



中国科技期刊卓越行动计划项目入选期刊

# 控制与决策

CONTROL AND DECISION



## 双目视觉辅助的激光惯导SLAM算法

刘辉, 张雪波, 李如意, 苑晶

引用本文:

刘辉,张雪波,李如意,苑晶. 双目视觉辅助的激光惯导SLAM算法[J]. 控制与决策, 2024, 39(6): 1787–1800.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2023.0158>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 无人系统视觉SLAM技术发展现状简析

A survey of visual SLAM in unmanned systems

控制与决策. 2021, 36(3): 513–522 <https://doi.org/10.13195/j.kzyjc.2019.1149>

#### 基于多层级特征的机械臂单阶段抓取位姿检测

Single-stage grasp pose detection of manipulator based on multi-level features

控制与决策. 2021, 36(8): 1815–1824 <https://doi.org/10.13195/j.kzyjc.2019.1840>

#### 基于滚动时域粒子群优化的视频去雾算法

Receding horizon particle swarm optimization based video defogging algorithm

控制与决策. 2021, 36(9): 2218–2224 <https://doi.org/10.13195/j.kzyjc.2019.1183>

#### 基于双边分解与 $L_1$ 暗通道的战场图像增强算法

A battlefield image enhancement algorithm based on bilateral decomposition and  $L_1$  dark channel prior

控制与决策. 2021, 36(5): 1165–1172 <https://doi.org/10.13195/j.kzyjc.2019.1213>

#### 一种基于稀疏系数匹配学习的图像去雾算法

An image dehazing method based on learning framework with sparse coefficient matching

控制与决策. 2020, 35(11): 2797–2802 <https://doi.org/10.13195/j.kzyjc.2018.1764>

# 双目视觉辅助的激光惯导SLAM算法

刘 辉<sup>1,2</sup>, 张雪波<sup>1,2†</sup>, 李如意<sup>1,2</sup>, 苑 晶<sup>1,2</sup>

(1. 南开大学 机器人与信息自动化研究所, 天津 300350; 2. 天津市智能机器人技术重点实验室, 天津 300350)

**摘要:** 激光同步定位与地图构建 (simultaneous localization and mapping, SLAM) 算法在位姿估计和构建环境地图时依赖环境结构特征信息, 在结构特征缺乏的场景下, 此类算法的位姿估计精度与鲁棒性将下降甚至运行失败. 对此, 结合惯性测量单元 (inertial measurement unit, IMU) 不受环境约束、相机依赖视觉纹理的特点, 提出一种双目视觉辅助的激光惯导 SLAM 算法, 以解决纯激光 SLAM 算法在环境结构特征缺乏时的退化问题. 即采用双目视觉惯导里程计算法为激光扫描匹配模块提供视觉先验位姿, 并进一步兼顾视觉约束与激光结构特征约束进行联合位姿估计. 此外, 提出一种互补滤波算法与因子图优化求解的组合策略, 完成激光里程计参考系与惯性参考系对准, 并基于因子图将激光位姿与 IMU 数据融合以约束 IMU 偏置, 在视觉里程计失效的情况下为激光扫描匹配提供候补的相对位姿预测. 为进一步提高全局轨迹估计精度, 提出基于迭代最近点匹配算法 (iterative closest point, ICP) 与基于图像特征匹配算法融合的混合闭环检测策略, 利用 6 自由度位姿图优化方法显著降低里程计漂移误差并构建环境地图. 最后, 将所提出方法在公开与自制数据集上进行实验验证, 并与主流开源的 SLAM 算法进行对比. 实验结果表明, 所提出算法可以在结构特征缺乏环境下稳定运行, 并且相较于对比算法具有更高的位姿估计精度和算法鲁棒性.

**关键词:** 激光 SLAM; 闭环检测; 结构特征缺乏环境; 位姿图优化; 视觉惯导里程计

**中图分类号:** TP273 **文献标志码:** A

**DOI:** 10.13195/j.kzyjc.2023.0158

**引用格式:** 刘辉, 张雪波, 李如意, 等. 双目视觉辅助的激光惯导 SLAM 算法 [J]. 控制与决策, 2024, 39(6): 1787-1800.

## Stereo vision aided lidar-inertial SLAM

LIU Hui<sup>1,2</sup>, ZHANG Xue-bo<sup>1,2†</sup>, LI Ru-yi<sup>1,2</sup>, YUAN Jing<sup>1,2</sup>

(1. Institute of Robotics and Automatic Information Systems, Nankai University, Tianjin 300350, China; 2. Key Laboratory of Intelligent Robotics, Tianjin 300350, China)

**Abstract:** Lidar-based simultaneous localization and mapping (SLAM) algorithms rely on the structural features to estimate the poses and construct the environmental map. Therefore, the pose estimation accuracy and robustness of such algorithms will decline or even fail in structureless environments. To solve this problem, combined with the characteristics that an inertial measurement unit (IMU) is not constrained by the environment and the camera depends on visual texture, a stereo vision aided lidar-inertial SLAM algorithm is proposed to solve the degradation problem of a pure lidar SLAM algorithm in structureless environments. This paper uses a stereo visual-inertial odometry algorithm to provide a priori pose for the lidar scan matching module, and further combines the visual pose constraint with the lidar structure feature constraints for joint pose estimation. In addition, a combined strategy of a complementary filtering algorithm and a factor graph optimization algorithm is proposed to align the lidar odometry reference frame with the inertial frame. Based on factor graph optimization, the lidar pose and IMU data are fused to constrain the IMU bias, so as to provide an alternative relative pose prediction for lidar scanning matching in the case of visual odometry failure. In addition, in order to further improve the accuracy of global trajectory estimation, this paper proposes a hybrid loop-closure detection strategy based on an iterative closest point matching algorithm (ICP) and an image feature matching algorithm, and the six degree of freedom pose graph optimization method is used to significantly reduce the odometry drift error and construct a global consistency environment map. Finally, experiments are carried out on public and self-made data sets, and the proposed method is compared with the mainstream open source SLAM algorithms. The experimental results show that the proposed algorithm can work stably in structureless environments, and achieve higher accuracy and robustness than the comparison algorithm.

**Keywords:** lidar-based SLAM; loop closure detection; structureless environment; pose graph optimization; visual-inertial odometry

收稿日期: 2023-02-13; 录用日期: 2023-07-04.

基金项目: 国家自然科学基金项目 (62073178, 62003176); 工业物联网与网络化控制教育部重点实验室开放基金项目; 天津市自然科学基金项目 (22JCZDJC00810); 天津市杰出青年基金项目 (20JCJQJC00140); 中央高校基本科研业务费专项资金项目 (ZB23003422); 中国博士后科学基金项目 (2020M670628).

责任编辑: 夏元清.

†通讯作者. E-mail: zhangxuebo@nankai.edu.cn.

## 0 引言

同步定位与地图构建(simultaneous localization and mapping, SLAM)是机器人领域的热点研究方向之一,其要求机器人通过自身携带的传感器对自身状态/位姿进行估计并同时构建环境地图<sup>[1-4]</sup>. 其中,激光雷达作为传感器,具有直接测量深度、不易受光照影响等优势,因此,激光SLAM/里程计是近年来的重要研究方向之一,并且发展出诸多有效的解决方案<sup>[5-10]</sup>. 相比于纯激光方案<sup>[5]</sup>,激光雷达与惯性测量单元(IMU)融合具备更为优良的性能. Zhang等<sup>[6]</sup>提出一种实时激光里程计与建图方法(lidar odometry and mapping, LOAM),其使用IMU数据辅助去除点云运动畸变并为激光扫描匹配部分提供运动先验,再基于点线误差与点面误差配准点云以估计载体运动. Shan等<sup>[7]</sup>采用相同的IMU数据融合策略辅助激光扫描匹配,针对无人车平台加入地面优化并采用两步优化策略以降低计算量;此外,加入后端位姿图优化策略,提出适用于地面车的轻量化激光SLAM方案LeGo-LOAM. 不同于上述激光雷达与IMU松耦合方案,紧耦合方案将激光雷达与IMU的测量数据进行联合优化,完成状态估计以获取更高精度的位姿估计. LIO-mapping (LIOM)<sup>[8]</sup>基于滑动窗口融合激光雷达与IMU测量数据以粗略估计载体位姿,而后采用带旋转约束的建图策略优化位姿并构建地图,可以得到与LOAM方案接近或更高的精度,但无法实时运行. 文献[9]基于迭代误差状态卡尔曼滤波器(iterated error-state Kalman filter, I-ESKF)提出实时激光-IMU里程计LINS,其位姿估计精度高于纯激光里程计. 与上述传感器观测值直接紧耦合策略不同,LIO-SAM<sup>[10]</sup>将激光扫描匹配的估计位姿与IMU数据融合以估计IMU状态量,并基于ICP匹配策略检测闭环,能够以较高的精度实时构建点云地图,但需要使用9轴IMU提供的姿态信息对齐激光里程计参考系与惯性参考系.

上述激光-IMU里程计/SLAM解决方案在一般情况下能够获得精度较高的位姿与地图,但在环境几何结构缺乏或结构相似的挑战环境下,例如在开阔区域或长直隧道中,由于激光雷达作为环境感知传感器无法获取足够的几何特征约束,此类算法的定位精度将显著降低,甚至运行失败.

与激光雷达相比,视觉传感器可以依赖环境中的纹理特征对自身状态进行估计,因此,在环境几何结构特征缺乏或结构相似环境下可以使用视觉里程计/SLAM进行辅助<sup>[11-14]</sup>. Zhang等<sup>[11]</sup>将采用激

光雷达辅助的视觉里程计作为LOAM的前端提出VLOAM算法框架,该算法通过高频的视觉帧间位姿估计去除点云运动畸变并为激光扫描匹配部分提供初值预测,对几何结构缺乏的场景具有一定的鲁棒性. 但此方案无闭环检测模块,无法消除系统长期运行导致的累积漂移误差. Shao等<sup>[13]</sup>使用紧耦合的双目视觉-IMU里程计提供位姿估计并消除点云运动畸变,后续基于激光扫描匹配对视觉位姿进一步优化. 在闭环检测模块中,该算法基于词袋模型(bag-of-words, BOW)<sup>[14]</sup>检测闭环,若检测到闭环则依次使用视觉PnP(perspective- $n$ -point)方法<sup>[15]</sup>与激光帧间ICP(iterative closest point)方法<sup>[16]</sup>由粗到精地估计闭环约束. 此方案仅采用视觉信息检测闭环,受限与相机观测视角和范围,在视角变化较大时往往无法有效检测闭环. 文献[17]基于VLOAM框架,将单目视觉-IMU里程计VINS-Mono与LOAM方案结合,在视觉或激光单一模块退化时,该方案引入切换机制以应对挑战环境. 但在视觉-IMU里程计失效时,该方案忽略了IMU信息,仅利用激光雷达数据进行位姿估计. 此外,单目视觉-IMU里程计系统在车载等平稳运动中初始化难度较大,且由于IMU无足够的运动激励导致尺度漂移问题,进而导致在激光模块失效情况下易产生错误的位姿估计. 除上述视觉前端辅助激光扫描匹配的方法外,文献[18]将独立运行的激光-IMU里程计和单目视觉-IMU里程计通过卡尔曼滤波算法进行位姿融合,并且加入重定位模块和闭环优化模块,可应用于MALIN<sup>[19]</sup>挑战赛等环境中.

最近一些工作<sup>[20-22]</sup>提出了激光雷达-视觉-IMU 3种传感器数据紧耦合联合状态估计的解决方案. Zuo等<sup>[21]</sup>提出LIC-Fusion,其基于MSCKF(multi-state constraint Kalman filter)框架,联合IMU测量值、激光几何特征和视觉特征进行位姿推算;此外为降低计算消耗,在LIC-Fusion 2.0<sup>[21]</sup>中引入基于平面特征跟踪的滑动窗口策略. 文献[22]提出VILENS算法,此算法基于图优化框架,联合优化3种传感器观测信息,并采用滑动窗口策略以平衡计算量和精度,具有较强的鲁棒性. 紧耦合优化方案能够以简洁的方式同时兼顾位姿估计精度和算法鲁棒性,但对硬件的依赖性较强,如传感器数据采集的同步性等.

本文提出一种双目视觉辅助的激光惯导定位和建图算法,通过融合相机-IMU-激光雷达传感器数据实现结构特征退化场景与普通场景下的位姿估计与地图构建功能. 具体而言,该算法采用双目视觉惯导里程计为激光扫描匹配模块<sup>[6-8]</sup>提供先验位姿约束,

完成当前激光数据帧的配准. 考虑到视觉里程计可能发生失效, 与文献[17]切换到纯激光模式相比, 本文通过维护激光-IMU融合模块为激光扫描匹配部分提供候补的相对位姿预测, 以达到最佳的配准效果. 此外, 所提出算法在闭环检测模块中结合激光结构信息与视觉图像信息同时检测闭环, 并采用6自由度位姿图优化策略降低位姿估计的漂移误差.

本文的贡献与创新如下:

- 1) 提出一种双目视觉辅助的激光惯导里程计算法, 在扫描匹配中额外引入视觉先验位姿约束, 提高算法的位姿估计精度与鲁棒性;
- 2) 提出互补滤波算法与因子图优化求解的组合策略, 快速实现激光里程计参考系与惯性参考系的初始对准与IMU状态初始化;
- 3) 提出激光ICP匹配与视觉闭环检测结合的混合闭环检测策略, 通过6自由度位姿图融合里程计约束与闭环约束, 显著减小轨迹漂移误差.

## 1 双目视觉-IMU-激光 SLAM

### 1.1 符号说明

本文在算法描述中定义5个参考坐标系, 即激光雷达坐标系  $\{L\}$ 、相机坐标系  $\{C\}$ 、IMU坐标系  $\{I\}$ 、激光里程计参考系  $\{W\}$  和重力方向为  $[0, 0, -1]^T$  的惯性参考系  $\{N\}$ , 其中  $\{W\}$  和  $\{N\}$  的原点由起始时刻激光雷达坐标系的原点确定, 且  $\{N\}$  系的  $xy$  平面

与重力方向垂直. 此外, 使用变换矩阵  ${}^A_B T_j \in SE(3)$  定义  $t_j$  时刻下坐标系  $\{B\}$  在  $\{A\}$  中的表示,  ${}^A_B R_j \in SO(3)$  和  ${}^A_B p_j \in R(3)$  分别为  ${}^A_B T_j \in SE(3)$  中的旋转矩阵和平移向量部分, 哈密顿四元数  ${}^A_B q_j$  与  ${}^A_B R_j$  对应且  $\otimes$  表示两个四元数之间的乘积<sup>[23]</sup>.

### 1.2 算法框架

图1为本文所提出方法的整体框图, 主要由视觉惯导里程计、点云预处理、扫描匹配、激光惯导融合和闭环优化模块组成, 其中视觉惯导里程计独立运行, 分别为激光惯导里程计与闭环优化模块提供先验位姿和闭环信息. 对于激光惯导里程计部分, 先由点云预处理模块利用IMU数据对每一帧激光雷达点云进行运动畸变处理, 再基于点云局部粗糙度<sup>[6]</sup>提取边缘和平面点特征, 所提取特征作为扫描匹配模块输入; 扫描匹配模块根据输入特征的时间戳获取相对位姿预测值, 该预测值可由视觉惯导里程计模块和激光惯导融合模块获取, 且优先使用视觉位姿, 若视觉位姿作为预测值, 则将其作为位姿先验约束与激光结构特征约束共同计算当前激光帧位姿; 激光惯导融合模块负责修正IMU的测量漂移并为扫描匹配部分提供位姿预测; 而闭环优化模块基于所输入的激光关键帧和视觉关键帧信息进行基于ICP匹配和基于视觉闭环检测的混合闭环检测流程, 若检测到闭环, 则执行全局位姿图优化, 降低历史轨迹的累积漂移误差.

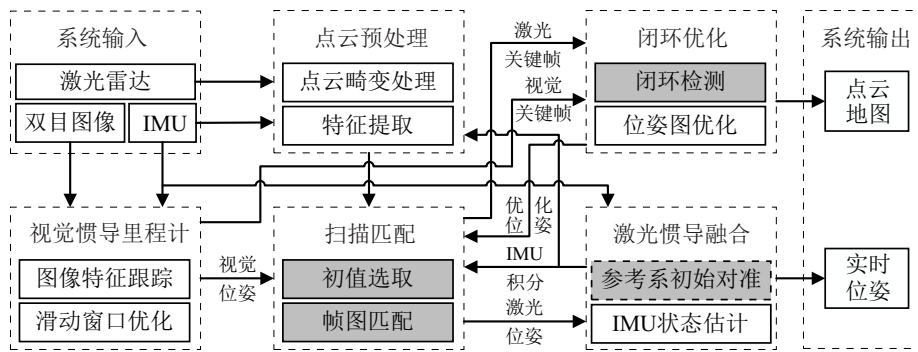


图1 视觉辅助的激光惯导SLAM框图

### 1.3 双目视觉惯导里程计

视觉惯导里程计模块采用VINS-Fusion<sup>[23]</sup>框架下的双目惯导里程计算法. 该算法可基于双目匹配获得三维特征点, 所估计位姿为真实尺度, 大大简化了单目视觉/单目视觉惯导等算法的初始化流程及对运动充分性的要求, 适用于车载运动.

本文使用视觉惯导里程计滑动窗口中次新帧  $c_{i-1}$  的状态量与IMU测量值积分得到高频位姿并作为输出, 如图2所示.

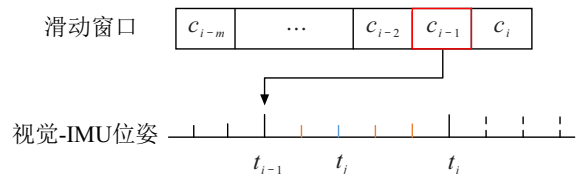


图2 视觉位姿输出

$t_j$  时刻的位姿  ${}^N_v T_j$  为

$${}^N_v p_j = {}^N_v p_{i-1} + \sum_{k=i-1}^{j-1} \left( {}^N_v v_k \Delta t + \frac{1}{2} g^N \Delta t^2 + \right.$$

$$\begin{aligned} & \frac{1}{2} {}^N v R_k (a_k - b_{i-1}^a) \Delta t^2), \\ & {}^N v v_j = {}^N v v_{i-1} + g^N \Delta t + \\ & \quad \sum_{k=i-1}^{j-1} {}^N v R_k (a_k - b_{i-1}^a) \Delta t, \\ & {}^N v q_j = {}^N v q_{i-1} \otimes \prod_{k=i-1}^{j-1} \left( \frac{1}{2} \Delta t (\omega_k - b_{i-1}^g) \right); \quad (1) \\ & {}^N v T_j = \begin{bmatrix} {}^N v R_j & {}^N v p_j \\ 0^T & 1 \end{bmatrix}. \quad (2) \end{aligned}$$

其中:  ${}^N v p$ 、 ${}^N v v$ 、 ${}^N v R$  (或  ${}^N v q$ )、 $g^N$  分别表示 IMU 在视觉惯导里程计起始时惯性参考系  $N_v$  下的位置、速度、姿态的旋转矩阵 (或四元数) 和重力加速度;  $a_k$  和  $\omega_k$  分别表示 IMU 在时刻  $t$  的加速度和角速度测量值;  $b_{i-1}^a$  和  $b_{i-1}^g$  分别表示 IMU 在  $t_{i-1}$  时刻的加速度计和陀螺仪偏置, 并且假设在  $t \in [t_{i-1}, t_i]$  内保持不变;  $\Delta t$  为 IMU 采样周期. 此外,  ${}^N v p_{i-1}$ 、 ${}^N v v_{i-1}$ 、 ${}^N v R_{i-1}$ 、 $b_{i-1}^a$  和  $b_{i-1}^g$  构成视觉帧  $c_{i-1}$  的状态量.

#### 1.4 视觉位姿辅助的激光惯导里程计

所提出的里程计算法通过预测位姿利用当前帧激光点云数据和历史帧数据构建结构特征约束, 与视觉惯导里程计模块提供的先验位姿约束联合估计当前激光帧位姿, 并将激光位姿与 IMU 数据融合产生高频的位姿输出. 该算法主要可分为点云预处理、扫描匹配和激光惯导融合 3 个模块.

##### 1.4.1 点云预处理

激光雷达在采集数据时随着载体运动, 导致采集的点云数据中包含运动畸变. 若激光雷达在  $t \in [t_{k-1}, t_k]$  时间内采集获得一帧点云数据, 则可通过将当前帧中每一激光点通过下式转换到  $t_k$  时刻的激光雷达坐标系下完成畸变去除:

$$\tilde{p}_i = \widetilde{{}^W T_k}^{-1} \widetilde{{}^W T_j} p_i. \quad (3)$$

其中:  $\tilde{p}_i$  和  $p_i$  分别为  $t_i$  时刻去除运动畸变的激光点和原始激光点,  $t_i \in [t_{k-1}, t_k]$ ;  $\widetilde{{}^W T_k}$  和  $\widetilde{{}^W T_j}$  由激光惯导模块输出的高频 IMU 积分位姿与预先标定的激光雷达-IMU 之间的外参数  ${}^L T$  计算得到.

与文献 [6] 类似, 本文通过计算每一个激光点的粗糙度从中提取“边缘”特征点集  $\mathcal{E}$  和“平面”特征点集  $\mathcal{P}$ . 待完成特征提取之后, 两类特征点集将用于配准以估计点云采集时刻的激光雷达位姿.

##### 1.4.2 扫描匹配

除使用激光数据构建的结构特征约束外, 扫描匹配模块融合视觉惯导里程计模块提供的先验位姿增

强约束信息, 提高算法在结构特征缺乏环境下的鲁棒性. 该模块包含当前帧的位姿初值预测、激光特征关联和优化方程构建等 3 个部分.

1) 位姿初值预测. 激光当前帧位姿初值预测值来源包括: ① 视觉惯导里程计提供的视觉-IMU 位姿; ② 激光惯导融合模块输出的 IMU 积分位姿.  $t_k$  时刻激光帧的预测位姿  ${}^W T_k$  计算如下式所示:

$$\widetilde{{}^W T_k} = {}^W T_{k-1} {}^L T^{-1} ({}^N T_{k-1}^{-1} {}^N T_k) {}^L T. \quad (4)$$

其中:  ${}^W T_k$  为  $t_{k-1}$  时刻激光雷达位姿,  ${}^N T_{k-1}$  和  ${}^N T_k$  分别为  $t_{k-1}$  和  $t_k$  时刻的 IMU 位姿. 当两类初值来源均正常时, 优先使用视觉惯导里程计提供的位姿作为位姿估计的初值以融合更多的信息.

2) 激光特征关联. 与文献 [10] 类似, 本文基于最近邻匹配规则, 建立当前激光帧特征点集与历史局部地图之间的数据关联, 其中局部地图由历史关键帧集合组成. 具体而言, 边缘特征点  $\varepsilon_i$  ( $\varepsilon_i \in \mathcal{E}_k$ ) 在局部地图中查找固定数量的最近邻点集并拟合局部直线段, 该直线段使用其经过的两点 ( $p_u^{\varepsilon_i}$ ,  $p_v^{\varepsilon_i}$ ) 表征; 类似地, 平面特征点  $\rho_i$  ( $\rho_i \in \mathcal{P}_k$ ) 的最近邻点集用于拟合平面, 使用该平面上的 3 点 ( $p_u^{\rho_i}$ ,  $p_v^{\rho_i}$ ,  $p_\omega^{\rho_i}$ ) 表征.

3) 优化方程构建. 本文通过同时最小化激光几何特征约束项与视觉先验位姿约束项的误差和来估计当前激光帧位姿, 即

$$\min_{{}^W T_k} \left\{ \sum_{\varepsilon_i \in \mathcal{E}_k} \psi(d_{\varepsilon_i}({}^W T_k)^2) + \sum_{\rho_i \in \mathcal{P}_k} \psi(d_{\rho_i}({}^W T_k)^2) + d_{T_k}({}^W T_k)^T \Omega_p d_{T_k}({}^W T_k) \right\}. \quad (5)$$

其中:  $\psi(\cdot)$  为双平方加权 (bisquare weight) 函数<sup>[24]</sup>;  $\Omega_p$  为先验项的信息矩阵; 激光特征约束项  $d_{\varepsilon_i}({}^W T_k)$  和  $d_{\rho_i}({}^W T_k)$  分别为  $\varepsilon_i$  和  $\rho_i$  与其关联特征的距离, 具体为

$$d_{\varepsilon_i}({}^W T_k) = \frac{|({}^W T_k \varepsilon_i - p_u^{\varepsilon_i}) \times ({}^W T_k \varepsilon_i - p_v^{\varepsilon_i})|}{|p_u^{\varepsilon_i} - p_v^{\varepsilon_i}|}, \quad (6)$$

$$d_{\rho_i}({}^W T_k) = \frac{|({}^W T_k \rho_i - p_u^{\rho_i}) \cdot (p_u^{\rho_i} - p_v^{\rho_i}) \times (p_u^{\rho_i} - p_\omega^{\rho_i})|}{|(p_u^{\rho_i} - p_v^{\rho_i}) \times (p_u^{\rho_i} - p_\omega^{\rho_i})|}. \quad (7)$$

当环境中的几何特征缺乏时, 仅通过式 (6) 和 (7) 建立优化方程将无法保证充足的约束来精确求解位姿, 如载体运行于开阔的环境中, 激光雷达的有效测量仅来自于地面部分, 平行于地面的平移和载体航向的不确定度较高. 与文献 [10] 相比, 本文通过添加视觉先验位姿约束项

$$d_{T_k}({}^W T_k) = \text{Log}(\widetilde{{}^W T_k}^{-1} {}^W T_k) \quad (8)$$

实现与视觉信息融合以获得额外约束,进而保证算法在几何特征缺乏环境下的鲁棒性,其中 $\text{Log}(\cdot)$ 表示SE(3)空间到向量空间的对数映射关系<sup>[23]</sup>. 特别地,当环境结构良好时,当前帧激光位姿估计值由式(6)和(7)所构建的几何特征约束为主导;当预测位姿为激光惯导模块预测提供(即视觉里程计模块失效)时, $\Omega_p$ 设置为0矩阵. 通过Levenberg-Marquardt算法迭代优化式(5),可以得到激光当前帧对应的位姿.

此外,为降低内存消耗,历史点云数据仅保存激光关键帧所对应的特征点集,其中若当前激光帧距上一关键帧之间的距离或姿态旋转超过一定阈值时(默认旋转角超过0.2 rad, 平移距离超过1 m),则当前帧被定义为关键帧.

### 1.4.3 参考系初始对准与IMU状态估计

为了将激光雷达位姿与IMU数据进行融合,需要确定里程计参考系 $W$ 在惯性参考系 $N$ 中的初始姿态(包括俯仰角pitch和横滚角roll). 为此,本文在数据融合前进行参考系对齐,具体采用互补滤波(complementary filter, CF)<sup>[25]</sup>算法提供的初值与因子图优化精确求解的估计策略,依据IMU原始测量值快速完成姿态估计(如2.3节实验所示).

对于IMU测量数据(线加速度与角速度),首先使用互补滤波算法估计当前IMU坐标系 $I$ 在 $N$ 系中的姿态,然后将测量数据与估计姿态作为输入,使用如图3(a)所示的因子图<sup>[26]</sup>结构精确化估计激光里程计参考系 $W$ 在惯性参考系 $N$ 下的旋转矩阵 ${}^N_W R$ (为便于采用统一形式描述,下文采用完整的变换矩阵 ${}^N_W T$ ,其中平移部分 ${}^N_W p = [0, 0, 0]^T$ ). 除 ${}^N_W T$ 外,该因子图包含IMU状态量 ${}^N_I x$ 和激光雷达位姿因子、IMU预积分因子<sup>[27]</sup>等. 为求解 ${}^N_W T$ 及 ${}^N_I x$ ,本文将此图结构所描述过程转换为非线性最小二乘优化问题,即

$$\chi^* = \arg \min_{\chi} \left\{ \sum_{i=1} \|r_L({}^N_W T, {}^N_I T)\|_{\Sigma_L}^2 + \sum_{i=0} \|r_I({}^N_I x, {}^N_{I_{i+1}} x)\|_{\Sigma_I}^2 \right\},$$

$$\chi = \{ {}^N_W T, {}^N_{I_1} x, {}^N_{I_2} x, {}^N_{I_3} x, \dots \}. \quad (9)$$

其中: $\Sigma_L$ 和 $\Sigma_I$ 分别为激光雷达位姿观测误差项 $r_L(\cdot)$ 和IMU预积分观测误差项 $r_I(\cdot)$ 的协方差, $\Sigma_L$ 与里程计估计位姿的不确定度有关, $\Sigma_I$ 的计算可参考文献<sup>[27]</sup>;  $r_L(\cdot)$ 和 $r_I(\cdot)$ 的定义分别为

$$r_L(\cdot) = \text{Log}({}^W_{I_i} \hat{T}^{-1} ({}^N_W T^{-1} {}^N_{I_i} T)), \quad (10)$$

$$r_I({}^N_{I_i} x, {}^N_{I_{i+1}} x) =$$

$$\begin{bmatrix} {}^{I_i}_{I_{i+1}} \hat{\alpha} - {}^N_{I_i} R^T ({}^N_{I_{i+1}} p - {}^N_{I_i} p - \frac{1}{2} g^N \Delta t^2 - {}^N_{I_i} v \Delta t) \\ {}^{I_i}_{I_{i+1}} \hat{\beta} - {}^N_{I_i} R^T ({}^N_{I_{i+1}} v - g^N \Delta t - {}^N_{I_i} v) \\ \text{Log}({}^{I_i}_{I_{i+1}} \hat{\gamma}^T ({}^N_{I_i} R^T {}^N_{I_{i+1}} R)) \\ b_{i+1}^a - b_i^a \\ b_{i+1}^g - b_i^g \end{bmatrix}. \quad (11)$$

${}^W_{I_i} \hat{T} = {}^W_{L_i} T^I T^{-1}$ 为扫描匹配模块输出的每一激光帧位姿经激光雷达-IMU外参转换得到; ${}^{I_i}_{I_{i+1}} \hat{\alpha}$ 、 ${}^{I_i}_{I_{i+1}} \hat{\beta}$ 和 ${}^{I_i}_{I_{i+1}} \hat{\gamma}$ 分别为 $t \in [t_i, t_{i+1})$ 时间内的平移、速度和旋转预积分项,其计算可参考文献<sup>[27]</sup>. 此外,为降低计算开销,本文采用逐渐增加数据的策略执行对齐过程,即每获得一个激光帧位姿便执行一次优化,直到 ${}^N_W T$ 收敛.

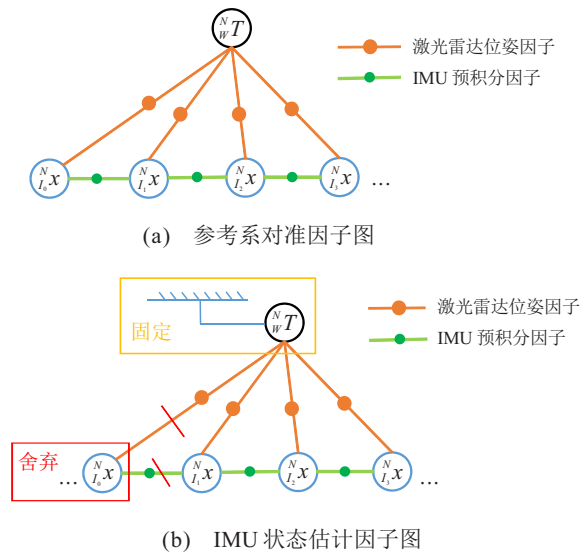


图3 激光雷达-IMU融合因子图

待参考系对齐完成后,将激光里程计参考系在惯性参考系的变换矩阵 ${}^N_W T$ 固定,即之后优化过程不改变 ${}^N_W T$ ,如图3(b)所示. 本文使用此图结构将激光雷达位姿数据与IMU测量数据进行融合来估计IMU状态量. 此外,最新的IMU状态量将按照类似于式(1)、(2)的过程用于产生IMU位姿. 当图结构中IMU状态量的数量达到固定阈值时,将最旧的状态量丢弃以保证运行速度.

### 1.5 激光视觉混合闭环检测与闭环优化

里程计模块在长时间运行之后需要加入闭环约束来消除历史轨迹的漂移误差. 多帧激光点云拼接形成的点云地图对环境的几何结构描述较为详细,因此,闭环检测可以朴素地采用基于欧氏距离的ICP匹配方法<sup>[7,10]</sup>,但受限于欧氏距离阈值,在漂移误差较大时此方法将失效. 为此,本文额外引入基于视觉图

像的场景识别方法DBOW2<sup>[14]</sup>来辅助检测闭环信息.

### 1.5.1 视觉闭环检测

视觉闭环检测可分为5个步骤: 1) 基于运动距离与旋转角度, 从视觉惯导里程计模块输出的图像帧及对应的视觉位姿中选取关键帧. 2) 在关键帧的左目图像中提取固定数量的角点(除窗口中原有特征点, 再额外提取500个新的角点)并生成BRIEF特征描述子<sup>[28]</sup>. 3) 使用当前关键帧在DBOW2数据库中进行闭环检索. 4) 若检测到闭环, 则基于RANSAC算法和基础矩阵剔除特征点匹配对的外点; 若剩余内点个数大于设定阈值(默认为25), 则再使用RANSAC算法和PnP算法<sup>[15]</sup>进一步筛除外点并计算当前关键帧与匹配到的历史关键帧之间的相对位姿变换. 5) 若步骤4)中所剩特征匹配对的数量仍大于设定阈值, 则认为检测到有效的视觉闭环. 当前关键帧和匹配关键帧的时间戳与相对位姿变换将进一步用于辅助检测激光闭环.

### 1.5.2 混合闭环检测

本文所提出的混合闭环检测流程如算法1所示.

#### 算法1 混合闭环检测流程.

输入: 视觉闭环信息队列  $\mathcal{V}_{loop}$ , 激光关键帧集合  $\mathcal{F}$ , 距离阈值  $d_n$ , ICP匹配误差阈值  $d_t$ , 时间差阈值  $t_{pre}$ ;  
输出: 回环信息集合  $\mathcal{L}_{loop}$ .

```

1: if  $\mathcal{V}_{loop} \neq \emptyset$  then
2:    $\{t_{v1}, t_{v2}, T_{init}^v\} = \mathcal{V}_{loop}.front();$ 
3:    $T_{init}^L = Transform(T_{init}^v);$ 
4:    $\{F_c^L, F_n^L\} = SearchWithTime(\mathcal{F}, t_{v1}, t_{v2});$ 
5:    $\mathcal{L}_{loop} \leftarrow \emptyset;$ 
6:    $ComputeLoopInfo(F_c^L, F_n^L, T_{init}^L, d_t, \mathcal{F}, \mathcal{L}_{loop});$ 
7:   if  $\mathcal{L}_{loop} \neq \emptyset$  then
8:      $\mathcal{V}_{loop} \leftarrow \emptyset;$ 
9:     return  $\mathcal{L}_{loop};$ 
10:  end if
11: end if
12:  $F_c^L \leftarrow \mathcal{F}.back();$ 
13:  $F_n^L = SearchWithDistance(F_c^L, d_n, t_{pre});$ 
14:  $T_{init} = GuessInit(F_c^L, F_n^L);$ 
15:  $ComputeLoopInfo(F_c^L, F_n^L, T_{init}, d_t, \mathcal{F}, \mathcal{L}_{loop});$ 
16: return  $\mathcal{L}_{loop}$ .
```

function  $ComputeLoopInfo(F_c, F_n, T_i, d, \mathcal{F}, \mathcal{L})$

```

1:  $M_n = ConstructLocalMap(F_n, \mathcal{F});$ 
2:  $\{e, T_{final}\} = ICP(F_c, M_n, T_i);$ 
3: if  $e < d$  then
4:    $\mathcal{L} \leftarrow \{c, n, T_{final}\};$ 
5: end if
6: return
```

首先检查是否存在视觉闭环信息, 若存在, 则解析最新的视觉闭环信息(视觉关键帧时间戳与相对位姿), 并依据时间戳分别查找时域临近的激光关键帧, 如第2~第4行所示. 其中:  $Transform()$  函数负责依据激光雷达-相机之间的外参将视觉关键帧的相对位姿  $T_{init}^v$  转换为激光雷达的相对位姿  $T_{init}^L$ , 而  $SearchWithTime()$  函数用于时域近邻查找. 然后执行  $ComputeLoopInfo()$  函数, 计算激光关键帧之间的闭环约束信息, 其具体负责: 1) 执行  $ConstructLocalMap()$ , 将关键帧  $F_n$  及其时域邻域内的关键帧组成局部点云地图  $M_n$ , 且  $M_n$  的参考坐标系为关键帧  $F_n$ ; 2) 执行ICP算法, 将较新的关键帧  $F_c$  与  $M_n$  进行配准, 其中位姿初值为  $T_i$ ; 3) 依据ICP配准误差判断激光闭环信息的有效性, 若ICP匹配误差  $e$  小于给定阈值  $d$  (默认为  $0.3 \text{ m}^2$ ), 则认为激光闭环有效. 具体流程如算法1中  $ComputeLoopInfo()$  函数第1~第6行所示.

为避免产生冗余闭环信息, 所提出算法依据视觉闭环信息产生有效激光闭环后, 将清空视觉闭环信息队列缓存, 并直接结束本次检测流程, 如第7~第10行所示. 若视觉闭环信息不存在或依据视觉闭环未产生有效激光闭环, 则基于欧氏距离的ICP匹配方法检测闭环, 如第12~第15行所示. 首先从激光关键帧集合中取出最新帧  $F_c^L$ ; 然后执行  $SearchWithDistance()$  函数, 依据距离阈值  $d_n$  (默认为20 m) 查找与  $F_c^L$  欧氏距离最近的关键帧  $F_n^L$ , 且两关键帧之间的时间差需大于设定阈值  $t_{pre}$  (默认为30 s); 最后执行  $ComputeLoopInfo()$  函数, 执行ICP配准计算激光闭环信息, 其中配准初值  $T_{init}$  由关键帧  $F_c^L$  与  $F_n^L$  的位姿计算得到, 如算法1中  $GuessInit()$  所示.

### 1.5.3 闭环优化

若产生有效的激光闭环信息, 则采用如图4所示的6自由度位姿图进行全局优化, 以消除历史轨迹的漂移误差. 该图中优化变量为激光关键帧位姿  $T_k$ , 包含里程计因子和闭环因子, 两种因子的相对变换观测值可统一表示为  ${}^u_v \widetilde{T}$ . 其中:  $u$  在里程计因子中表示前一激光关键帧下标, 而在闭环因子中表示闭环检测中找到的历史关键帧下标;  $v$  统一表示当前帧下标. 此外, 闭环检测和优化比较占用计算资源, 因此该模块以较低频率运行于单独线程.

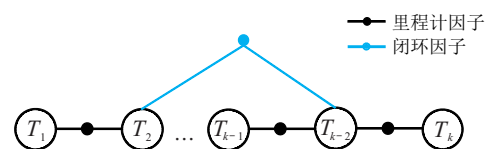


图4 闭环优化位姿图

## 2 实验评价

为验证所提出算法的有效性, 本文使用公开数据集MVSEC<sup>[29]</sup>的outdoor车载序列和自制数据集分别针对激光扫描匹配、激光-惯导参考系对齐以及混合闭环检测等3部分内容进行测试验证, 且与开源算法进行对比以验证本文算法的性能. 图5为自制数据采集平台及相应的传感器, 包括SUMMIT移动平台、Velodyne-16多线激光雷达、Realsense D435i相机和Xsens MTi-30 IMU, 各传感器与载体坐标系如图5所示(红色轴, 绿色轴, 蓝色轴, 且符合右手定则); 计算平台为微星笔记本, CPU为Intel i7-9750H 2.6 GHz, 24 G内存, 软件平台为Ubuntu 16.04系统与ROS kinetic软件框架.



图5 数据采集平台

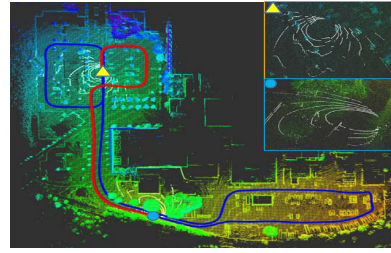
### 2.1 不同配置下激光扫描匹配测试

为了验证1.4.2节所提出的视觉里程计位姿辅助激光扫描匹配方法的性能, 本节实验选择MVSEC数据集中的outdoor\_day\_1序列和自制数据集的seq\_sim序列进行测试. 结合本文所针对的几何结构稀疏场景, 对两个序列进行预处理, 即滤除部分时间段内载体周围障碍物形成的激光点云数据. 相应序列的全局点云地图如图6所示. 其中: 蓝色线为全局轨迹, 红色线为预处理数据对应轨迹, 黄色和蓝色方框标记了两处预处理后的激光点云帧数据, 即主要来自地面的观测数据(白色标记).

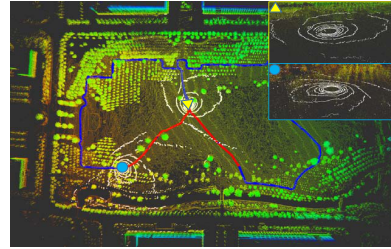
具体采用如下对比算法进行验证:

- 1) VIO: 双目视觉-IMU里程计VINS Fusion;
- 2) LIO: 激光-IMU模块提供初值预测的激光惯导里程计方法;
- 3) VLIO\*: 视觉-IMU模块提供初值预测的视觉激光惯导里程计方法;
- 4) VLIO: 即本文最终的方法, 视觉-IMU模块提供初值预测且添加视觉先验位姿约束项的视觉激光惯导里程计方法.

评价指标选择绝对轨迹误差 (absolute trajectory



(a) outdoor\_day\_1+



(b) seq\_sim+

图6 预处理后缺乏几何结构特征的点云图

error, ATE)<sup>[30]</sup>的均方根误差 (root mean square error, RMSE) 形式来评估不同配置下算法的全局轨迹精度, 实验结果如表1所示. 所提出VLIO算法除在seq\_sim序列上的结果略逊于VLIO\*, 取得次优精度外, 在其他3个序列上的实验精度均为最优. 由outdoor\_day\_1+和seq\_sim+相较于outdoor\_day\_1和seq\_sim的结果差异可见, 经预处理后序列上的实验结果更优, 展示了数据集预处理操作滤除障碍物点云的必要性以及VLIO算法的优越性. 此外, 光照剧烈变化情况也可能使outdoor\_day\_1+与outdoor\_day\_1序列上的结果产生差异, 如图7所示, 图7(a)~(d)依次为行驶过程中某4s内的相机视野图像. 可见, VLIO较其他对比方法有明显的精度提升. 各轨迹使用轨迹评估工具EVO<sup>[31]</sup>与参考真值进行对齐, 各算法在预处理数据集的轨迹如图8所示.

表1 不同算法绝对轨迹误差 单位: m

序列	VIO	LIO	VLIO*	VLIO
outdoor_day_1+	9.583	—	13.904	<b>9.064</b>
seq_sim+	39.148	—	8.418	<b>7.365</b>
outdoor_day_1	9.586	1.810	1.933	<b>1.495</b>
seq_sim	39.192	6.408	<b>6.264</b>	6.403

注: “+”表示数据集经过预处理, “—”表示无有效结果.

### 2.2 几何结构稀疏场景下的SLAM算法运行实验

为验证所提出算法在几何结构缺乏环境下的鲁棒性, 在如图9(a)所示的测试场景中, 遥控移动机器人在中间空旷区域以0.5 m/s的速度运行, 获得数据集序列seq\_dege, 其轨迹如图9(b)卫星图中的绿色曲线所示, 其中红色星状物标识了起始位置.

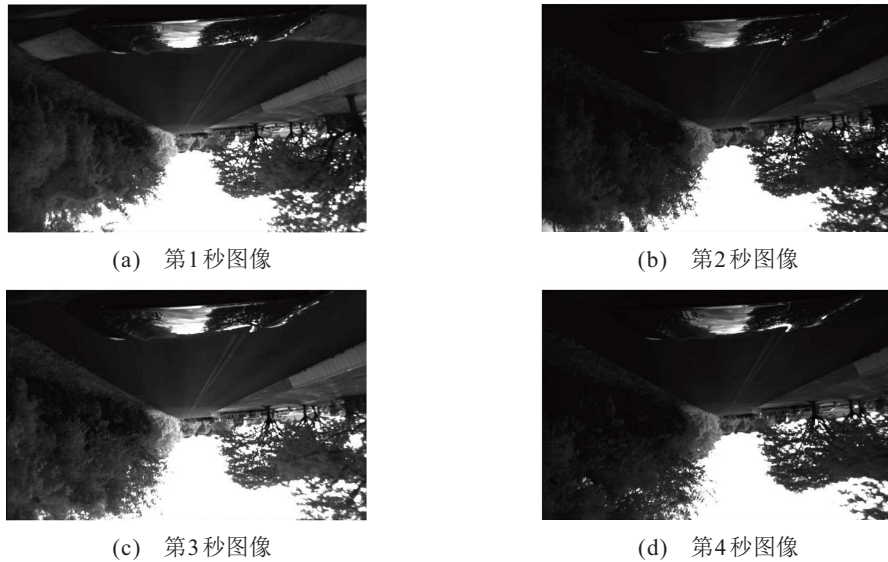


图7 outdoor\_day\_1序列光照变化情况示例

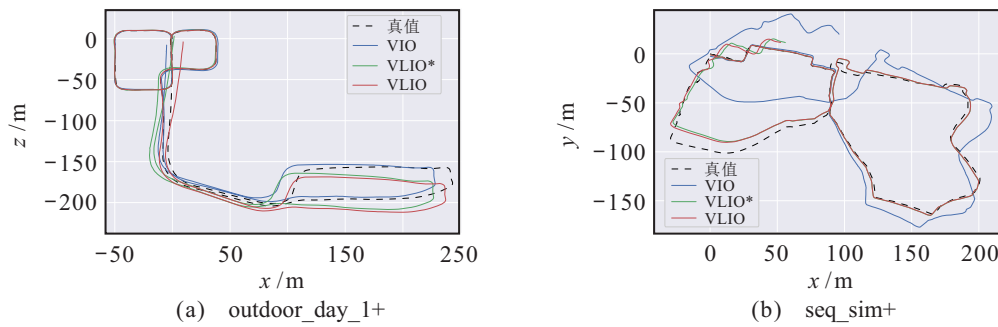


图8 缺乏几何结构场景下不同算法的位姿轨迹



图9 测试场景

将所出提算法与 LIO-SAM<sup>[10]</sup>、LIOM<sup>[8]</sup>和 A-LOAM<sup>[32]</sup>进行对比,其中本文算法和 LIO-SAM 均关闭闭环检测模块.从录制数据集中选择8个不同的起始位置运行各算法,所得轨迹如图10所示.其中:红色星状物标记了各启动位置;灰色虚线标识了 RTK-GPS 轨迹(误差为厘米级)并作为参考真值,各算法轨迹使用 EVO 与参考真值进行对齐.该实验中启动位置1~8将不断靠近空旷区域的中心部分,起始时刻的激光雷达数据愈加稀疏.如图10(a)~(d)所示:在起始时刻激光数据相对良好时,LIO-SAM 和 A-LOAM 起始轨迹处呈现不同程度的“抖动”现象,LIOM 仅在位置1处启动可以得到有效的结果;当起

始激光雷达数据相对稀疏时,得到如图10(e)~(h)所示轨迹,LIO-SAM 和 LIOM 无法正常启动,A-LOAM 在起始阶段可以大致估计运动轨迹,但之后运行失败.而本文所提出的算法 VLIO 在以上测试中所得轨迹较为平滑且与参考真值贴合良好.当环境中几何特征缺乏时,激光雷达数据无法提供足够的约束,LIO-SAM 和 LIOM 等算法将无法对系统状态进行有效估计,此时使用 IMU 进行位姿预测将导致系统状态快速发散.A-LOAM 算法使用纯激光雷达数据进行位姿估计,系统状态发散较慢.而本文算法 VLIO 通过结合视觉信息可以保证充足的环境特征约束,有效增强了所提出算法的鲁棒性.

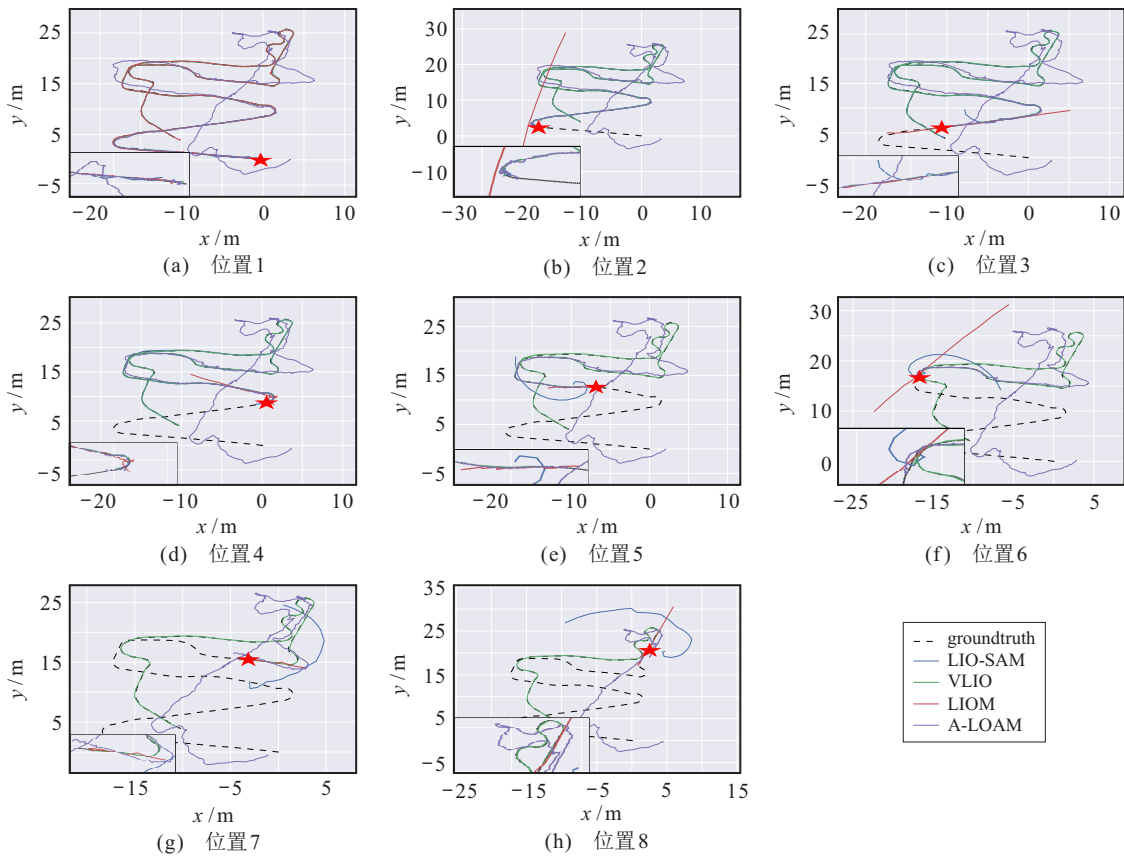


图 10 不同位置启动

由于对比算法的起始阶段存在不稳定问题,本节采用相对位姿误差(relative pose error, RPE)<sup>[33]</sup>的RMSE形式作为定量评价指标以评估各算法所估计轨迹的局部精度,结果如表2所示,其中仅展示了图10(a)~(h)等环境几何结构相对良好情况时的算法轨迹误差. 所提出算法与其他算法相比,误差降低了42.36%~81.53%. 此外,LIO-SAM和LOAM算法的轨迹误差变化显著,而VLIO算法在不同位置启动时轨迹误差近似且较低,体现了本文算法VLIO在几何结构特征缺乏时具有较好的鲁棒性.

表2 各算法的相对位姿误差

算法	位置							
	1	2	3	4	5	6	7	8
VLIO	0.049	0.048	0.049	0.050	0.050	0.052	0.048	0.051
LIO-SAM	0.142	0.083	0.267	0.120	-	-	-	-
LIOM	0.095	-	-	-	-	-	-	-
A-LOAM	0.196	0.209	0.233	0.212	-	-	-	-

注:“-”表示无有效结果.

### 2.3 里程计参考系与惯性参考系对准实验

本节实验用于验证所提出参考系对齐流程的有效性,采用手持图5中传感器的方式采集数据,实验

场景如图11所示,其中所提出算法VLIO关闭视觉惯导里程计模块. 实验中以不同姿态启动算法进行验证,并将Xsens MTi-30 IMU输出的姿态作为参考,分别对比以下3种参考系对齐流程:

- 1) 互补滤波算法(CF): 将IMU的加速度计和陀螺仪测量作为输入,取起始时刻的姿态作为估计值;
- 2) CF\_Init: 本文1.4.3节所提出方法,使用互补滤波算法与因子图优化进行参考系对齐;
- 3) Raw\_Init: 直接采用因子图优化进行对齐.

其中后两种算法取启动后的前20s数据作为输入,最后一次的优化值作为输出.

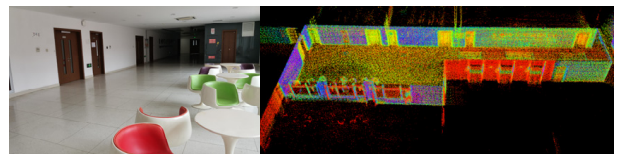


图 11 测试场景

本节采用欧拉角形式来展示结果,由于航向角(yaw)与重力方向无关,所示结果仅包括横滚角(roll)和俯仰角(pitch),4组结果如表3所示. 误差计算方式采用四元数形式计算角度差得到. 实验结果显示:A组中起始姿态与重力方向夹角较小,CF\_Init与Raw\_Init结果相近且较CF估计结果误差小;当起始

姿态与重力方向夹角依次增大,如B、C、D组所示,Raw\_Init的估计结果随初始姿态角增大而误差不断增加,本文提出的CF\_Init以互补滤波估计姿态为初值,可以有效降低估计误差.此外,上述参考系对齐流

程的误差变化曲线如图12所示,本文提出的CF\_Init采用互补滤波算法提供初值,相比于Raw\_Init的估计误差收敛速度较快,能够在5s内将误差降低于1°左右,保证系统快速运行.

表3 参考系对齐结果

单位: (°)

序号	参考值		CF			CF_Init			Raw_Init		
	roll	pitch	roll	pitch	error	roll	pitch	error	roll	pitch	error
A	-4.761	1.707	-1.558	3.799	3.823	-4.274	2.309	0.816	-4.011	1.639	<b>0.764</b>
B	-0.092	8.675	-1.742	10.623	2.553	-0.298	8.841	<b>0.265</b>	0.160	6.939	1.757
C	5.810	10.302	4.778	6.509	3.931	5.208	9.368	<b>1.153</b>	4.343	8.193	2.604
D	1.696	-21.652	2.149	-20.821	0.951	1.753	-21.497	<b>0.168</b>	1.364	-18.054	3.628

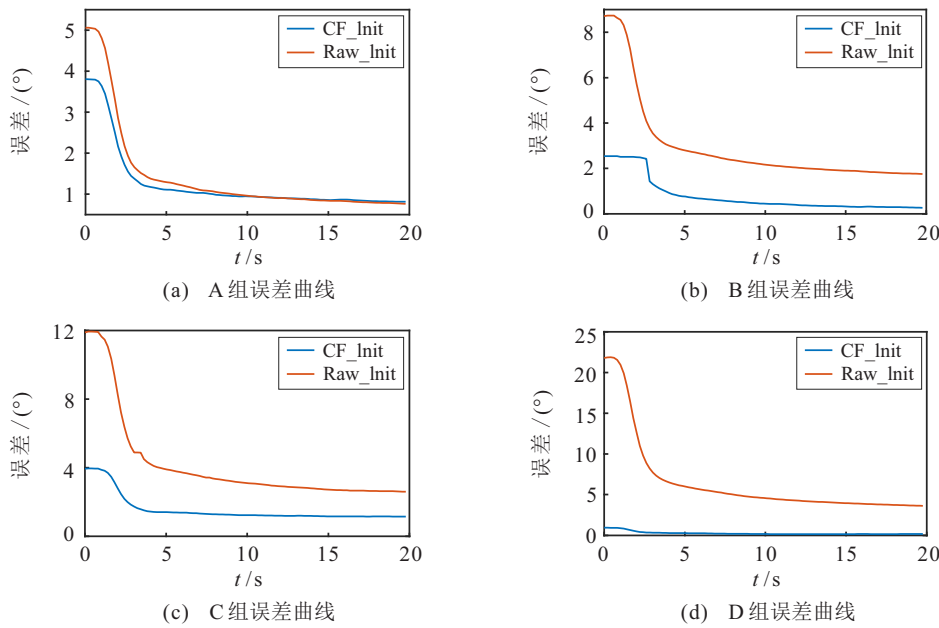


图12 误差收敛曲线

### 2.4 混合闭环检测

本节采用序列seq\_loop\_1、seq\_loop\_2和数据集MVSEC中的outdoor\_day\_2序列进行闭环模块验证.其中:seq\_loop\_1序列由移动平台在草坪旁的人行小路采集数据形成,全长约760m;seq\_loop\_2序列在seq\_loop\_1序列场景四周的马路上采集得到,约1170m;而outdoor\_day\_2序列为典型城市道路环境,包含较多的闭环场景.

为验证所提出的混合闭环策略,本节采用3种不同闭环策略的算法配置进行测试,包括:1) VLIO-Lp-V:采用视觉信息检测闭环,如算法1中第1~第11行所示;2) VLIO-Lp-L:采用激光信息检测闭环,如算法1中第12~第15行所示;3) VLIO-Loop:采用本文所提出的混合闭环策略,如算法1所示.在不同配置下,所提出算法在seq\_loop序列中的轨迹与点云地图分

别如图13、图14所示.其中:VLIO-Lp-L算法仅采用基于欧氏距离的ICP策略检测闭环,由于里程计累积误差较大,导致无法有效检测到闭环;VLIO-Lp-V与VLIO-Loop算法采用视觉信息正确检测到闭环,显著降低了轨迹的漂移误差.此外,经闭环优化后,所建点云地图可以与Google地图良好对齐.所提出算法在outdoor\_day\_2序列下的轨迹结果如图15所示.其中:VLIO-Lp-V算法由于视觉图像视角差异,如载体双向行驶经过同一场景时,无法检测到有效闭环,导致部分场景的漂移误差仍保留在全局轨迹中;而VLIO-Lp-L和VLIO-Loop均有激光闭环检测机制,更容易检测到有效闭环,且相对于VLIO-Lp-V更贴近于参考真值.各配置算法的定量评估结果如表4所示,其中评价指标为ATE的RMSE形式.

此外,为进一步评估所提出闭环检测策略的有

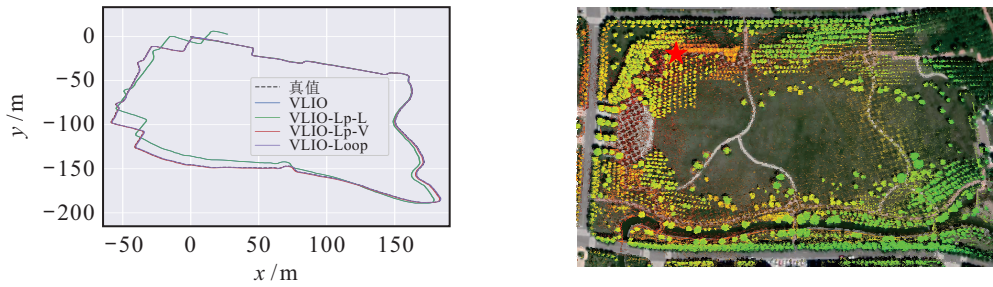


图 13 seq\_loop\_1测试结果

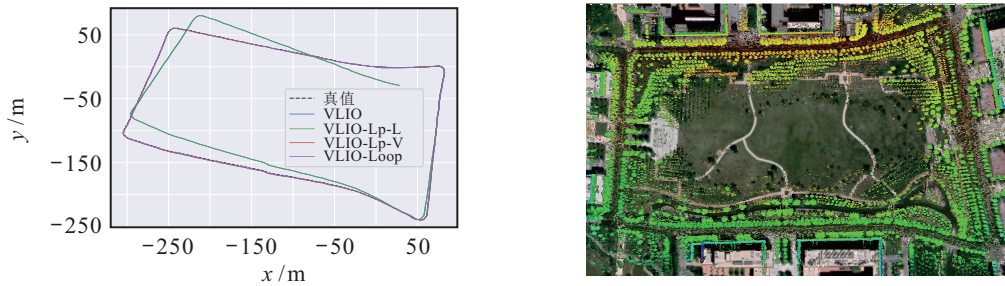


图 14 seq\_loop\_2测试结果

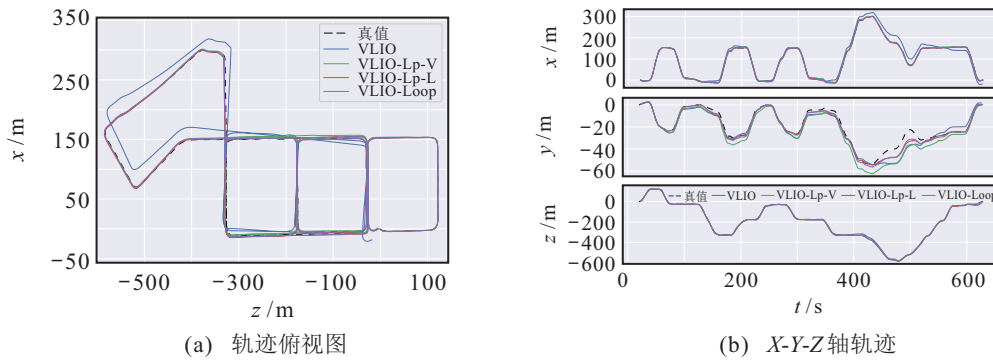


图 15 outdoor\_day\_2轨迹结果

表 4 不同闭环策略下的算法轨迹误差 单位: m

序列	VLIO	VLIO-Lp-V	VLIO-Lp-L	VLIO-Loop
seq_loop_1	7.085	0.426	7.085	<b>0.414</b>
seq_loop_2	15.937	3.061	15.937	<b>2.926</b>
outdoor_day_2	15.475	7.006	3.818	<b>3.401</b>

效性,本节分别统计 VLIO-Loop 算法在测试序列下的检测时间消耗与闭环约束误差,如表 5 所示. 其中:所提出算法闭环检测时间的平均值和中位数均在 300 ms 以内,在实际应用中,闭环检测模块频率设定为 1 Hz,充分满足算法运行需求;而闭环约束误差用于评估算法计算得到的闭环约束精度,即所估计的相对位姿与真值之间的 RMSE 统计误差,包含平移和旋转两项. 与里程计算法漂移导致的米级误差相比,闭环约束的平移误差保持在亚米级,可有效降低历史轨迹误差. 此外, MVSEC 数据集参考真值的频率较低,误差统计时采用线性插值算法提高真值频率,导致所提出算法在 outdoor\_day\_2 序列上的闭环约束统计误

差相对较高.

表 5 闭环检测时间消耗与闭环约束误差

序列	时间消耗/ms		闭环约束误差	
	平均值	中位数	平移/m	旋转/rad
seq_loop_1	170.2	148.3	0.205	0.017
seq_loop_2	187.1	176.1	0.066	0.012
outdoor_day_2	234.7	211.1	0.657	0.197

### 2.5 不同配置下的算法性能分析

本文所提出的算法由视觉惯导里程计、扫描匹配、闭环优化和激光惯导融合模块组成,其中算法精度主要受前 3 部分模块影响,而激光惯导融合模块主要负责提高算法输出的位姿频率. 鉴于此,本节对前 3 部分模块组成的不同算法配置进行测试,结果如表 6 所示. 由 MVSEC 和自制集等结构良好场景序列下的结果可知,激光惯导里程计算法 LIO 相比于视觉惯导里程计算法 VIO 具有更佳的精度与稳定性. 而 VLIO 算法通过融合激光视觉惯导信息,进一步提升

了算法整体的鲁棒性与精度,如模拟集测试结果所示.此外,完整算法 VLIO-Loop 通过闭环检测模块添加闭环约束,并采用6自由度位姿图优化算法有效降低了历史轨迹误差,提高了算法轨迹估计精度.

表6 不同配置下的算法绝对轨迹误差 单位: m

数据集	序列	VIO	LIO	VLIO	VLIO-Loop
MVSEC	outdoor_day_1	9.586	1.810	<b>1.495</b>	1.541
	outdoor_day_2	11.963	18.689	15.475	<b>3.401</b>
	outdoor_night_1	96.481	3.599	1.836	<b>1.836</b>
	outdoor_night_2	74.392	11.939	9.179	<b>3.669</b>
	outdoor_night_3	42.366	6.569	3.880	<b>3.737</b>
	seq_loop_1	36.223	7.072	7.085	<b>0.414</b>
自制集	seq_loop_2	13.953	15.815	15.937	<b>2.926</b>
	seq_sim	39.192	6.408	6.403	<b>0.591</b>
模拟集	outdoor_day_1+	8.639	—	9.064	<b>6.681</b>
	seq_sim+	39.148	—	7.365	<b>0.601</b>

注: “—” 表示无有效结果.

此外,本节统计所提出算法在 outdoor\_day\_2 序列上的时间消耗,其中主要模块计算时间的统计值如表7所示.所提出算法采用多线程方式实现,包括视觉惯导模块部分的“特征跟踪”线程和“滑动窗口优化”线程、激光“扫描匹配”线程、“激光惯导融合”线程和“闭环优化”线程.除闭环优化模块外,其他模块的运行存在一定串联顺序关系,构成的里程计算法

以间隔1帧的激光雷达点云作为输入,可实现5 Hz 频率运行,并通过将扫描匹配位姿与IMU融合,能够实现IMU频率的高频位姿输出.

表7 各算法模块时间消耗 单位: ms

算法模块	运行时间	
	平均值	中位数
特征跟踪	16	14.6
滑动窗口	34.1	32.6
扫描匹配	94.8	80.1
激光惯导融合	1.9	1.3
闭环优化	234.7	211.1

2.6 不同算法整体性能对比实验

本节将所提出算法 VLIO-Loop 与 A-LOAM、LIO-SAM、LIOM、R<sup>2</sup>LIVE<sup>[34]</sup> 和 LVI-SAM<sup>[35]</sup> 等算法进行对比,其中LIO-SAM和LVI-SAM等算法需依赖IMU姿态信息,为使其能够在MVSEC数据集上运行,本文采用互补滤波算法进行IMU姿态解算.各算法在测试序列中的绝对轨迹误差结果如表8所示,其中MVSEC数据集为典型城市场景,环境结构与图像纹理信息丰富,各算法的轨迹误差结果近似.在 outdoor\_day 序列中,所提出算法的绝对轨迹误差最小;而在 outdoor\_night 序列中,所提出算法仍具有较佳的性能,并且在 outdoor\_night\_2 序列中成功检测到有效闭环约束,显著降低了轨迹误差.

表8 不同算法的绝对轨迹误差对比

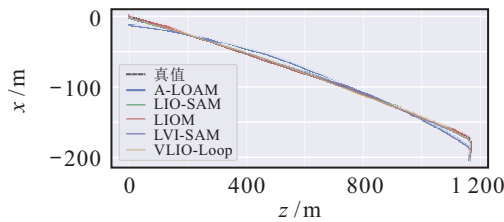
数据集	序列	最大运动速率/(m/s)	A-LOAM <sup>[32]</sup> /m	LIO-SAM <sup>[10]</sup> /m	LIOM <sup>[8]</sup> /m	R <sup>2</sup> LIVE/m	LVI-SAM <sup>[35]</sup> /m	VLIO-Loop/m
MVSEC	outdoor_day_1	7.6	2.247	2.054	1.826	—	2.630	<b>1.541</b>
	outdoor_day_2	12	18.108	4.019	6.268	—	10.584	<b>3.401</b>
	outdoor_night_1	10.4	6.433	1.790	2.944	—	<b>1.729</b>	1.836
	outdoor_night_2	11.2	13.579	12.186	11.646	—	11.928	<b>3.669</b>
	outdoor_night_3	10	7.344	4.338	5.563	—	<b>3.309</b>	3.737
	seq_loop_1	0.5	25.284	7.161	9.587	4.861	6.989	<b>0.414</b>
自制集	seq_loop_2	0.5	18.624	16.688	22.279	6.619	16.297	<b>2.926</b>
	seq_sim	0.5	8.638	0.608	7.882	0.712	0.750	<b>0.591</b>
模拟集	outdoor_day_1+	7.6	46.226	—	11.737	—	16.488	<b>6.681</b>
	seq_sim+	0.5	—	—	8.038	0.601	<b>0.477</b>	2.984

注: “—” 表示无有效结果.

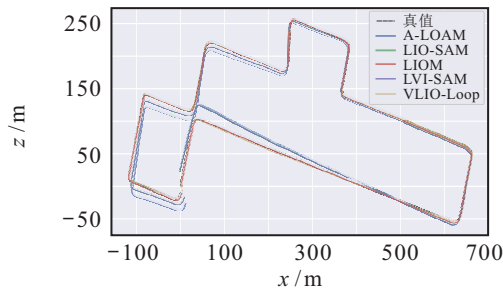
在MVSEC数据集上 outdoor\_night\_1 ~ outdoor\_night\_3 序列上的轨迹结果如图16所示.在这3个数据集上,A-LOAM的轨迹均偏离真值最为明显.在 outdoor\_night\_1 序列上,除A-LOAM外,其他方法的表现较为近似.在 outdoor\_night\_2 序列上,本文方法所得轨迹与真值的重合程度最高.由于VLIO-Loop

检测到了有效的闭环约束,轨迹误差大大降低.在 outdoor\_night\_3 序列上,从轨迹结果图像上看,VLIO-Loop 与 LVI-SAM 的轨迹误差最小,且轨迹的起点与终点基本重合.在自制数据集序列中,seq\_loop 序列仅包含起始-终止位置处闭环场景,所提出算法成功检测到闭环,轨迹精度最高.而在 seq\_sim 序列

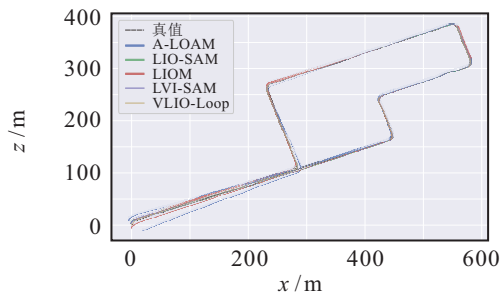
中, LIO-SAM、R2LIVE和LVI-SAM算法亦检测到较多有效闭环, 达到了与本文算法相近的结果. 在模拟数据集中, 由于部分时间段内环境结构特征缺乏, 各算法表现差异较大. 与A-LOAM、LIO-SAM和R2LIVE算法出现失败情况相比, LIOM、LVI-SAM和本文算法VLIO-Loop均具有更加鲁棒的位姿估计效果. 在outdoor\_day\_1+序列中, 各算法结果与原序列结果相比, 精度均有所下降, 而本文算法的相对下降程度最低. 在seq\_sim+序列中, 所提出算法取得了与LVI-SAM算法近似的结果, 并均通过闭环优化实现了比LIOM算法更加精确的轨迹估计.



(a) outdoor\_night\_1 轨迹俯视图



(b) outdoor\_night\_2 轨迹俯视图



(c) outdoor\_night\_3 轨迹俯视图

图 16 MVSEC 序列轨迹结果

### 3 结 论

本文提出了一种采用双目视觉信息辅助的激光惯导SLAM算法. 该算法通过视觉位姿为激光扫描匹配部分提供先验位姿约束, 与激光几何结构特征所构建的约束共同用于位姿估计, 使其可以应用于普通环境与结构特征缺乏环境; 其次, 为了避免视觉模块失效后仅采用激光数据进行位姿估计可能导致算法精度下降问题, 本文引入激光雷达与IMU数据融合模块进行相对位姿预测, 基于互补滤波算法与因子图优化策略快速对准激光里程计参考系和惯性参考

系, 保证激光-IMU数据正确融合并基于因子图策略约束IMU偏置; 此外, 本文提出了视觉激光混合闭环检测策略, 利用激光雷达观测视角范围大以及视觉检测机制不受距离限制的优势同时检测闭环, 并采用6自由度位姿图优化方法显著降低了轨迹误差, 完成了全局连续地图的构建.

### 参考文献(References)

- [1] Cadena C, Carlone L, Carrillo H, et al. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age[J]. IEEE Transactions on Robotics, 2016, 32(6): 1309-1332.
- [2] Gao H M, Zhang X B, Wen J, et al. Autonomous indoor exploration via polygon map construction and graph-based SLAM using directional endpoint features[J]. IEEE Transactions on Automation Science and Engineering, 2019, 16(4): 1531-1542.
- [3] Wen J, Zhang X B, Gao H M, et al. CAE-RLSM: Consistent and efficient redundant line segment merging for online feature map building[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(7): 4222-4237.
- [4] Sun Q X, Yuan J, Zhang X B, et al. Plane-edge-SLAM: Seamless fusion of planes and edges for SLAM in indoor environments[J]. IEEE Transactions on Automation Science and Engineering, 2021, 18(4): 2061-2075.
- [5] Velas M, Spänzel M, Herout A. Collar Line Segments for fast odometry estimation from Velodyne point clouds[C]. 2016 IEEE International Conference on Robotics and Automation. Stockholm, 2016: 4486-4495.
- [6] Zhang J, Singh S. Low-drift and real-time lidar odometry and mapping[J]. Autonomous Robots, 2017, 41(2): 401-416.
- [7] Shan T X, Englot B. LeGO-LOAM: Lightweight and ground-optimized lidar odometry and mapping on variable terrain[C]. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems. Madrid, 2018: 4758-4765.
- [8] Ye H Y, Chen Y Y, Liu M. Tightly coupled 3D lidar inertial odometry and mapping[C]. 2019 International Conference on Robotics and Automation. Montreal, 2019: 3144-3150.
- [9] Qin C, Ye H Y, Pranata C E, et al. LINS: A lidar-inertial state estimator for robust and efficient navigation[C]. 2020 IEEE International Conference on Robotics and Automation. Paris, 2020: 8899-8906.
- [10] Shan T X, Englot B, Meyers D, et al. LIO-SAM: Tightly-coupled lidar inertial odometry via smoothing and mapping[C]. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems. Las Vegas, 2020: 5135-5142.
- [11] Zhang J, Singh S. Visual-lidar odometry and mapping: Low-drift, robust, and fast[C]. 2015 IEEE International Conference on Robotics and Automation. Seattle, 2015:

- 2174-2181.
- [12] 彭道刚, 欧阳海林, 威尔江, 等. 结合改进密集模块深度估计网络和多视几何的视觉里程计[J]. 控制与决策, 2023, 38(4): 980-988.  
(Peng D G, Ouyang H L, Qi E J, et al. Visual odometry combined with depth estimation network of improved dense block and multi-view geometry[J]. Control and Decision, 2023, 38(4): 980-988.)
- [13] Shao W Z, Vijayarangan S, Li C, et al. Stereo visual inertial LiDAR simultaneous localization and mapping[C]. 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems. Macau, 2019: 370-377.
- [14] Galvez-López D, Tardos J D. Bags of binary words for fast place recognition in image sequences[J]. IEEE Transactions on Robotics, 2012, 28(5): 1188-1197.
- [15] Lepetit V, Moreno-Noguer F, Fua P. EPnP: An accurate  $O(n)$  solution to the PnP problem[J]. International Journal of Computer Vision, 2009, 81(2): 155-166.
- [16] Besl P J, McKay N D. A method for registration of 3-D shapes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(2): 239-256.
- [17] Wang Z Y, Zhang J H, Chen S Y, et al. Robust high accuracy visual-inertial-laser SLAM system[C]. 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems. Macau, 2019: 6636-6641.
- [18] Alliez P, Bonardi F, Bouchafa S, et al. Real-time multi-SLAM system for agent localization and 3D mapping in dynamic scenarios[C]. Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE, 2020: 4894-4900.
- [19] Zhu N, Ortiz M, Renaudin V, et al. Dataset of the intermediate competition in challenge MALIN: Indoor-outdoor inertial navigation system data for pedestrian and vehicle with high accuracy references in a context of firefighter scenario[J]. Data in Brief, 2021, 34: 106626.
- [20] Zuo X X, Geneva P, Lee W, et al. LIC-fusion: LiDAR-inertial-camera odometry[C]. 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems. New York: ACM, 2019: 5848-5854.
- [21] Zuo X X, Yang Y L, Geneva P, et al. LIC-fusion 2.0: LiDAR-inertial-camera odometry with sliding-window plane-feature tracking[C]. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems. Las Vegas, 2020: 5112-5119.
- [22] Wisth D, Camurri M, Das S, et al. Unified multi-modal landmark tracking for tightly coupled lidar-visual-inertial odometry[J]. IEEE Robotics and Automation Letters, 2021, 6(2): 1004-1011.
- [23] Qin T, Pan J, Cao S Z, et al. A general optimization-based framework for local odometry estimation with multiple sensors[J/OL]. 2019, arXiv: 1901.03638.
- [24] Andersen R. Modern methods for robust regression[M]. Los Angeles: Sage Publications, 2008.
- [25] Valenti R, Dryanovski I, Xiao J Z. Keeping a good attitude: A quaternion-based orientation filter for IMUs and MARGs[J]. Sensors, 2015, 15(8): 19302-19330.
- [26] Carlone L, Kira Z, Beall C, et al. Eliminating conditionally independent sets in factor graphs: A unifying perspective based on smart factors[C]. 2014 IEEE International Conference on Robotics and Automation. Hong Kong, 2014: 4290-4297.
- [27] 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.
- [28] Calonder M, Lepetit V, Strecha C, et al. BRIEF: Binary robust independent elementary features[C]. 2010 European Conference on Computer Vision. Berlin, 2010: 778-792.
- [29] Zhu A Z, Thakur D, Özaslan T, et al. The multivehicle stereo event camera dataset: An event camera dataset for 3D perception[J]. IEEE Robotics and Automation Letters, 2018, 3(3): 2032-2039.
- [30] Sturm J, Engelhard N, Endres F, et al. A benchmark for the evaluation of RGB-D SLAM systems[C]. 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. Vilamoura-Algarve, 2012: 573-580.
- [31] Michael G. EVO: Python package for the evaluation of odometry and SLAM[EB/OL]. (2022-03-29)[2022-10-15]. <https://github.com/MichaelGrupp/evo>.
- [32] Qin T, Cao S Z. A-LOAM: Advanced implementation of LOAM[EB/OL]. (2022-01-21)[2022-10-20]. <https://github.com/HKUST-Aerial-Robotics/A-LOAM>.
- [33] Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? The KITTI vision benchmark suite[C]. 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, 2012: 3354-3361.
- [34] Lin J R, Zheng C R, Xu W, et al. R<sup>2</sup>LIVE: A robust, real-time, LiDAR-inertial-visual tightly-coupled state estimator and mapping[J]. IEEE Robotics and Automation Letters, 2021, 6(4): 7469-7476.
- [35] Shan T X, Englot B, Ratti C, et al. LVI-SAM: Tightly-coupled lidar-visual-inertial odometry via smoothing and mapping[C]. 2021 IEEE International Conference on Robotics and Automation. Xi'an, 2021: 5692-5698.

## 作者简介

刘辉(1997—), 男, 工程师, 硕士, 从事移动机器人定位与建图、多传感器融合SLAM等研究, E-mail: liuhui0@mail.nankai.edu.cn;

张雪波(1984—), 男, 教授, 博士生导师, 从事移动机器人学、视觉伺服、运动规划、无人车自主导航、强化学习与智能博弈、多摄像机网络等研究, E-mail: zhangxuebo@nankai.edu.cn;

李如意(2001—), 女, 硕士生, 从事移动机器人定位与建图的研究, E-mail: ruyi@mail.nankai.edu.cn;

苑晶(1980—), 男, 教授, 博士生导师, 从事移动机器人控制、移动机器人导航、SLAM、多机器人系统、目标跟踪等研究, E-mail: yuanj@nankai.edu.cn.