

DOI: 10.19650/j.cnki.cjsi.J2107793

一种可参数自标定的鲁棒视觉/惯性定位方法^{*}

杨子寒, 赖际舟, 吕品, 袁诚, 刘建业
(南京航空航天大学自动化学院 南京 211106)

摘要:以视觉/惯性里程计为代表的视觉/惯性定位方法近年来被广泛应用。传统的视觉/惯性里程计通过离线标定方法得到固定的相机畸变参数,当相机畸变参数标定不准确或发生变化时定位精度会下降。针对于此,提出了一种面向相机畸变参数在线自标定的鲁棒视觉/惯性里程计方法:首先,将相机畸变参数加入到视觉/惯性里程计的待优化变量中,推导了视觉重投影误差关于待优化变量的雅可比矩阵;然后通过因子图优化技术,实现相机畸变参数的在线自标定与载体导航信息的实时优化求解;最后,通过 EuRoC 数据集试验和实际试验验证了所提方法的有效性。实际试验结果表明:相对传统的视觉/惯性里程计方法,所提方法在室外开阔场景中精度提升了 65.40%。

关键词:视觉/惯性里程计;因子图优化;在线自标定

中图分类号: TH86 TP242.6 **文献标识码:** A **国家标准学科分类代码:** 510.80

A robust visual/inertial positioning method with parameter self-calibration

Yang Zihan, Lai Jizhou, Lyu Pin, Yuan Cheng, Liu Jianye

(College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: The visual/inertial positioning method represented by the visual/inertial odometry has been widely used recently. The traditional visual/inertial odometry obtains the fixed distortion parameters of camera through offline calibration methods. When the calibration of camera distortion parameters is inaccurate or the camera distortion parameters are changed, the positioning accuracy will decrease. Aiming at this problem, a robust visual/inertial odometry method facing to online self-calibration of camera distortion parameters is proposed. Firstly, the camera distortion parameters are added to the variables to be optimized in the visual/inertial odometry. The Jacobian matrix of the visual reprojection error with respect to the variables to be optimized is derived. Then through the factor graph optimization method, online self-calibration of camera distortion parameters and real-time optimization solution of carrier navigation information are realized. Finally, the effectiveness of the proposed method was verified through the EuRoC dataset and real-world experiments. The real-world experiment result demonstrates that compared with the traditional visual/inertial odometry method, the accuracy of the proposed method is improved by 65.40% in an outdoor wide open scene.

Keywords: visual/inertial odometry; factor graph optimization; online self-calibration

0 引言

高精度定位技术是机器人导航与自动驾驶等领域的重要研究方向。在室外环境下,可以通过全球卫星导航系统(global navigation satellite system, GNSS)进行定位;在室内等 GNSS 拒止环境下,同时定位与地图构建(simultaneous localization and mapping, SLAM)是目前常

用的导航方式,主要有视觉 SLAM^[1-7]和激光雷达 SLAM^[8-12]。视觉传感器相比于激光雷达传感器具有感知信息丰富、体积小、功耗低和成本低等优点。然而,单一的视觉传感器易受外界环境干扰,无法有效跟踪载体的快速运动,且单目相机无法感知环境的真实深度。惯性传感器的引入可以有效改善上述问题,因此视觉/惯性里程计(visual inertial odometry, VIO)得到了广泛关注。

VIO 是指融合视觉传感器信息、惯性测量单元

收稿日期:2021-04-21 Received Date: 2021-04-21

^{*} 基金项目:国家自然科学基金(61973160)、航空科学基金(2018ZC52037)、工信部民机专项(2018-S-36)资助

(inertial measurement unit, IMU) 信息进行六自由度导航参数解算的方法, 可以分为基于滤波的 VIO^[13-16] 和基于优化的 VIO^[17-23] 两大类。而根据惯性信息与视觉的融合程度, 还可以将 VIO 分为松耦合和紧耦合^[24] 两类。松耦合是将 IMU 与视觉各自位姿估计的结果进行融合, 紧耦合则联合 IMU 约束与视觉约束优化求解载体位姿与特征点位置。基于滤波的 VIO 方法主要采用扩展卡尔曼滤波 (extended kalman filter, EKF) 对状态量进行预测和更新, 目前的经典方案有多状态约束卡尔曼滤波器^[13-14] (multi-state constraint kalman filter, MSCKF)、鲁棒 VIO^[15] (robust visual inertial odometry, ROVIO) 等。相比于传统的基于 EKF 的 VIO 算法, MSCKF 没有在系统中维护数量庞大的特征点坐标, 而是通过零空间投影使重投影误差与视觉特征点无关, 降低了系统的运算量。基于优化的 VIO 方法通过联合视觉约束与惯性约束, 优化求解载体位姿与特征点坐标, 主要有基于关键帧的开放视觉/惯性 SLAM^[17] (open keyframe-based visual-inertial SLAM, OKVIS)、单目视觉惯性系统^[18] (monocular visual-inertial system, VINS-Mono) 等方案。OKVIS 和 VINS-Mono 都是基于关键帧优化的 VIO 方法, 视觉误差项为重投影误差, 惯性误差项为预积分误差。但 VINS-Mono 在系统级别上更为完整, 具备自动初始化、回环检测和全局位姿图优化等功能。

相机畸变参数反应了相机所成的像相对于真实物体的失真程度以及图像的形变程度。畸变主要有径向畸变和切向畸变两种。为了增大相机的感知范围, 相机的前方会加入透镜, 透镜的形状会引起径向畸变; 切向畸变是在相机的生产过程中不能保证透镜和成像平面严格平行而引起的。在 VIO 系统中, 相机畸变参数会影响导航信息的解算精度, 错误的畸变参数会导致解算精度的降低。目前, 相机畸变参数通常采用离线标定方法估计, 经典的离线标定方法有张正友标定法^[25] 和 Kalibr 标定工具箱, 这两种方法均需要标定板 (棋盘格、Aprilgrid 等)。然而, 相机在长期使用过程中畸变参数可能会发生变化, 从而影响 VIO 系统精度。

为了解决此问题, 文献[26] 在 EKF 框架下实现了相机畸变参数的在线标定, 然而该算法只进行了一次线性化, 影响了定位精度。文献[27-28] 在图优化框架下实现了相机畸变参数的在线标定, 但都只建立了相机畸变的一阶近似模型, 对于畸变的描述精度不足。

针对此问题, 本文提出了一种面向相机畸变参数在线自标定的鲁棒 VIO 方法。主要贡献如下:

1) 构建了相机高阶畸变参数因子, 并将其加入视觉重投影误差项中;

2) 改进了基于优化的紧耦合 VIO 方法, 实现在 VIO 系统框架下完成相机畸变参数的自标定, 解决了相机畸

变参数标定不准确或发生变化时 VIO 系统定位精度降低的问题。

1 相机畸变参数在线自标定的 VIO 架构

本文提出的相机畸变参数在线自标定的改进 VIO 构架与已有的 VIO 方法框架基本一致, 其改进主要在于对相机的畸变误差进行了实时估计与补偿, 提高了传统 VIO 算法的精度。改进后 VIO 框架分为前端和后端两个主要模块。前端模块用于对图像和 IMU 信息进行预处理, 包括图像信息的特征提取与跟踪以及 IMU 信息的预积分, 同时对特征点进行去畸变处理; 后端模块联合边缘化的先验信息、IMU 预积分残差、视觉重投影残差, 这些误差被用来估计和优化载体状态, 此外还用来实时估计相机畸变参数。本文提出的改进 VIO 框架的主要区别是, 现有的 VIO 方法将相机畸变参数作为常数。而本文提出的方法在前端对图像特征点进行去畸变操作, 并将相机畸变参数作为优化变量建模到视觉重投影误差中, 在后端联合视觉约束与 IMU 约束优化求解载体状态、特征点坐标以及相机畸变参数, 改进 VIO 的前端与后端共同实现了对相机畸变错误标定带来误差的实时估计与标定补偿。

2 改进 VIO 传感器数据预处理

2.1 改进 VIO 前端视觉处理

对于每一帧新采集的图像, 本文采用 KLT 稀疏光流算法^[29] 跟踪上一帧的图像特征, 同时通过随机抽样一致 (random sample consensus, RANSAC) 算法剔除错误跟踪的特征点。当跟踪的特征点数量小于设定的阈值时, 添加由 Shi-Tomasi 角点检测器^[30] 检测到的新特征。特征检测和匹配完成后, 将特征点投影到归一化图像平面上, 但此时的特征点是存在畸变的, 需要对其进行去畸变操作。随着相机制造工艺的提升, 切向畸变引入的畸变量大大减小, 因此径向畸变是图像产生畸变的主要因素, 所以本文仅考虑径向畸变。

径向畸变是随着与图像中心之间距离的增加而增加的, 因此可以用与图像中心距离有关的二次多项式函数来修正径向畸变, 传统在线标定 VIO 算法仅考虑一阶近似模型, 本文使用了径向畸变的高阶模型, 具体公式如下:

$$\begin{cases} x_d = x(1 + k_1 r^2 + k_2 r^4) \\ y_d = y(1 + k_1 r^2 + k_2 r^4) \end{cases} \quad (1)$$

其中, (x, y) 是去畸变后归一化平面点的坐标; (x_d, y_d) 是畸变点的坐标; k_1 和 k_2 为径向畸变参数; $r^2 = x^2 + y^2$ 。

采用逆畸变模型消除畸变:

$$\begin{cases} x = \frac{1 + 3k_1 r_d^2 + 5k_2 r_d^4}{1 + 4k_1 r_d^2 + 6k_2 r_d^4} x_d \\ y = \frac{1 + 3k_1 r_d^2 + 5k_2 r_d^4}{1 + 4k_1 r_d^2 + 6k_2 r_d^4} y_d \end{cases} \quad (2)$$

其中, $r_d^2 = x_d^2 + y_d^2$ 。

此外,如果从上一个关键帧跟踪到的特征点数量小于阈值或跟踪到的特征点平均视差大于阈值,则将该帧作为新的关键帧。

2.2 改进 VIO 的 IMU 预积分

IMU 可以测量载体的三轴加速度和角速度,原始加速度和角速度测量值 $\hat{a}_t^b$ 、 $\hat{\omega}_t^b$ 受加速度计和陀螺仪白噪声、随机游走的影响:

$$\begin{aligned} \hat{a}_t^b &= a_t^b + b_{a_t} + R_t^b g^w + n_a \\ \hat{\omega}_t^b &= \omega_t^b + b_{\omega_t} + n_{\omega} \end{aligned} \quad (3)$$

其中, n_a 和 n_{ω} 分别为加速度计和陀螺仪白噪声; b_{a_t} 和 b_{ω_t} 分别为加速度计和陀螺仪的随机游走,其导数为白噪声; g^w 为导航坐标系下的重力; R_t^b 为采样时刻导航坐标系到 IMU 坐标系的旋转矩阵。

相邻的两个图像帧第 k 帧和第 $k+1$ 帧之间存在多帧惯性测量值,相邻两惯性帧之间的预积分过程如下:

$$\begin{aligned} \hat{\alpha}_{i+1}^{b_k} &= \hat{\alpha}_i^{b_k} + \hat{\beta}_i^{b_k} \Delta t + \frac{1}{2} R(\hat{\gamma}_i^{b_k}) (\hat{a}_i^b - b_{a_i}) \Delta t^2 \\ \hat{\beta}_{i+1}^{b_k} &= \hat{\beta}_i^{b_k} + R(\hat{\gamma}_i^{b_k}) (\hat{a}_i^b - b_{a_i}) \Delta t \end{aligned} \quad (4)$$

$$\hat{\gamma}_{i+1}^{b_k} = \hat{\gamma}_i^{b_k} \otimes \left[\frac{1}{2} (\hat{\omega}_i^b - b_{\omega_i}) \Delta t \right]$$

其中, i 和 $i+1$ 为第 k 帧和第 $k+1$ 帧图像之间的惯性测量帧; $\hat{\alpha}_{i+1}^{b_k}$ 为位置预积分估计值; $\hat{\beta}_{i+1}^{b_k}$ 为速度预积分估计值; 旋转用四元数 γ 表示; $\hat{\gamma}_{i+1}^{b_k}$ 为旋转预积分, $R(\gamma)$ 表示将四元数转化为旋转矩阵; $\otimes$ 表示四元数的乘法运算; Δt 为第 i 帧惯性测量与第 $i+1$ 帧惯性测量之间的时间差。

初始时 $\hat{\alpha}_{b_k}^{b_k}$ 和 $\hat{\beta}_{b_k}^{b_k}$ 为 0, $\hat{\gamma}_{b_k}^{b_k}$ 为单位四元数,可以通过式(4),采用迭代的方式对第 k 帧和第 $k+1$ 帧图像之间所有的惯性传感器测量值进行预积分,最终得到相邻两图像帧之间的惯性预积分估计值 $\hat{\alpha}_{b_{k+1}}^{b_k}$ 、 $\hat{\beta}_{b_{k+1}}^{b_k}$ 和 $\hat{\gamma}_{b_{k+1}}^{b_k}$ 。

3 改进 VIO 算法设计

为了在保证定位精度的同时提高算法的实时性,本文采用了基于滑动窗口的紧耦合因子图框架来估计优化

VIO 的所有待优化变量。如图 1 所示为改进 VIO 流程的示意图,包括滑动窗口内的状态变量、视觉测量、惯性测量以及视觉特征点,以及采用因子图优化的方式联合视觉约束、IMU 约束优化求解状态变量以及特征点位置。

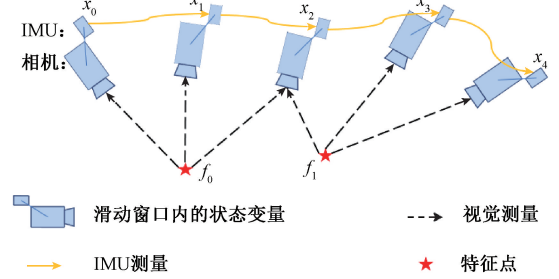


图 1 改进 VIO 流程示意图

Fig. 1 Schematic diagram of the improved VIO process

改进 VIO 因子图如图 2 所示,圆形节点表示需要优化的状态变量和特征点位置,带有方框的边表示视觉约束和 IMU 约束。与传统 VIO 不同,改进 VIO 的状态变量不仅包括 IMU 状态,还包括相机畸变参数。

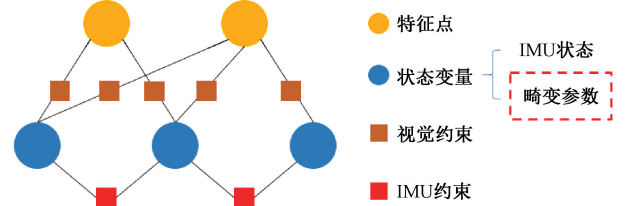


图 2 改进 VIO 因子图

Fig. 2 Factor graph of the improved VIO

传统 VIO 系统中滑动窗口内的所有待优化变量定义为:

$$\begin{aligned} X &= [x_0, x_1, \dots, x_n, P_0, P_1, \dots, P_m] \\ x_k &= [p_{b_k}^w, v_{b_k}^w, q_{b_k}^w, b_{a_{b_k}}, b_{\omega_{b_k}}], k \in [0, n] \end{aligned} \quad (5)$$

其中, x_k 为采集第 k 帧图像时对应的 IMU 状态,包括 IMU 在导航坐标系下的位置 $p_{b_k}^w$ 、速度 $v_{b_k}^w$, IMU 坐标系到导航坐标系的旋转的四元数 $q_{b_k}^w$ 以及加速度计和陀螺仪的随机游走 $b_{a_{b_k}}$ 、 $b_{\omega_{b_k}}$, $P_l (l \in [0, m])$ 为特征点位置。

本文所提方法在传统 VIO 系统中滑动窗口内所有待优化变量的基础上加入了相机畸变参数,改进 VIO 系统中滑动窗口内所有待优化变量定义如下:

$$\begin{aligned} X &= [x_0, x_1, \dots, x_n, x_c, P_0, P_1, \dots, P_m] \\ x_k &= [p_{b_k}^w, v_{b_k}^w, q_{b_k}^w, b_{a_{b_k}}, b_{\omega_{b_k}}], k \in [0, n] \\ x_c &= [k_1, k_2] \end{aligned} \quad (6)$$

其中, x_c 为相机畸变参数,包括径向畸变参数 k_1 和 k_2 。

本文通过最小化所有测量的残差之和来优化求解滑动窗口内的所有状态变量,目标函数定义为:

$$\min_X \left\{ \| \mathbf{r}_p - \mathbf{H}_p \mathbf{X} \|^2 + \sum_{k \in B} \| \mathbf{r}_B(\mathbf{z}_{k+1}^k, \mathbf{X}) \|^2_{P_{k+1}^k} + \sum_{(i,j) \in C} \| \mathbf{r}_C(\mathbf{z}_i^j, \mathbf{X}) \|^2_{P_i^j} \right\} \quad (7)$$

其中, $\| \mathbf{r}_p - \mathbf{H}_p \mathbf{X} \|^2$ 为边缘化约束; $\{ \mathbf{r}_p, \mathbf{H}_p \}$ 为边缘化的先验信息。在优化求解时,只有部分测量和状态变量保存在滑动窗口内,其它关键的状态和测量被边缘化并保存成先验信息,以提高运算速度的同时保留尽可能多的约束来提高解算精度。 $\mathbf{r}_B(\mathbf{z}_{k+1}^k, \mathbf{X})$ 为 IMU 预积分残差, B 是滑窗内所有 IMU 测量的集合, P_{k+1}^k 为 IMU 预积分误差的协方差。 $\mathbf{r}_C(\mathbf{z}_i^j, \mathbf{X})$ 为视觉残差, C 是滑窗内至少被两帧图像观测到的特征点集合, P_i^j 为视觉重投影误差的协方差。IMU 预积分残差与视觉残差需要通过它们各自协方差的逆将其转换成马氏距离进行加权求和。通过列文伯格-马夸尔特 (LM) 算法优化求解式(7)。

3.1 改进 VIO 的 IMU 预积分残差建立

IMU 预积分残差为预积分预测值与预积分估计值的差值,滑动窗口内第 k 帧和第 $k+1$ 帧图像之间的 IMU 预积分残差可以定义为:

$$\mathbf{r}_B(\mathbf{z}_{k+1}^k, \mathbf{X}) = \begin{bmatrix} \mathbf{R}_W^b(\mathbf{p}_{b_{k+1}}^W - \mathbf{p}_{b_k}^W - \mathbf{v}_{b_k}^W \Delta t_k + \frac{1}{2} \mathbf{g}^W \Delta t_k^2) - \hat{\alpha}_{b_{k+1}}^{b_k} \\ \mathbf{R}_W^b(\mathbf{v}_{b_{k+1}}^W - \mathbf{v}_{b_k}^W + \mathbf{g}^W \Delta t_k) - \hat{\beta}_{b_{k+1}}^{b_k} \\ 2[(\hat{\gamma}_{b_{k+1}}^{b_k})^{-1} \otimes (\mathbf{q}_{b_k}^W)^{-1} \otimes \mathbf{q}_{b_{k+1}}^W]_{xyz} \\ \mathbf{b}_{a_{b_{k+1}}}^{b_k} - \mathbf{b}_{a_{b_k}}^{b_k} \\ \mathbf{b}_{\omega_{b_{k+1}}}^{b_k} - \mathbf{b}_{\omega_{b_k}}^{b_k} \end{bmatrix} \quad (8)$$

其中, $\mathbf{R}_W^b(\mathbf{p}_{b_{k+1}}^W - \mathbf{p}_{b_k}^W - \mathbf{v}_{b_k}^W \Delta t_k + \frac{1}{2} \mathbf{g}^W \Delta t_k^2)$ 为位置预积分预测值; $\mathbf{R}_W^b(\mathbf{v}_{b_{k+1}}^W - \mathbf{v}_{b_k}^W + \mathbf{g}^W \Delta t_k)$ 为速度预积分预测值; $(\mathbf{q}_{b_k}^W)^{-1} \otimes \mathbf{q}_{b_{k+1}}^W$ 为旋转预积分预测值; $\mathbf{R}_W^b$ 为第 k 帧图像采集时刻导航坐标系到 IMU 坐标系的旋转矩阵; $\mathbf{p}_{b_k}^W$ 、 $\mathbf{p}_{b_{k+1}}^W$ 分别为第 k 帧和第 $k+1$ 帧图像采集时刻 IMU 在导航坐标系下的位置; $\mathbf{v}_{b_k}^W$ 、 $\mathbf{v}_{b_{k+1}}^W$ 分别为第 k 帧和第 $k+1$ 帧图像采集时刻 IMU 在导航坐标系下的速度; Δt_k 为第 k 帧和第 $k+1$ 帧图像之间的时间间隔; $\mathbf{q}_{b_k}^W$ 、 $\mathbf{q}_{b_{k+1}}^W$ 分别为第 k 帧和第 $k+1$ 帧图像采集时刻 IMU 坐标系在导航坐标系下旋转的四元数; $[\gamma]_{xyz}$ 表示提取四元数 γ 的向量部分,用于三维旋转残差的表示。 $\hat{\alpha}_{b_{k+1}}^{b_k}$ 、 $\hat{\beta}_{b_{k+1}}^{b_k}$ 和 $\hat{\gamma}_{b_{k+1}}^{b_k}$ 分别为位置预积分、速度预积分与旋转预积分的估计值,可通过式(4)计算得到。

3.2 改进 VIO 的视觉残差建立

在传统的特征点法视觉里程计方法中,视觉残差为重投影误差。特征点在视觉重投影误差中有两种参数化方法:导航系下特征点三维位置参数化和图像帧下特征点逆深度参数化。传统视觉重投影误差没有考虑相机的畸变参数,无法抑制离线标定不准确引入的误差,本部分构建了相机高阶畸变参数因子并通过两种特征点参数化方法将其建模到传统视觉重投影误差中,实现了相机畸变参数的在线自标定优化。

1) 特征点三维位置参数化:在传统参数化方法下,特征点被参数化为导航坐标系下的三维位置:

$$\mathbf{P}_l = [X_l, Y_l, Z_l]^T \quad (9)$$

在归一化图像平面上,传统视觉重投影误差为:

$$\mathbf{r}_C(\mathbf{z}_l^j, \mathbf{X}) = \mathbf{z}_l^j - \frac{1}{Z_l} (\mathbf{R}_l^c(\mathbf{R}_W^{b_j}(\mathbf{P}_l - \mathbf{t}_{b_j}^W) - \mathbf{t}_c^b)) \quad (10)$$

$$\mathbf{z}_l^j = [x_l^j, y_l^j]^T$$

其中, $\mathbf{z}_l^j$ 为特征点 l 在第 j 帧相机归一化图像平面上的坐标; $(\mathbf{R}_{b_j}^W, \mathbf{t}_{b_j}^W)$ 为第 j 帧图像采集时刻 IMU 在导航坐标系下的位姿; $(\mathbf{R}_c^b, \mathbf{t}_c^b)$ 为相机和 IMU 之间的外参。

根据式(2),本部分将相机畸变参数建模到视觉重投影误差中:

$$\mathbf{r}_C(\mathbf{z}_l^j, \mathbf{X}) = \mathbf{z}_l^j(k_1, k_2) - \frac{1}{Z_l} (\mathbf{R}_l^c(\mathbf{R}_W^{b_j}(\mathbf{P}_l - \mathbf{t}_{b_j}^W) - \mathbf{t}_c^b)) \quad (11)$$

$$\mathbf{z}_l^j(k_1, k_2) = \frac{1+3k_1r_d^2+5k_2r_d^4}{1+4k_1r_d^2+6k_2r_d^4} [x_l^j, y_l^j]^T$$

通过优化 IMU 位姿 $(\mathbf{R}_{b_j}^W, \mathbf{t}_{b_j}^W)$ 、特征点位置 $\mathbf{P}_l$ 和相机畸变参数 k_1 、 k_2 使重投影误差最小,本部分将相应的雅可比矩阵进行了推导:

$$\mathbf{J}^0 = \begin{bmatrix} \frac{\partial \mathbf{r}_C}{\partial \mathbf{t}_{b_j}^W} & \frac{\partial \mathbf{r}_C}{\partial \mathbf{q}_{b_j}^W} \end{bmatrix} =$$

$$[\mathbf{K}(\mathbf{R}_l^c \mathbf{R}_W^{b_j}) - \mathbf{K} \{ \mathbf{R}_l^c [\mathbf{R}_W^{b_j}(\mathbf{P}_l - \mathbf{t}_{b_j}^W)]^\wedge \}]$$

$$\mathbf{J}^1 = \begin{bmatrix} \frac{\partial \mathbf{r}_C}{\partial \mathbf{P}_l} \end{bmatrix} = [-\mathbf{K}(\mathbf{R}_l^c \mathbf{R}_W^{b_j})] \quad (12)$$

$$\mathbf{J}^2 = \begin{bmatrix} \frac{\partial \mathbf{r}_C}{\partial k_1} \end{bmatrix} = \frac{-(r_d^2 + 2k_2r_d^6)}{(1 + 4k_1r_d^2 + 6k_2r_d^4)^2} \begin{bmatrix} x_l^j \\ y_l^j \end{bmatrix}$$

$$\mathbf{J}^3 = \begin{bmatrix} \frac{\partial \mathbf{r}_C}{\partial k_2} \end{bmatrix} = \frac{-r_d^4 + 2k_1r_d^6}{(1 + 4k_1r_d^2 + 6k_2r_d^4)^2} \begin{bmatrix} x_l^j \\ y_l^j \end{bmatrix}$$

其中, $\mathbf{K}$ 是重投影误差关于归一化坐标系下特征点的导数,具体为:

$$\mathbf{K} = \begin{bmatrix} \frac{1}{Z_l} & 0 & -\frac{X_l}{Z_l^2} \\ 0 & \frac{1}{Z_l} & -\frac{Y_l}{Z_l^2} \end{bmatrix} \quad (13)$$

$[\cdot]^\wedge$ 运算符表示把三维实向量映射为 3×3 反对称阵。

2) 特征点逆深度参数化: 在传统参数化方法下, 特征点被参数化为某一图像帧下特征深度的倒数, 假设在第 i 帧图像上特征点 l 的逆深度为 λ_l , 则第 i 帧图像与第 j 帧图像关于特征点 l 的重投影误差为:

$$\begin{aligned} r_c(z_l^j, X) &= z_l^j - \frac{1}{Z_l^{c_j}} \cdot \\ R_b^c(R_w^{b_j}(R_{b_i}^w(R_c^b \frac{1}{\lambda_l} z_l^i + t_c^b) + t_{b_i}^w - t_{b_j}^w) - t_c^b) \\ z_l^i &= [x_l^i, y_l^i]^T, z_l^j = [x_l^j, y_l^j]^T \end{aligned} \quad (14)$$

其中, z_l^i 为特征点 l 在第 i 帧相机归一化图像平面上的坐标; $Z_l^{c_j}$ 为特征点 l 从第 i 帧经坐标变换到第 j 帧后在相机坐标系下 z 轴的分量; $(R_{b_i}^w, t_{b_i}^w)$ 为第 i 帧图像采集时刻 IMU 在导航坐标系下的位姿。

与式 (10) 类似, 本部分将相机畸变参数建模到式 (14) 中:

$$\begin{aligned} r_c(z_l^j, X) &= z_l^j(k_1, k_2) - \frac{1}{Z_l^{c_j}} \cdot \\ R_b^c(R_w^{b_j}(R_{b_i}^w(R_c^b \frac{z_l^i(k_1, k_2)}{\lambda_l} + t_c^b) + t_{b_i}^w - t_{b_j}^w) - t_c^b) \\ z_l^i(k_1, k_2) &= \frac{1 + 3k_1 r_d^2 + 5k_2 r_d^4}{1 + 4k_1 r_d^2 + 6k_2 r_d^4} [x_l^i, y_l^i]^T \\ z_l^j(k_1, k_2) &= \frac{1 + 3k_1 r_d^2 + 5k_2 r_d^4}{1 + 4k_1 r_d^2 + 6k_2 r_d^4} [x_l^j, y_l^j]^T \end{aligned} \quad (15)$$

通过优化两帧 IMU 的位姿 $(R_{b_i}^w, t_{b_i}^w)$ 、 $(R_{b_j}^w, t_{b_j}^w)$ 、特征点 l 的逆深度 λ_l 和相机畸变参数 k_1, k_2 使重投影误差最小, 本部分将相应的雅可比矩阵进行了推导:

$$\begin{aligned} J^0 &= \begin{bmatrix} \frac{\partial r_c}{\partial t_{b_i}^w} & \frac{\partial r_c}{\partial q_{b_i}^w} \end{bmatrix} = \\ &[-K(R_b^c R_w^{b_j}) \quad KR_b^c R_w^{b_j} R_{b_i}^w [Z_l^{b_i}]^\wedge] \\ J^1 &= \begin{bmatrix} \frac{\partial r_c}{\partial t_{b_j}^w} & \frac{\partial r_c}{\partial q_{b_j}^w} \end{bmatrix} = \\ &[K(R_b^c R_w^{b_j}) \quad -KR_b^c [R_w^{b_j}(R_{b_i}^w Z_l^{b_i} + t_{b_i}^w - t_{b_j}^w)]^\wedge] \end{aligned} \quad (16)$$

$$\begin{aligned} J^2 &= \begin{bmatrix} \frac{\partial r_c}{\partial \lambda_l} \end{bmatrix} = KR_b^c R_w^{b_j} R_{b_i}^w R_c^b \frac{1}{\lambda_l^2} z_l^i(k_1, k_2) \\ J^3 &= \begin{bmatrix} \frac{\partial r_c}{\partial k_1} \end{bmatrix} = K_1' \begin{bmatrix} x_l^j \\ y_l^j \end{bmatrix} - KR_b^c R_w^{b_j} R_{b_i}^w R_c^b \frac{1}{\lambda_l} K_1' \begin{bmatrix} x_l^i \\ y_l^i \end{bmatrix} \\ J^4 &= \begin{bmatrix} \frac{\partial r_c}{\partial k_2} \end{bmatrix} = K_2' \begin{bmatrix} x_l^j \\ y_l^j \end{bmatrix} - KR_b^c R_w^{b_j} R_{b_i}^w R_c^b \frac{1}{\lambda_l} K_2' \begin{bmatrix} x_l^i \\ y_l^i \end{bmatrix} \end{aligned}$$

其中,

$$\begin{aligned} Z_l^{b_i} &= R_c^b \frac{1}{\lambda_l} z_l^i(k_1, k_2) + t_c^b \\ K_1' &= \frac{-(r_d^2 + 2k_2 r_d^6)}{(1 + 4k_1 r_d^2 + 6k_2 r_d^4)^2} \\ K_2' &= \frac{-r_d^4 + 2k_1 r_d^6}{(1 + 4k_1 r_d^2 + 6k_2 r_d^4)^2} \end{aligned} \quad (17)$$

通过上述推导, 可以将相机畸变参数加入 VIO 系统进行统一优化求解, 得到更高精度的位姿信息。在实际应用中, 逆深度更加符合高斯分布, 同时也具有更高的数值稳定性, 因此在本文的试验中采用了逆深度参数化方法。

4 试验测试与分析

4.1 试验设置

为了验证本文算法的性能及有效性, 本文在 EuRoC 数据集^[31]与实际无人车平台上进行了试验验证。通过与 VINS-Mono 算法的对比, 验证当相机畸变参数标定不准确或发生变化时本文算法的优越性。本文试验采用的处理平台为 Inter CORE i7-8750H 处理器, 主频 2.2 GHz。

4.2 EuRoC 数据集试验与分析

EuRoC 数据集是苏黎世联邦理工学院开发的 VIO 测试数据集, 由微型无人机在两个室内场景中采集: 包括 20 Hz 的双目图像信息, 与图像信息同步的 200 Hz 的 IMU 测量信息, 由激光跟踪扫描仪、运动捕捉系统获取的无人机 3D 位置、6D 位姿信息, 以及传感器的标定参数。在本文测试中, 仅使用了左目相机的图像信息。图 3 为 EuRoC 数据集 MH_01_easy、MH_05_difficult 序列场景图。

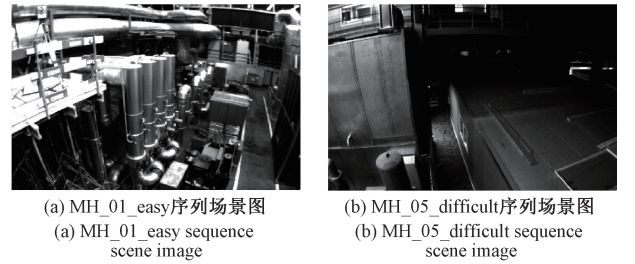


图3 EuRoC 数据集场景图

Fig. 3 Scene images of EuRoC dataset

EuRoC 数据集提供了较为精确的相机畸变参数作为参考值, 本文分别将其用于 VINS-Mono 和本文提出的方法, 对比定位精度。均方根误差 (root mean square error, RMSE) 常被用来评估 VIO 的定位精度, 开源工具包 evo 可以用来评估对比不同的 VIO 方法, 本文使用 evo 工具对比了 VINS-Mono 与本文所提方法的定位精度。

表 1 为 VINS-Mono 与本文所提方法在几个数据集序列上的定位均方根误差。从表 1 的结果中可以看出,即使 EuRoC 数据集提供了较为精确的相机畸变参数,但仍存在一定的误差,错误标定参数导致在进行特征点重投影时会引入额外的像素误差,造成后端优化时精度的下降。本文所提的方法由于在线优化补偿了相机畸变参数,因而定位精度优于 VINS-Mono。

表 1 VINS-Mono 与本文所提方法的定位均方根误差
Table 1 The positioning RMSE of VINS-Mono and the proposed method

数据集序列	VINS-Mono/m	本文所提方法/m	精度提升/%
MH_01_easy	0.154 6	0.129 9	15.98
MH_02_easy	0.178 2	0.155 4	12.79
MH_03_medium	0.195 2	0.193 4	0.92
MH_04_difficult	0.347 6	0.326 0	6.21
MH_05_difficult	0.302 5	0.290 2	4.07

图 4 是 VINS-Mono 与本文所提的方法在 MH_02_easy 序列上的解算轨迹与基准轨迹的对比图,解算轨迹颜色越红表明估计误差越大,颜色越紫表明估计误差越小。表 2 为 MH_02_easy 序列上 VINS-Mono 与本文所提的方法的定位误差数据,本文所提方法在定位误差的最大值、平均值、最小值以及标准差上均小于 VINS-Mono,因此本文所提方法的定位精度要优于 VINS-Mono。图 5 为本文所提的方法在 MH_02_easy 序列试验中相机畸变参数误差估计值的变化曲线,从图中可以看出,估计的误差值会在 20 s 内迅速收敛到稳定值。

4.3 无人车试验与分析

为了进一步验证算法的可行性,本文利用实际无人车平台进行实际验证,图 6 为无人车试验平台,其采用百度 Apollo 车载底盘,搭载了小觅深度版双目相机和 RTK 传感器。小觅深度版双目相机的图像分辨率为 640×480,频率为 30 Hz,IMU 频率为 200 Hz,在试验中仅使用左目相机的图像信息;RTK 提供定位基准参考,其定 2 cm,用于评估 VIO 方法的导航定位结果。试验全程长 358.9 m,用时 231.9 s,试验场景为室外开阔场景,如图 7 所示。

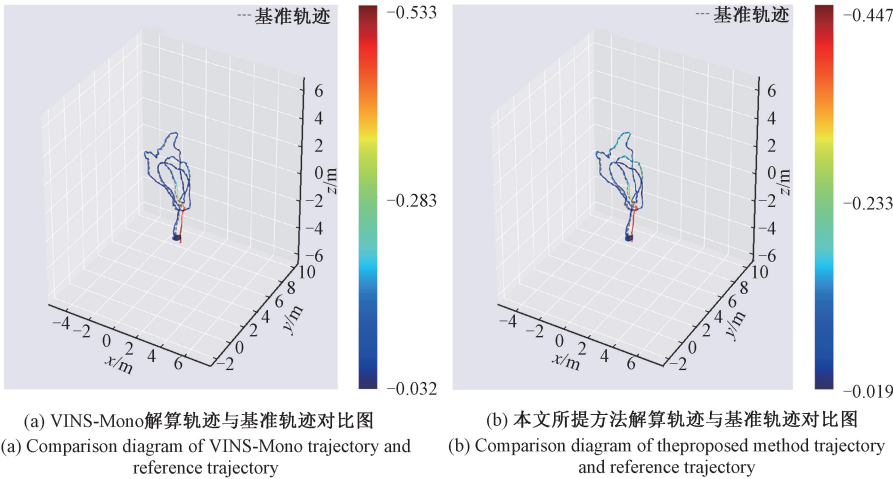


图 4 VINS-Mono 与本文所提方法对比图(MH_02_easy)

Fig. 4 Comparison diagrams of VINS-Mono and the proposed method (MH_02_easy)

表 2 VINS-Mono 与本文所提方法的定位误差(MH_02_easy)
Table 2 The positioning errors of VINS-Mono and the proposed method (MH_02_easy) m

误差类型	VINS-Mono	本文所提方法
最大值	0.533 2	0.447 4
平均值	0.133 7	0.121 1
最小值	0.032 4	0.018 6
标准差	0.117 8	0.097 3

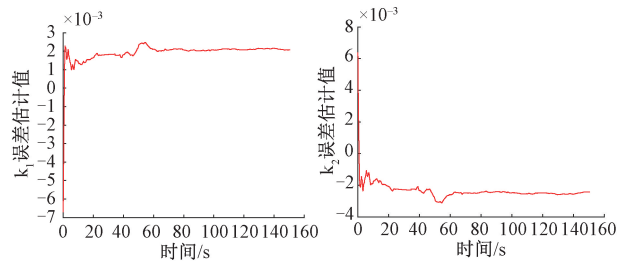


图 5 相机畸变参数误差估计值(MH_02_easy)

Fig. 5 Error estimation values of camera distortion parameters (MH_02_easy)



图6 无人车试验平台

Fig. 6 The experiment platform of unmanned vehicle



图7 无人车行驶场景

Fig. 7 Driving scene images of the unmanned vehicle

位精度优于无人车平台试验中,首先用传统相机畸变参数标定方法对其进行离线标定,然后将标定结果用于 VINS-Mono 和本文提出的方法,将两者的定位结果分别与 RTK 进行对比分析。图 8 为无人车试验 RTK、VINS-Mono 与本文所提方法轨迹对比图,实线为 RTK 定位结果,虚线为 VINS-Mono 的定位结果,点划线为本文方法的定位结果。本文所提方法与 VINS-Mono 的定位误差如图 9 所示,虚线为 VINS-Mono 的定位误差,点划线为本文所提方法的定位误差。表 3 为 VINS-Mono 与本文所提的方法的定位误差数据,VINS-Mono 定位误差的平均值、标准差与均方根误差分别为 17.59、9.11 和 19.80 m,本文所提方法定位误差的平均值、标准差与均方根误差分别为 6.26、2.80 和 6.85 m,均低于 VINS-Mono,本文所提方法的定位精度相比于 VINS-Mono 提升了 65.40%。图 10 为本文所提的方法在无人车试验中相机畸变参数误差估计值的变化曲线,从图 10 中可以看出,估计的误差值会在 30 s 内收敛到稳定值。可见,本文所提的方法能够对相机畸变参数进行在线优化标定,有效提高了算法的定位精度与鲁棒性。

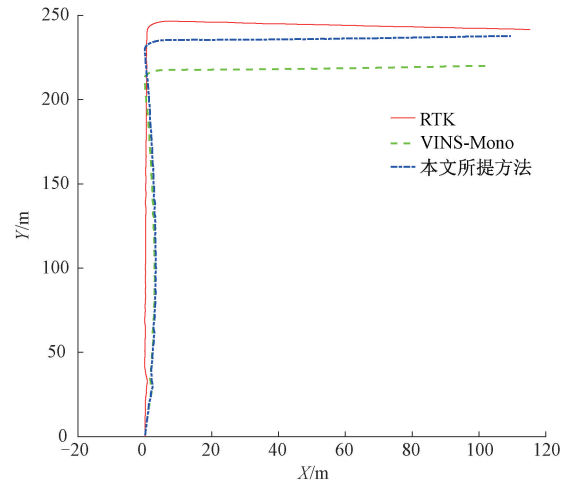


图8 无人车试验 RTK、VINS-Mono 与本文所提方法轨迹
Fig. 8 The trajectories of RTK, VINS-Mono and the proposed method in unmanned vehicle experiment

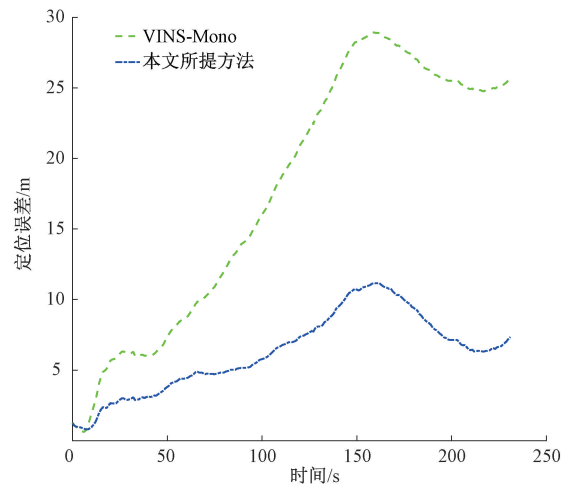


图9 无人车试验 VINS-Mono 与本文所提方法定位误差
Fig. 9 The positioning errors of VINS-Mono and the proposed method in unmanned vehicle experiment

表3 无人车试验 VINS-Mono 与本文所提方法的定位误差
Table 3 The positioning errors of VINS-Mono and the proposed method in unmanned vehicle experiment m

误差类型	VINS-Mono	本文所提方法
最大值	28.93	11.17
平均值	17.59	6.26
最小值	0.62	0.80
标准差	9.11	2.80
均方根误差	19.80	6.85

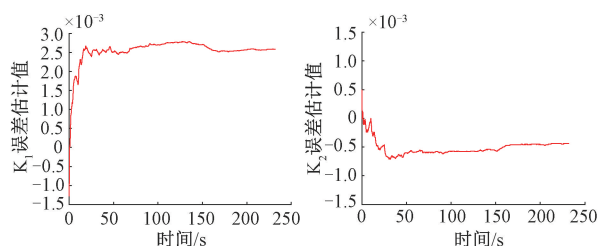


图10 无人车试验相机畸变参数误差估计值

Fig. 10 Error estimation values of camera distortion parameters in unmanned vehicle experiment

5 结 论

本文针对相机畸变参数标定不准确或因机械冲击等因素发生变化时 VIO 定位精度下降的问题,提出了一种面向相机畸变参数在线自标定的鲁棒 VIO 方法。该方法将相机畸变参数建模到视觉重投影误差中,在基于因子图优化的 VIO 后端与 IMU 状态和视觉特征一起优化求解,可以对相机畸变参数进行实时优化补偿。在 EuRoC 数据集和实际的无人车平台上进行了试验验证,并将本文提出的方法与 VINS-Mono 进行对比分析,试验结果表明,本文所提方法提高了相机畸变参数标定不准确或发生变化时 VIO 的定位精度与鲁棒性,且在实际的应用环境中更具有意义。

参考文献

- [1] ZHAO Y, YE W, VELA P A. Low-latency visual SLAM with appearance-enhanced local map building[C]. 2019 International Conference on Robotics and Automation (ICRA), IEEE, 2019: 8213-8219.
- [2] YUAN X, CHEN S. SaD-SLAM: A visual slam based on semantic and depth information [C]. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020: 4930-4935.
- [3] ZHANG T, ZHANG H, LI Y, et al. Flowfusion: Dynamic dense rgb-d slam based on optical flow[C]. 2020 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2020: 7322-7328.
- [4] CANOVAS B, ROMBAUT M, NÈGRE A, et al. Speed and memory efficient dense RGB-D SLAM in dynamic scenes[C]. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2020: 4996-5001.
- [5] CAMPOS C, ELVIRA R, RODRÍGUEZ J J G, et al. ORB-SLAM3: An accurate open-source library for visual, visual-inertial and multi-map SLAM[J]. Arxiv Preprint, 2020, ArXiv:2007.11898.
- [6] 孙龙龙,江明,焦传佳. 基于运动矢量的改进视觉 SLAM 算法[J]. 电子测量与仪器学报,2020,34(9): 23-31.
- SUN L L, JIANG M, JIAO CH J. Improved visual SLAM algorithm based on the motion vector [J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(9): 23-31.
- [7] LIM H, KIM Y, JUNG K, et al. Avoiding degeneracy for monocular visual SLAM with point and line features[J]. Arxiv Preprint, 2020, ArXiv:2103.01501, 2021.
- [8] SHAN T, ENLOT 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 (IROS), IEEE, 2018: 4758-4765.
- [9] RUAN J, LI B, WANG Y, et al. GP-SLAM+: Real-time 3D lidar SLAM based on improved regionalized Gaussian process map reconstruction[J]. Arxiv Preprint, 2020, Arxiv:2008.00644.
- [10] ZHAO J, ZHAO L, HUANG S, et al. 2D laser SLAM with general features represented by implicit functions[J]. IEEE Robotics and Automation Letters, 2020, 5(3): 4329-4336.
- [11] LIN J, ZHANG F. Loam livox: A fast, robust, high-precision LiDAR odometry and mapping package for LiDARs of small FoV [C]. 2020 IEEE International Conference on Robotics and Automation (ICRA), 2020: 3126-3131.
- [12] OELSCH M, KARIMI M, STEINBACH E. R-LOAM: Improving LiDAR odometry and mapping with point-to-mesh features of a known 3D reference object[J]. IEEE Robotics and Automation Letters, 2021, 6(2): 2068-2075.
- [13] MOURIKIS A I, ROUMELIOTIS S I. A multi-state constraint Kalman filter for vision-aided inertial navigation[C]. Proceedings 2007 IEEE International Conference on Robotics and Automation, IEEE, 2007: 3565-3572.
- [14] LI M, MOURIKIS A I. High-precision, consistent EKF-based visual-inertial odometry [J]. The International Journal of Robotics Research, 2013, 32(6): 690-711.
- [15] 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), 2015: 298-304.
- [16] 齐乃新,张胜修,杨小冈,等. 基于相机状态方程多模增广的改进 MSCKF 算法[J]. 仪器仪表学报,2019,

- 40(5):89-98.
- QI N X, ZHANG SH X, YANG X G, et al. An improved MSCKF algorithm based on multi-mode augmentation method for the camera state equation[J]. Chinese Journal of Scientific Instrument, 2019, 40(5): 89-98.
- [17] 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.
- [18] 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.
- [19] YUAN C, LAI J, LYU P, et al. A novel fault-tolerant navigation and positioning method with stereo-camera/micro electro mechanical systems inertial measurement unit (MEMS-IMU) in hostile environment[J]. Micromachines, 2018, 9(12): 626.
- [20] QIN T, PAN J, CAO S, et al. A general optimization-based framework for local odometry estimation with multiple sensors[J]. Arxiv Preprint, 2019, Arxiv: 1901.03638.
- [21] USENKO V, DEMMEL N, SCHUBERT D, et al. Visual-inertial mapping with non-linear factor recovery[J]. IEEE Robotics and Automation Letters, 2019, 5(2): 422-429.
- [22] 潘林豪,田福庆,应文健,等. 单目相机-IMU 外参自动标定与在线估计的视觉-惯导 SLAM[J]. 仪器仪表学报, 2019, 40(6): 56-67.
- PAN L H, TIAN F Q, YING W J, et al. VI-SLAM algorithm with camera-IMU extrinsic automatic calibration and online estimation[J]. Chinese Journal of Scientific Instrument, 2019, 40(6): 56-67.
- [23] 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.
- YU H SH, GUO F, GUO L F, et al. Robust monocular visual-inertial SLAM based on the improved SuperPoint network[J]. Chinese Journal of Scientific Instrument, 2021, 42(1): 116-126.
- [24] CORKE P, LOBO J, DIAS J. An introduction to inertial and visual sensing[J]. International Journal of Robotics Research, 2007, 26(6): 519-535.
- [25] ZHANG Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
- [26] LI M, YU H, ZHENG X, et al. High-fidelity sensor modeling and self-calibration in vision-aided inertial navigation[C]. 2014 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2014: 409-416.
- [27] KEIVAN N, SIBLEY G. Online SLAM with any-time self-calibration and automatic change detection[C]. 2015 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2015: 5775-5782.
- [28] SCHNEIDER T, LI M, BURRI M, et al. Visual-inertial self-calibration on informative motion segments[C]. 2017 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2017: 6487-6494.
- [29] LUCAS B D, KANADE T. An iterative image registration technique with an application to stereo vision[C]. In Proceedings of the 7th International Joint Conference on Artificial Intelligence, 1981: 674-679.
- [30] SHI J. Good features to track[C]. 1994 Proceedings of IEEE conference on computer vision and pattern recognition, 1994: 593-600.
- [31] BURRI M, NIKOLIC J, GOHL P, et al. The EuRoC micro aerial vehicle datasets[J]. International Journal of Robotics Research, 2016, 35(10): 1157-1163.

作者简介



杨子寒, 2019 年于南京航空航天大学获得学士学位, 现为南京航空航天大学硕士研究生, 主要研究方向为视觉/惯性组合导航。

Email: yangzh@nuaa.edu.cn



赖际舟(通信作者), 分别于 2000 年和 2005 年于南京航空航天大学获得学士和博士学位, 现为南京航空航天大学教授、博士生导师, 主要研究方向为惯性导航、多信息可靠融合导航及感知技术。

Email: laijz@nuaa.edu.cn

Lai Jizhou (Corresponding author) received his B. Sc. degree in 2000 and Ph. D. degree in 2005 both from Nanjing University of Aeronautics and Astronautics (NUAA). He is currently a professor and Ph. D. supervisor in NUAA. His main research interest include inertial navigation, multi-information reliable fusion navigation and perception technology.