

基于改进Vins-Mono的INS/视觉组合定位方法研究

吴浩¹, 陶为戈^{1*}, 孙志刚²

¹江苏理工学院电气信息工程学院, 江苏 常州

²哈尔滨工业大学电气工程及自动化学院, 黑龙江 哈尔滨

收稿日期: 2024年4月28日; 录用日期: 2024年5月24日; 发布日期: 2024年5月31日

摘要

针对现有Vins-Mono算法在惯性导航系统(Inertial Navigation System, INS)与视觉组合定位中存在图像特征点提取分布不均匀、特征点匹配准确率低和系统定位精度不高的问题, 提出一种改进的Vins-Mono算法。首先, 采用二叉树算法划分图像特征点疏密不同的区域, 在各个区域内同时进行特征点提取, 加快了提取速率; 然后, 采用RANSAC算法进行特征点匹配, 提高了匹配准确率, 并结合汉明距离加快了匹配时间; 最后, 采用LM算法取代高斯牛顿算法, 添加信赖区域控制更新步长, 提高了系统定位精度。相比于传统Vins-Mono算法, 本文算法在特征点提取分布均匀度上提高了54%, 在特征点匹配准确率上提高了28.4%, 在不同难度数据集上的平均定位精度提高了34%。

关键词

改进Vins-Mono, 二叉树算法, RANSAC算法, LM算法, 组合定位

Research on INS/Vision Combination Localization Method Based on Improved Vins-Mono

Hao Wu¹, Weige Tao^{1*}, Zhigang Sun²

¹School of Electrical and Information Engineering Jiangsu University of Technology, Changzhou Jiangsu

²School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin Heilongjiang

Received: Apr. 28th, 2024; accepted: May 24th, 2024; published: May 31st, 2024

*通讯作者。

文章引用: 吴浩, 陶为戈, 孙志刚. 基于改进 Vins-Mono 的 INS/视觉组合定位方法研究[J]. 计算机科学与应用, 2024, 14(5): 255-264. DOI: 10.12677/csa.2024.145134

Abstract

Aiming at the problems of the existing Vins-Mono algorithm in combined positioning of Inertial Navigation System (INS) and vision, such as uneven distribution of image feature points extraction, low matching accuracy, and low positioning accuracy, an improved Vins-Mono algorithm is proposed. Firstly, the quadtree algorithm is used to divide regions with different densities of feature points and extract feature points simultaneously in each region to speed up the extraction rate. Then, the RANSAC algorithm is used to match the feature points, which improves the matching accuracy and speeds up the matching time by combining the Hamming distance. Finally, the LM algorithm is used to replace Gauss Newton algorithm, and trust region control is added to update the step size, which improves the positioning accuracy of the system. Compared with the traditional Vins-Mono algorithm, the algorithm in this paper improves the distribution uniformity of feature point extraction by 54%, the accuracy of feature point matching by 28.4%, and the average positioning accuracy on different difficulty datasets by 34%.

Keywords

Improved Vins-Mono, Quadtree Algorithm, RANSAC Algorithm, LM Algorithm, Combined Location

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着自动驾驶技术的兴起，人们对定位精度的要求越来越高。传统的定位方法依靠全球卫星导航系统(Global Navigation Satellite System, GNSS)和惯性导航系统(Inertial Navigation System, INS)，二者组合后能弥补各自不足，可以满足大多数场景下的定位精度要求。然而，处于封闭环境下，GNSS 信号受到屏蔽，组合定位系统退化为单一 INS，误差会随时间累积，无法长时间高精度定位。由于相机成本低、体积小、受封闭环境影响小等优点，利用视觉进行定位的视觉里程计技术得到了快速发展，视觉里程计和 INS 组合后的视觉 - 惯性里程计系统抗干扰性强，定位精度高，成为当下研究热点。

Leutenegger [1]等提出了一种基于 IMU 预积分处理办法的 VIO 算法 OKVIS，相较于 MSCKF，提高了系统运行速率，以非线性优化算法取代滤波算法，使得系统定位精度更高；我国香港科技大学沈劭劼 [2]团队提出了 Vins-Mono 算法，在 OKVINS 的基础上加入回环检测大大提高了系统定位精度，减少了轨迹漂移。MSCKF、OKVIS、Vins-Mono 这三种为目前主流的 INS/视觉组合定位方法，其中 Vins-Mono 算法的定位精度最高。

Vins-Mono 算法中对图像特征点提取、匹配采用的是 ORB 算法和汉明距离，存在特征点提取不均匀、匹配准确率低的问题。季莘翔[3]等在 ORB 算法中引入密度峰值聚类算法，通过对提取出的特征点的密度峰值计算，优化特征点提取的有效性；余正强[4]等先通过预设条件，使得特征点稀疏化，再引入特征点数量调整机制，改善了特征点分布过于集中的现象，提高了视觉里程计的定位精度；骆大森[5]等在特征点匹配之前，先用相位相关法求取连续图像间的像素偏移，再采用最速邻近算法进行特征点匹配，相位相关法计算出的像素偏移作为限制条件，提高了匹配的准确率；陈坚炜[6]等采用粗匹配和精匹配结合的策略，粗匹配中匹配限制阈值为 0.6，精匹配中利用单位球的约束除去误匹配的特征点；赵磊[7]等在后

端优化中, 选用 g2o 图优化的方式进行位姿估计, 提高了系统定位的精度; 卢艳军[8]等在传统图优化算法的基础上, 在代价函数中引入鲁棒核函数, 对环境引起的传感器误差进行过滤, 提高了系统的定位精度。以上这些方法都改进了 INS/视觉组合定位的精度, 但其结果仍有提升空间。

文本针对 Vins-Mono 算法中存在图像特征点提取不均匀、匹配准确率低、定位精度低等问题, 提出了一种改进的 INS/视觉组合定位方法。主要从上述三个方面进行改进, 在特征点提取中引入二叉树算法, 改善特征点提取后不均匀的情况; 采用 RANSAC 算法进行特征点匹配, 提高匹配正确率, 并结合汉明距离加快了匹配时间; 使用 LM 优化算法提高了系统定位精度。

2. Vins-Mono 算法原理

Vins-Mono 算法结构[9]如图 1 所示, 由输入端、前端视觉里程计、回环检测、后端非线性优化、输出端组成。

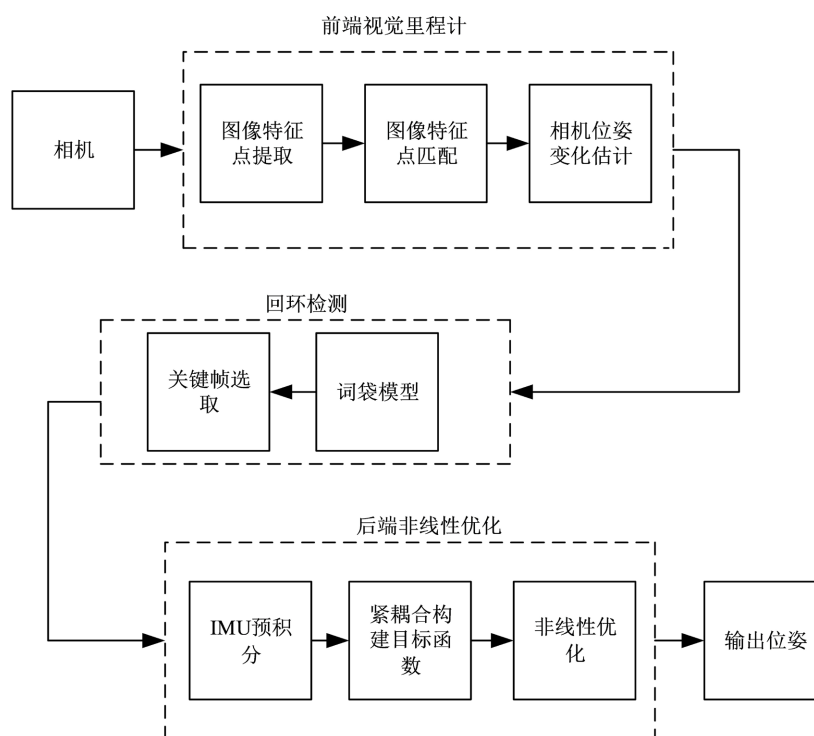


Figure 1. Vins-Mono structure
图 1. Vins-Mono 结构图

输入端通过相机采集图片, 前端视觉里程计接受到图片后, 采用 ORB 算法对图像进行特征点提取, 以汉明距离作为阈值, 匹配出连续两帧图像中相同的特征点, 根据匹配后的特征点的像素坐标变化, 采用对极几何算法估计出相机实际中的位姿变化。前端视觉里程计完成对相机位姿估计后进入回环检测模块, 基于词袋模型判断采集的图像帧是否为关键帧, 如果连续的图像帧都是关键帧, 则触发了轨迹回环, 将作为约束条件引入到后端非线性优化中。在后端非线性优化中, INS 通过惯性测量单元获取数据, 将 IMU 数据进行预积分处理减少计算冗余, 根据前端视觉里程计中相机位姿的变化结果、IMU 预积分解算结果和回环检测的约束条件, 以紧耦合的方式构建目标函数, 再通过高斯牛顿算法进行优化, 最后将优化后的位姿结果输出。

3. Vins-Mono 算法改进

针对传统 Vins-Mono 算法中存在图像特征点提取不均匀、特征点匹配准确率低和定位精度不高的问题, 本文分别在前端视觉里程计中的特征点提取、特征点匹配和后端非线性优化中的优化环节进行了改进。

3.1. 基于四叉树 ORB 算法的特征点提取

Vins-Mono 算法中特征点提取采用了 ORB 算法, ORB 算法将 FAST 关键点 BRIEF 描述子结合。FAST 关键点是比邻像素间的差异得出来的点[10], 对于图中某一像素点 p , 其亮度为 I_p , 将 p 与以 p 为圆心某一半径内的所有像素的亮度进行对比, 像素的点亮度与点 p 的亮度的差超过阈值且区域内满足条件的点的数量超过 N , 则点 p 被认定为特征点, 一般情况下, 圆半径为 3, N 为 12。接着使用 BRIEF 对这些特征点进行描述, BRIEF 是二进制编码描述子, BRIEF 描述向量由 0 和 1 组成, 0 和 1 表示的是相邻两个像素点的灰度值的大小关系。

不经处理提取出的特征点在整个图像中会出现“扎堆”现象, 特征点冗余度过高加大了计算量, 降低了系统定位精度, 因此本文引入四叉树算法[11]。四叉树算法将图像帧划分为 2×2 的网格, 在每个网格中, 如果检测到的特征点数量过多, 则可以根据设置的阈值判断是否需要继续分割网格, 通过这种方式, 可以减少区域内特征点的密集程度, 最后将图像拼接从而均匀分布特征点。但是如果特征点数量过多, 会导致四叉树的分割次数过多, 降低了检索效率, 因此本文将分割最大次数限制为 4 次, 不仅提高了特征点提取的速度, 还通过这种方式提高了特征点质量, 四叉树 ORB 算法流程如图 2 所示。

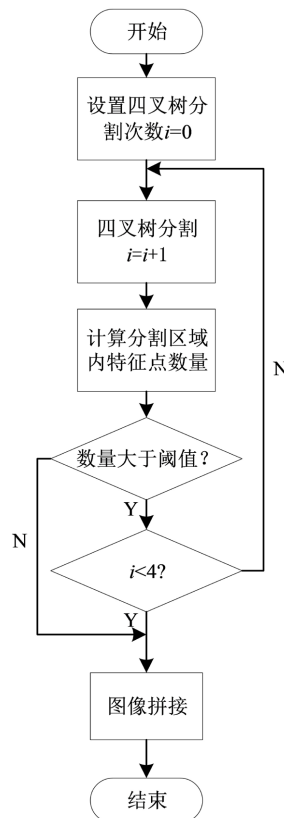


Figure 2. Quadtree ORB algorithm structure
图 2. 四叉树 ORB 算法流程图

3.2. 基于 RANSAC 算法的特征点匹配

Vins-Mono 算法中特征点通过设置汉明距离的阈值并将计算结果与之比较来进行匹配，设连续两帧图像特征点集合为 $\{m_1, m_2, \dots\}$ 、 $\{n_1, n_2, \dots\}$ ，将特征点二进制描述子不同位的数量作为汉明距离，汉明距离 $D(k_1, k_2)$ 反应了特征点间的相似关系，该算法的评价指标单一且计算量大，匹配的准确率低，汉明距离的具体表达式为：

$$D(k_1, k_2) = \sum_{i=1}^n x_i \oplus y_i \quad (1)$$

本文采用结合汉明距离的 RANSAC 算法。RANSAC [12] 算法通过随机选择部分数据子集来估计模型参数，并使用估计的模型对数据进行测试。RANSAC 算法主要过程为：首先，从特征点集中随机选取一定数量的点，然后使用最小二乘法拟合模型并得到模型的参数，再计算其他特征点与拟合模型的拟合度，如果小于阈值则为内点。不断重复上过程，当迭代达到一定次数后，或者最佳模型出现时，选取保存内点数最多的那一组数据作为最终结果。其中迭代次数 s 应满足：

$$s = \frac{\ln(1-p)}{\ln(1-\alpha^n)} \quad (2)$$

其中， α 表示外点的比例， p 表示置信度。为了抑制因误差而导致迭代次数过大，收敛过程耗时较长的情况。将两倍的汉明距离作为约束条件，提前对特征点进行质量排序，每次迭代只在高质量特征点中进行模型估计，大大提高了匹配的准确性和速度。

3.3. 基于 LM 算法的非线性优化

Vins-Mono 算法在后端优化中，非线性优化算法采用高斯牛顿算法。紧耦合将视觉的残差项和 IMU 预积分残差项组合作为目标函数：

$$Target = \arg \min (E_{proj} + E_{INS}) \quad (3)$$

视觉重投影误差指的是相机观测得到的投影位置与按估计位姿变换得到的投影之间的像素误差，IMU 预积分误差主要由零偏引起的测量值与观测值的误差，将其目标函数简写为 $f(x)$ ，高斯牛顿法将其进行一阶泰勒展开：

$$f(x + \Delta x) \approx f(x) + J(x)\Delta x \quad (4)$$

其中， $J(x)$ 是 $f(x)$ 关于 x 的导数，即雅克比矩阵，为了求 Δx 使得 $\|f(x + \Delta x)\|^2$ 达到最小，构建一个线性的最小二乘方程：

$$\begin{aligned} \Delta x^* &= \arg \min_{\Delta x} \frac{1}{2} \|f(x) + J(x)\Delta x\|^2 \\ &= \frac{1}{2} (f(x) + J(x)\Delta x)^T (f(x) + J(x)\Delta x) \\ &= \frac{1}{2} (\|f(x)\|^2 + 2f(x)^T J(x)\Delta x + \Delta x^T J(x)^T J(x)\Delta x) \end{aligned} \quad (5)$$

将上式对 Δx 求导，令导数为零，则：

$$\begin{aligned} 2J(x)^T f(x) + 2J(x)^T J(x)\Delta x &= 0 \\ J(x)^T J(x)\Delta x &= -J(x)^T f(x) \end{aligned} \quad (6)$$

把左边系数定义为 H ，右边定义为 g ，则上式为：

$$H\Delta x = g \quad (7)$$

高斯牛顿法用 J 矩阵的转置乘以 J 矩阵作为牛顿法中二阶 H 矩阵的近似, 从而省略了计算 H 矩阵的过程。而实际计算中, J 矩阵的转置乘以 J 矩阵并非正是正的, 会出现奇异值, 即使 H 非奇异值, 如果求得的 Δx 非常大, 也会导致局部近似不够准确, 不能保证收敛, 还有可能让目标函数更大, 导致系统的定位精度降低, 因此本文采用 LM 算法取代高斯牛顿法。

LM 算法[13]引入了信赖半径, 在给定初始值 x_0 和信赖半径 u 后的每一次迭代时, 求解最小二乘方程:

$$\min_{\Delta x_k} \frac{1}{2} \|f(x_k) + J(x_k)\Delta x_k\|^2, \text{ s.t. } \|D\Delta x_k\|^2 \leq u \quad (8)$$

上式为带不等式约束条件的优化问题, 采用拉格朗日法将其转化为一个无条件约束的问题:

$$\min_{\Delta x_k} \frac{1}{2} \|f(x_k) + J(x_k)\Delta x_k\|^2 + \frac{1}{2} \lambda \|D\Delta x_k\|^2 \quad (9)$$

上式中左半边和右半边分别对 Δx 求梯度, 则有:

$$(H + \lambda D^T D)\Delta x = g \quad (10)$$

相比于高斯牛顿算法, LM 算法多了一个阻尼系数 λ , $D^T D$ 可简写为单位矩阵 I , 当 $\lambda > 0$ 时, $(H + \lambda I)$ 正定, 保证了梯度下降的方向; 当 λ 很大时, $\Delta x \approx \frac{1}{\lambda} g$ 也是最速下降方向, 当 λ 较小时, LM 算法接近于高斯牛顿算法。利用增益 p 来选取阻尼系数, p 的表达式为:

$$p = \frac{f(x + \Delta x) - f(x)}{J(x)\Delta x} \quad (11)$$

当 p 变大时, 则说明实际目标函数变化超过了预期变化, 此时应当减小 λ , 反之则提高 λ 。信赖半径 u 的变化与 p 的变化呈正相关, 由此控制更新步长, 加快目标函数的收敛速度, 使得系统的定位精度更高。

4. 实验与分析

本文的实验运行环境为 Ubuntu18.04, CPU 为 Intel (R) Core (TM) i5-11400H。用于实验验证的数据集为 EuRoc, EuRoc 数据集通过无人机在特定室内场景采集完成, 包含了多种场景下的轨迹信息和图像, 能有效验证算法的可靠性。

4.1. 特征点提取对比

为验证本文图像特征点提取算法的有效性, 选取 EuRoc 数据集中一组连续图像, 图像 1 和图像 2 是同一场景下的连续图像。将运用本文算法提取的特征点结果与运用传统 ORB 算法的结果进行对比。

从图 3、图 4 可以看出, 每种算法在连续帧中提取的特征点数量都一致, 两种算法特征点提取的准确性得到验证。本文算法由于引入四叉树分割法, 提高了特征点提取的质量, 在图像 1 和图像 2 中提取到的特征点数量为 134, 而 ORB 算法在两张图像中提取到的特征点数量为 500。本文算法的提取速度分别为 0.48 s、0.52 s, ORB 算法的提取速度为 0.47 s、0.53 s, 由于四叉树最大分裂次数为 4, 所以本文算法的提取速度没有受到影响, 二者相差不大。

图 3(b)、图 4(b)可以看出, 图像里红框中提取出的特征点出现了“扎堆”现象, 这会影响到后续对相机位姿的估计, 加大了计算量容易造成系统定位精度的降低, 本文采用分布均匀度 U [14]作为特征点分布均匀性的客观评价指标, 计算方法如式(12)所示。

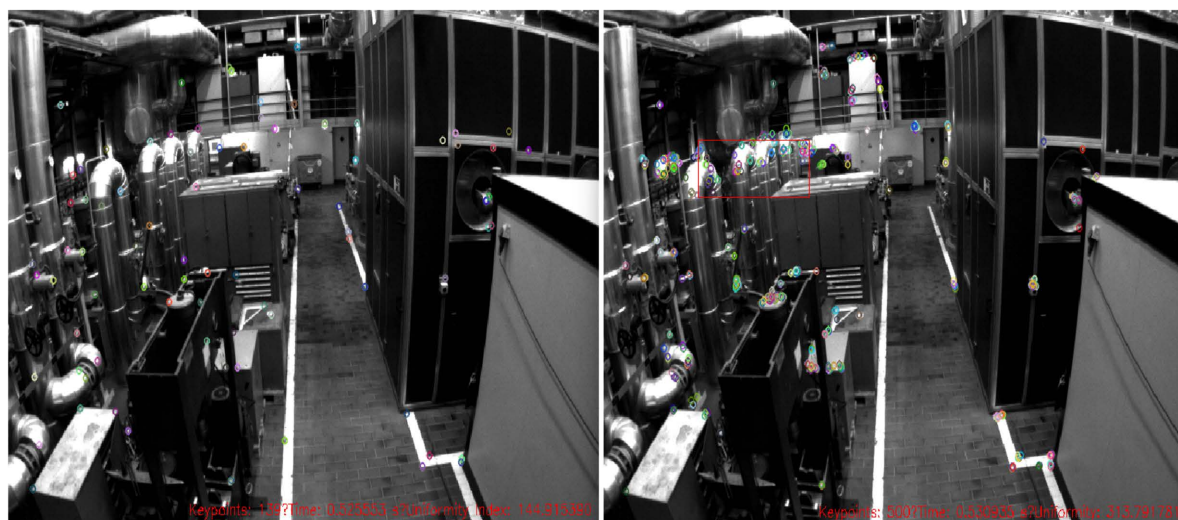
$$U = 101 \log(v) \quad (12)$$

其中, v 为区域统计分布向量的方差, 区域统计分布向量指各区域特征点数量值形成的向量。在特征点提取准确性得到保证的前提下, U 值越小说明提取特征点分布性越好, 从图 3、图 4 可以看出, 本文算法计算出的分布均匀度为 141、144, ORB 算法计算出的分布均匀度为 311、313, 证明本文算法在提取特征点的分布均匀度上提高了 54%。



(a) 本文算法

(b) ORB 算法

Figure 3. Image 1 results comparison diagram**图 3.** 图像 1 特征点提取结果对比图

(a) 本文算法

(b) ORB 算法

Figure 4. Image 2 results comparison diagram**图 4.** 图像 2 特征点提取结果对比图

4.2. 特征点匹配对比

为了直观感受本文特征点匹配算法的有效性, 将 4.1 中图像 1 和图像 2 采用 ORB 算法提取后分别用本文特征点匹配算法和汉明距离匹配算法进行对比, 两种算法的匹配结果如表 1 所示。

Table 1. The two algorithms matched the results

表 1. 两种算法匹配结果

| 算法 | 匹配时间 | 总匹配点数 | 正确匹配点数 | 准确率 |
|------|---------|-------|--------|-------|
| 汉明距离 | 0.042 s | 500 | 287 | 57.4% |
| 本文算法 | 0.034 s | 500 | 429 | 85.8% |

表 1 可以看出, 本文采用的 RANSAC 算法在结合汉明距离后, 加快了匹配时间, 用两倍汉明距离作为阈值, 减少了冗余匹配且 RANSAC 算法在不断迭代过程中, 提高了匹配的准确率, 本文算法匹配的准确率为 85.8%, 比只采用汉明距离的匹配算法提高了 28.4%。两种算法匹配的结果如图 5 所示。

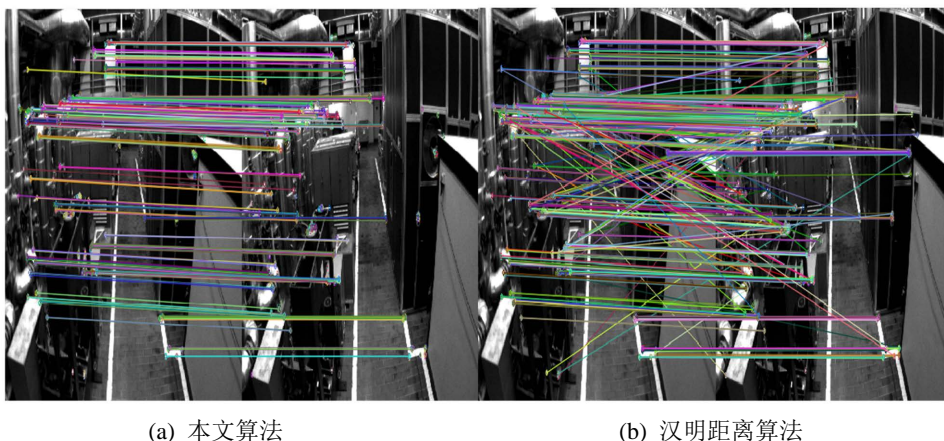


Figure 5. Comparison of the matching results of two algorithms

图 5. 两种算法特征点匹配结果对比图

4.3. 定位精度对比

本文采用 LM 算法取代了原 Vins-Mono 中的高斯牛顿算法, 对目标函数进行非线性优化, 并结合 4.1 和 4.2 的改进, 在 EuRoc 中不同场景下的数据集与原 Vins-Mono 算法进行对比, 场景选为实验室(MH)和工厂(VH)且分别选取两种场景内不同难度的数据集进行验证, 两种算法在不同场景下的不同难度上数据集的轨迹对比如图 6、图 7 所示。

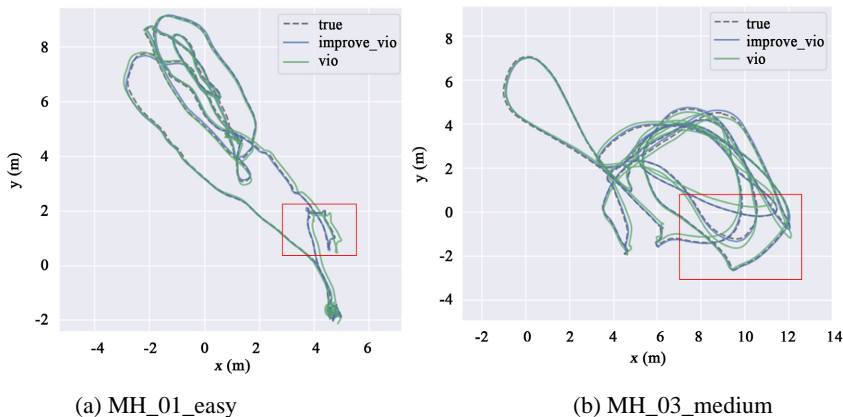


Figure 6. Comparison of two difficulty trajectories in the laboratory environment

图 6. 实验室环境下两种难度轨迹对比图

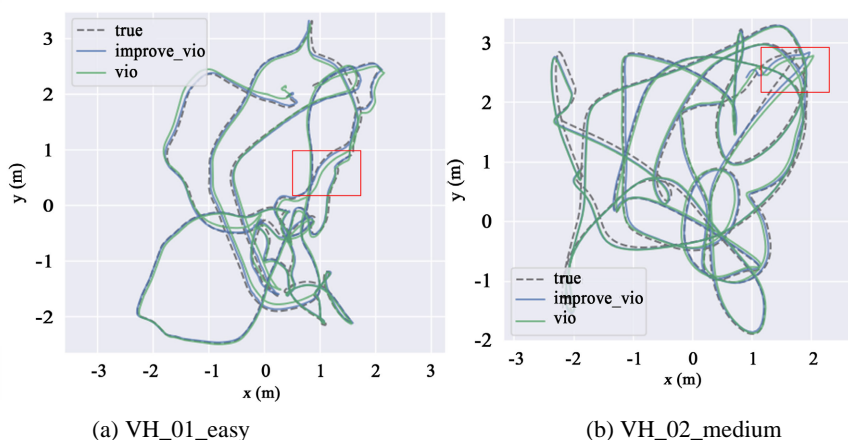


Figure 7. Comparison of two difficulty trajectories in a factory environment
图 7. 工厂环境下两种难度轨迹对比图

从图 6 和图 7 看出, 本文改进后的 Vins-Mono 算法在不同数据集上的轨迹更加接近真实值, 图中红框中可以看出, 在部分轨迹段传统 Vins-Mono 算法出现了明显的轨迹漂移, 而本文改进后算法的定位精度更高

为了进一步量化本文算法的定位精度, 选择均方根误差(RMSE) [15]作为定位精度的评价标准, 对不同场景下各轨迹精度进行量化对比, 具体参数如表 2 所示。

Table 2. Comparison of positioning accuracy of different data sets

表 2. 不同数据集定位精度对比

| 数据集 | VINS-Mono | 本文算法 |
|--------------|-----------|-------|
| MH_01_easy | 0.070 | 0.046 |
| MH_03_medium | 0.068 | 0.039 |
| VH_01_easy | 0.087 | 0.045 |
| VH_02_medium | 0.073 | 0.064 |

从表 2 可以看出, 本文算法在不同难度数据集上的定位精度分别提高了 34%、42%、48%、12%, 平均定位精度提高了 34%, 验证了本文改进算法在 INS/视觉组合定位中的有效性。

5. 结论

本文针对现有 Vins-Mono 算法在 INS/视觉组合定位中, 存在图像特征点提取不均匀、匹配准确率低、定位精度不高的问题, 提出了一种改进的 Vins-Mono 算法。在 EuRoc 公开数据集上进行验证, 采用基于四叉树法的 ORB 特征点提取算法, 使得特征点分布均匀化, 将 RANSAC 算法结合汉明距离进行特征点匹配, 提高了匹配的准确性, 将 Vins-Mono 算法后端优化中原高斯牛顿法替换为 LM 算法, 提高了定位精度。实验结果表明, 改进后的 Vins-Mono 算法在图像特征点提取均匀度上提高了 54%、特征点匹配准确率上提高了 28.4%、不同难度数据集上的平均定位精度提高了 34%, 验证了本文改进算法在 INS/视觉组合定位中的有效性。

基金项目

常州市 5G+工业互联网融合应用重点实验室(CM20223015)。

参考文献

- [1] 曾庆化, 罗怡雪, 孙克诚, 等. 视觉及其融合惯性的 SLAM 技术发展综述[J]. 南京航空航天大学学报, 2022, 54(6): 1007-1020.
- [2] Qin, T., Li, P. and Shen, S. (2018) VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator. *IEEE Transactions on Robotics*, **34**, 1004-1020. <https://doi.org/10.1109/TRO.2018.2853729>
- [3] 季莘翔, 王宇钢, 林一鸣. 改进 ORB 特征提取算法的 ORB-SLAM2 定位研究[J]. 组合机床与自动化加工技术, 2023(11): 60-64.
- [4] 余正强, 蒋林, 郭宇飞. 基于融合 BEBLID 和改进 ORB 算法的单目视觉里程计研究[J]. 组合机床与自动化加工技术, 2024(3): 6-11.
- [5] 骆大森, 吴英, 陈燕苹, 等. 基于相位相关法改进的 ORB 图像匹配算法[J]. 电子质量, 2023(8): 22-26.
- [6] 陈坚炜, 何元烈, 何铭臻, 等. 多目鱼眼视觉 SLAM 特征点误匹配的剔除优化[J]. 大连工业大学学报, 2024, 43(1): 61-72.
- [7] 赵磊, 李振伟, 杨晓利, 等. 一种基于改进闭环检测的视觉 SLAM 算法[J]. 计算机应用与软件, 2023, 40(4): 263-268+328.
- [8] 卢艳军, 周宠, 张晓东, 等. 基于改进图优化的惯性/卫星/视觉组合导航方法[J]. 沈阳航空航天大学学报, 2023, 40(6): 21-27.
- [9] 朱立成, 韩振浩, 赵博, 等. 基于 VINS-MONO 和改进 YOLOv4-Tiny 的果园自主寻筐方法[J]. 农业机械学报, 2023, 54(8): 97-109.
- [10] 于航, 杨维鑫, 李光雷, 等. 基于改进 ORB-SLAM2 的单目视觉算法[J]. 南方农机, 2023, 54(13): 16-19.
- [11] Rodrigo, C.A., Carlos, L.T.B. and Cristiano, C.D.L. (2023) Concept Drift Detection with Quadtree-Based Spatial Mapping of Streaming Data. *Information Sciences*, **625**, 578-592.
- [12] Arunima, S., Kushwaha, S.K.P., Nandy, S. and Padalia, H. (2022) An Approach for Tree Volume Estimation Using RANSAC and RHT Algorithms from TLS Dataset. *Applied Geomatics*, **14**, 785-794. <https://doi.org/10.1007/s12518-022-00471-x>
- [13] Zhao, J.H., Qian, X.X., Jiang, B. and Wang, B. (2024) Lithium Battery Model Parameter Identification Based on the GA-LM Algorithm. *International Journal of Green Energy*, **21**, 1147-1160. <https://doi.org/10.1080/15435075.2023.2242926>
- [14] 赵宇琦, 连晓峰, 谭励, 等. 一种均匀抗噪自适应的 ORB 特征提取算法[J]. 计算机仿真, 2022, 39(4): 321-327.
- [15] 孙国祥, 黄银锋, 汪小昆, 等. 基于 LIO-SAM 建图和激光视觉融合定位的温室自主行走系统[J]. 农业工程学报, 2024, 40(3): 227-239.