

视觉退化场景下以 IMU 为中心的多传感器自适应消元融合 SLAM 方法

万佳颖, 胡新荣

(武汉纺织大学 计算机与人工智能学院, 湖北 武汉 430200)

摘要: 针对视觉感知退化环境下, 自动驾驶车辆定位精度下降甚至定位失效的问题, 提出一种以惯性测量单元 (IMU) 为中心的视觉-LiDAR-IMU-GNSS 多传感器融合系统。该系统整体框架由 3 个子系统构成: 视觉-IMU 子系统 (VIS)、LiDAR-IMU 子系统 (LIS) 以及 GNSS-IMU 子系统 (GIS)。首先, 设计 IMU 机械化处理向 VIS 和 LIS 提供精确先验姿态信息, 以 IMU 为中心辅助其他子系统完成初始化过程; 其次, 将 IMU、视觉、LiDAR 与 GNSS 的观测量转化为误差因子, 构建多传感器融合因子图模型; 最后, 针对后端优化过程中计算量大的问题, 在滑动窗口内采用自适应消元方法, 将因子图转换为贝叶斯网络进行优化计算。在 KITTI 和 M2DGR 两个数据集中测试的平均轨迹误差分别为 2.379 m 和 2.699 m, CPU 平均占用率 31.9%, 相较于 VINS-Mono、LIO-SAM 和 LVI-SAM 算法, 本文方法平均定位精度分别提高了 77.26%、19.78% 和 45.75%。

关键词: 自动驾驶; 多传感器融合; SLAM; 因子图优化; 变量消元

DOI: 10.11907/rjdx.251806

中图分类号: TP319

文献标识码: A

文章编号: 1672-7800(XXXX)0XX-0001-09

扫描二维码阅读全文:



An IMU-Centric Multi-Sensor Fusion SLAM Method with Adaptive Elimination for Visually Degraded Scenarios

WAN Jiaying, HU Xinrong

(College of Computer and Artificial Intelligence, Wuhan Textile University, Wuhan 430200, China)

Abstract: To address the issue of degraded localization accuracy or even positioning failure in autonomous vehicles under visually degraded environments, this paper proposes an IMU-centric Global Navigation Satellite System (GNSS)-visual-LiDAR-IMU multi-sensor fusion SLAM system (ICAE-SLAM). The overall system framework is designed with three subsystems: the Visual-IMU system (VIS), the LiDAR-IMU system (LIS), and the GNSS-IMU system (GIS). First, the IMU mechanization process is designed to provide precise prior attitude information to both VIS and LIS, where the IMU-centric assistance in the initialization of other subsystems. Subsequently, observations from IMU, visual, LiDAR, and GNSS are transformed into error factors for constructing a multi-sensor fusion factor graph model. Finally, to tackle the high computational load in the back-end optimization process, an adaptive elimination method is adopted within the sliding window to convert the factor graph into a Bayesian network. Tested on the KITTI and M2DGR datasets, the proposed method achieves average trajectory errors of 2.379 m and 2.699 m, respectively, with an average CPU utilization rate of 31.9%. Compared to VINS-Mono, LIO-SAM, LVI-SAM algorithms, the localization accuracy is improved by 77.26%, 19.78% and 45.75% on average, respectively.

Key Words: autonomous driving; multi-sensor fusion; simultaneous localization and mapping (SLAM); factor graph optimization; variables elimination

0 引言

随着计算机技术、人工智能算法以及传感器感知技术

的飞速发展, 自动驾驶汽车已经实现从限定场景测试到城市道路运营的跨越式发展。高精度定位作为实现车辆自动驾驶的核心技术, 通过多传感器融合 SLAM (Simultaneous Localization and Mapping) 实现车辆厘米级精度的环境

收稿日期: 2026-02-04

作者简介: 万佳颖 (2000-), 女, 武汉纺织大学计算机与人工智能学院硕士研究生, 研究方向为多传感器融合 slam; 胡新荣 (1973-), 女, 博士, 武汉纺织大学计算机与人工智能学院教授、硕士生导师, 研究方向为图像处理与媒体计算、3D 计算机图形学、计算机视觉、虚拟现实和机器学习等。本文通讯作者: 胡新荣。

感知与位姿估计,可有效弥补单一传感器在复杂动态环境下的感知局限性。例如,LiDAR SLAM在缺乏明显几何结构的环境(如长走廊)中易发生退化^[1-3]。同样地,视觉SLAM在涉及激烈运动、快速变化的光线条件和无纹理环境的场景中面临挑战^[4-6]。

为解决上述问题,研究者们已经开发了各种LiDAR-视觉SLAM方法,包括V-LOAM^[7]、R²LIVE^[8]、FAST-LIVO^[9]、R³LIVE^[10]、LVI-SAM^[11]、FAST-LIVO2^[12]等,均集成了来自LiDAR和相机的信息。例如,Wen等^[13]提出LIVER算法,它通过视觉惯性辅助方法来弥补LiDAR退化造成的影响,有效提高了在相似区域和照明条件不佳情况下的定位鲁棒性和准确性。Cheng等^[14]提出一种视觉-LiDAR-IMU融合方法,通过LiDAR提供视觉特征的深度信息,提高建图精度,但在动态环境中关键帧筛选策略可能因运动模糊导致匹配失效。Lee等^[15]提出了Switch-SLAM算法,它通过利用当检测LiDAR里程计退化时,从LiDAR转换到视觉里程计的结构来实现高稳定性和准确性。然而,强调LiDAR退化时的切换策略,却未明确涉及视觉同样退化时的应对机制。此外,当面临超出系统处理能力的持续性退化时,上述方法均存在局限性。主要包括它们对感知传感器的依赖,以及对使用最大后验概率(MAP)估计的融合方式的依赖,如迭代卡尔曼滤波器和因子图优化。对于后者而言,由于图优化需要求解大规模非线性最小二乘问题,其计算复杂度随优化窗口内节点数量的增加而呈立方级增长,对系统的实时性构成严重挑战。因此,在这种融合框架下,长期累积的故障信息可能会对整体系统性能产生不利影响。

为解决上述问题,本文提出了一种高精度、强鲁棒以及低占用率的视觉-LiDAR-IMU-GNSS多传感器融合SLAM方法(ICAE-SLAM)。首先,设计IMU机械化处理向视觉-IMU子系统(VIS)和LiDAR-IMU子系统(LIS)提供精确先验姿态信息,以IMU信息为基准实现对其他子系统的快速初始化过程;其次,将IMU、视觉、LiDAR与GNSS的多个观测量转化为误差因子,构建新的多传感器融合因子图模型;最后,在后端优化过程中结合滑动窗口和自适应变量消元算法,将因子图转换为贝叶斯网络进行优化计算。本文主要贡献如下:

(1)构建了一个以IMU为中心的视觉-LiDAR-IMU-GNSS多传感器融合框架。该框架采用紧耦合的融合方式,即使在视觉传感器退化的极端条件下,依然能够依靠IMU子系统输出稳定且高精度的定位信息。

(2)提出了一种基于自适应策略的变量消元优化方法。通过动态调整因子权重与消元顺序,以优先保留关键状态变量并抑制异常观测的影响,大幅提升了定位系统的稳定性与优化效率。

(3)在公共复杂数据集KITTI^[16]和M2DGR^[17]上对提出的方法进行大量实验。实验结果表明,与SOTA方法相

比,该方法在复杂环境下具有更好的鲁棒性和准确性。

1 方法

1.1 状态变量定义

首先,定义在本文中使用的坐标系和符号。将世界坐标系定义为 W ,IMU主体坐标系定义为 B ,LiDAR坐标系定义为 L ,相机坐标系定义为 C 。为便于分析,假设IMU坐标系和车辆的机体坐标系保持一致。车辆的状态可以定义如公式(1)所示:

$$x_k = [p_{b_k}^w, v_{b_k}^w, q_{b_k}^w, b_a, b_g], k \in [0, n] \quad (1)$$

其中, x_k 表示第 k 个车辆状态变量,包括位置 $p_{b_k}^w$ 、速度 $v_{b_k}^w$ 、方向 $q_{b_k}^w$ 以及加速度计和陀螺仪偏差 b_a, b_g 。

1.2 系统框架

本文的系统框架如图1所示。整体框架以IMU为中心,围绕3个子系统构建:①VIS子系统;②GIS子系统;③LIS子系统。一旦GIS子系统(GNSS-IMU组合)初始化,IMU机械化则被用于为VIS子系统和LIS子系统提供准确的先验姿态信息。VIS子系统直接使用IMU姿态进行初始化,采用Lukas-Kanade(LK)光流法对检测到的Shi-Tomasi角点进行逐帧跟踪。对于LIS子系统,使用IMU姿态消除点云的偏移,为LiDAR-IMU里程计的优化提供初始值。在消元图优化中,通过自适应消元策略将IMU预积分残差、视觉-IMU里程计残差、LiDAR-IMU里程计残差和GNSS残差在因子图优化框架内紧密融合。

1.3 系统初始化

融合系统首要任务是对IMU进行准确初始化,如图2所示。采用基于因子图优化的GIS子系统实现IMU初始化,分别执行以下两种策略:①在静止状态下,通过检测零速条件可估计IMU的滚转角、俯仰角和陀螺仪零偏,为后续状态估计提供初值;②在动态状态下,GNSS提供高频、低漂移的绝对位姿,用于初始化IMU中的位姿、航向角与尺度信息。在VIS子系统中,通过GIS提供的先验姿态实现初始估计。首先,利用IMU的姿态信息粗略对齐多帧图像的坐标系,提高匹配效率;其次,采用Shi-Tomasi角点检测器提取特征点,通过LK光流法实现特征点的帧间关联;再次,利用匹配的特征点,结合IMU提供的初始姿态,估计相机的初始位姿;最后,通过三角测量计算空间特征点的初始三维位置,并利用关键帧优化视觉-IMU里程计,进一步精确化姿态估计。在LIS子系统中,为了提升点云配准精度,可利用IMU输出的姿态信息对初始位姿进行对齐,并对LiDAR扫描间的旋转畸变进行运动补偿。同时,为了减少计算量,可通过曲率筛选从点云中提取边缘和平面特征,并依赖IMU预积分提供相邻帧间的初始位姿变换。

1.4 因子图优化模型

构建的系统因子图模型如图3所示,主要包含3种类型的因子:①GNSS-IMU里程计因子;②视觉-IMU里程计

因子; ③LiDAR-IMU 里程计因子。

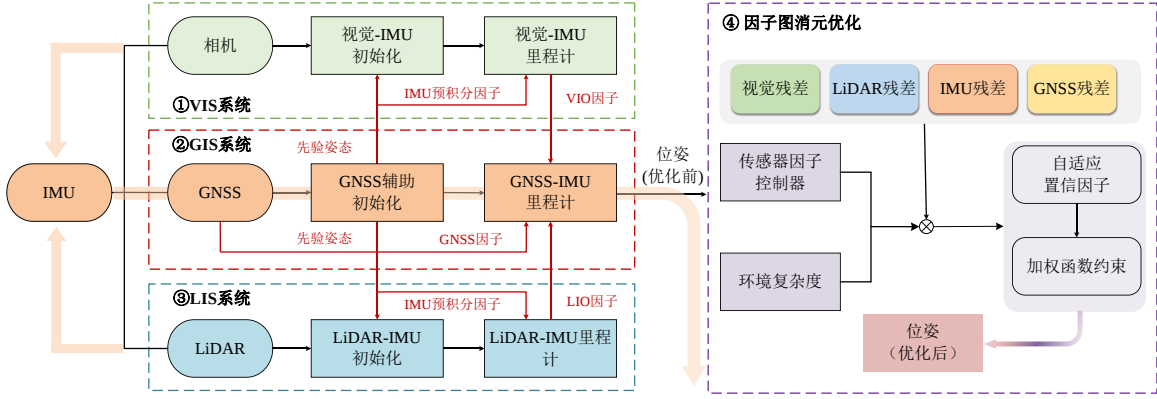


Fig. 1 The ICAE-SLAM algorithm framework

图1 ICAE-SLAM算法框架

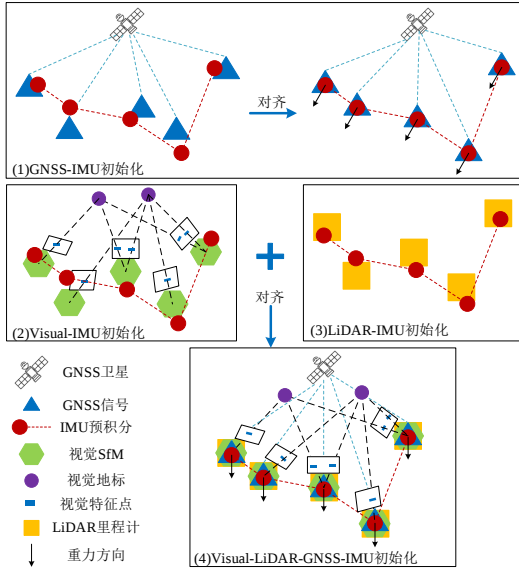


Fig. 2 System initialization

图2 系统初始化

1.4.1 GNSS-IMU 里程计因子

GNSS-IMU 里程计因子主要由 VIO 因子、LIO 因子、IMU 预积分因子、先验因子以及 GNSS 测量因子组成, 每个因子图节点均与车辆状态关联。节点间的相对位姿运动由 IMU 预积分获取, 同时, 融合视觉和 LiDAR 的相对状态估计作为局部约束来校正 IMU 预积分偏差。

由文献[18]可知, IMU 预积分因子可以被描述如公式(2)所示:

$$e_{ij}^{imu} = \begin{bmatrix} \left(R_{b_i}^w \right)^T \begin{pmatrix} p_{b_j}^w - p_{b_i}^w - v_{b_i}^w \Delta t_{ij} \\ -0.5 g^w \Delta t_{ij}^2 + \Delta p_{glcor, ij}^w \end{pmatrix} - \Delta \hat{p}_{ij}^{pre} \\ \left(R_{b_i}^w \right)^T \begin{pmatrix} v_{b_j}^w - v_{b_i}^w - g_{b_i}^w \Delta t_{ij} \\ + \Delta v_{glcor, ij}^w \end{pmatrix} - \Delta \hat{v}_{ij}^{pre} \\ 2 \left[\left(q_{b_j}^w \right)^{-1} \otimes q_{w_i}^w(t_i) \otimes q_{b_i}^w \otimes \hat{q}_{ij}^{pre} \right]_v \\ b_{g_j} - b_{g_i} \\ b_{a_j} - b_{a_i} \end{bmatrix} \quad (2)$$

其中, $\Delta p_{glcor, ij}^w$ 和 $\Delta v_{glcor, ij}^w$ 分别表示位置和速度预积分的科氏力; $\Delta \hat{p}_{ij}^{pre}$, $\Delta \hat{v}_{ij}^{pre}$ 和 $\hat{q}_{ij}^{pre}$ 分别表示位置、速度和姿态的预积分测量值。

GNSS-RTK 测量值的因子可以表示如公式(3)所示:

$$e_{GNSS} = p_{b_i}^w + R_{b_i}^w l_{GNSS}^b - \hat{p}_{GNSS, h}^w \quad (3)$$

其中, l_{GNSS}^b 表示 GNSS 与 IMU 之间的杆臂向量。

由于 LiDAR 和相机不是全局参考, 且里程计估计是基于车辆的第一姿态, 因此使用相对姿态 $[\Delta p_j^i, \Delta q_j^i]$ 作为局部约束, 来约束 IMU 预积分的位置 $\hat{\alpha}_j^i$ 和旋转 $\hat{\gamma}_j^i$ 。相对姿态因子 e_{ij}^{vio} , e_{ij}^{lio} 被定义如公式(4)所示:

$$e_{ij}^{lio}, e_{ij}^{vio} = \begin{bmatrix} \Delta p_j^i \\ \Delta q_j^i \end{bmatrix} \Theta \begin{bmatrix} \alpha_j^i \\ \gamma_j^i \end{bmatrix} \quad (4)$$

针对每一帧数据, 构建一个由相对姿态因子、GNSS 残差 e_{ij}^{gnss} 和边缘化先验残差 E_{prior} 组成的非线性优化问题, 其目标函数如公式(5)所示:

$$E = E_{prior} + \sum_{(i,j) \in B} e_{ij}^{vioT} W_{ij}^{-1} e_{ij}^{vio} + \sum_{(i,j) \in B} e_{ij}^{lioT} W_{ij}^{-1} e_{ij}^{lio} + \sum_{(i,j) \in B} e_{ij}^{gnssT} W_{ij}^{-1} e_{ij}^{gnss} \quad (5)$$

其中, B 表示所有关键帧的集合; W_{ij}^{-1} 表示协方差矩阵, 根据观测的可靠性来计算, 可用于衡量 VIO、LIO 和 GNSS 测量值对 IMU 预积分约束的置信度。

1.4.2 视觉-IMU 里程计因子

视觉-IMU 里程计因子在像素平面中观察到的特征可以表示为 $\tilde{p}_p$, 对于在第一个关键帧 i 和另一个观察到的关键帧 j 中具有逆深度 δ_l 的地标 l , 视觉重投影误差可以描述如公式(6)所示:

$$e_v = [b_1 \quad b_2]^T \cdot \left(\frac{\hat{p}_{c_j}}{\|\hat{p}_{c_j}\|} - \pi_c^{-1}(\tilde{p}_{p_i}) \right) \quad (6)$$

其中, π_c^{-1} 表示相机反投影函数, 将像素平面 p_p 中的特征转换为单位相机坐标系; b_1 和 b_2 表示跨越 $\hat{p}_{c_j}$ 的切平面的两个正交基。

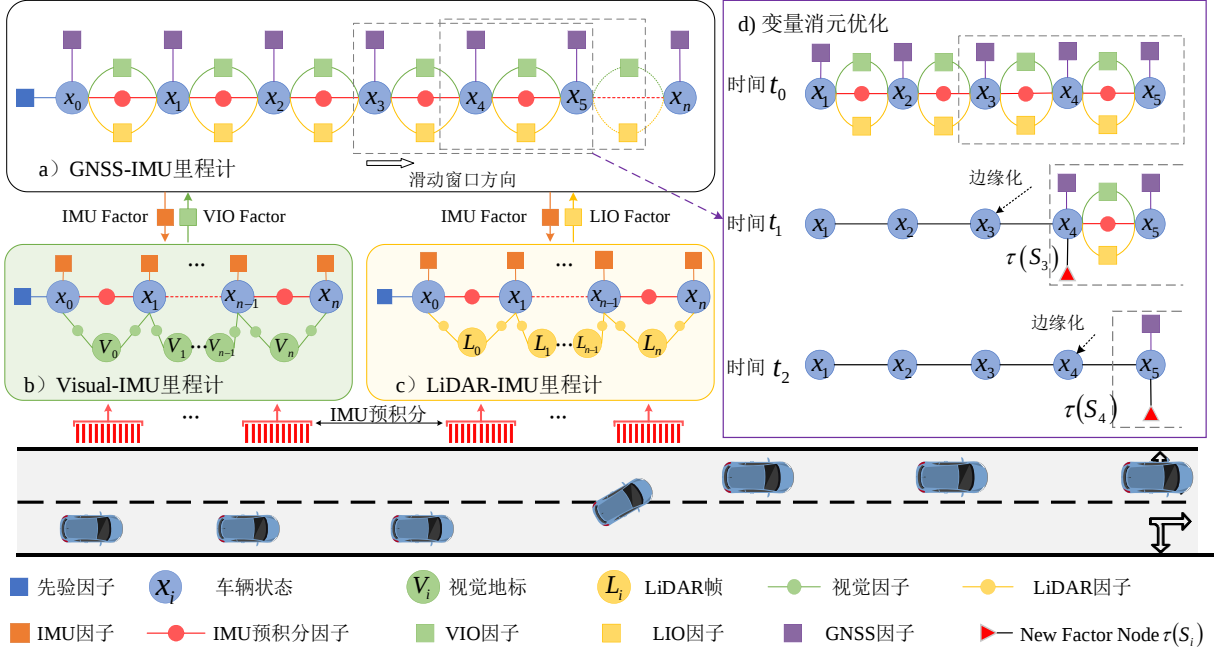


Fig. 3 The factor graph framework of ICAE-SLAM

图3 ICAE-SLAM的因子图框架

在VIS系统中,IMU预积分因子提供相邻帧的运动信息,GNSS-IMU里程计则提供绝对位姿信息。其视觉-IMU里程计最小化的非线性优化问题由视觉重投影因子 e_V 、IMU预积分因子 e_{imu} 、边缘化因子 e_m 和GNSS-IMU里程计的姿态先验 e^{prior} 组成。其表达式如公式(7)所示:

$$\min_{T_{i+1}} \left\{ \sum_{i \in obs(i)} e_V^T W_v^{-1} e_V + \sum_{(i, i+1) \in C} e_{imu}^T W_{imu}^{-1} e_{imu} + e_m + e^{prior} \right\} \quad (7)$$

其中, $i \in obs(i)$ 表示所有跟踪到特征点 i 的集合; $(i, i+1) \in C$ 表示IMU预积分因子连接的所有视觉观测节点的集合; W_v 、 W_{imu} 表示视觉重投影因子和IMU预积分因子的协方差矩阵。

1.4.3 LiDAR-IMU里程计因子

LiDAR-IMU里程计中,对LiDAR扫描的边缘和平面特征进行提取,并利用IMU因子进行运动预测。通过计算特征点到对应边缘和平面的距离,构建LiDAR误差因子。其表达式如公式(8)—公式(9)所示:

$$e_{e_i} = \frac{\left| (p_{i+1,k}^e - p_{i,u}^e) \times (p_{i+1,k}^e - p_{i,v}^e) \right|}{\left| p_{i,u}^e - p_{i,v}^e \right|} \quad (8)$$

$$e_{p_i} = \frac{\left| \begin{pmatrix} p_{i+1,k}^p - p_{i,u}^p \\ (p_{i,u}^p - p_{i,v}^p) \times (p_{i,u}^p - p_{i,w}^p) \end{pmatrix} \right|}{\left| (p_{i,u}^p - p_{i,v}^p) \times (p_{i,u}^p - p_{i,w}^p) \right|} \quad (9)$$

其中, k, u, v 和 w 表示LiDAR集合中的特征索引。边缘特征 $p_{i+1,k}^e$ 、 $p_{i,u}^e$ 和 $p_{i,v}^e$ 是形成体素地图中对应边缘线的点,平面特征 $p_{i+1,k}^p$ 、 $p_{i,u}^p$ 、 $p_{i,v}^p$ 和 $p_{i,w}^p$ 是形成体素地图的对应平面。

因此,构建一个由LiDAR因子、IMU预积分因子 e_{imu} 和GNSS-IMU里程计先验因子 e^{prior} 组成的非线性优化问题。

其表达式如公式(10)所示:

$$\min_{T_{i+1}} \left\{ \sum_{p_{i+1}^e \in F_{i+1}^e} W_i e_e + \sum_{p_{i+1}^p \in F_{i+1}^p} W_i e_p + \sum_{(i, i+1) \in R} W_i e_{imu} + e^{prior} \right\} \quad (10)$$

其中, $T_{i+1} = [R \quad t]$ 表示相邻帧特征点的最优变换矩阵, F_{i+1}^e 和 F_{i+1}^p 分别表示第 $i+1$ 个关键帧的边缘特征和平面特征。 W_i 表示LiDAR信息矩阵,用于评估当前LiDAR扫描的可观性。

1.5 自适应消元优化方法

1.5.1 滑动窗口优化

假设,滑动窗口中的状态变量 $N=3$,按照状态变量 x_3, x_4 依次消元转化为贝叶斯网络。首先,在 t_1 时刻对 x_3 进行消元,将历史状态 x_2 的信息作为先验信息传递至 x_3 ,从而生成新的因子节点 $\tau(S_3)$ 。随后,在 t_2 时刻,将状态变量 x_3 边缘化,对 x_4 进行消元并产生新因子节点 $\tau(S_4)$ 。由于窗口外的历史状态被压缩为先验因子,这一消元过程会保留所有相关变量以构成舒尔补,从而导致雅可比矩阵的维度显著增加。因此,本文提出一种基于QR分解的自适应滑动窗口消元方法。该方法以IMU状态为核心,在边缘化窗口内历史状态时,动态评估各传感器观测质量,同时,结合其状态与IMU之间的耦合程度,自适应决定消元顺序。

1.5.2 自适应消元策略

将IMU状态变量作为核心,优先保留在分离器 S_j 中,并延迟其消元。相应贝叶斯网络分解如公式(11)所示:

$$\begin{aligned} p(X) &= p(x_{imu} | S_{imu}) p(x_1 | S_1) p(x_2 | S_2) \cdots p(x_n | S_n) \\ &= p(x_{imu} | S_{imu}) \prod_j p(x_j | S_j) \end{aligned} \quad (11)$$

其中, x_{imu} 表示IMU的状态变量, S_{imu} 表示IMU的分离器, S_j 表示与变量 x_j (其他传感器) 相关的分离器。其他传

传感器变量的消元以IMU状态为条件,使得各传感器的信息仅与IMU状态相关,而不会直接相互影响。

为了实现自适应消元机制,在每个条件概率中引入置信权重 $\lambda_i \in [0, 1]$,用于控制传感器观测对估计结果的影响程度,其表达式如公式(12)所示:

$$p(x_i | S_i) \propto \exp\left(-\frac{1}{2} \lambda_i \cdot \|r_i(x_i, S_i)\|_{\Sigma_i^{-1}}^2\right) \quad (12)$$

其中, $r_i(\cdot)$ 为除IMU以外的传感器的残差项, λ_i 根据观测质量动态设定。以视觉为例,当特征点数量过少时,降低视觉惯性因子的置信权重 $\lambda_{visual} \rightarrow 0$,使其被弱化甚至剔除,从而抑制低质量视觉观测对IMU状态估计的干扰。

此时,IMU预积分的形式保持不变,但其影响相对增强,如公式(13)所示:

$$\varphi_{imu}(X_{imu}) = \exp\left\{-\frac{1}{2} \|A_{imu} X_{imu} - b_{imu}\|_{\Sigma_{imu}}^2\right\} \quad (13)$$

其中, X_{imu} 表示IMU的状态变量; A_{imu} 和 b_{imu} 表示与IMU相关的矩阵和向量; Σ_{imu} 表示协方差矩阵,其逆(信息矩阵)显著大于其他传感器,从而确保优化以IMU为核心。

从因子图 $\varphi_{j:n}$ 中消元变量 x_j 的算法步骤如算法1所示。

算法1 自适应消元优化算法

输入:状态变量 x_j ,传感器观测数据,传感器退化程度

输出:新因子 $\tau(S_j)$

1. 自适应选择消元变量 x_j ,优先选择非IMU状态变量,并根据各传感器实时退化程度,计算其置信权重 λ_{visual} 、 λ_{lidar} 、 λ_{gnss}
2. 根据变量 x_j 的综合置信度 Λ_j 升序排列,优选消去综合置信度低的变量
3. 构建分离器变量集合 $S(x_j) \leftarrow \phi_i(X_i)$,其中保留所有与IMU状态相关的变量在分离器中
4. 构造加权乘积因子 $\psi(x_j, S_j) \leftarrow \prod_i \lambda_i \phi_i(X_i)$
5. 对 x_j 边缘化并构成条件因子 $p(x_j | S_j) \leftarrow \psi(x_j, S_j)$,生成一个新因子 $\tau(S_j)$
6. 将新的因子 $\tau(S_j)$ 添加到因子图末尾,完成一次消元,进入下一轮消元

在视觉感知退化时(如特征缺失),通过增加IMU因子的置信度,从而抑制其他传感器的噪声影响。此时,贝叶斯网络的条件概率修改如公式(14)所示:

$$\begin{aligned} \psi(x_j, S_j) &\leftarrow \prod_i \varphi_i(X_i) \\ &= \exp\left\{-\frac{1}{2} \|\bar{A}_j[x_j, S_j] - \bar{b}_j\|_2^2\right\} \\ &= \exp\left\{-\frac{1}{2} \lambda_j \|R_j x_j + T_j S_j - d_j\|_2^2\right\} \\ &\quad \cdot \exp\left\{-\frac{1}{2} \lambda_{imu} \|\bar{A}_{imu} S_j - \bar{b}_{imu}\|_2^2\right\} \end{aligned} \quad (14)$$

其中, A_j 和 b_j 表示与传感器相关的矩阵与向量; $\bar{A}_j$ 和 $\bar{b}_j$ 分别表示所有 A_j 和 b_j 的合并; $\bar{A}_{imu}$ 和 $\bar{b}_{imu}$ 表示与IMU相关的矩阵与向量; R_j 和 T_j 表示其他传感器到IMU的旋转与平

移; d_j 表示对应的观测。通过将IMU作为多传感器融合中的中心节点,不仅能够提供高精度和高频的状态估计,而且能够增强系统的鲁棒性,尤其在感知退化的环境下。IMU的引入使得各个传感器的观测能够在IMU状态的基础上进行融合,从而优化整体状态估计的效率和精度。

1.5.3 自适应置信度的设计

由于传感器自身异常及环境中的复杂干扰,测量信息极易受到严重影响。为了在视觉感知退化环境下提高定位精度和系统稳定性,设计了自适应置信度权重。视觉退化因素反映了特征匹配的质量,并评估了相机、LiDAR和GNSS的性能下降程度,如公式(15)—公式(17)所示:

$$\lambda_{visual} = u/w \quad (15)$$

$$\lambda_{lidar} = \exp(-d_{mse}/\gamma^2) \quad (16)$$

$$\lambda_{gnss} = 1 - \exp\left(-\frac{\overline{C/N_0}}{\eta}\right) \quad (17)$$

其中, λ_{visual} 表示反映视觉传感器的退化程度。假设,在滑动窗口中跟踪的视觉特征越多,视觉信息越准确。 u 表示视觉特征跟踪的数量, w 表示滑动窗口的大小。 λ_{lidar} 表示反映LiDAR传感器的退化程度。 d_{mse} 表示LiDAR点云中参考特征和变换特征之间匹配的平均误差。 γ 表示距离尺度调整的阈值,其作用是确保误差处于一定的尺度范围内。 λ_{gnss} 表示反映GNSS传感器的退化程度。 $\overline{C/N_0}$ 表示所有跟踪卫星的平均载噪比,载噪比越高,函数值越接近1,表示退化程度越低。 η 表示一个归一化的参数。本文将 u 、 w 分别设置为8和15, γ 设置为0.2, η 设置为35。

为将上述置信度应用于因子图优化过程,本文在自适应消元策略中引入综合置信度,以确定消元变量的优先级。对于每个状态变量 x_j ,若其被多个传感器同时观测,则其综合置信度定义如公式(18)所示:

$$\Lambda_j = \min(\lambda_{visual}, \lambda_{lidar}, \lambda_{gnss}) \quad (18)$$

该定义体现了“木桶效应”,即变量的整体可靠性由最不可靠的传感器观测所决定。在自适应消元优化算法中,优先消去综合置信度 Λ_j 最低的状态变量,从而在优化过程中降低低质量观测对系统整体精度的影响。具体执行流程如下:①计算所有待消元状态变量(优先非IMU变量)对应的综合置信度 Λ_j ;②按 Λ_j 升序排列候选变量,形成消元队列;③依次对队列中的变量进行消元,每次消元后更新因子图结构;④重复以上步骤直至完成所有非IMU变量的消元,最终构建出轻量化的贝叶斯网络。

2 实验与分析

为验证本文ICAE-SLAM系统的综合性能与核心优势,在公开数据集KITTI与M2DGR上开展了全面且系统的实验评估。首先,在两组数据集上对系统开展性能测试,并从多维度评估子系统模块的性能分析;其次,在视觉感知退化场景下对系统进行鲁棒性测试,并通过轨迹精度衡

量系统的整体表现;最后,选取当前SLAM领域主流开源算法作为对比基线,在复杂真实场景下开展横向性能对比,充分凸显了本文算法的优越性。

为确保评估过程的客观与公平,所有算法实验均在Ubuntu 18.04的ROS平台上进行开发与运行。实验硬件平台为一台配备Intel Core i7-14650HX处理器和NVIDIA GeForce RTX 4060 GPU的笔记本电脑。本文ICAE-SLAM系统的核心子模块构建基于成熟开源框架,VIS采用VINS-Mono的基础框架,LIS部分基于LIO-SAM的核心逻辑实现,以确保系统基准性能的可靠性。

2.1 系统性能测试

本文对系统每个模块的定位性能进行测试,包含LIS、VIS、GIS以及有/无自适应消元的全组合系统的精度测试,同时,对各模块CPU资源占用率和计算耗时进行测试。本文选择在KITTI-07数据集上进行定位精度测试,实验结果如表1所示。各个单独子系统测试的绝对轨迹误差(Absolute Trajectory Error, APE)分别为0.926 m、1.839 m和1.479 m,相较于全组合系统分别增加了38.8%、175.7%和121.7%的误差。在全组合系统中,无自适应消元时APE为1.554 m,而引入自适应消元优化后APE显著下降了57.1%,相对轨迹误差(Relative Pose Error, RPE)也降低了49.7%。这充分验证了自适应消元方法在提升多传感器融合系统定位精度方面的有效性。

Table 1 Performance test of system modules

表1 系统各模块性能测试

	LIS	VIS	GIS	全组合模式	
				无消元	有消元
APE/m	0.926	1.839	1.479	1.554	0.667
RPE/m	1.982	1.217	0.101	0.195	0.098

在各模块CPU资源占用率和计算耗时方面,实验结果如表2所示。LIS、VIS和GIS模块的CPU占用率分别为5.2%、6.6%和4.3%,对应的计算耗时分别为4.8 ms、5.3 ms和3.8 ms,体现出较高的资源利用率和良好的实时性。初始化模块和自适应消元优化模块的CPU资源占用率分别为3.5%和12.3%,计算耗时分别为2.8 ms和8.1 ms。系统整体的CPU占用率为31.9%,时间耗时为24.8 ms,表明本文算法在保持精度的同时,具备良好的实时性能和计算效率。

Table 2 System CPU utilization and computation time

表2 系统运行CPU占用率和计算耗时

	初始化	前端位姿估计			后端消元优化	合计
		LIS	VIS	GIS		
CPU占用率/%	3.5	5.2	6.6	4.3	12.3	31.9
计算耗时/ms	2.8	4.8	5.3	3.8	8.1	24.8

为验证本文方法在高速运动导致视觉退化情况下的鲁棒性,将车辆速度设置为1×、3×和5×(对应原始速度为1.53 m/s)进行对比实验,实验结果如图4所示。随着速度的增加,图像序列间的帧间运动幅度显著增大,导致运动

模糊加剧,图像细节清晰度下降,纹理特征点匹配显著减少。在1倍速条件下,图像质量保持清晰,特征检测算法能够稳定提取大量角点,基于LK光流法的跟踪器可在连续帧间维持大量稳定的特征轨迹。然而,当播放速度提升至5倍速时,因帧间位移过大,图像模糊程度显著增加,特征点在不同帧中的表现差异剧增,导致光流法跟踪匹配失败。此时,光流跟踪出现大规模误匹配与跟踪丢失,VIS的特征跟踪前端失效,进而引起视觉里程计的退化(对应图4虚线框标注的“视觉退化区域”),但该方法仍能保持稳定的定位和建图。在x、y和z三维轨迹的对比中,该系统未显示出明显的累积误差,并始终与参考轨迹保持一致。结果表明,即使在视觉特征严重退化且相机快速运动导致视觉匹配不可靠的情况下,该系统仍能利用IMU提供的稳定运动约束来实现连续且稳定的位姿估计。

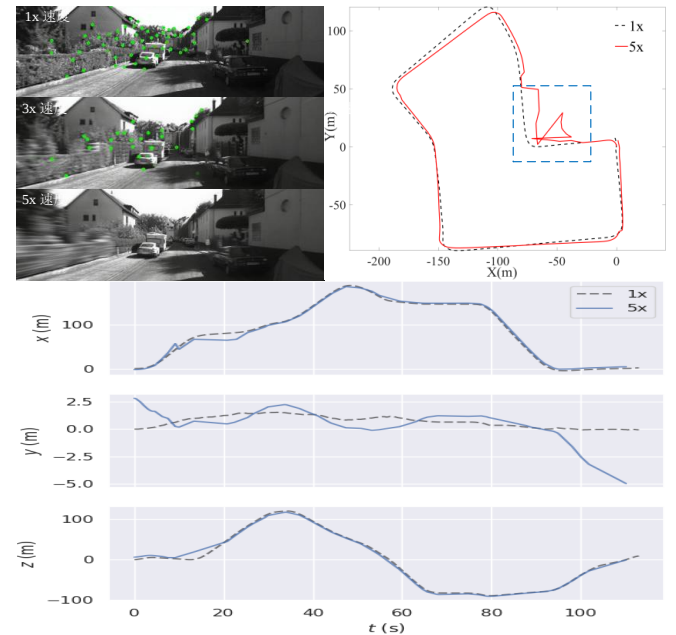


Fig. 4 Image frames and trajectory plots at different speeds

图4 不同速度下的图像帧与轨迹图

本文在KITTI序列05上对ICAE-SLAM和LIO-SAM的点云地图进行了比较,结果如图5所示。尽管LIO-SAM表现良好,但其性能受限于IMU数据的质量,在某些场景中会出现不稳定的情况。相比之下,本文方法生成的地图干净且对齐良好,没有重叠或重影现象,清晰地保留了道路边缘、车辆、树木。这得益于本文提出的自适应消除优化策略,该策略能够在复杂的动态场景中实现稳健的定位和出色的建图质量。

本文进一步比较了多个数据集序列的点云建图效果和定位精度,点云地图构建效果如图6所示,定位精度结果如表3所示。由图6可知,所生成的地图在几何结构还原和细节保留方面具有显著优势,尤其在M2DGR数据集的夜间低光照场景下,本文方法在特征点提取与轨迹呈现方面仍能保持较好的建图质量。由表3可知,本文方法在两个数据集中APE的平均值分别达到了2.379 m和2.699

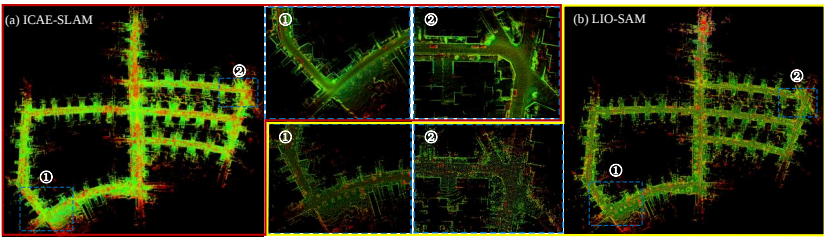


Fig. 5 The comparison results of ICAE-SLAM and LIO-SAM on sequence 05 of the KITTI dataset

图5 ICAE-SLAM和LIO-SAM在KITTI数据集序列05上的比较结果

m,RPE的平均值分别达到了0.191 m和0.227 m,表现出较 高的定位精度。

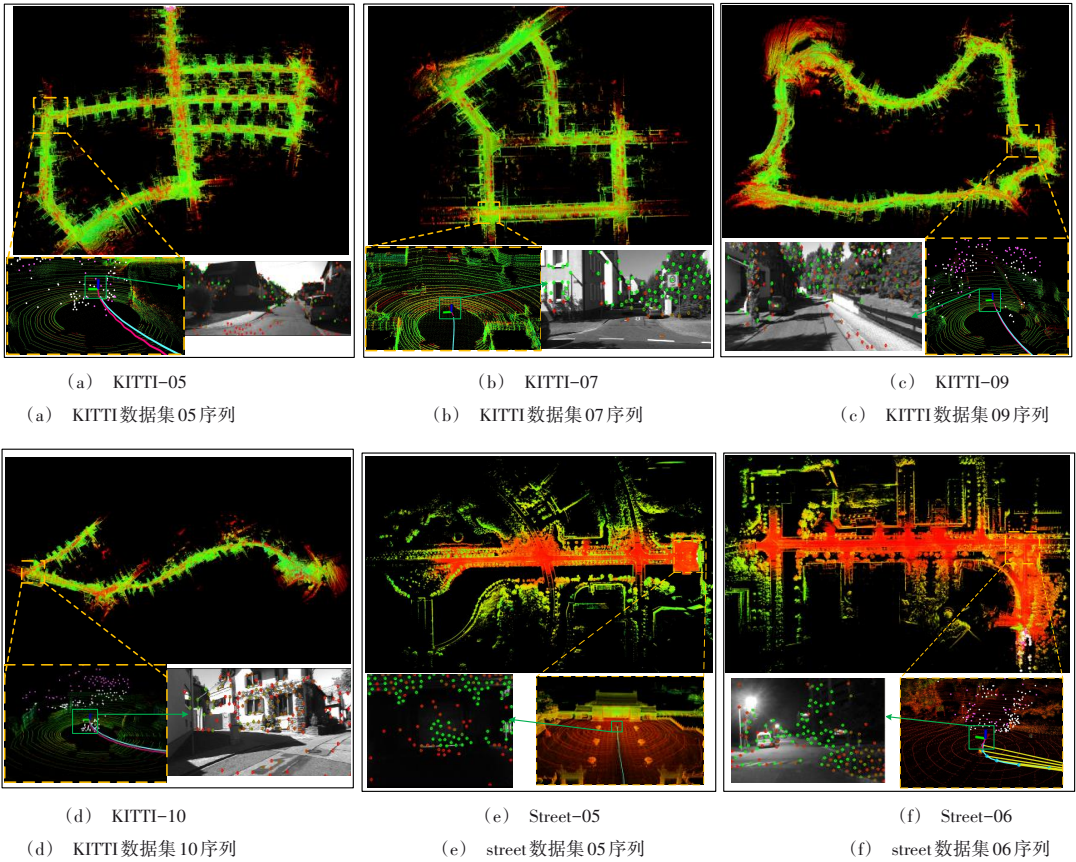


Fig. 6 The maps generated by the method proposed in this paper on the KITTI and M2DGR datasets

图6 本文方法在KITTI、M2DGR数据集中生成的地图

Table 3 The APE and RPE of the method proposed in this paper on the KITTI and M2DGR datasets

表3 本文方法在KITTI和M2DGR数据集集中的APE与RPE			
	序列	APE/m	RPE/m
KITTI	04	0.115	0.649
	05	1.210	0.098
	07	0.667	0.101
	09	7.771	0.078
	10	2.136	0.029
	平均值	2.379	0.191
	gate-01	2.190	0.177
M2DGR	gate-02	3.083	0.209
	gate-03	2.059	0.226
	street_05	4.014	0.211
	street_06	2.148	0.311
	平均值	2.699	0.227

2.2 对比实验

为验证本文方法的优越性,在不同数据集序列下与LIO-SAM、LVI-SAM和VINS-Mono进行比较,轨迹对比情况如图7所示。由图7可知,VINS-Mono的轨迹用灰色点划钱表示,LIO-SAM用灰色点线表示,LVI-SAM用灰色虚线表示,地面真值用黑色虚线表示,本文方法用黑色实线表示。为增强各种算法获得的轨迹对比效果,使用黑框放大轨迹内的特定细节。在放大的区域可以明显看出,本文方法的轨迹最接近真值,这表明本文方法可以产生比其他算法更准确的定位结果。其定位量化结果如表4所示。本文算法在各个序列上均取得最高定位精度。相较于VINS-Mono、LIO-SAM和LVI-SAM,本文方法在APE方面分别提高了约77.26%、19.78%和45.75%,在RPE方面分别

提高了约83.25%、76.30%和57.38%。

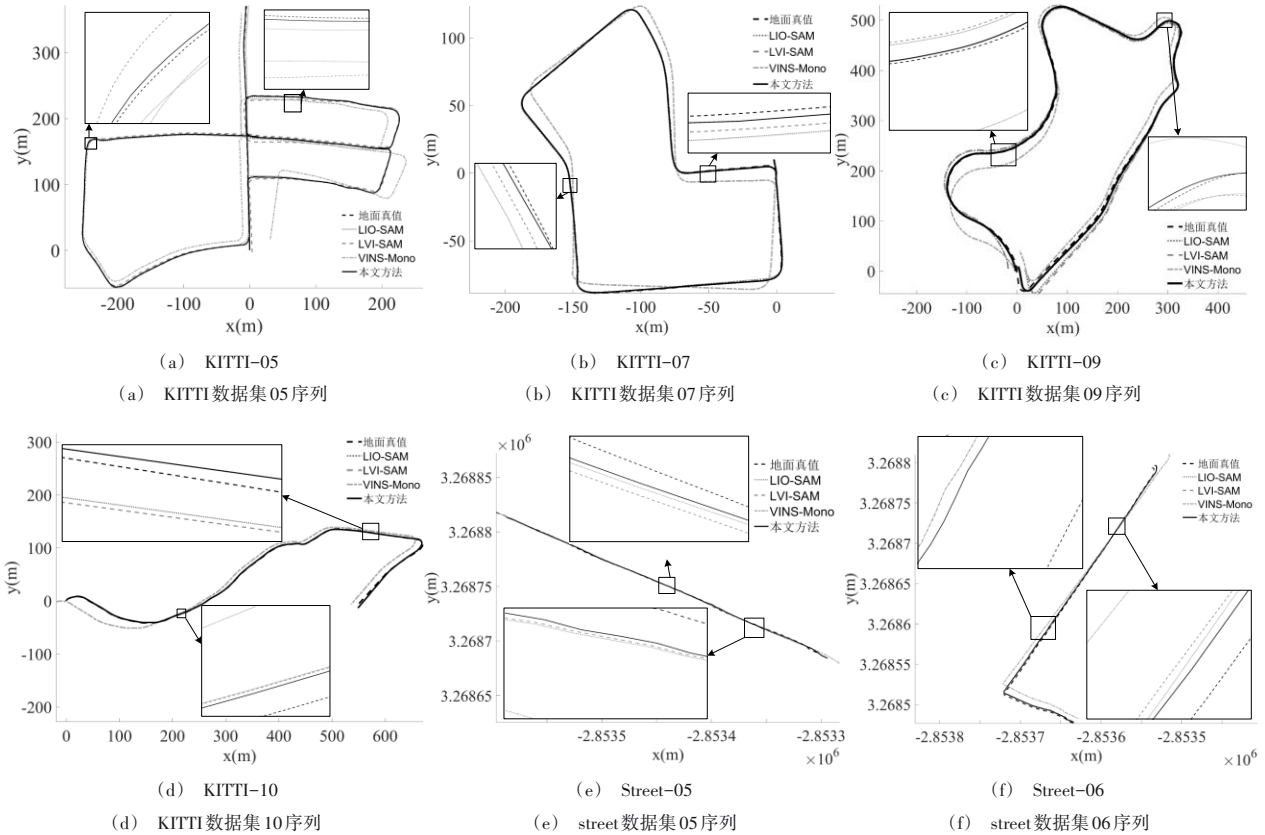


Fig. 7 Comparison of trajectories of different algorithms on the KITTI and M2DGR datasets

图7 不同算法在KITTI、M2DGR数据集上的轨迹对比

Table 4 Comparison of APE and RPE of different methods

表4 不同方法的绝对轨迹和相对位姿轨迹误差的对比

序号	VINS-Mono		LIO-SAM		LVI-SAM		本文方法	
	APE/m	RPE/m	APE/m	RPE/m	APE/m	RPE/m	APE/m	RPE/m
KITTI-04	15.479	2.141	0.296	0.823	9.232	2.359	0.115	0.649
KITTI-05	10.856	2.114	2.197	1.619	8.427	0.213	1.210	0.098
KITTI-07	11.839	1.217	0.947	1.981	1.554	0.395	0.667	0.101
KITTI-09	14.640	1.592	9.642	1.199	9.746	0.145	7.771	0.078
KITTI-10	13.832	1.260	2.184	1.457	2.389	0.466	2.136	0.029
gate-01	6.022	0.266	2.508	0.228	2.561	0.227	2.190	0.177
gate-02	7.333	0.509	3.436	0.246	3.465	0.248	3.083	0.209
gate-03	10.583	0.703	2.200	0.618	2.370	0.237	2.059	0.226
street-05	7.798	1.472	5.669	0.317	4.315	0.222	4.014	0.211
street-06	13.276	1.196	2.576	0.330	2.750	0.389	2.148	0.311
平均值	11.166	1.247	3.165	0.882	4.681	0.490	2.539	0.209

3 结语

针对视觉感知退化场景,本文提出一种以IMU为中心的视觉-LiDAR-IMU-GNSS多传感器融合系统。通过GNSS-IMU子系统对IMU进行精确初始化,并为VIS子系统和LIS子系统提供统一且高精度的先验姿态信息,从而有效提升了系统整体的一致性与稳定性。为解决复杂因

子图模型的计算效率问题,本文设计了一种基于自适应策略的消元优化算法。该算法能够动态调整因子权重与消元顺序,将因子图转化为贝叶斯网络,在保证定位精度的同时,显著降低了计算量,进而有效提升了系统的实时性能。在KITTI和M2DGR数据集上的实验表明,与VINS-Mono、LIO-SAM和LVI-SAM相比,本文方法在定位精度方面展现出显著优势,CPU平均占用率31.9%,验证了所提框架的有效性。在自动驾驶应用中,系统在GNSS拒止(如隧道)或激光雷达感知退化(如雨雪)等极端场景下仍存在一定的局限性。后续研究将深入挖掘多传感器融合的协同优势,增强系统在极端场景下的感知适应能力与定位可靠性,为自动驾驶提供更加稳定、安全的环境感知支撑。

参考文献:

- [1] ZHANG J, SINGH S. Low-drift and real-time lidar odometry and mapping [J]. Autonomous Robots, 2017, 41(2): 401-416.
- [2] SHAN T, ENLOT B, MEYERS D, et al. LIO-SAM: tightly-coupled lidar inertial odometry via smoothing and mapping [C]//IEEE/RSJ International Conference on Intelligent Robots and Systems, 2020: 5135-5142.
- [3] XU W, CAI Y, HE D, et al. FAST-LIO2: fast direct LiDAR-inertial odometry [J]. IEEE Transactions on Robotics, 2022, 38(4): 2053-2073.
- [4] MUR-ARTAL R, MONTIEL J M M, TARDOS J D. ORB-SLAM: a ver-

- satellite and accurate monocular SLAM system[J]. IEEE Transactions on Robotics, 2015, 31(5): 1147–1163.
- [5] ENGEL J, SCHÖPS T, CREMERS D. LSD-SLAM: large-scale direct monocular SLAM[C]//European Conference on Computer Vision, 2014: 834–849.
- [6] 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.
- [7] ZHANG J, SINGH S. Visual-lidar odometry and mapping: low-drift, robust, and fast[C]//IEEE International Conference on Robotics and Automation, 2015: 2174–2181.
- [8] LIN J, ZHENG C, XU W, et al. R2 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.
- [9] ZHENG C, ZHU Q, XU W, et al. FAST-LIVO: fast and tightly-coupled sparse-direct LiDAR-inertial-visual odometry[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems, 2022: 4003–4009.
- [10] LIN J, ZHANG F. R3LIVE: a robust, real-time, RGB-colored, LiDAR-inertial-visual tightly-coupled state estimation and mapping package[C]//International Conference on Robotics and Automation, 2022: 10672–10678.
- [11] SHAN T, ENGLT B, RATTI C, et al. LVI-SAM: tightly-coupled LiDAR-visual-inertial odometry via smoothing and mapping[C]//IEEE International Conference on Robotics and Automation, 2021: 5692–5698.
- [12] ZHOU B, ZHENG C, WANG Z, et al. FAST-LIVO2: LiDAR-inertial-visual odometry with efficient memory and computation[J]. IEEE Robotics and Automation Letters, 2025, 10(8): 7931–7938.
- [13] WEN T, FANG Y, LU B. LIVER: a tightly coupled LiDAR-inertial-visual state estimator with high robustness for underground environments[J]. IEEE Robotics and Automation Letters, 2024, 9(3): 2399–2406.
- [14] CHENG J, ZHANG L, CHEN Q, et al. High precision and robust vehicle localization algorithm with visual-LiDAR-IMU fusion[J]. IEEE Transactions on Vehicular Technology, 2024, 73(8): 11029–11043.
- [15] LEE J, KOMATSU R, SHINOZAKI M, et al. Switch-slam: switching-based LiDAR-inertial-visual slam for degenerate environments[J]. IEEE Robotics and Automation Letters, 2024, 9(8): 7270–7277.
- [16] GEIGER A, LENZ P, STILLER C, et al. Vision meets robotics: the KITTI dataset[J]. The International Journal of Robotics Research, 2013, 32(11): 1231–1237.
- [17] YIN J, LI A, LI T, et al. M2DGR: a multi-sensor and multi-scenario SLAM dataset for ground robots[J]. IEEE Robotics and Automation Letters, 2021, 7(2): 2266–2273.
- [18] 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.

(责任编辑:陈维捷)