

引用格式: 陈路, 谢维斯, 谭杰, 等. 面向移动机器人的多传感器紧耦合导航定位方法 [J]. 电子科技大学学报, 2026, 55(1): 109-115.
CHEN L, XIE W S, TAN J, et al. A tightly coupled multi-sensor fusion navigation and localization method for mobile robots[J]. Journal of University of Electronic Science and Technology of China, 2026, 55(1): 109-115.

面向移动机器人的多传感器紧耦合 导航定位方法



陈路, 谢维斯, 谭杰, 陈丽竹, 高勇*

(电子科技大学电子科学与工程学院, 成都 611731)

摘要: 移动机器人依赖单一传感器往往难以克服光照变化、外部干扰、反射表面影响以及累积误差等问题, 限制了环境感知能力和自身位姿测量的精度与可靠性。该文采用一种非线性优化的方法, 实现 (IMU、红外相机、RGB 相机、激光雷达) 数据层面紧耦合组合定位建图系统 IIVL-LM。提出一种基于 RGB 图像信息的实时照度值转换模型, 系统根据不同照度值通过非线性插值法输入视觉 SLAM 模型中进行实时建图, 然后通过动态加权法对红外相机与 RGB 相机的关键帧特征提取融合。在模拟的室内救援场景数据集下, 与多种主流融合定位方法相比, IIVL-LM 在照度变化的苛刻条件下尤其是在低照度下性能提升明显, 平均 RMSE ATE 提升了 23%~39% (0.006~0.013)。IIVL-LM 保证了系统始终会在不少于 3 个传感器有效的状态下进行, 在确保精度的同时对未知开放场景有更强的鲁棒性, 尤其对于室内救援这种复杂场景的应用具有一定的价值。

关键词: 移动机器人; 多传感器融合; 照度转换; 非线性紧耦合; SLAM

中图分类号: TP309.7

文献标志码: A

DOI: 10.12178/1001-0548.2024113

A tightly coupled multi-sensor fusion navigation and localization method for mobile robots

CHEN Lu, XIE Weisi, TAN Jie, CHEN Lizhu, and GAO Yong*

(School of Electronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: Mobile robots relying solely on a single sensor often struggle to overcome challenges such as illumination change, external disturbances, effects of reflective surfaces, and cumulative errors, which limit their environmental perception capabilities as well as the accuracy and reliability of pose estimation. This article adopts a nonlinear optimization method to achieve a tightly coupled integrated localization and mapping system, IIVL-LM, at the data level (IMU, infrared camera, RGB camera, LiDAR). A real-time luminance conversion model based on RGB image information is proposed. The system incorporates varying luminance values into the visual SLAM model through nonlinear interpolation for real-time mapping, then fuses the feature extraction of key frames from the infrared camera and the RGB camera through dynamic weighting. In a simulated indoor rescue scenario dataset, compared to various mainstream fusion positioning methods, the IIVL-LM system exhibits a notable performance improvement under challenging luminance conditions, especially in low-light environments. The average Root Mean Square Error (RMSE) of the Absolute Trajectory Error (ATE) improved by 23% to 39% (0.006 to 0.013). The IIVL-LM system ensures that it operates with at least three active sensors at all times, thereby enhancing its robustness in unknown and open environments while maintaining precision. This capability is particularly valuable for applications in complex settings such as indoor rescue scenarios.

Key words: mobile robots; multi-sensor fusion; illuminance conversion; nonlinear tight coupling; SLAM

随着机器人技术的飞速发展, 搭载视觉等多传感器的机器人在各个领域的应用越来越广泛^[1-3], 如

工业、军事、灾害救援、太空探索及家政服务等^[4-5]。为了实现机器人的自主导航和智能交互, 精确的建

收稿日期: 2024-05-05

基金项目: 人才启航计划 (2025QHXM24)

作者简介: 陈路, 博士生, 主要从事自动控制、计算机视觉检测方面的研究。

*通信作者 E-mail: gaoyong@uestc.edu.cn

图和定位技术, 即 SLAM (simultaneous localization and mapping) 显得尤为重要。该问题主要包括两个领域: 定位、建图。定位指的是精确地确定自身在当前环境中的位置和姿态; 建图是将在周围环境中得到的部分观测值整合在一起的过程。研究者最初的研究中是将定位与建图分开的, 但后来发现它们二者紧密耦合, 即移动机器人在当前环境中定位需要一个精确的地图, 同时为了构建一个好的地图则需要准确的位置。但在现实环境中存在大量不同类型的测量噪声及干扰, 这对解决 SLAM 问题造成很大麻烦, 因此, 通过多传感器融合来解决 SLAM 中的不确定性问题十分关键^[6-7]。近些年来, LOMA 采用了 GPS、激光雷达和 IMU 集成融合对机器人进行状态估计, 通过引入 IMU 来消除激光雷达倾斜问题^[8-10]。MonoSLAM^[11] 是较早提出用单目视觉进行定位导航的方法, 文献 [12] 在此基础上采用滤波算法 EKF 来保持跟踪图像中的特征一致性。PTAM^[13] 则通过提取关键帧来获得更准确的 BA 优化。ORB-SLAM-VI^[14] 提出了一种视觉惯性 SLAM 系统, 可以在长中短期数据中关联, 但 IMU 初始化速度较慢, 损害了鲁棒性。ORB-SLAM3 系统^[15] 融合了 RGB-D 相机并提出了多地图系统, 有效提高了召回率。R3LIVE++^[16] 提出了激光雷达惯性里程计 (LIO) 和视觉惯性里程计 (VIO) 同步实时运行框架, 同时获取雷达测量结果与图像中的几何结构的亮度信息, 并将数据有效紧耦合。

本文通过对多传感器信息进行非线性插值及加权融合多类特征的方法, 实现三维空间环境感知、建模、标定、匹配、测量等。重点提出了在光学变

化强烈, 尤其是低照度甚至黑暗场景下的解决方法, 最终实现移动机器人在未知环境的自主导航。论文的主要贡献可以概括为以下 4 点。

1) 提出了一种基于 RGB 图像的照度值转化模型, 通过对模拟的室内救援场景对最大光照强度和最小光照强度下的 RGB 图多次采样均值后通过转化作为归一化的标准值。

2) 基于实时照度值对视觉传感器的图像进行非线性插值, 并将两种特征进行加权融合, 权值分配根据照度值设置的阈值动态变化, 保证多传感器数据紧耦合。

3) 基于 R3LIVE++ 框架提出了一种适用于复合机器人的 IMU、视觉 (红外和 RGB)、激光雷达紧密耦合的组合定位建图系统 IIVR-LM, 同时结合红外相机具备深度信息的获取能力, 优化了其中的 VIO 融合模块。

4) 在模拟室内的救援场景下, 对比同类框架进行了较为充分的实验。

1 多传感器导航定位紧耦合方法

1.1 IIVL-LM 系统框架

本文在 R3LIVE++ 融合框架的基础上对 VIO 进行细化, 针对完全黑暗导致 RGB 相机的特征失效问题, 将 VIO 模块分成红外和 RGB 相机两个子模块, 通过照度转换模块实时计算照度达到红外帧与 RGB 帧的有效紧耦合, 并通过插值法引入关键帧, 有效地将不同照度值环境划分为 3 种非线性融合阶段。在中间阶段采用加权融合的方法提取图像特征, 同时采用 Ceres Solver^[15] 方法求解该模块非线性优化插值问题, 系统如图 1 所示。

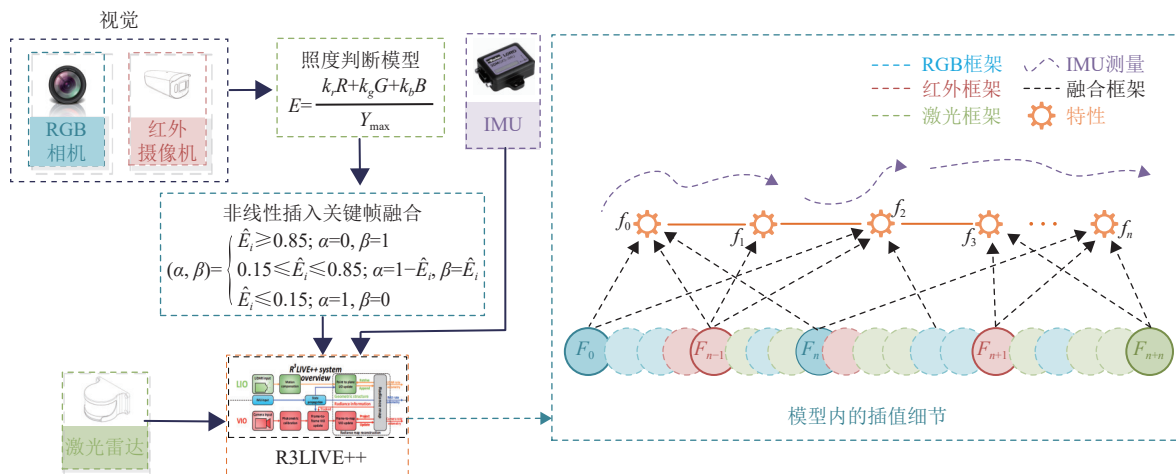


图 1 IIVL-LM 系统各模块构成示意图

1.2 多传感器融合基本理论

激光雷达融合方法是基于当前迭代的位姿估计, 将各帧点云转换到全局坐标系下进行融合:

$${}^G P_s = {}^G \tilde{P}_{I_k} ({}^L R_L {}^L P_s + {}^L P_L) + {}^G \tilde{P}_{I_k} \quad (1)$$

式中, ${}^G P_s$ 为全局坐标系中点 P 的位置; ${}^G \tilde{P}_{I_k}$ 为从迭代坐标系到全局坐标系的旋转矩阵; ${}^L R_L$ 为从局部坐标系到迭代坐标系的旋转矩阵; ${}^L P_s$ 为局部坐标系中点 P 的位置; ${}^L P_L$ 为迭代坐标系中局部坐标的位置; ${}^G \tilde{P}_{I_k}$ 为在全局坐标系中第 k 个迭代的中间坐标系位置。

为了将点云精准匹配, 找到最近邻点拟合一个具有法向和质心的平面, 使用激光雷达测量残差:

$$r_1(\tilde{x}_k, {}^L P_s) = \mathbf{u}_s^T ({}^G P_s - Q_s) \quad (2)$$

式中, \tilde{x}_k 为点云坐标; Q_s 为点 P 在目标点云中找到的最近邻匹配点或其拟合平面投影点坐标; \mathbf{u}_s 表示参考平面 (或目标点云中对应点) 的法向量。投影误差的计算采用 Lucas Kanade 光流来定位这些点在当前图像中的位置。然后迭代地最小化点 P 的重投影误差:

$$r_c(\tilde{x}_k, \rho_{s_k}, {}^G P_s) = \rho_{s_k} - \pi({}^G P_s, \tilde{x}_k) \quad (3)$$

$$\pi({}^G P_s, \tilde{x}_k) = \pi_{\text{ph}}({}^G P_s, \tilde{x}_k) + \frac{\mathbf{I}_{\tilde{c}_k}}{\Delta t_{k-1,k}} (\rho_{s_k} - \rho_{s_{k-1}}) \quad (4)$$

式中, $r_1(\tilde{x}_k, {}^L P_s)$ 表示使用激光雷达计算的残差; $r_c(\tilde{x}_k, \rho_{s_k}, {}^G P_s)$ 表示重投影误差; $\pi({}^G P_s, \tilde{x}_k)$, $\pi_{\text{ph}}({}^G P_s, \tilde{x}_k)$ 是投影函数, 用于计算从 3D 点到 2D 图像平面的映射; $\mathbf{I}_{\tilde{c}_k}$ 表示从中间坐标系到相机坐标系的平移向量。

同理在视觉融合模块的像素跟踪误差及地图点定位误差为:

$${}^G P_s = {}^G P_s^{\text{gt}} + n_{p_s} \quad n_{p_s} \sim N(0, \sum n_{p_s}) \quad (5)$$

$$P_{s_k} = P_s^{\text{gt}} + n_{p_{s_k}} \quad n_{p_{s_k}} \sim N(0, \sum n_{p_{s_k}}) \quad (6)$$

式 (5)、式 (6) 表示噪声模型, 通常假设为高斯分布; n_{p_s} 对应位置噪声; $n_{p_{s_k}}$ 对应测量噪声。在经过误差校正后的真实系统状态下, 观测残差应满足零残差条件:

$$0 = r_c(x_k, P_{s_k}^{\text{gt}}, {}^G P_s^{\text{gt}}) \approx r_c(\tilde{x}_k, \rho_{s_k}, {}^G P_s) + \mathbf{H}_s^y \delta \tilde{x}_k + \beta_s \quad (7)$$

式中, \mathbf{H}_s^y 是雅可比矩阵或系统矩阵; $\delta \tilde{x}_k$ 是状态向

量的变化量; β_s 是误差项。

则 frame-to-map 辐射误差为:

$$r_c(\tilde{x}_k, {}^G P_s, \gamma_s) = \Phi_s - \gamma_s, \Phi_s = \tilde{\epsilon}_k \Gamma_k(\tilde{\rho}_{s_k}) \quad (8)$$

式中, Φ_s 为目标函数; γ_s 为估计值。

最后通过贝叶斯更新将地图中的辐射矢量与新观测到的辐射矢量进行融合:

$$\sum_{n_{\tilde{\gamma}_s}} = \left(\left(\sum_{n_{\gamma_s}} + \sigma_{\text{ic}}^2 \Delta t_{\gamma_s} \right)^{-1} + \sum_{n_{\Phi_s}}^{-1} \right)^{-1} \quad (9)$$

$$\tilde{\gamma}_s = \left(\left(\sum_{n_{\gamma_s}} + \sigma_{\text{ic}}^2 \Delta t_{\gamma_s} \right)^{-1} \gamma_s + \sum_{n_{\Phi_s}}^{-1} \Phi_s \right)^{-1} \sum_{n_{\tilde{\gamma}_s}} \quad (10)$$

$$\gamma_s = \tilde{\gamma}_s, \sum_{n_{\gamma_s}} = \sum_{n_{\tilde{\gamma}_s}} \quad (11)$$

式中, $\sum_{n_{\tilde{\gamma}_s}}$ 是协方差更新公式。

1.3 照度转换模型

考虑到实时性及照度不会突然短期内急速变化, 照度检测可以根据应用场景选择固定间隔帧数抽检, 或者根据回环定位检测时抽取的关键帧同步进行判断, 转化公式为:

$$E = \frac{k_r R + k_g G + k_b B}{Y_{\text{max}}} \quad (12)$$

式中, R, G, B 分别是图像 RGB 通道的值; k 为通道系数, 其中 k_r, k_g, k_b 分别设置为 0.299、0.587、0.114。 Y_{max} 是 RGB 通道的最大值, 8 位 RGB 图像下为 255。

本文红外相机可以同时输出两种图像, 因此自动时间戳帧对齐。设置和两个相机帧率相同的每 60 帧检测一次 RGB 图像照度值, 并取前后各 2 帧去掉最大最小值后求和取平均, 那么检测帧的照度为:

$$E_i = \frac{\sum_{j \in \{i-2, i-1, i, i+1, i+2\}} E_j - \max(E_j) - \min(E_j)}{3} \quad (13)$$

采用样本采集室内最大、最小照度进行归一化处理, 在平均最大照度 50 287 lux, 最小 21 lux 的情况下, 共计 48 帧得出 E_{max} 和 E_{min} , 那么归一化后的检测帧的照度值为:

$$\hat{E}_i = \frac{E_i - E_{\text{min}}}{E_{\text{max}} - E_{\text{min}}} \quad (14)$$

1.4 非线性特征权重分配方法

根据 \hat{E}_i 值动态非线性分配红外图像权重值 α 与 RGB 图像权重值 β , 如式 (15) 所示:

$$(\alpha, \beta) = \begin{cases} \widehat{E}_i \geq 0.85 & \alpha = 0, \beta = 1 \\ 0.15 \leq \widehat{E}_i \leq 0.85 & \alpha = 1 - \widehat{E}_i, \beta = \widehat{E}_i \\ \widehat{E}_i \leq 0.15 & \alpha = 1, \beta = 0 \end{cases} \quad (15)$$

当照度转化值大于 0.85 时, 红外图像在视觉定位模块的作用显著降低, 因此系统直接舍弃红外, 当照度转化值小于 0.15 时, RGB 图像也在失效边缘, 在此条件下舍弃使用。当照度转化值在中间范围时, 系统采取线性分配特征值。

图像平均亮度偏差特征 AvgDev 与融合纹理特征 Contrast 的计算方式为:

$$\text{AvgDev} = \frac{1}{WH} \sum_{i=1}^W \sum_{j=1}^H (I(i, j) - 128) \quad (16)$$

$$\text{Contrast} = \sum_{i,j=0}^{G-1} (i-j)^2 P(i, j) \quad (17)$$

式中, W, H 代表图片长宽值; i, j 为像素坐标; I 为像素值; P 为共生矩阵。

则加权后的特征融合公式为:

$$F_f = \alpha (\text{AvgDev}_{\text{IR}} + \text{Contrast}_{\text{IR}}) + \beta (\text{AvgDev}_{\text{VL}} + \text{Contrast}_{\text{VL}}) \quad (18)$$

本文依赖动态权值对 3 个传感器插帧值进行优化, 考虑到照度不影响 LiDAR 的效果, 那么

LiDAR 插值比例不变, 而 RGB 与红外插值则根据权值的变化而变化, 如图 2 所示。

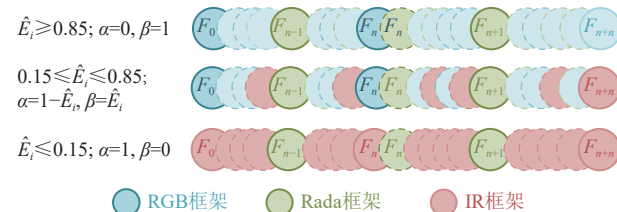


图 2 基于权值的非线性插值帧方法

1.5 基于深度信息的视觉与 IMU 融合 (VIO) 框架

在完成红外与 RGB 相机的有效融合并完成视觉前端的工作以后, 基于红外相机具备深度信息获取能力对 R3LIVE++ 中的 VIO 模块进行优化。根据移动机器人的移动特点, 并结合 IMU 和 VO 各自的特点建立系统模型, 优化后的视觉惯导里程计模型如图 3 所示。

根据 IMU 瞬时反应快的特性, 将 IMU 的测量值用来驱动过程模型以适应移动机器人的机动性。同时, 利用立体视觉不会产生积累误差的特性, 将立体视觉的定位估计结果作为观测模型的观测值以校正 IMU 的误差。最后, 根据机器人在运动时可能产生的瞬时滑动和跳跃的约束条件, 将 y 轴和 z 轴上的观测速度建模为零均值的噪声。

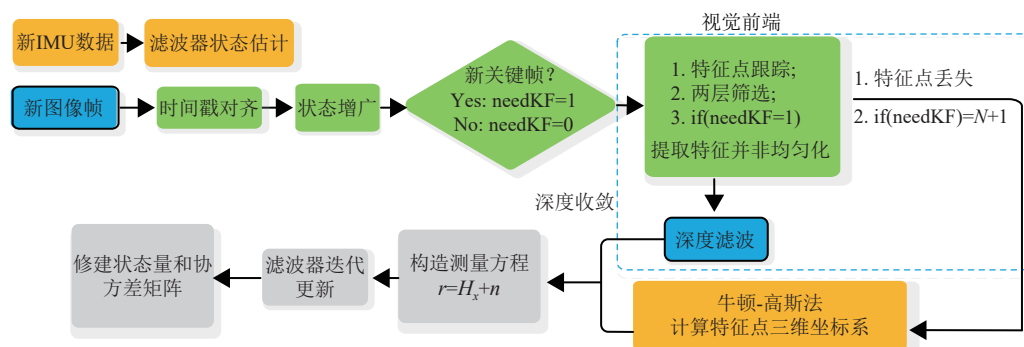


图 3 优化后的视觉-惯导里程计 (VIO)

2 实验与分析

2.1 实验平台

本文实验硬件平台为多功能移动机器人, 其中包括 AGVroute 底盘、Duco 协作机械臂、组合建图定位系统, 底盘控制为 FX 嵌入式处理器, 控制指令采用 API 通信协议传输。测试硬件实况如图 4 所示, 核心单元为 LIVOX 的 LiDAR 传感器、内嵌 IMU 的 RGB 相机和红外相机构成的 IIVL-LM 系统, 在三维空间坐标系下进行地图重构和位姿估

计。实际运行该平台完成建图及定位测试。各传感器具体参数如表 1 所示。

软件环境为 Ubuntu 20.04 操作系统, 并集成了 ROS (机器人操作系统) 以协调和整合各个模块框架。此外, 机载电脑硬件为 8 G 显存的 NVIDIA GeForce GTX 1650 Ti GPU, 以及 Intel (R) Core (TM) i5-10300H CPU, 主频为 2.50 GHz, 以验证本文提出的方法在一般通用硬件条件下的实时性与稳定性。

2.2 实验环境及数据集

实验在模拟的室内救援场景下进行, 场地面积 484.3 m², 非完全封闭, 有外部日照光源透入。本文通过分时段测试来实现照度变化的实验条件, 时间均匀地分布在 0~24 时区间内, 此外还通过室内灯光控制来实现局部照度变化的测试条件, 最终通过照度计计算当前照度, 照度范围为 10~50 000 lux。移动机器人累积测试行驶里程 4 792 m。

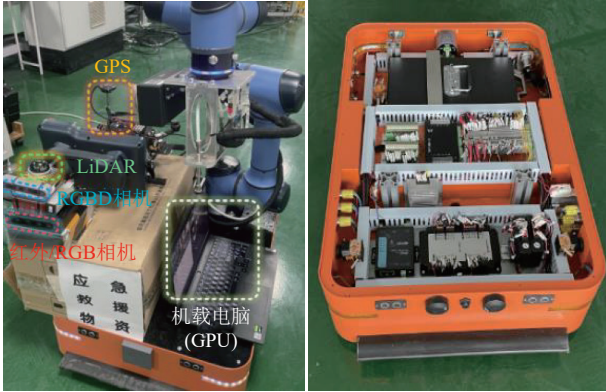


图4 部署在移动机器人上的 IIVL-LM 系统及传感器示意图

表1 传感器主要参数

传感器	参数	参数值
LiDAR	波长/nm	905
	FOV/(°)	水平360, 垂直-7~52
	点云输出及帧率	200 000点/秒, 10 Hz
RGB相机	分辨率	752*480
	最大帧率/FPS	60
IMU	频率/Hz	100~500
红外相机	IR有效距离/m	3
	帧率/FPS	60

2.3 实验评价标准

除了对理论定位曲线与系统实际定位曲线 X/Y 轴误差及方差进行评估外, 对比参照实验主要采用 ATE (absolute trajectory error) 与 RMSE (root mean square error) 结合^[16] 来评估系统精度。该方法直接计算传感器位姿的真实值与定位系统估计值之间的差。根据位姿的时间戳将真实值和估计值进行对齐, 然后计算每对位姿之间的差值, 该标准非常适合于评估定位系统的性能。第 i 帧的 ATE 定义为:

$$F_i = Q_i^{-1} S P_i \quad (19)$$

式中, Q 为真实位姿; S 为转换矩阵; P 为估计位姿。

结合均方根 RMSE 可以更精确地统计误差:

$$\text{RMSE}(F_{1:n}, \Delta) = \left(\frac{1}{m} \sum_{i=1}^m \|\text{trans}(F_i)\|^2 \right)^{\frac{1}{2}} \quad (20)$$

式中, Δ 为时间间隔; n 为总帧数; $m = n - \Delta$;

$\text{trans}(F_i)$ 代表相对位姿误差中的平移部分。

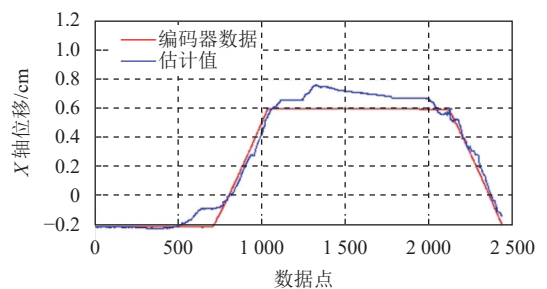
2.4 实验结果

由于测试地面不光滑, 不考虑轮子打滑的因素, 轮式编码器采集的定位数据曲线可视为理论真实曲线。本文首先采用轮式编码器曲线与 IIVL-LM 系统下的自身估计的位置曲线进行比较, 在标准室内光照照度值 30 500 lux 的条件下其中之一的测试数据如图 5a~5b 所示, 经过 10 次测试平均 X 轴误差为 4.26 cm, Y 轴误差为 4.57 cm, 均方误差为 X 轴 0.33%, Y 轴 0.54%。

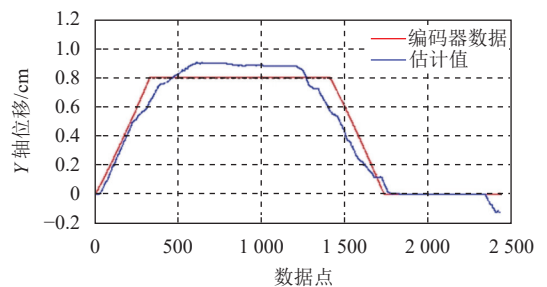
随后, 本文采用上述相同的方法, 在不同时段照度不同的情况下将 IIVL-LM 系统与常见的 ORB-SLAM^[17]、VINS-Mono^[18]、R3LIVE^[19]、DSO^[20]、SVO^[21] 定位建图算法进行对比, 具体结果如表 2 所示, 运行线路如图 5c 所示。

2.5 实验分析

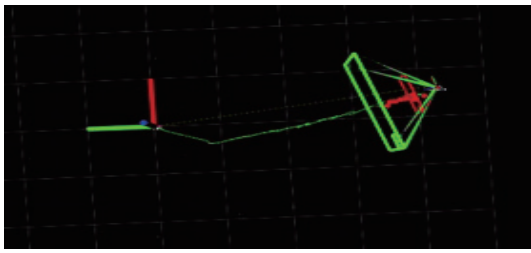
本文主要侧重红外图像与 RGB 图像的融合, 应对光照变化或者黑暗条件下的定位及建图, 从表 2 可以看出, 与主流的方法相比, 在夜间照度不足 20~8 000 lux 的情况下均取得了最好的效果, RMSE ATE 值领先第 2 0.001~0.016。不结合激光雷达的算法在黑暗情况下仅仅依靠 IMU 等单一传感器对精度影响很大, 其 RMSE ATE 值在 910 m 较短的行程下最大达到了 0.044, 这是一个较大的误差, 因为仅仅依靠激光雷达和惯性导航可能会有偏振及累计误差带来的影响。反之, 在光照充足稳定的条件下, 各算法的差距并不明显, 验证了本文方法在光照不足或者变化下的鲁棒性。



a. X轴数据对比



b. Y轴数据对比



c. 移动机器人实际测试运行轨迹

图 5 IIVL-LM 系统下复核机器人的 X,Y 轴数据对比及实际轨迹

表 2 各方法在不同照度下的定位性能对比 (RMSE ATE)

时段	平均照度/lux	距离/m	ORB-SLAM	VINS-Mono	R3LIVE	DSO	SVO	IIVL-LM
0~4 am	20	910	0.044	0.042	0.037	0.028	0.033	0.018
4~8 am	1 200	800	0.036	0.040	0.016	0.021	0.018	0.016
8~12 am	17 000	657	0.019	0.015	0.011	0.020	0.018	0.012
12~16 pm	30 000	745	0.026	0.021	0.023	0.026	0.032	0.021
16~20 pm	20 000	800	0.031	0.027	0.030	0.041	0.039	0.032
20~24 pm	8 000	880	0.042	0.044	0.038	0.036	0.037	0.020
Avg RMSE ATE 与总体平均值比较 (0.028)			0.033	0.032	0.026	0.029	0.030	0.020
			-18%	-14%	+7%	-4%	-7%	+40%

3 结束语

本文提出一个多传感器非线性紧耦合的建图定位框架 IIVL-LM。系统结合红外相机、IMU、RGB、激光雷达的融合可以有效地在黑暗、光照变化的条件下同时保证至少 3 个传感器是有效的,同时可以消除其他传感单元在长时间或者样本不足的情况下的累积误差及漂移问题,有效提升了系统的鲁棒性。采用 RGB 图像分析计算出实时照度,通过转换与最大照度和最小照度的对比生成照度转换模型。为了在数据层面紧耦合各传感器,采用根据照度变化的动态非线性插值方法插入关键帧实现系统的回环检测,并根据实时照度值对红外帧及 RGB 帧进行加权融合,有效保证了提取特征的效率。但是还有以下值得深入研究的方向。

1) 多传感器融合的数据量庞大,如何在保证实时性的前提下,对高分辨率 RGB 图像与其他传感器观测数据进行快速计算和有效融合,是后续值得研究的问题之一。

2) 该方法仅在地面移动机器人进行了研究验证,但是救援往往是立体的,该系统红外相机支

持 3D 深度检测。未来三位一体救援、自动化工厂应用等领域必然需要在无人机结合地面机器人上进行三维建图实现 3D 建图定位及导航。

3) 在环境建模与认知方面继续深入研究。传统的环境建模方法已经无法满足机器人在未知复杂环境中自主导航对环境信息的要求,在后续对环境理解方面,自顶向下,采用基于模型的手段(如马尔可夫随机场、条件随机场等)表达场景的位置信息、尺度信息、物体之间的关系信息以及场景中存在某一物体的概率信息等,无需对场景中的物体仔细辨识,就能像人一样根据少量的信息快速对场景判断。

4) 拓展行为的自学习研究,设计行为控制器时在特征层上融合关于地形的理解,采用强化学习等机器学习方法在线学习,提高行为控制器对未知环境的适应性和灵活性。

参考文献

- [1] AOYAMA Y, SARAVANOS A D, THEODOROU E A. Receding horizon differential dynamic programming under parametric uncertainty[EB/OL]. [2023-10-10]. <https://arxiv.org/pdf/2104.10836>.
- [2] SHI Y L, ZHANG W M, YAO Z, et al. Design of a hybrid indoor location system based on multi-sensor fusion for robot navigation[J]. *Sensors*, 2018, 18(10): 3581.
- [3] SHEN S J, MULGAONKAR Y, MICHAEL N, et al. Multi-sensor fusion for robust autonomous flight in indoor and outdoor environments with a rotorcraft MAV[C]// Proceedings of the IEEE International Conference on Robotics and Automation. New York: IEEE, 2014: 4974-4981.
- [4] CHEN L, LI G, ZHAO K Q, et al. A perceptually adaptive long-term tracking method for the complete occlusion and disappearance of a target[J]. *Cognitive Computation*, 2023, 15(6): 2120-2131.
- [5] CHEN L, LI G, XIE W S, et al. A survey of computer vision detection, visual SLAM algorithms, and their applications in energy-efficient autonomous systems[J]. *Energies*, 2024, 17(20): 5177.
- [6] CHEN C, ZHU H, LI M G, et al. A review of visual-inertial simultaneous localization and mapping from filtering-based and optimization-based perspectives[J]. *Robotics*, 2018, 7(3): 45.
- [7] LE GENTIL C, VIDAL-CALLEJA T, HUANG S D. IN2LAMA: INertial lidar localisation and MAPPING[C]// Proceedings of the International Conference on Robotics and Automation. Montreal, QC: IEEE, 2019: 6388-6394.
- [8] ZHANG J, SINGH S. Low-drift and real-time lidar odometry and mapping[J]. *Autonomous Robots*, 2017, 41(2): 401-416.
- [9] LYNEN S, ACHELIK M W, WEISS S, et al. A robust

- and modular multi-sensor fusion approach applied to MAV navigation[C]//Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. New York: IEEE, 2013: 3923-3929.
- [10] HENING S, IPPOLITO C A, KRISHNAKUMAR K S, et al. 3D LiDAR SLAM integration with GPS/INS for UAVs in urban GPS-degraded environments[C]//Proceedings of the AIAA Information Systems-AIAA Infotech @ Aerospace. Reston, Virginia: AIAA, 2017: 0448.
- [11] DAVISON A J. Real-time simultaneous localisation and mapping with a single camera[C]//Proceedings of the Proceedings 9th IEEE International Conference on Computer Vision. New York: IEEE, 2003: 1403-1410.
- [12] CIVERA J, DAVISON A J, MARTÍNEZ MONTIEL J M. Inverse depth parametrization for monocular SLAM[J]. *IEEE Transactions on Robotics*, 2008, 24(5): 932-945.
- [13] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces[C]//Proceedings of the 6th IEEE and ACM International Symposium on Mixed and Augmented Reality. New York: IEEE, 2007: 225-234.
- [14] CAMPOS C, ELVIRA R, RODRÍGUEZ J J G, et al. ORB-SLAM3: An accurate open-source library for visual, visual-inertial, and multimap SLAM[J]. *IEEE Transactions on Robotics*, 2021, 37(6): 1874-1890.
- [15] AGARWAL S, KEIR M. Ceres solver: Tutorial & reference[EB/OL]. [2023-10-11]. https://www.helloandroid.cn/android/4.3_r1/download/external/ceres-solver/docs/ceres-solver.pdf.
- [16] KASAR A. Benchmarking and comparing popular visual SLAM algorithms[EB/OL]. [2024-11-24]. <https://arxiv.org/abs/1811.09895>.
- [17] MUR-ARTAL R, TARDÓS J D. Visual-inertial monocular SLAM with map reuse[J]. *IEEE Robotics and Automation Letters*, 2017, 2(2): 796-803.
- [18] QIN T, LI P L, SHEN S J. VINS-mono: A robust and versatile monocular visual-inertial state estimator[J]. *IEEE Transactions on Robotics*, 2018, 34(4): 1004-1020.
- [19] LIN J R, ZHANG F. R³LIVE++: A Robust, real-time, radiance reconstruction package with a tightly-coupled LiDAR-Inertial-Visual state estimator[EB/OL]. [2024-11-08]. <https://arxiv.org/abs/2209.03666>.
- [20] ENGEL J, KOLTUN V, CREMERS D. Direct sparse odometry[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(3): 611-625.
- [21] FORSTER C, ZHANG Z C, GASSNER M, et al. SVO: Semidirect visual odometry for monocular and multicamera systems[J]. *IEEE Transactions on Robotics*, 2017, 33(2): 249-265.

责任编辑 叶 芳