

Tightly-Coupled Vision and Wheel Odometer Localization for Ground Vehicle^{*}

ZHU Daojun^{1,2}, XU Zhannan¹, MA Tingting¹, CAO Pingguo¹, SONG Quanjun^{1*}

(1. Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei Anhui 230031, China;

2. Science Island Branch, Graduate School of USTC, Hefei Anhui 230026, China)

Abstract: The Localization of ground vehicle in unknown environments is a problem in robotics that has undergone extensive research but still needs further improvement. VINS-Mono is currently one of the best effective algorithms, but its application on wheeled robots will cause degradation of localization accuracy due to degradation. By making full use of the coupling relationship between the wheel odometer and the vision system, this paper proposes a wheeled robot localization system estimator based on the tight coupling of vision and wheel odometer. Aiming at the fusion of low-frequency visual information and high-frequency wheel speedometer information, the pre-integration measurement and error propagation process of wheel speedometer are deduced in this paper. Aiming at the scale-free initialization of monocular vision and the initialization degradation of VINS-Mono on wheeled robots, this paper proposes a fast and accurate system initialization process that combines vision and wheel odometer information. Through actual verification in representative indoor and outdoor environments, the method proposed in this paper not only has better localization accuracy than VINS-Mono, but its computational complexity is only about one-tenth of VINS-Mono.

Key words: SLAM; robot localization; sensor fusion; ground vehicle

EEACC: 7230 **doi:** 10.3969/j.issn.1004-1699.2021.07.008

基于视觉和轮速计紧耦合的轮式机器人定位算法研究^{*}

朱道俊^{1,2}, 徐湛楠¹, 马婷婷¹, 曹平国¹, 宋全军^{1*}

(1. 中国科学院合肥物质科学研究院, 安徽 合肥 230031; 2. 中国科学技术大学研究生院科学岛分院, 安徽 合肥 230026)

摘 要: 轮式机器人在未知环境下的定位是机器人学中一个经过广泛研究但仍然需要进一步提高的问题。VINS-Mono 是目前效果最好的算法之一, 然而其应用在轮式机器人上会因退化问题导致定位精度下降。本文通过充分利用轮速计与视觉系统之间的耦合关系, 提出了一种基于视觉和轮速计紧耦合的轮式机器人定位系统估计器。针对低频视觉信息和高频轮速计信息的融合问题, 文中推导了轮速计预积分测量和误差传播过程。针对单目视觉无尺度初始化和 VINS-Mono 在轮式机器人上的初始化退化问题, 文中提出了一种联合视觉和轮速计信息的快速精确系统初始化流程。通过在代表性室内及室外环境的实际验证, 本文所提出的方法不但比 VINS-Mono 具有更好的定位精度, 而且其计算复杂度仅为 VINS-Mono 的约十分之一。

关键词: 同时建图与定位; 机器人定位; 传感器融合; 轮式机器人

中图分类号: TP242.6

文献标识码: A

文章编号: 1004-1699(2021)07-0896-08

随着机器人技术的发展和计算资源能力的提高, 轮式机器人的应用逐渐得到普及。定位技术是实现轮式机器人自主导航的重要技术环节。对于室内外小场景范围内的轮式机器人应用, 如巡逻机器人, 低成本、高精度、高鲁棒性、易用性成为了定位系统的研究目标。

同时建图与定位(SLAM)技术^[1]是研究机器人在未知环境下定位的主流技术方案。其中基于视觉传感器的 SLAM(V-SLAM)技术由于其低成本、测量信息丰富得到了广泛研究。经典的 VSLAM 工作有 PTAM^[2]、ORB-SLAM^[3]、ORB-SLAM2^[4]等, Mur-Artal R 等提出的 ORB-SLAM 第一次将 SLAM 分为

项目来源: 安徽省重点研究与开发计划项目(201904d07020007, 202004a07020051)

收稿日期: 2021-01-29 **修改日期:** 2021-04-15

三个运行线程进行处理,实现了 V-SLAM 大场景下的实时运行能力^[3]。然而实际环境中由于场景特征的单一(如走廊环境)和传感器的快速运动等,很容易导致 V-SLAM 定位失败,鲁棒性较差。

视觉-惯性融合 SLAM(VI-SLAM)是通过融合视觉和 IMU 传感器的 SLAM 方案,由于 IMU 具有很好的动态特性,VI-SLAM 能有效处理机体的快速运动和场景单一的情况。经典工作有 MSCKF^[5]、S-MSCKF^[6]、OKVIS^[7]、VINS-Mono^[8]等。香港科技大学秦通等提出的 VINS-Mono 因其快速精确初始化流程、视觉-惯性紧耦合非线性滑动窗口优化估计器、四自由度位姿图优化等特点成为当前 VI-SLAM 中定位效果最好的算法之一,实现了在无人航空飞行器上的精准位姿轨迹估计,然而其在轮式移动机器人上的应用却会出现很多退化问题,导致定位精度下降甚至定位失败。为了实现 VINS-Mono 精确有效的初始化,需要保证初始化过程中机体有充分的三轴运动激励,而轮式机器人由于平面运动的限制,很难做到这点。Wu 等通过数学和实验证明了当轮式机器人在某些特定运动场景下,VINS 会出现额外的不可观自由度,会带来很大的估计偏差^[9],比如纯直线运动、常量局部加速度运动等,而这些运动情形在轮式机器人上是很常见的。

针对轮式机器人,轮速计传感器可以很好的替代 IMU,轮速计也能够提供高频率的机体状态量测量,同时对于特殊运动场景,轮速计不会带来退化问题。Wu 等也证明了通过融合轮速计信息能够避免尺度不可观的问题,提高定位精度^[9]。Zhang 等^[10]、Liu 等^[11]、Lee 等^[12]通过融合视觉-IMU-轮速计信息来实现轮式机器人定位,结果相比于 VINS 都有了很大的提高。Zheng 等提出了一种基于视觉-轮速计紧耦合的定位系统,它基于 ORB-SLAM 框架进行改进实现,提出了一种新的视觉重投影观测误差形式^[13]。

为了同时追求系统的运行效率和定位精度,本文发现在室内外小范围平面运动场景下,视觉-IMU-轮速计定位系统相比于视觉-轮速计定位系统,会引入大量的额外系统负担,比如,IMU 的融合大大增加了系统的状态维度、增加了系统初始化的难度^[8],然而定位精度不会有很大的提高,甚至对于低成本 IMU,会带来更大的定位误差。因此,本文中提出紧耦合视觉和轮速计定位方案,提高系统定位精度的同时提升系统运行效率,主要贡献如下:①提出了一种基于非线性滑动窗口优化的视觉-轮速计定位系统估计器;②提出改进了一种联合视觉-轮

速计的快速精确系统初始化方案。③轮速计-相机坐标系变换外参在线标定。

1 数学记号

在本文中,我们使用大写字母表示向量/矩阵,使用小写字母表示标量, ${}^A_B\mathbf{R} \in \text{SO}(3)$ 表示坐标系 B 到坐标系 A 的旋转变换, ${}^A_B\mathbf{q}$ 为其单位四元数表示,这里使用 Hamilton 表达形式^[14]; ${}^A\mathbf{p}_B \in \mathbb{R}$ 表示坐标系 B 原点在坐标系 A 下的坐标; ${}^A(\cdot)$ 表示坐标系 A 下的某量, $(\hat{\cdot})$ 表示状态估计值或者测量值, $\delta(\cdot)$ 表示状态的误差形式,本文使用右乘扰动来定义旋转误差,其形式为:

$$\delta\theta = 2(\hat{\mathbf{q}}^{-1} \otimes \mathbf{q})_{\text{img}} \quad (1)$$

\otimes 表示四元数乘法, $(\cdot)_{\text{img}}$ 表示四元数的虚数部分; $\lfloor \cdot \rfloor_{\times}$ 表示向量对应的反对称矩阵,比如向量 $\mathbf{a} = [a_1 \ a_2 \ a_3]^T$ 对应的反对称矩阵为:

$$\lfloor \mathbf{a} \rfloor_{\times} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}.$$

\mathbf{e}_i 表示 $R^{3 \times 3}$ 单位矩阵的第 i 列, $\mathbf{A} = [\mathbf{e}_1^T \ \mathbf{e}_2^T]$ 。SO(3) 上的李代数表示为 $\text{so}(3)$, $\log(\cdot)$ 获得 SO(3) 对应的 $\text{so}(3)$,反之, $\exp(\cdot)$ 将 $\text{so}(3)$ 转换为 SO(3)。向量 \mathbf{v} 的第 i 个元素表示为 $v(i)$ 。

本文涉及到三个不同的坐标系,分别是世界坐标系、轮速计坐标系(机体坐标系)和相机坐标系,我们将它们依次表示为 w 、 b 和 c 。另外 b_k 、 c_k 表示时间顺序上第 k 个机体坐标系和对应的相机坐标系。

2 系统描述

如图 1 所示。系统输入为相机和轮速计观测信息,对于新的图像帧,先使用 KLT 光流法跟踪上一帧的特征点,再提取新的 Shi-Tomasi 角点。对于新帧和上一帧之间的轮速计测量值,进行预积分处理,详见 4.4 节。

若系统未完成初始化,则进行相机轮速计联合

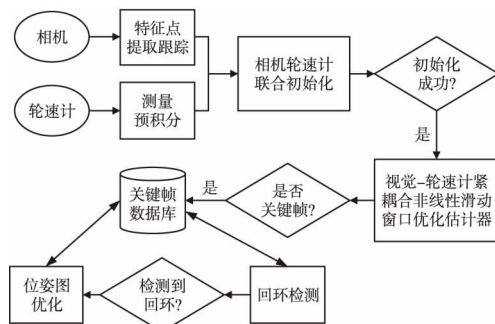


图1 系统流程图

初始化,详见 3 节。若初始化成功,则进行视觉-轮速计紧耦合非线性滑动窗口优化,详见 4 节,若窗口中帧数达到最大,则判断次新帧是否为关键帧,若是,则选择窗口中的最老关键帧进行边缘化,否则边缘化次新帧^[8]。

为消除长时间累积误差,对新关键帧进行回环检测,在历史关键帧数据库中查找当前关键帧的回环帧。回环检测基于视觉词袋库 DBoW2^[15] 进行实现。若检测到回环,则进行六自由度位姿图优化,保证长时间位姿轨迹估计的一致性^[8]。

3 视觉-轮速计联合初始化

与 VINS-Mono 不同的是,本文提出使用两帧快速初始化,将第一帧 F_0 设置为世界坐标系原点,为保证比较好的初始化,初始化两帧之间需要有足够的视差,设置初始化视差阈值为 T_p ,当前帧 F_c 与 F_0 的视差小于 T_p 时,则丢弃该帧,反之,设置当前帧为初始化第二帧 F_1 。由于轮速计能够提供比较精确的短时间位姿估计,因此可通过轮速计积分测量来预测 F_1 的位姿,如图 2 所示,对 F_0 和 F_1 跟踪匹配的特征点进行三角化,最后进行一次集束调整(BA)优化,所需最小化的误差函数为:

$$\min_{x_0} \{ \|r_w(x_0, \hat{z}_{F_1}^{F_0})\|_{\Omega_w}^2 + \sum_{(l,j) \in C} \rho(x_l, y_l r_f(\|r_f(x_0, \hat{z}_{F_j}^{F_j})\|_{\Omega_f}^2)) \} \quad (2)$$

式中:优化状态 $x_0 = \{x_{F_0}, x_{F_1}, L_0\}$, x_{F_j} , ($j=0,1$) 表示第 j 帧的位姿, L_0 表示所能跟踪观测到的路标点集合, $\hat{z}_{F_1}^{F_0}$ 表示轮速计积分测量值, $\hat{z}_{F_j}^{F_j}$ 表示路标点 l 在 F_j 帧上的观测值, C 表示路标点观测对集合, $\rho(\cdot)$ 为鲁棒核函数,用于降低错误特征点匹配的优化影响权重, Ω_w, Ω_f 分别表示轮速计积分测量和特征点观测的噪声协方差矩阵, $\|\cdot\|_{\Omega}$ 表示马氏距离。

为求解式(2),我们采用迭代求解的方式,每次迭代通过求解增量方程 $H\delta x = b$ 得到更新增量 δx ,然后更新到当前状态估计值上: $x' = x \oplus \delta x$,对于平移状态, $p' = p \oplus \delta p = p + \delta p$,对于旋转状态, $q' = q \oplus$

$\delta\theta = q \otimes \left[\frac{1}{2} \delta\theta \right]$ 。本文中使用 Ceres Solver 库^[16] 来

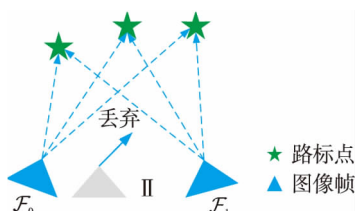


图2 系统初始化

求解上述迭代优化问题。

4 视觉-轮速计紧耦合非线性滑动窗口优化估计器

非线性滑动窗口优化估计器中的状态及测量约束关系可用因子图进行表示,如图 3 所示,其中圆形和椭圆形表示系统优化的状态,方块表示测量误差因子,与方块相连接的连接线表示该测量因子所关联到的状态。优化的状态包括关键帧位姿、轮速计相机外参和特征点位置,详见 4.1 节。传感器测量约束因子包括特征点重投影误差因子和轮速计预积分测量误差因子,详见 4.2 和 4.3 节。为了处理低频视觉信息和高频轮速计信息的融合,本文进行了相邻关键帧间轮速计测量预积分和误差传播,详见 4.4 节。

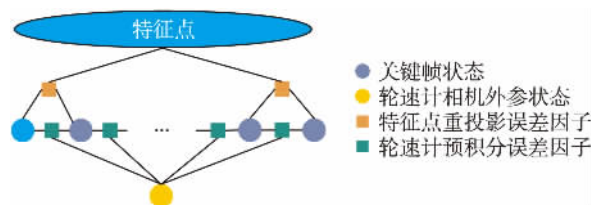


图3 非线性滑动窗口优化估计器因子图

4.1 系统状态

滑动窗口中估计状态为:

$$\chi = \{x_k, x_{k+1}, \dots, x_{k+N-2}, x_{k+N-1}, \lambda_1, \lambda_2, \dots, \lambda_m, x_{ex}\} \quad (3)$$

式中: N 为滑动窗口大小, m 为滑动窗口中所能观测到的路标点数量 $x_i = \{^w q_i, ^w p_i\}$, $i = k, \dots, k+N-1$ 表示第 i 帧位姿状态, $\lambda_l, l = 1, \dots, m$ 表示第 l 个路标点逆深度状态,参数化表示在锚坐标系(观测首帧相机坐标系)下^[8],定义如下:

$$^w p_l = ^w R \left(^c R \frac{1}{\lambda_l} \begin{bmatrix} ^c \bar{u}_l \\ 1 \end{bmatrix} + ^c p_c \right) + ^w p_{b_a} \quad (4)$$

$$^c \bar{u}_l = \begin{bmatrix} ^c \bar{u}_l \\ ^c \bar{v}_l \end{bmatrix} = \pi_c^{-1}(^c u_l)$$

式中: b_a/c_a 分别表示观测首帧机体/相机坐标系, $^c u_l$ 表示路标点 l 在 c_a 帧下的像素观测坐标值,将其通过相机内参和畸变参数转换到归一化平面坐标 $^c \bar{u}_l$,这个转换过程即为 π_c^{-1} 。 $x_{ex} = \{^b q, ^b p_c\}$ 表示轮速计坐标系和相机坐标系之间的外参变换,这里将其作为系统状态进行在线估计,可以减小不精确的外参标定结果对定位精度的影响。如果有比较精确的外参线下标定值,可以将其从系统状态估计中移除。

4.2 视觉观测误差因子

如图 4 所示,路标点 l 在 c_j 帧的重投影误差定义为观测坐标和重投影坐标之间的坐标差值,重投

影坐标计算方法为首先将路标点 l 通过其逆深度和观测首帧位姿估计转换到世界坐标系下,再通过 c_j 帧位姿估计投影到 c_j 帧图像平面上,如下所示:

$$r_{c_j} = {}^{c_j}\bar{\mathbf{u}}_l - {}^{c_j}\hat{\mathbf{u}}_l \quad (5.1)$$

$${}^{c_j}\bar{\mathbf{u}}_l = \pi_p({}^{c_j}\mathbf{p}_l) = \mathbf{A} \frac{{}^{c_j}\mathbf{p}_l}{{}^{c_j}\mathbf{p}_l(2)} \quad (5.2)$$

$${}^{c_j}\mathbf{p}_l = {}^b\mathbf{R}^T \{ [{}^b\mathbf{R}^T({}^w\mathbf{p}_l - {}^w\mathbf{p}_{b_j})] - {}^b\mathbf{p}_c \} \quad (5.3)$$

$${}^{c_j}\hat{\mathbf{u}}_l = \pi_c^{-1}({}^{c_j}\hat{\mathbf{u}}_l) \quad (5.4)$$

其中式(5.1)表示重投影误差, ${}^{c_j}\bar{\mathbf{u}}_l$ 为路标点 l 在 c_j 帧上预测归一化平面坐标, ${}^{c_j}\hat{\mathbf{u}}_l$ 为观测归一化平面坐标,式(5.2)将相机坐标系下坐标点投影到相机归一化平面上, ${}^{c_j}\mathbf{p}_l$ 为路标点 l 在相机坐标系 c_j 下的坐标, ${}^{c_j}\mathbf{p}_l(2)$ 表示向量的第三个元素,式(5.3)将路标点 l 的世界坐标转换到 c_j 帧相机坐标系下, ${}^w\mathbf{p}_l$ 通过式(4)计算获得。

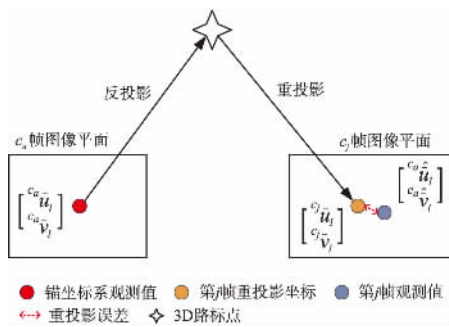


图 4 特征点重投影观测误差

由式(4)、式(5)可以看出,路标点 l 一次视觉重投影误差分别关联到锚机体坐标系位姿状态、当前机体坐标系位姿状态和路标点的逆深度状态,因此需要分别对三个状态求误差雅可比矩阵,分别设为 $\mathbf{J}_{b_a}^l$ 、 $\mathbf{J}_{b_j}^l$ 、 \mathbf{J}_{λ}^l 。

$\mathbf{J}_{b_a}^l$ 解析式为:

$$\begin{aligned} \mathbf{J}_{b_a}^l &= \mathbf{J}_{\pi_p} [\mathbf{J}_{b_a}^r \quad \mathbf{J}_{b_a}^t] \\ \mathbf{J}_{b_a}^r &= -{}^b\mathbf{R} {}^b\mathbf{R}^T {}^w\mathbf{R} {}^w\mathbf{R}^T \frac{1}{\lambda_l} \begin{bmatrix} {}^c\mathbf{u}_l \\ 1 \end{bmatrix} + {}^b\mathbf{p}_c \downarrow_x \\ \mathbf{J}_{b_a}^t &= {}^b\mathbf{R} {}^b\mathbf{R}^T {}^w\mathbf{R} \end{aligned} \quad (6)$$

式中: $\mathbf{J}_{b_a}^r$ 表示式(5.3) ${}^{c_j}\mathbf{p}_l$ 对 b_a 帧位姿中旋转部分的导数, $\mathbf{J}_{b_a}^t$ 表示式(5.3) ${}^{c_j}\mathbf{p}_l$ 对 b_a 帧位姿中平移部分的导数, \mathbf{J}_{π_p} 表示式(5.2) ${}^{c_j}\bar{\mathbf{u}}_l$ 对 ${}^{c_j}\mathbf{p}_l$ 的导数,下面的表示方法类似。

$\mathbf{J}_{b_j}^l$ 解析式为:

$$\begin{aligned} \mathbf{J}_{b_j}^l &= \mathbf{J}_{\pi_p} [\mathbf{J}_{b_j}^r \quad \mathbf{J}_{b_j}^t] \\ \mathbf{J}_{b_j}^r &= {}^b\mathbf{R} {}^b\mathbf{R}^T {}^w\mathbf{R} {}^w\mathbf{R}^T \mathbf{p}_l \downarrow_x \end{aligned}$$

$$\mathbf{J}_{b_j}^t = -{}^b\mathbf{R} \quad (7)$$

\mathbf{J}_{λ}^l 解析式为:

$$\mathbf{J}_{\lambda}^l = \mathbf{J}_{\pi_p} {}^b\mathbf{R} {}^b\mathbf{R}^T {}^w\mathbf{R} {}^w\mathbf{R}^T \begin{bmatrix} {}^c\bar{\mathbf{u}}_l \\ 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{\lambda^2} \end{bmatrix} \quad (8)$$

式中: \mathbf{J}_{π_p} 解析式为:

$$\mathbf{J}_{\pi_p} = \begin{bmatrix} \frac{1}{{}^{c_j}\mathbf{p}_l(2)} & 0 & \frac{-{}^{c_j}\mathbf{p}_l(0)}{{}^{c_j}\mathbf{p}_l(2)^2} \\ 0 & \frac{1}{{}^{c_j}\mathbf{p}_l(2)} & \frac{-{}^{c_j}\mathbf{p}_l(1)}{{}^{c_j}\mathbf{p}_l(2)^2} \end{bmatrix} \quad (9)$$

4.3 轮速计误差因子

两个相邻关键帧 F_k 和 F_{k+1} 之间的轮速计积分测量为:

$$\hat{\mathbf{z}}_{k+1}^k = \begin{bmatrix} \hat{\mathbf{d}}_{k+1}^k \\ \hat{\theta}_{k+1}^k \end{bmatrix} \quad (10)$$

式中: $\hat{\mathbf{d}}_{k+1}^k$ 为 x 轴和 y 轴方向的平移测量值, $\hat{\theta}_{k+1}^k$ 为偏航角测量值。则对应的轮速计测量误差为:

$$\mathbf{r}_{k+1}^k = \begin{bmatrix} r_d \\ r_\theta \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{b_k} {}^w\mathbf{R}^T ({}^w\mathbf{p}_{b_{k+1}} - {}^w\mathbf{p}_{b_k}) - \hat{\mathbf{d}}_{k+1}^k \\ \mathbf{e}_3^T \log({}^b\mathbf{R} {}^b\mathbf{R}^T {}^w\mathbf{R}) - \hat{\theta}_{k+1}^k \end{bmatrix} \quad (11)$$

式中: r_d 表示轮速计平移测量误差, r_θ 表示轮速计航向角测量误差。

从式(11)可以看出, \mathbf{r}_{k+1}^k 关联到第 k 、 $k+1$ 关键帧位姿估计,因此需要对这两帧位姿计算雅可比矩阵,设它们分别为: \mathbf{J}_{b_k} 、 $\mathbf{J}_{b_{k+1}}$ 。 \mathbf{J}_{b_k} 的解析表达式为:

$$\begin{aligned} \mathbf{J}_{b_k} &= \begin{bmatrix} \mathbf{J}_{r_d}^d & \mathbf{J}_{r_\theta}^d \\ \mathbf{J}_{r_d}^\theta & \mathbf{J}_{r_\theta}^\theta \end{bmatrix} = \\ &= \begin{bmatrix} -\mathbf{A}_{b_k} {}^w\mathbf{R}^T & \mathbf{A}_{b_k} {}^w\mathbf{R}^T ({}^w\mathbf{p}_{b_{k+1}} - {}^w\mathbf{p}_{b_k}) \downarrow_x \\ -\mathbf{e}_3^T {}^b\mathbf{R} {}^b\mathbf{R}^T {}^w\mathbf{R} & \mathbf{0} \end{bmatrix} \end{aligned} \quad (12)$$

式中: $\mathbf{J}_{r_d}^d$ 、 $\mathbf{J}_{r_\theta}^d$ 、 $\mathbf{J}_{r_d}^\theta$ 、 $\mathbf{J}_{r_\theta}^\theta$ 分别表示 r_d 和 r_θ 相对于第 k 帧关键帧位姿旋转部分和平移部分的导数,下式同理。

$\mathbf{J}_{b_{k+1}}$ 的解析表达式为:

$$\mathbf{J}_{b_{k+1}} = \begin{bmatrix} \mathbf{J}_{r_d}^{d_{k+1}} & \mathbf{J}_{r_\theta}^{d_{k+1}} \\ \mathbf{J}_{r_d}^{\theta_{k+1}} & \mathbf{J}_{r_\theta}^{\theta_{k+1}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{A}_{b_k} {}^w\mathbf{R}^T \\ \mathbf{e}_3^T & \mathbf{0} \end{bmatrix} \quad (13)$$

4.4 轮速计预积分

由于相邻关键帧之间有多次轮速计测量,为了获得相邻关键帧之间轮速计测量的噪声协方差矩阵,需要将轮速计测量进行积分。如图5所示,用 i 、 j 表示

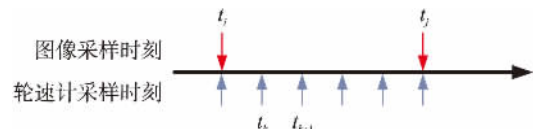


图 5 轮速计预积分

图像测量时间索引,用 k 表示轮速计测量时间索引,我们将图像测量时刻 t_i 和 t_j 之间的轮速计测量进行积分和误差递推。

轮速计 k 时刻的测量为:

$$\hat{\mathbf{v}}_k = {}^k\mathbf{v}_{k+1} + \mathbf{n}_{vk}, \mathbf{n}_{vk} = \begin{bmatrix} \mathbf{n}_{dk} \\ \mathbf{n}_{\theta k} \end{bmatrix} \sim N(0, \Sigma_{vk}) \quad (14)$$

根据 $SS\mathcal{E}(2)$ 上的递推公式,有:

$${}^i\mathbf{v}_{k+1} = \begin{bmatrix} {}^i\mathbf{d}_{k+1} \\ {}^i\theta_{k+1} \end{bmatrix} = \begin{bmatrix} {}^i\mathbf{d}_k + \Phi({}^i\theta_k)({}^i\hat{\mathbf{d}}_k - \mathbf{n}_{dk}) \\ {}^i\theta_k + \hat{\theta}_k - \mathbf{n}_{\theta k} \end{bmatrix} \quad (15)$$

式中: $\Phi(\cdot)$ 表示对应的二维旋转矩阵, ${}^i\mathbf{v}_{k+1}$ 的噪声协方差由 ${}^i\mathbf{v}_k$ 噪声协方差和 \mathbf{n}_{vk} 测量噪声协方差递推得到,因此需要求得式(15)对 ${}^i\mathbf{v}_k$ 和 \mathbf{n}_{vk} 的雅可比矩阵,设分别为 \mathbf{J}_{vk} 、 \mathbf{J}_{nk} ,解析表达式为:

$$\mathbf{J}_{vk} = \begin{bmatrix} \mathbf{I}_2 & \Phi({}^i\theta_k) \mathbf{G} \hat{\mathbf{d}}_k \\ \mathbf{0} & 1 \end{bmatrix} \quad (16)$$

$$\mathbf{J}_{nk} = \begin{bmatrix} -\Phi({}^i\theta_k) & \mathbf{0} \\ \mathbf{0} & -1 \end{bmatrix}$$

式中: $\mathbf{G} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, \mathbf{I}_2 表示 2×2 的单位矩阵。因此 ${}^i\mathbf{v}_{k+1}$ 的噪声协方差矩阵为:

$${}^i\Sigma_{k+1} = \mathbf{J}_{vk} {}^i\Sigma_k \mathbf{J}_{vk}^T + \mathbf{J}_{nk} \Sigma_{vk} \mathbf{J}_{nk}^T \quad (17)$$

以此递推,可以得到 t_i 和 t_j 时刻图像帧之间的轮速计积分测量 ${}^i\mathbf{v}_j$ 和噪声协方差矩阵 ${}^i\Sigma_j$ 。

5 实验及分析

5.1 实验配置

本部分对本文所提出算法进行了实验验证,实验平台如图 6 所示,四轮车上搭载了一个 PC 和一个相机。PC 用于读取相机和轮速计传感器测量信息和进行算法处理,机器人操作系统(ROS)用来进行四轮车、相机和 PC 的信息通信传输,使用操作系统为 Ubuntu18.0。

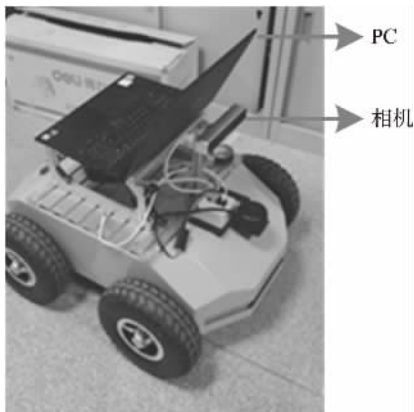


图 6 实验平台配置

实验设备的具体型号配置如表 1 所示。本实验代码基于开源 VINS-Mono^[17] 代码框架修改。

表 1 实验设备具体配置信息

| PC | Intel Core I7-9750H CPU; 20GB RAM |
|-----|-----------------------------------|
| 相机 | 小觅双目相机标准版,分辨率 752x480 |
| 编码器 | 400 线 |

注:实验中仅使用了双目相机的左目信息。

为了验证所提算法的精确度,另外进行了两组对比实验,其一是基于视觉-IMU 融合的 VINS-Mono 方案;其二是仅使用轮速计的定位方案。为了量化评估轨迹误差,实验数据采集时确保了起始位置和终止位置相同,通过计算终点、起始点位置估计距离差和轨迹总长度的比例来定义估计误差,如式(18)所示,其中 $\hat{\mathbf{p}}_{\text{end}}$ 为终止点估计位置, $\hat{\mathbf{p}}_{\text{start}}$ 为起始点位姿估计, \hat{S} 为轨迹总长度估计。因为起始点设置为世界坐标系原点,所以 $\hat{\mathbf{p}}_{\text{start}} = 0$ 。

$$e = \frac{\|\hat{\mathbf{p}}_{\text{end}} - \hat{\mathbf{p}}_{\text{start}}\|}{\hat{S}} = \frac{\|\hat{\mathbf{p}}_{\text{end}}\|}{\hat{S}} \times 100\% \quad (18)$$

传感器之前的外参数通过线下进行标定,相机内参和相机-IMU 外参数通过 Kalibr^[18] 进行标定。相机-轮速计外参数通过 CamOdomCalibraTool^[19] 进行标定,该标定参数为粗略标定,会对该参数进行在线估计优化。

5.2 轨迹误差分析

实验中录制了三种不同场景的数据集进行测试,分别是实验室、会议室和走廊环境,如图 7 所示。对每个实验场景采集的数据集,分别对本文所提算法和比较算法运行了 3 次,并取结果平均值作为误差结果。如图 8 所示,分别为本文所提算法、VINS-Mono 和轮速计在三个实验场景下其中一次运行结果所估计的运动轨迹。从图中可以看出,VINS-Mono 和轮速计都产生了很大的轨迹误差,VINS-Mono 甚至在实验室场景中产生了发散,原因在于以下几个方面,一是本实验采用了低价的 IMU 传感器,有比较大的误差,二是 VINS-Mono 对相机和 IMU 之前的外参标定和时间同步参数有很高的精度要求,而这对低价的传感器配置来说是很难实现,三是本实验的场景主要是平面运动,且大部分是直线运动和匀加速运动,这给 VINS-Mono 的初始化以及后面的估计带来了很大的不可观问题^[9]。另外,仅用轮速计的轨迹估计随着时间增长误差将不断增大,原因是轮速计是通过积分得到位姿估计的,因为测量带有误差,所以长时间的轮速计积分将不断积累误差。而本文所提算法通过融合相机和轮速计信

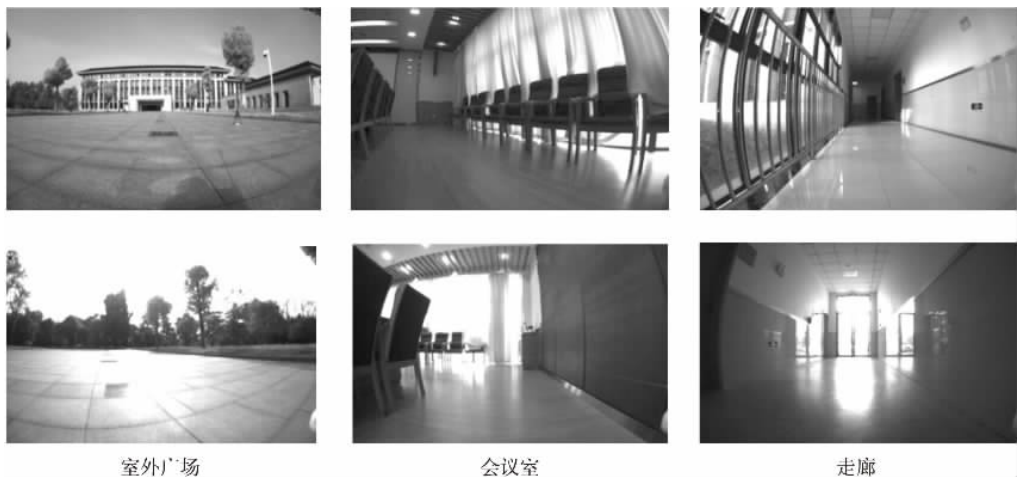


图 7 实验数据采集环境,从左至右分别为实验室、会议室和走廊

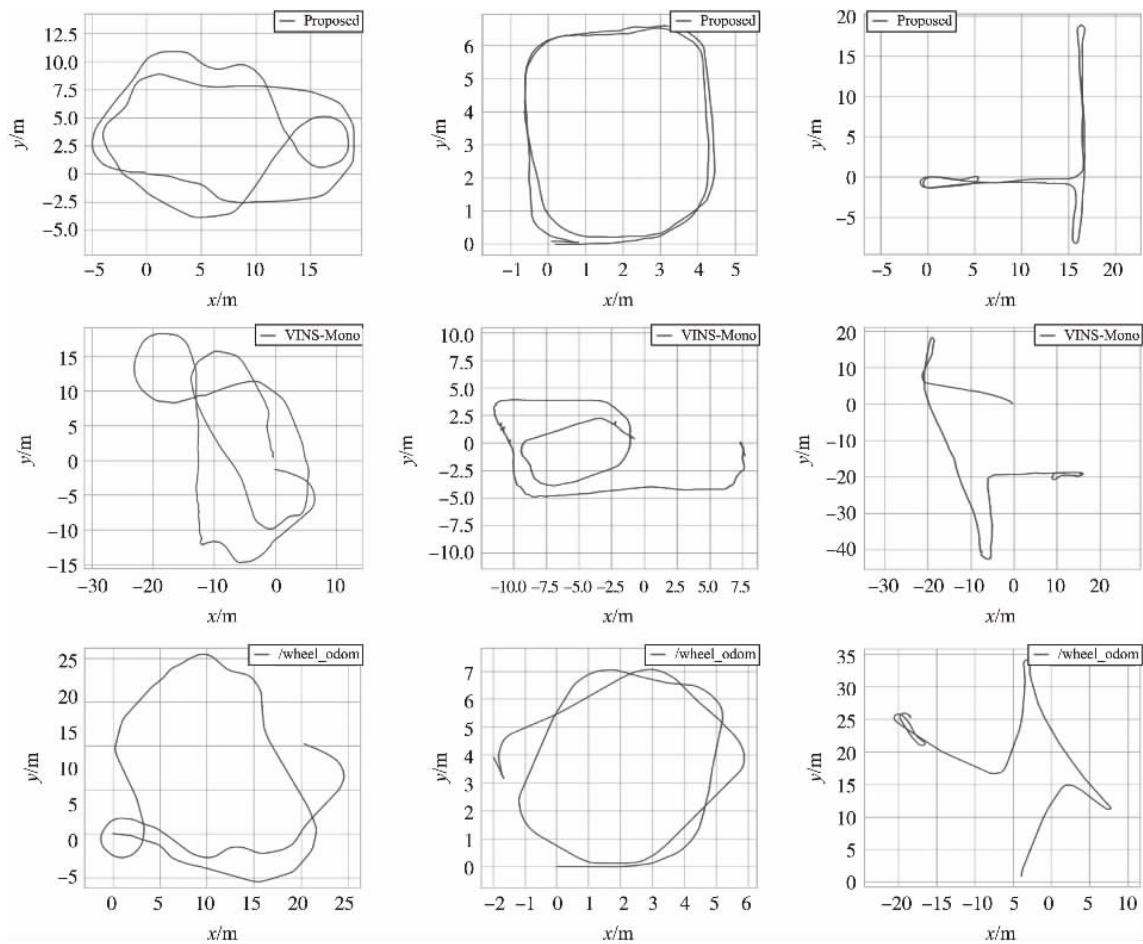


图 8 所提算法和其他两种对比算法分别在三种实验场景下的估计轨迹对比图

息获得了很精确的定位表现,通过视觉特征点、观测测量来矫正轮速计积分误差,同时轮速计积分测量也对相邻帧间的位姿起到较精确的约束作用。相比 IMU,视觉-轮速计融合能够获得更鲁棒的表现,机体终点位置估计能很好的回到初始点,其终点位置估计误差如表 2 所示。如图 9 所示为所提算法和其他两种对比算法分别在三种实验场景下的估计轨迹误差,其中水平横轴表示不同的估计算法,水平纵轴表示不同的实验场景,垂直轴表示算法估计误差。

可以看出所提算法相比于 VINS-Mono 和纯轮速计估计都有很大程度的精度提高。

表 2 所提算法在三个实验场景下的估计误差

| 实验场景 | 终点位置估计/m | 轨迹总长度估计/m | 终点位置估计误差/% |
|------|----------------------|-----------|------------|
| 室外广场 | (-0.098,0.123,0.037) | 123 | 0.131 |
| 会议室 | (0.105,0.081,0.047) | 41 | 0.343 |
| 走廊 | (0.183,0.030,0.148) | 102 | 0.232 |

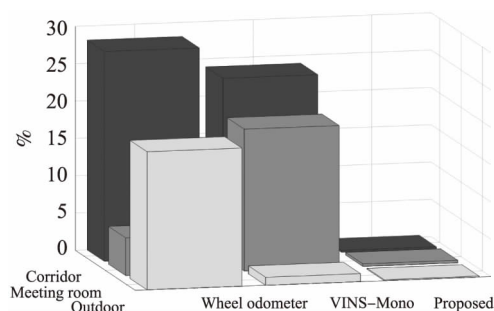


图 9 所提算法和对比算法分别在三种场景下估计轨迹误差对比

5.3 轮速计-相机外参在线标定实验分析

实验中对轮速计-相机外参进行了粗略线下标定,为了进一步优化该参数,可对其进行在线估计。轮速计-相机外参数有 6 个自由度,分别是 3 自由度平移和 3 自由度旋转,然而由于平面运动,只有 5 个自由度可观, z 轴平移不可观^[20],因此只能估计出轮速计-相机外参数中的 x 、 y 轴平移和 roll、pitch、yaw 旋转角。如图 10 和图 11 所示,分别为轮速计-相机平移外参和旋转外参估计值,可以看出,参数估计在经历初始阶段的稍微震荡后,最后都趋于收敛稳定。

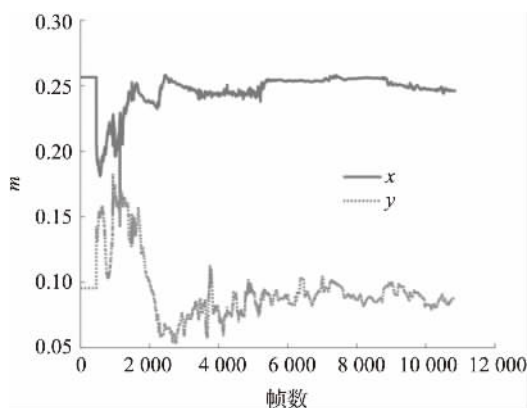


图 10 轮速计-相机平移外参在线估计结果

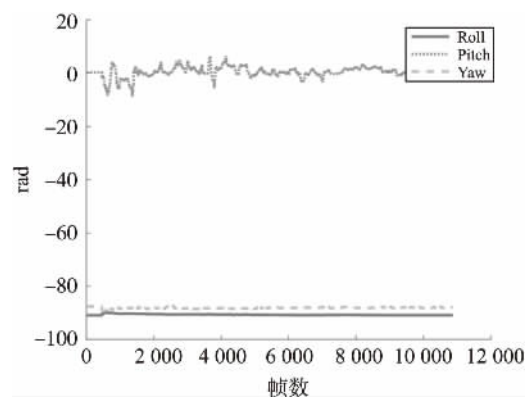


图 11 轮速计-相机旋转外参在线估计结果

5.4 系统运行效率实验分析

本部分实验对比了所提算法与 VINS-Mono 的运行效率。保持其他条件参数相同,如滑动窗口大

小(实验中设置为 10 帧),优化完成的最长时间(实验中设置为 0.04 s)和最大迭代次数(实验中设置为 8)。以会议室场景为例,如图 12 所示,为两个系统在初始化完成后的 30 s 内每次迭代优化所需的时间

间,可以看出本文所提算法一次优化耗时大约是 VINS-Mono 的 1/10,效率得到了很大提高。原因在于轮速计信息的融合并没有引入额外的系统状态,每帧的状态维度为 6 维,相反,VINS-Mono 系统中,每帧需要估计的状态维度为 15 维,因此大大增加了优化迭代时需要的时间。

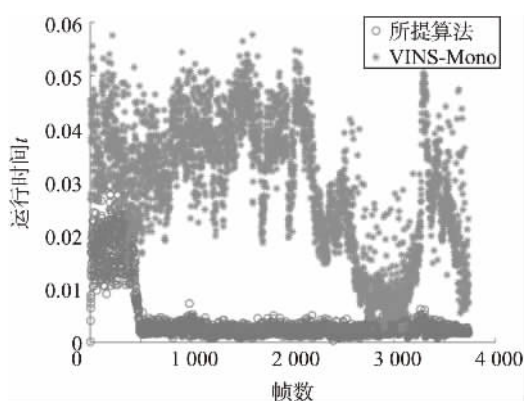


图 12 本文所提算法与 VINS-Mono 运行效率对比

6 结论

本文提出了一个基于视觉-轮速计紧耦合的轮式移动机器人定位算法。采用了基于滑动窗口的非线性优化框架,最小化特征点观测重投影误差和帧间轮速计积分测量误差,来保证系统的实时性和精确度。此外,本文提出了基于视觉-轮速计联合的系统初始化流程,能够快速精确的完成初始化。本文所提系统对相机-轮速计坐标系变换外参进行了在线标定,在粗略的线下标定外参数下,对其进行在线估计优化,减少了对传感器线下高精度标定的要求。通过实验表明,在低成本实验传感器配置的情况下,本文所提算法在小范围室内外轮式机器人应用场景下相比于 VINS-Mono 获得了更鲁棒更精确的定位表现,同时运行效率也有比较大的提高。本文实验场景为室内外规则平面环境,未考虑轮速计打滑造成的影响,下一步将加入对轮速计打滑情况的处理。

参考文献:

- [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] Klein G, Murray D. Parallel Tracking and Mapping for Small AR Workspaces [C] // IEEE & ACM International Symposium on Mixed & Augmented Reality. ACM, 2008.
- [3] Mur-Artal R, Montiel J M M, Tardos J D. ORB-SLAM: A Versatile and Accurate Monocular SLAM System [J]. IEEE Transactions on Robotics, 2015, 31(5) : 1147-1163.
- [4] Mur-Artal R, Tardós J D. Orb-Slam2: An Open-Source Slam System for Monocular, Stereo, and rgb-d Cameras [J]. IEEE Transactions on Robotics, 2017, 33(5) : 1255-1262.
- [5] 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.
- [6] Sun K, Mohta K, Pfrommer B, et al. Robust Stereo Visual Inertial Odometry for Fast Autonomous Flight [J]. IEEE Robotics and Automation Letters, 2018, 3(2) : 965-972.
- [7] 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.
- [8] 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.
- [9] Wu K J, Guo C X, Georgiou G, et al. Vins on Wheels [C] // 2017 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2017: 5155-5162.
- [10] Zhang M, Chen Y, Li M. Vision-Aided Localization for Ground Robots [C] // 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2019: 2455-2461.
- [11] Liu J, Gao W, Hu Z. Visual-Inertial Odometry Tightly Coupled with Wheel Encoder Adopting Robust Initialization and Online Extrinsic Calibration [C] // 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2019: 5391-5397.
- [12] Lee W, Eickenhoff K, Yang Y, et al. Visual-Inertial-Wheel Odometry with Online Calibration [C] // Proceedings of the 2020 International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA. 2020: 24-30.
- [13] Zheng F, Liu Y H. Visual-Odometric Localization and Mapping for Ground Vehicles Using SE(2) -XYZ Constraints [C] // 2019 International Conference on Robotics and Automation (ICRA). IEEE, 2019: 3556-3562.
- [14] Sola J. Quaternion Kinematics for the Error-State Kalman Filter [J]. Laboratoire d'Analyse et d'Architecture des Systèmes-Centre National de la Recherche Scientifique (LAAS-CNRS), Toulouse, France, Tech Rep, 2012.
- [15] Gálvez-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.
- [16] Agarwal S, Mierle K. Ceres Solver [CP/OL]. <http://ceres-solver.org>, 2021-03-15.
- [17] Qin Tong, Cao Shaozhu. VINS-Mono [CP/OL]. <https://github.com/HKUST-Aerial-Robotics/VINS-Mono>, 2019-03-21.
- [18] Paul Furgale, Hannes Sommer, Jérôme Maye, et al. Kalibr [CP/OL]. <https://github.com/ethz-asl/kalibr>, 2021-03-11.
- [19] ZHU Dongfu, CHAI Zheng. CamOdoCalibraTool [CP/OL]. <https://github.com/MegviiRobot/CamOdoCalibraTool>, 2019-07-19.
- [20] He Y, Guo Y, Ye A, et al. Camera-Odometer Calibration and Fusion Using Graph Based Optimization [C] // 2017 IEEE International Conference on Robotics and Biomimetics (ROBIO). IEEE, 2017: 1624-1629.



朱道俊(1996—),男,中国科学技术大学在读研究生,主要研究方向为视觉SLAM,多传感器融合,daojun@mail.ustc.edu.cn;



宋全军(1972—),通讯作者,男,博士,研究员,中国人工智能学会智能机器人专委会副主任委员,主要研究领域为智能机器人,人机接触交互,qjsong@iim.ac.cn.