%,无论是平均误差或者最大误差,本文改进的算法相比原算法都有一定降低。

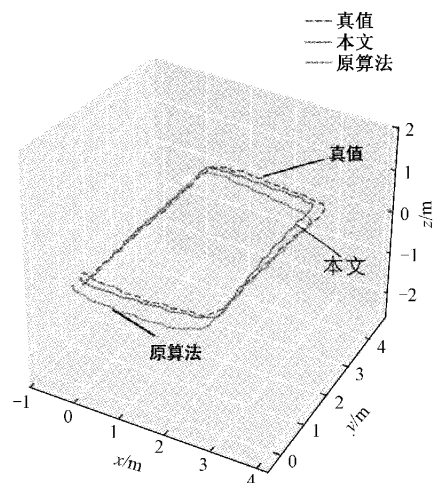


图8 原算法与本文改进算法的室内轨迹比较图

表2 原算法 VINS-Mono 与本文改进算法的各种误差表  
m

算法	均方根误差 (RMSE)	平均误差 (mean)	最大误差 (max)	标准差 (std)
VINS-Mono	0.329 371	0.312 843	0.512 475	0.120 797
本文算法	0.228 559	0.211 314	0.323 142	0.094 798

#### 2) 室外场景实验结果与分析

在室外的实验中,本文使用手持相机传感器、GPS 传感器、实时动态载波相位差分(real time kinematic, RTK)传感器和笔记本组成的套件运行定位算法。其中,高精度 RTK 传感器通过载波相位差分技术能获取实时三维定位结果,可以达到厘米级定位精度,本文将其录制的位置数据当作地面真值,以此来定量评估本文的算法定位精度,

高精度 RTK 传感器硬件如图 9 所示,其中图 9(a)是 RTK 的天空端,图 9(b)是 RTK 的地面端。

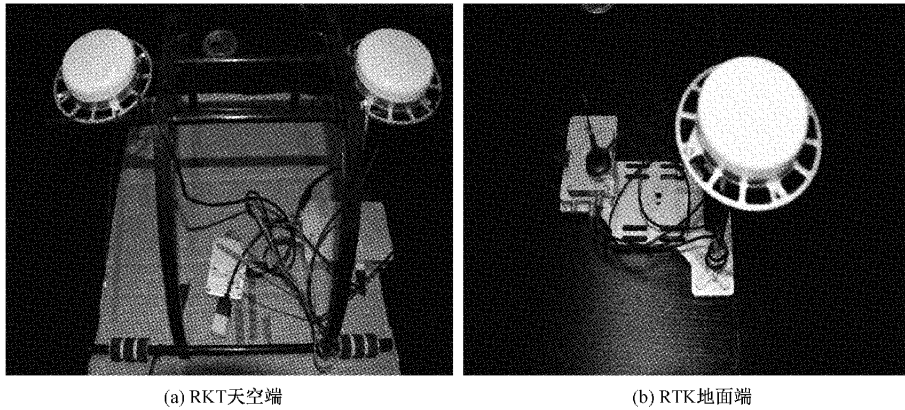


图 9 RTK 传感器硬件图

本文的室外实验环境如图 10 和 11 所示,算法运行的真实轨迹为图 10 中的矩形轨迹和图 11 中的圆形轨迹,其中起点为标注位置,矩形轨迹总长度在 150 m 左右,圆形轨迹总长度在 170 m 左右。

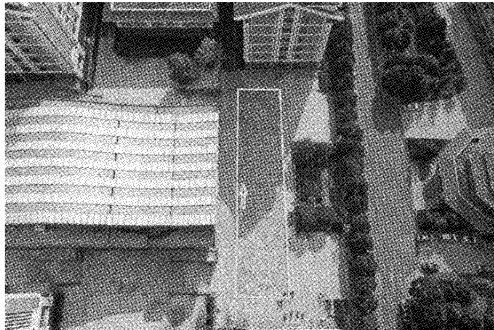


图 10 室外实验场景 1 图

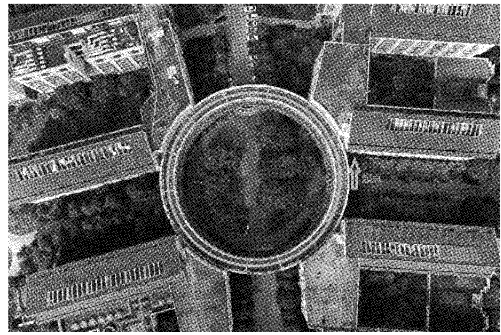


图 11 室外实验场景 2 图

原算法与改进算法的室外场景 1 轨迹比较图如图 12(a)所示,场景 2 轨迹比较图如图 12(b)所示,两图中,轨迹 3 为原算法 VINS-Mono 估计的轨迹,轨迹 1 为加入了改进线特征的算法估计轨迹,轨迹 4 为加入了改进线特

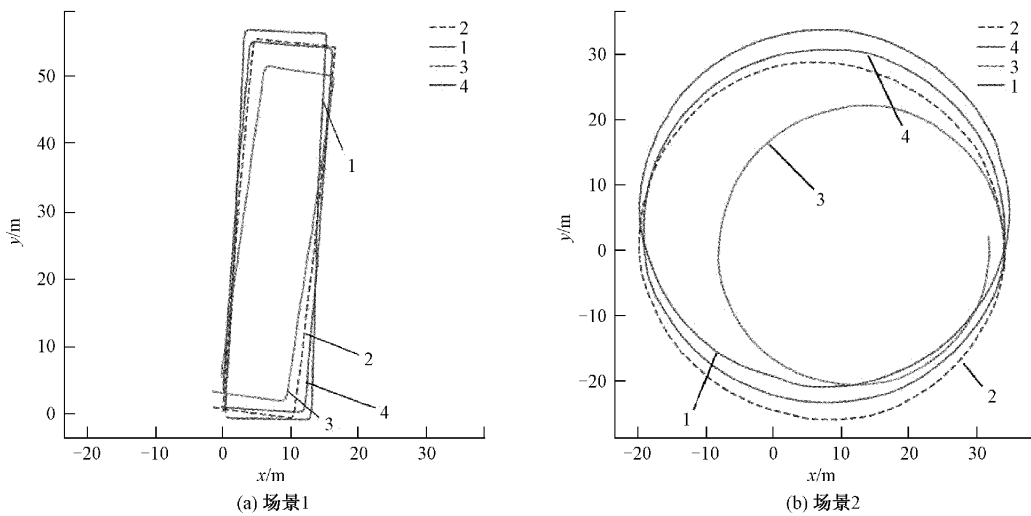


图 12 原算法与改进算法的室外轨迹比较图

征并且融合了 GPS 观测数据的算法估计轨迹,轨迹 2 为通过 RTK 传感器观测数据绘制的真值轨迹。由于绘图坐标

系的原因,图中轨迹略微倾斜。从图中可以看出,加入线特征的改进算法轨迹相较于原算法更接近真实轨迹,在加

入线特征的基础上,融合了GPS观测数据的算法轨迹,相较于前两种算法,更加接近真实轨迹,定位误差更小。表3是3种算法均方根误差比较表,从表中可以看出,在室外环境下,加入线特征的改进算法,相较于原算法,均方根误差平均降低了33.58%。实验表明,相较于室内丰富纹理场景,加入线特征的算法精度在室外空旷环境中能提升更多。在加入线特征的基础上,融合了GPS观测数据后的改进算法,相较于原算法,均方根误差降低了53.99%。实验表明,在室外环境下,融合了GPS传感器数据的视觉惯性里程计算法定位精度有进一步提升。

表3 原算法与本文改进算法的均方根误差

均方根 误差	VINS-Mono	改进 VINS	改进 VINS-GPS
场景1均方根误差	2.346 170	1.435 761	1.283 542
场景2均方根误差	6.403 646	4.375 776	2.742 355

## 5 结 论

针对弱纹理非结构化环境中,基于点特征的纯视觉SLAM容易出现特征点丢失而引起定位精度下降的问题,以及在长期大范围环境下运行存在累积漂移的问题,本文首先提出了一种基于点线特征的改进视觉惯性SLAM算法。采用基于锚点的主动分组策略对仅仅适用于结构化场景的LSD线特征算法进行改进,提出了一种的长度阈值筛选策略,提升了算法精度。相较于VINS-Mono,定位误差在仿真实验中降低了39.14%,在室内场景降低了23.48%,在室外场景降低了33.58%。最后,利用全局传感器GPS约束校正视觉惯性里程计的局部位姿,使定位精度进一步提升,相较于原算法,定位误差降低了53.99%。

## 参考文献

- [1] 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, 2007, 2(6): 3565-3572.
- [2] 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.
- [3] CAMPOS C, ELVIRA R, RODRÍGUEZ J, 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.
- [4] ANGELINO C V, BARANIELLO V R, CICALA L. UAV position and attitude estimation using IMU, GNSS and camera [C]. 2012 15th International Conference on Information Fusion, IEEE, 2012, 10(7): 735-742.
- [5] LYNEN S, ACHELNIK M W, WEISS S, et al. A robust and modular multi-sensor fusion approach applied to mav navigation [C]. 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2013, 13(9): 3923-3929.
- [6] 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, 14(7): 4974-4981.
- [7] QIN T, CAO S, PAN J, et al. A general optimization-based framework for global pose estimation with multiple sensors [J]. ArXiv Preprint, 2019, ArXiv: 1901.03638.
- [8] CAO S, LU X, SHEN S. Gvins: Tightly coupled gnss-visual-inertial fusion for smooth and consistent state estimation [J]. ArXiv Preprint, 2021, ArXiv: 2103.07899.
- [9] LU Y, SONG D. Robust RGB-D odometry using point and line features [C]. Proceedings of the IEEE International Conference on Computer Vision, 2015.
- [10] VAKHITOV A, FUNKE J, MORENO-NOGUER F. Accurate and linear time pose estimation from points and lines [C]. European Conference on Computer Vision. Springer, Cham, 2016, 17(2): 583-599.
- [11] 严纪强, 吴一琳, 胡剑凌. 基于点线特征的改进视觉惯性SLAM方法 [J]. 电视技术, 2022, 46(3): 3-10.
- [12] VON GIOI R G, JAKUBOWICZ J, MOREL J M, et al. LSD: A line segment detector [J]. Image Processing on Line, 2012, 2(1): 35-55.
- [13] HARRIS C, STEPHENS M. A combined corner and edge detector [C]. Alvey Vision Conference, 1988, 15(50): 5244.
- [14] 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.
- [15] 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.
- [16] WEN H, TIAN J, LI D. PLS-VIO: Stereo vision-inertial odometry based on point and line features [C]. 2020 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS),



- IEEE, 2020.
- [17] 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.
- [18] ZHANG Y, WEI D, LI Y. AG3line: Active grouping and geometry-gradient combined validation for fast line segment extraction[J]. Pattern Recognition, 2021, 113(4): 107834.
- [19] FU Q, WANG J, YU H, et al. PL-VINS: Real-time monocular visual-inertial SLAM with point and line[J]. ArXiv Preprint, 2020; ArXiv: 2009.07462.
- [20] 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.
- [21] UMEYAMA S. Least-squares estimation of transformation parameters between two point patterns [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1991, 13(4): 376-380.

#### 作者简介

胡敏森(通信作者),硕士研究生,主要研究方向为无人系统导航与控制。

E-mail:2682826764@qq.com

周思达,博士,教授,主要研究方向为无人系统导航与控制。