doi:10.6041/j.issn.1000-1298.2026.05.035

基于点线特征融合改进 IMU 初始化的双目视觉惯性 SLAM 方法

陈久朋[1,2]　杨　旺[1]　伞红军[1,3]　冯金祥[1]　伞　亮[1]

(1.昆明理工大学机电工程学院,昆明 650500;2.玉溪新天力农业装备制造有限公司,玉溪 652699
3.云南省先进装备智能制造技术重点实验室,昆明 650500)

摘要:针对基于特征点的 SLAM 系统在弱纹理场景下存在特征提取不足、易跟踪丢失等问题,为提高在复杂场景中的系统初始化精度和鲁棒性,本文在 ORB-SLAM3 框架的基础上加入了线特征,并对视觉惯性初始化进行了改进。首先在前端视觉里程计部分融入了 LSD 算法和 LBD 描述子进行线特征的提取和匹配,建立点、线特征重投影误差模型,并用基于非线性优化的 BA 方法来最小化重投影误差,同时引入自适应因子动态调整线特征权重。接着通过扩展双目 MNEC 约束构建陀螺仪偏差估计器,采用旋转-平移解耦优化策略,并引入残差评估机制确保视觉惯性初始化可靠性,同时将 IMU 残差、特征点重投影误差以及直线重投影误差共同作为非线性优化的约束条件对相机位姿进行估计。在 euroc 数据集和真实场景中进行实验,结果表明与改进前 ORB-SLAM3 算法相比,在数据集下改进算法定位精度提高 22.9%,真实环境中偏移量减少 1.4 m,从而验证了改进算法的可行性和有效性。

关键词:视觉 SLAM;点线特征融合;初始化优化;弱纹理场景

中图分类号:TP242;TH112　　**文献标识码**:A　　**文章编号**:1000-1298(2026)05-0373-14　　**OSID**:

Binocular Vision SLAM Algorithm for Improving IMU Initialization Based on Point-line Feature Fusion

CHEN Jiupeng[1,2]　YANG Wang[1]　SAN Hongjun[1,3]　FENG Jinxiang[1]　SAN Liang[1]

(1. *Faculty of Mechanical and Electrical Engineering*, *Kunming University of Science and Technology*, *Kunming* 650500, *China*
2. *Yuxi Xintianli Agricultural Equipment Manufacturing Co.*, *Ltd.*, *Yuxi* 652699, *China*
3. *Key Laboratory of Advanced Equipment Intelligent Manufacturing Technology of Yunnan Province*, *Kunming* 650500, *China*)

Abstract: Aiming at the problems of insufficient feature extraction and easy tracking loss of feature point-based SLAM systems in texture-less scenarios, as well as improving the initialization accuracy and robustness of the system in challenging scenarios, the line features on the basis of the ORB-SLAM3 framework were incorporated and the visual inertia initialization was improved. Firstly, the LSD algorithm and LBD descriptor were incorporated into the front-end visual odometry part for line feature extraction and matching, the point and line feature reprojection error models were established, and the BA method based on nonlinear optimization was used to minimize the reprojection error, while the adaptive factor was introduced to dynamically adjust the weights of the line features. Then the gyroscope bias estimator was constructed by extending the binocular MNEC constraints, adopting the rotation-translation decoupling optimization strategy and introducing the residual evaluation mechanism to ensure the reliability of visual inertia initialization, while the IMU residuals, the feature point reprojection errors, and the line reprojection errors were jointly used as the constraints of nonlinear optimization for the estimation of the camera position. Experiments were conducted in the euroc dataset and real scenes, and the results showed that compared with the pre-improved ORB-SLAM3 algorithm, the improved algorithm improved the localization accuracy by 22.9% under the dataset and reduced the offsets by 1.4 m in the real environment, thus verifying the feasibility and effectiveness of the improved algorithm.

Key words: visual SLAM; point-line feature fusion; initialization optimization; texture-less scenarios

收稿日期:2025-07-19　修回日期:2025-09-16

基金项目:国家自然科学基金项目(52565057)、云南省基础研究计划项目(202301AU070059)和云南彩云博士后项目

作者简介:陈久朋(1993—),男,讲师,博士,主要从事机器人技术和农业机械研究,E-mail:18314490225@163.com

通信作者:伞红军(1976—),男,副教授,博士,主要从事机器人技术和农业机械研究,E-mail:sanhjun@163.com

0 引言

同时定位与建图(Simultaneous localization and mapping, SLAM)技术作为机器人自主导航和增强现实等领域的核心技术,近年来取得了显著进展。根据获取环境信息的传感器不同,SLAM 技术可分为激光 SLAM 和视觉 SLAM[1]。视觉惯性 SLAM 系统通过融合相机和惯性测量单元(IMU)的互补优势,有效解决了纯视觉 SLAM 中的尺度模糊性和运动退化问题[2]。然而,传统的基于点特征的 SLAM 系统在弱纹理环境、光照变化剧烈和动态场景等复杂场景下仍存在特征提取不足、跟踪丢失等问题[3-4]。

近年来,ORB - SLAM3 等系统在视觉惯性 SLAM 领域取得了突破性进展,支持单目、双目和 RGB - D 多种配置。OpenVINS[5]和 LVI - SAM[6]等框架分别通过扩展卡尔曼滤波和多传感器融合进一步提升了系统性能。在点线融合方面,PL - VINS 首次实现了基于优化的实时单目视觉惯性 SLAM 点线融合,处理速度比原始算法提升 3 倍,定位精度提高 12% ~16%[7]。PLPF - VSLAM 提出了自适应特征选择策略[8],EPL - VINS 引入 LK - RG 线段跟踪方法并采用 2 自由度优化的空间直线表示,降低了计算开销[9]。UPLP - SLAM 实现了点线面特征的统一优化框架[10],PLMOT - SLAM 将点线融合与动态目标跟踪相结合[11]。

IMU 初始化是视觉惯性 SLAM 系统的关键环节,需要准确估计尺度、重力方向、速度和传感器偏差等参数。传统方法通常需要复杂的参数调优,在动态环境中容易失效[12]。PLE - SLAM 提出了扩展的点线特征初始化方法,通过解析求解加速度计偏差和重力方向,紧耦合视觉惯性观测优化陀螺仪偏差,并将动态特征消除、闭环检测与深度神经网络相结合[13]。CAMPOS 等[14]提出的仅惯性优化方法实现了 2 s 内的一致性初始化,尺度误差约为 5%。Stereo - NEC 将法向极线约束扩展至双目系统,直接求解陀螺仪偏差,提升了初始旋转估计的鲁棒性[15]。在线标定框架通过可观测性分析识别影响标定参数的退化运动[16],Kalibr 工具箱支持全面的视觉惯性标定[17]。LAMP2.0 在极端环境中实现了鲁棒的多机器人 SLAM[18],DROID - SLAM 通过深度学习提升了系统鲁棒性[19]。国内学者在 SLAM 领域也取得了重要进展。陈劭等[20]提出了 RGB - D 视觉 SLAM 算法,采用改进的 RE - RANSAC 算法和 GICP 优化,提高了定位精度和实时性。董蕊芳等[21]在 ORB - SLAM2 基础上进行了多层次优化,构建多层地图和导航系统,提高了室内自主导航能力。齐咏生等[22]提出了基于点线特征融合的延迟边缘化视觉惯性 SLAM 方法,采用延迟边缘化策略处理非线性优化问题,提高了复杂环境下的定位精度。现有点线融合方法主要针对单目或 RGB - D 系统设计,对双目立体视觉在线特征提取和匹配方面的固有优势挖掘不足,且普遍存在计算开销大的问题,严重限制了在资源受限环境中的实时性能;IMU 初始化技术对动态环境的适应性不足,需要复杂的参数调优过程,限制了系统的部署灵活性,同时现有研究主要关注单目系统的初始化,双目初始化方法研究相对薄弱,未能充分利用双目视觉系统的可观测性优势。

基于上述分析,本文提出一种基于点线特征融合改进 IMU 初始化的双目视觉惯性 SLAM 方法。该方法通过融合 LSD 和 LBD 算法提取匹配线特征,构建点线特征重投影误差模型,采用带自适应因子的动态线权重 BA 优化。在初始化阶段,扩展双目 MNEC 约束构建陀螺仪偏差估计器,通过旋转-平移解耦优化策略及残差评估机制提升鲁棒性。联合优化 IMU 残差、特征点及直线重投影误差,显著提升弱纹理等复杂场景下的初始化精度和系统鲁棒性。

1 算法框架

本文基于 ORB - SLAM3 开源框架扩展,构建了一个双目实时点线融合视觉 SLAM 系统。其核心为通过线特征辅助机制和改进的 IMU 初始化,显著提升了系统在复杂场景下的鲁棒性与效率。如图 1 所示,系统架构延续了三线程并行架构设计:①跟踪线程中,融入了 LSD 算法和 LBD 描述子,进行线特征的提取和匹配,并引入自适应因子来调整线特征的误差权重。在跟踪过程中,初始位姿估计通过最小化当前帧与前一帧之间检测到的特征点和线匹配的重投影误差来计算;随后,通过在当前帧和局部地图之间找到点和线特征匹配,并再次执行仅运动光束平差法,对该位姿进行优化。②局部建图线程中负责局部地图的构建和优化。在初始化阶段,对视觉惯性初始化进行了改进,通过旋转-平移解耦优化策略和残差评估机制,提高了初始化精度和鲁棒性。在局部 BA 优化阶段,则联合优化关键帧位姿、点特征及线特征的空间位置,进一步提升了地图精度。③回环检测线程在原始架构上扩展了线特征处理模块,通过相似性变换与位姿图优化校正线特征,从而消除长期累积的漂移,确保全局一致性。系统的各个线程通过一个共视图连接,

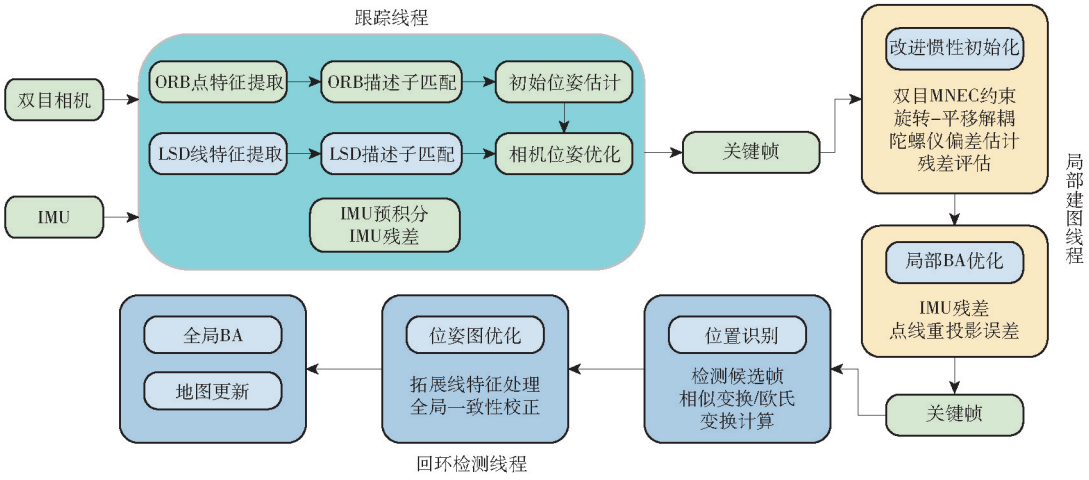


图 1　系统框架图

Fig. 1　System framework diagram

用于共享相同点和线的关键帧，以及一个最小生成树，将所有关键帧连接在一起，共同协同工作以实现高精度的实时定位。

2　点线特征融合方法

2.1　线特征提取与描述

在视觉里程计阶段引入线特征，以增强系统的稳定性和鲁棒性。在线特征提取方面，采用经典的 LSD(Line segment detector)线段检测算法[23]。该算法采用了一种基于梯度的启发式方法，无需进行复杂的参数估计，直接在图像空间中检测直线段。而且 LSD 算法计算速度快和提取的线段具有方向性，因此在图像匹配时可以减少误匹配，同时良好的实时性和准确率使得 LSD 算法可以更好地适用于 SLAM 算法。同时，采用 LBD 描述子来对线特征进行描述[24]。LBD 描述符具备尺度不变性，能在不同尺度空间有效提取线段特征，并且计算简单高效，通过对像素梯度统计来生成描述子，可快速处理大量图像数据，同时局部特征描述能力强，能精准刻画线段局部外观并结合几何约束构建关系对，以区分不同线特征，以及抗噪能力较好，对像素梯度的统计处理使其能在一定程度上可抑制噪声影响，不仅便于存储和传输，还能利用高效算法快速进行匹配计算。

2.2　特征融合的 BA 优化

光束平差法(Bundle adjustment，BA)是一种用于优化相机位姿和三维点坐标的非线性优化方法，是 SLAM 中非常重要的一个优化方法。其核心思想是通过比较实际观测到的特征点和线与模型预测的投影位置之间的差异，建立误差函数并最小化，然后通过非线性优化不断调整相机位姿和地图点位置，使重投影误差最小化，从而抑制噪声累积、提高轨迹和地图的全局一致性。

2.2.1　点特征重投影误差模型

考虑三维空间特征点的投影过程，假设某空间点在世界坐标系下的坐标为 $\boldsymbol{P}_j=(X_j,Y_j,Z_j)$，其投影到第 i 帧中的像素坐标为 $\boldsymbol{z}_j=(\boldsymbol{u}_j,\boldsymbol{v}_j)$，相机内参为 $\boldsymbol{K}$，位姿为旋转矩阵 $\boldsymbol{R}$ 和平移向量 $\boldsymbol{t}$，它的李群表示为 $\boldsymbol{T}$，则像素位置与空间点位置的关系为

$$\boldsymbol{s}_j\begin{bmatrix}\boldsymbol{u}_j\\ \boldsymbol{v}_j\\ 1\end{bmatrix}=n\left(\boldsymbol{KT}\begin{bmatrix}X_j\\ Y_j\\ Z_j\\ 1\end{bmatrix}\right) \tag{1}$$

式中　$n(\cdot)$——齐次坐标到非齐次坐标的转换函数

$\boldsymbol{s}_j$——尺度因子

将式(1)写成相应的矩阵形式为

$$\boldsymbol{s}_j\boldsymbol{u}_j=n(\boldsymbol{K}(\boldsymbol{RP}_j+\boldsymbol{t}))=n(\boldsymbol{KTP}_j) \tag{2}$$

由于相机位姿未知以及存在观测点的噪声，因此该等式存在一个误差，如图 2 所示，其中 p_1、p_2 是空间点 P 投影到参考帧和当前帧中的观测点，而 $\hat{p}_2$ 是 P 的重投影点。而该等式的误差项是将空间点的观测位置与重投影位置作差，称为重投影误差 $\boldsymbol{e}_{i,j}$，计算公式为

$$\boldsymbol{e}_{i,j}=\boldsymbol{u}_j-\frac{1}{\boldsymbol{s}_j}\boldsymbol{KTP}_j \tag{3}$$

式中　$\boldsymbol{e}_{i,j}$——第 i 帧图像中观测到的第 j 个地图点的重投影误差

2.2.2　线特征重投影误差模型

在 ORB－SLAM3 的框架中，传统的点特征重投影误差计算方法并不适用于线段特征。这主要由于线段端点存在不稳定性即帧间运动过程中容易发生遮挡或丢失。

设 $\overline{E}_{iw}\in SE(3)$ 表示第 i 帧的相机位姿，$P_{w,k}\in\mathbf{R}^3$，$Q_{w,k}\in\mathbf{R}^3$ 是第 i 帧观测到的第 k 条线段的三维

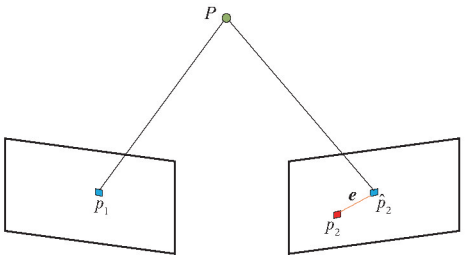


图 2 点特征重投影误差示意图

Fig. 2 Schematic of reprojection error of point features

端点，$\boldsymbol{p}_{i,k} \in \mathbf{R}^2$、$\boldsymbol{q}_{i,k} \in \mathbf{R}^2$ 是图像平面中对应于第 i 帧中第 k 条线段的 2 个观测端点，其中 w 表示世界坐标系。

欧氏距离：该误差函数使用三维线段在图像平面中的观测线端点与其重投影线段在图像平面上的欧氏距离作为评判标准。观测线归一化的线性参数 $\boldsymbol{l}_{i,k}$ 为

$$\boldsymbol{l}_{i,k} = \frac{{}^{h}\boldsymbol{p}_{i,k} \times {}^{h}\boldsymbol{q}_{i,k}}{\| {}^{h}\boldsymbol{p}_{i,k} \times {}^{h}\boldsymbol{q}_{i,k} \|} \tag{4}$$

式中${}^{h}\boldsymbol{p}_{i,k} \in \mathbf{R}^3$、${}^{h}\boldsymbol{q}_{i,k} \in \mathbf{R}^3$ 表示 $\boldsymbol{p}_{i,k}$、$\boldsymbol{q}_{i,k}$的齐次坐标。因此误差函数可以定义为

$$\boldsymbol{e}_{i,k} = \begin{bmatrix} \boldsymbol{l}_{i,k} \cdot {}^{h}\boldsymbol{\pi}(E_{iw}, P_{wk}) \\ \boldsymbol{l}_{i,k} \cdot {}^{h}\boldsymbol{\pi}(E_{iw}, Q_{wk}) \end{bmatrix} \tag{5}$$

其中${}^{h}\boldsymbol{\pi}(E_{iw}, P_{wk})$、${}^{h}\boldsymbol{\pi}(E_{iw}, Q_{wk})$ 表示的是 $\boldsymbol{\pi}(E_{iw}, P_{wk})$、$\boldsymbol{\pi}(E_{iw}, Q_{wk})$ 的齐次坐标，而 $\boldsymbol{\pi}(E_{iw}, P_{wk})$、$\boldsymbol{\pi}(E_{iw}, Q_{wk})$表示重投影线段在图像平面上的二维端点，可表示为

$$\boldsymbol{\pi}(E_{iw}, P_{wk}) = \begin{bmatrix} f_x \dfrac{m_x}{m_z} + c_x \\ f_y \dfrac{m_y}{m_z} + c_y \end{bmatrix} \tag{6}$$

$$\boldsymbol{m} = \begin{bmatrix} m_x & m_y & m_z \end{bmatrix}^{\mathrm{T}} = \boldsymbol{R}_{iw} P_{w,k} + \boldsymbol{t}_{iw} \tag{7}$$

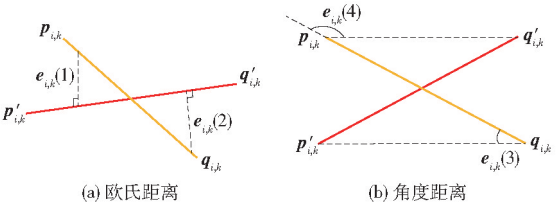
式中$\boldsymbol{R}_{iw} \in SO(3)$、$\boldsymbol{t}_{iw} \in \mathbf{R}^3$ 分别为与 E_{iw}相关的旋转和平移矩阵，(f_x, f_y) 为焦距，(c_x, c_y) 代表像素与成像之间原点的平移量，两者可以从相机的标定中得到。同理可以计算出 $\boldsymbol{\pi}(E_{iw}, Q_{wk})$。在光束平差过程中，$\boldsymbol{\Sigma}_{e_{i,k}}$为协方差单位矩阵。

角度距离：该误差函数使用观测线与重投影线之间的角度距离作为评判标准。假设定义相同，如果 $\boldsymbol{p}'_{i,k} \in \mathbf{R}^2$、$\boldsymbol{q}'_{i,k} \in \mathbf{R}^2$ 代表重投影线段在图像平面上的二维端点与 $\boldsymbol{\pi}(E_{iw}, P_{wk})$，$\boldsymbol{\pi}(E_{iw}, Q_{wk})$类似，那么误差函数为

$$\boldsymbol{e}_{i,k} = \begin{bmatrix} \dfrac{\boldsymbol{l}_{q_{i,k}p'_{i,k}} \cdot \boldsymbol{l}_{q_{i,k}p_{i,k}}}{\| \boldsymbol{l}_{q_{i,k}p'_{i,k}} \| \, \| \boldsymbol{l}_{q_{i,k}p_{i,k}} \|} - 1 \\ \dfrac{\boldsymbol{l}_{p_{i,k}q'_{i,k}} \cdot \boldsymbol{l}_{q_{i,k}p_{i,k}}}{\| \boldsymbol{l}_{p_{i,k}q'_{i,k}} \| \, \| \boldsymbol{l}_{q_{i,k}p_{i,k}} \|} + 1 \end{bmatrix} \tag{8}$$

该误差函数表示向量 $\boldsymbol{l}_{q_{i,k}p'_{i,k}}$，$\boldsymbol{l}_{q_{i,k}p_{i,k}}$和 $\boldsymbol{l}_{p_{i,k}q'_{i,k}}$，$\boldsymbol{l}_{q_{i,k}p_{i,k}}$之间的余弦夹角。

本文融合欧氏 + 角度距离构造出来新的误差函数，如图 3 所示，新的误差函数包含了欧氏和角度距离的误差函数，从而提高系统的精度。在这里，采用了一种稍微不同的策略来丢弃离群值。主要问题是融合的 2 个函数表示不同的距离，因为它们的值不在一个数量级上。为了解决这个问题，通过使用协方差矩阵的逆矩阵（它实际上起到了加权矩阵的作用），根据阈值对误差函数的每一个分量进行归一化处理，使其在成为离群值时达到前 2 种情况设定的值 1。这样不仅可以通过使用像素距离和角度距离的信息更有效地剔除离群值，而且 2 个单独的误差函数也会相应地加权，从而使它们在光束平差的过程中几乎同等地发挥作用。



(a) 欧氏距离 (b) 角度距离

图 3 线特征重投影误差示意图

Fig. 3 Schematics of line feature reprojection error

2.3 自适应权重点线融合优化

为了在光束平差中融合点和线特征，引入了一个新的自适应因子，用于权衡线特征在这个过程中的参与程度。这是因为与线特征相关的误差函数和用于点特征的误差函数相比存在更大偏差。特别是通过最小化重投影的三维线段的端点与其在图像平面上对应的无限直线之间的欧氏距离，一般会错误地认为重投影的线段端点相对应的点是其对应无限直线上距离最近的点。事实上，由于没有一个线段检测器能在不同角度观察场景时重现具有稳定端点的相同线段，因此无法完全找回正确的对应点。此外，在检测到的点匹配较少的稀疏纹理场景中，ORB - SLAM3 的精度会明显降低，这与系统在能确定足够多点对应关系的环境中其卓越的效果形成对比。因此，自适应因子的理念是，随着内点数量的增加，线的参与度降低；当内点数量较少时，线与点的权重相同。对于所有类型的光束平差优化，本文使用了在 g2o[25] 中实现的列文伯格-马夸尔特方法。以下是系统中对纯运动光束平差法和局部光束平差法的改进。

2.3.1 运动光束平差法

从点特征重投影误差模型中可知，$\boldsymbol{e}_{i,j}$表示第 i 帧中第 j 个地图点的投影误差，$\boldsymbol{\omega}$ 是包含待优化变量的向量，该向量第 i 帧的位姿为$\{\boldsymbol{R}_{iw}, \boldsymbol{t}_{iw}\}$，那么对于点线特征融合算法，相机的位姿可估计为

$$\begin{cases}\boldsymbol{\omega}^{*}=\underset{\boldsymbol{\omega}}{\operatorname{argmin}}(\boldsymbol{A}+\boldsymbol{B})\\ F(\alpha(P))=W^{-\alpha(P)\operatorname{div}T}\end{cases}\tag{9}$$

其中

$$\boldsymbol{A}=\sum_{j\in\boldsymbol{P}}\rho\left(\boldsymbol{e}_{i,j}^{T}\sum_{\boldsymbol{e}_{i,j}}^{-1}\boldsymbol{e}_{i,j}\right)\tag{10}$$

$$\boldsymbol{B}=\sum_{k\in L}\rho\left(F(\alpha(P))\boldsymbol{e}_{i,k}^{T}\sum_{i,k}^{-1}\boldsymbol{e}_{i,k}\right)\tag{11}$$

式中　$\boldsymbol{\omega}$——包含待优化变量的向量

$\boldsymbol{\omega}^{*}$——通过最小化得到的最优位姿

P——点匹配集合组

L——线匹配集合组

$\rho(\cdot)$——鲁棒的 Huber 代价函数

$\alpha(P)$——集合 P 的元素总数

div——取除法整数部分的运算符

$F(\alpha(P))$——自适应因子

W、T——权重和阈值，其值的选取很大程度上影响系统的定位精度和稳定性

2.3.2　局部光束平差法

在局部光束平差中，向量 $\boldsymbol{\omega}$ 包含一组共视关键帧 k 的位姿以及在这些关键帧中看到的所有特征的位置。如果该组共视关键帧以及与它们观测到相同点线特征的关键帧集合 K_L 包含这组关键帧，同时也包含那些与它们存在公共点和公共线的关键帧，分别将这些关键帧记为 P_L 和 L_L，并且 $E_{i,j}$、$E_{i,k}$ 是公式(9)中给出的误差，那么点线方法的优化问题为

$$\boldsymbol{\omega}^{*}=\underset{\omega}{\operatorname{argmin}}\sum_{i\in K_L}\left(\sum_{j\in P_L}\rho(E_{i,j})+\sum_{k\in L_L}\rho(E_{i,k})\right)\tag{12}$$

3　视觉惯性初始化

在视觉惯性 SLAM 系统中视觉和 IMU 非联合初始化的精度很大程度上依赖于视觉信息的精度。即使像 ORB－SLAM3 这样最先进的立体 VI－SLAM 系统中，也只有在连续帧之间具有足够基线和轻度旋转的场景中才能得到精确的相机轨迹估计。然而在诸如纯旋转或剧烈旋转等复杂的情况下，ORB－SLAM3 的初始化可能会导致精度和鲁棒性下降。为了克服这些限制，提高在复杂场景中的初始化精度和鲁棒性，通过考虑陀螺仪偏差估计对旋转精度的重大影响，以及惯性参数与视觉信息之间的联系对 ORB－SLAM3 系统视觉惯性初始化进行改进。主要是将单目法向极线约束(MNEC)扩展到了双目中，并用 MNEC 方法将法向量与相对旋转估计和视觉信息相结合，同时使用考虑陀螺仪偏差影响的一阶泰勒级数来近似旋转，在去除陀螺仪偏差后获得精确的旋转估计后，利用它通过 3－DoF 光束平差法来改进平移估计，从而提高初始化精度和鲁棒性。

3.1　单目 MNEC 原理

单目法向极线约束(MNEC)反映了单个移动相机 2 个位姿下的方向向量与相应极平面(Epipolar plane)法线的几何关系。如图 4 所示，当考虑 2 个连续时刻的特征对应关系时，单位方向向量对 $\boldsymbol{f}_i$ 和 $\boldsymbol{f}_i'$ 分别从时刻 k 和 $k+1$ 的相机中心 O_k 和 O_{k+1} 出发，指向三维点 P_i。每对方向向量与相关的法向量 $\boldsymbol{n}_i=\boldsymbol{f}_i\times\boldsymbol{R}_{c_{k+1}}^{c_k}\boldsymbol{f}_i'$ 定义一个极线平面，其中 $\boldsymbol{R}_{c_{k+1}}^{c_k}$ 表示从坐标系 c_k 到坐标系 c_{k+1} 的相对旋转。这些极线平面的交线是两帧之间的基线 $t_{c_{k+1}}^{c_k}$，同时也表示两帧之间的相对平移，在图中用虚线表示。同时法向向量共同形成相应的极线法平面，该平面垂直于基线 $t_{c_{k+1}}^{c_k}$。假设观察到 n 个三维点，将 n 个极线平面的法向量堆叠成一个矩阵 $\boldsymbol{N}=[\boldsymbol{n}_1\quad\boldsymbol{n}_2\quad\cdots\quad\boldsymbol{n}_n]$。而要证明这些法向量共面，即等价于矩阵 $\boldsymbol{M}=\boldsymbol{N}\boldsymbol{N}^{\mathrm{T}}$ 的最小特征值为零。

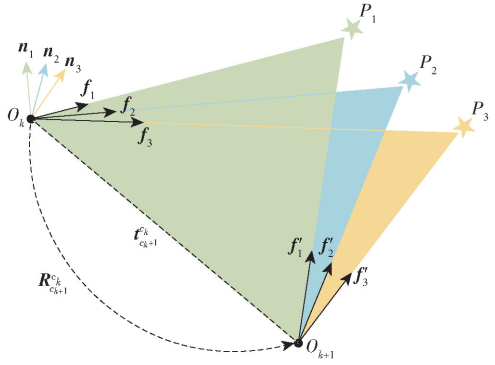


图 4　单目法极线约束示意图

Fig. 4　Schematic of monocular epipolar constraint

而根据定义，MNEC 的残差 e_i 为

$$e_i=|\boldsymbol{n}_i^{T}t_{c_{k+1}}^{c_k}|\tag{13}$$

而 MNEC 在 SLAM 中的两种应用为：

(1)旋转：KNEIP 等[26]利用了提出的极线约束来计算 2 个校准图像之间的相对位姿。通过强制极线平面法向量共面，直接约束了 2 个相机视图之间相对旋转的 3 个自由度，且与平移无关。即主要目标是确定能使最小特征值 $\lambda_{min}(\boldsymbol{M}_{k,k+1})$ 最小化的相对旋转 $\boldsymbol{R}_{c_{k+1}}^{c_k}$，从而达到优化相机位姿的效果，相关公式为

$$\begin{cases}\boldsymbol{R}_{c_{k+1}}^{c_k*}=\underset{\boldsymbol{R}_{c_{k+1}}^{c_k}}{\operatorname{argmin}}\lambda_{\min}(\boldsymbol{M}_{k,k+1})\\ \boldsymbol{M}_{k,k+1}=\sum_{i=1}^{n}(\boldsymbol{f}_i\times\boldsymbol{R}_{c_{k+1}}^{c_k}\boldsymbol{f}_i')(\boldsymbol{f}_i\times\boldsymbol{R}_{c_{k+1}}^{c_k}\boldsymbol{f}_i')^{\mathrm{T}}\end{cases}\tag{14}$$

其中 c_k 是拍摄第 k 幅图像时的左目相机坐标系，矩阵 $\boldsymbol{M}_{k,k+1}$ 是实对称半正定矩阵，此外由于法向量的共面性约束，其秩为 2。

(2)陀螺仪偏差：受 KNEIP 等[26]工作的启发，

HE 等[27]采用 MNEC 直接优化陀螺仪偏差,将视觉信息和相机与 IMU 外在校准矩阵[$\boldsymbol{R}_b^c \mid \boldsymbol{t}_b^c$]结合起来,这将目标函数从式(14)修改为

$$\begin{cases} b_g^* = \underset{b_g}{\operatorname{argmin}} \lambda_{\min}(\boldsymbol{M}_{k,k+1}) \\ \boldsymbol{n}_i = \boldsymbol{f}_i \times \boldsymbol{R}_b^c \hat{\gamma}_{b_{k+1}}^{b_k} \boldsymbol{R}_\alpha^b \boldsymbol{f}' \\ \boldsymbol{M}_{k,k+1} = \sum_{i=1}^{n} (\boldsymbol{f}_i \times \boldsymbol{R}_b^c \hat{\gamma}_{b_{k+1}}^{b_k} \boldsymbol{R}_\alpha^b \boldsymbol{f}')(\boldsymbol{f}_i \times \boldsymbol{R}_b^c \hat{\gamma}_{b_{k+1}}^{b_k} \boldsymbol{R}_\alpha^b \boldsymbol{f}')^{\mathrm{T}} \\ \hat{\gamma}_{b_{k+1}}^{b_k} = \gamma_{b_{k+1}}^{b_k} \exp(J_{b_g}^{\gamma} b_g) \end{cases} \tag{15}$$

式中 b_g——IMU 坐标系中的陀螺仪偏差

b_k——拍摄第 k 幅图像时的 IMU 坐标系

$\gamma_{b_{k+1}}^{b_k}$——从 b_k 到 b_{k+1}的旋转预积分

$J_{b_g}^{\gamma}$——预积分因陀螺仪偏差估算的微小差异而发生的变化量

考虑到陀螺仪偏差在第 k 个关键帧和第 $k+1$ 个关键帧之间的持续时间内的影响,使用 $\gamma_{b_{k+1}}^{b_k}$的一阶泰勒近似值 $\hat{\gamma}_{b_{k+1}}^{b_k}$估算。

3.2 扩展至双目 MNEC

由单目法向极线约束可知,通过将视觉信息和相机与 IMU 外在校准矩阵[$\boldsymbol{R}_b^c \mid \boldsymbol{t}_b^c$]结合起来,并通过求解最小特征值 $\lambda_{\min}(\boldsymbol{M}_{k,k+1})$的最小值,可以优化相机位姿和陀螺仪偏差。将单目法向极线约束扩展到了双目法向极线约束,如图 5 所示,从而能够结合更多的视觉信息来提高陀螺仪偏差估计的鲁棒性和准确性。

为了在利用立体视觉信息的同时估计初始陀螺仪偏差,通过最小化公式(11)中的最小特征值 $\lambda_{\min}$来初始化陀螺仪偏差

$$\begin{cases} b_g^* = \underset{b_g}{\operatorname{argmin}} \lambda_{\min} \\ \lambda_{\min} = \sum_{(k,k+1)\in\varepsilon} (\lambda_{\min}({}^{L}\boldsymbol{M}_{k,k+1}) + \lambda_{\min}({}^{R}\boldsymbol{M}_{k,k+1})) \\ {}^{L}\boldsymbol{M}_{k,k+1} = \sum_{i=1}^{n_L} \boldsymbol{C}\boldsymbol{C}^{\mathrm{T}} \\ {}^{R}\boldsymbol{M}_{k,k+1} = \sum_{i=1}^{n_R} \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}} \\ \hat{\gamma}_{b_{k+1}}^{b_k} = \gamma_{b_{k+1}}^{b_k} \exp(J_{b_g}^{\gamma} b_g) \end{cases} \tag{16}$$

其中
$$\boldsymbol{C} = \boldsymbol{f}_{L_i} \times \boldsymbol{R}_b^{c_L} \hat{\gamma}_{b_{k+1}}^{b_k} \boldsymbol{R}_{c_L}^b \boldsymbol{f}'_{L_i} \tag{17}$$

$$\boldsymbol{E} = \boldsymbol{f}_{R_i} \times \boldsymbol{R}_b^{c_R} \hat{\gamma}_{b_{k+1}}^{b_k} \boldsymbol{R}_{c_R}^b \boldsymbol{f}'_{R_i} \tag{18}$$

式中 c_L、c_R——左目相机和右目相机

$\boldsymbol{f}_{L_i}$ $\boldsymbol{f}'_{L_i}$——左目相机中的一对方向向量

$\boldsymbol{f}_{R_i}$ $\boldsymbol{f}'_{R_i}$——右目相机中的一对方向向量

n_L、n_R——左目相机和右目相机观测到的特征点数目

$\boldsymbol{R}_b^{c_L}$、$\boldsymbol{R}_b^{c_R}$——左目相机与 IMU 外参校准旋转矩阵和右目相机与 IMU 外参校准旋转矩阵

${}^{L}\boldsymbol{M}_{k,k+1}$、${}^{R}\boldsymbol{M}_{k,k+1}$——左目相机和右目相机的法向量构成的实对称半正定矩阵

ε——2 个连续时间的一组关键帧对

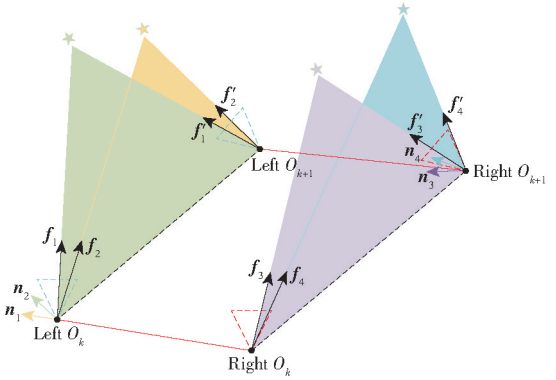


图 5 双目法极线约束示意图

Fig. 5 Schematic of binocular epipolar constraint

3.3 旋转-平移解耦优化

这一步骤是将旋转和平移分开优化,这主要是因为通过实验观察发现在去除陀螺仪偏差后 IMU 旋转积分得到的旋转比从纯视觉信息中得到的旋转更精确。因此获得最优陀螺仪偏差后,通过 IMU 旋转积分更新滑动窗口内的每个旋转 $\boldsymbol{R}_{b_k}^w$,并使用 3-DoF BA 优化每个平移$\boldsymbol{t}_w^{c_k}$,计算公式为

$$\begin{cases} \boldsymbol{R}_{c_k}^w = \boldsymbol{R}_{b_k}^w \boldsymbol{R}_{c_L}^b \\ \boldsymbol{t}_w^{c_k *} = \underset{\boldsymbol{t}_w^{c_k}}{\operatorname{argmin}} \sum_{i\in M} \boldsymbol{G} \end{cases} \tag{19}$$

其中
$$\boldsymbol{G} = \rho(\| \boldsymbol{x}^i - \pi(\boldsymbol{R}_w^{c_k} \boldsymbol{X}^i + \boldsymbol{t}_w^{c_k}) \|_{\boldsymbol{\Sigma}}^2) \tag{20}$$

式中 $\boldsymbol{X}^i$——通过三角测量得到的世界坐标系中的三维空间特征点

$\boldsymbol{x}^i$——$\boldsymbol{X}^i$ 对应的二维特征

$\pi(\cdot)$——重投影函数

$\boldsymbol{\Sigma}$——对应于与金字塔中关键点的尺度级别相关的协方差

$\boldsymbol{R}_{c_L}^b$——与 $\boldsymbol{R}_b^{c_L}$ 互逆的旋转矩阵

$\boldsymbol{R}_{c_k}^w$——相机在第 k 帧时的姿态方向

4 视觉惯性紧耦合优化

视觉惯性紧耦合(Visual-inertial tightly coupled)是一种通过融合 IMU(惯性测量单元)数据与立体视觉信息来联合构建运动方程和观测方程,从而实现较高精度的状态估计。在传统的 ORB-SLAM3 框架中,视觉惯性紧耦合仅将 IMU 残差和特征点的重投影误差作为优化约束,虽然能够实现较好的定

位效果，但在缺乏纹理或特征点稀疏的场景中，系统的鲁棒性和精度可能受到限制。因此为了进一步提升性能，在确保视觉测量坐标系与惯性测试坐标系之间已完成精确标定的前提下，在原有框架的基础上引入了直线特征（LSD 检测器提取的线段），将 IMU 残差、特征点重投影误差以及直线重投影误差共同作为非线性优化的约束条件。精确的标定，包括相机内参以及相机与 IMU 之间的外参，是实现两者数据时空同步与深度融合的关键前提，为后续的优化提供了精确的先验条件。通过这种多特征的融合，系统能够更充分地利用环境中的几何结构信息，尤其在低纹理或动态干扰场景下，直线特征的加入可以弥补点特征的不足，从而提高状态估计的精度和鲁棒性。

在进行紧耦合之前，视觉惯性初始化成功与否对系统影响重大，成功的初始化能够为系统提供准确的初始位置、姿态和地图信息，使定位精度更高，构建出的地图能准确反映真实环境，让系统在面对各种复杂情况时更稳定，还可提升实时性能，使视觉和惯性数据有效融合，充分发挥各自优势；而初始化失败则会使定位偏差大且误差不断累积，地图构建扭曲错乱，系统易受环境干扰而不稳定甚至崩溃，实时性大打折扣，数据融合也会出现偏差，无法准确感知环境，严重影响系统的整体性能和应用效果。因此在完成旋转和平移优化后，本文还加入了评估初始化过程是否成功的步骤，而这一步骤的主要过程是计算法向极线约束的平均残差，公式为

$$\bar{e}=\frac{1}{N}\sum_{k=0}^{N-1}\bar{e}_{k,k+1} \tag{21}$$

其中

$$\bar{e}_{k,k+1}=\frac{1}{\eta}\sum_{i=1}^{\eta}\left(\left|\boldsymbol{n}_i^{\mathrm{T}}\boldsymbol{t}_w^{c_k*}\right|\right)$$

$$\boldsymbol{n}_i=\boldsymbol{f}_i\times\boldsymbol{R}_b^{c_L}\hat{\boldsymbol{\gamma}}_{b_{k+1}}^{b_k}\boldsymbol{R}_{c_L}^{b}\boldsymbol{f}'$$

式中　$\bar{e}$——总平均残差

$\bar{e}_{k,k+1}$——从时间 k 到时间 $k+1$ 法向极线约束的平均残差

N——初始化期间滑动窗口中的关键帧数量

η——2 个连续关键帧中共同可见的方向向量对数

如果 $\bar{e}$ 低于阈值，则认为系统初始化成功，并继续进行视觉惯性光束平差法（VI－BA）。这样做是因为前面的步骤不仅提供了精确的初始估计作为联合 VI－BA 的输入值，加速其收敛，还加快了 VI－BA 的运行时间。

在完成视觉惯性初始化后，跟踪线程中的局部地图跟踪将与 IMU 一起执行视觉惯性光束平差法，以优化姿态、速度和 IMU 当前帧的偏差，其中 $\boldsymbol{\omega}_{VI}$ 是包含以上待优化变量的向量，可表示为

$$\boldsymbol{\omega}_{VI}^{*}=\boldsymbol{H}+\boldsymbol{I}+\boldsymbol{J} \tag{22}$$

其中

$$\boldsymbol{H}=\underset{\boldsymbol{\omega}_{VI}}{\operatorname{argmin}}\sum_{j\in P}\rho(E_{i,j}) \tag{23}$$

$$\boldsymbol{I}=\sum_{k\in L}\rho(F(\alpha(P))E_{i,k}) \tag{24}$$

$$\boldsymbol{J}=\sum_{k\in\boldsymbol{B}}\boldsymbol{r}_{I_i}^{\mathrm{T}}\sum_{I_{i,i+1}}^{-1}\boldsymbol{r}_{I_i} \tag{25}$$

式中　$\boldsymbol{B}$——惯性测量单元

$\boldsymbol{r}_{I_i}$——惯性残差

同时在局部光束平差中，视觉惯性 SLAM 也表示为一个基于关键帧的最小化问题。一组关键帧 K_L 的位姿和在这些关键帧中看到的所有点与线的位置，以及 IMU 参数，都包含在向量 $\boldsymbol{\omega}_{VI}^{k}$ 中，用方程表示为

$$\boldsymbol{\omega}_{VI}^{k}=\boldsymbol{S}+\boldsymbol{L}+\boldsymbol{M} \tag{26}$$

其中

$$\boldsymbol{S}=\underset{}{\operatorname{argmin}}\sum_{i\in K_L}\sum_{j\in P_L}\rho(e_{i,j}) \tag{27}$$

$$\boldsymbol{L}=\sum_{k\in L_L}\rho(e_{i,k}) \tag{28}$$

$$\boldsymbol{M}=\sum_{k\in\boldsymbol{B}}\boldsymbol{r}_{I_i}^{\mathrm{T}}\sum_{I_{i,i+1}}^{-1}\boldsymbol{r}_{I_i} \tag{29}$$

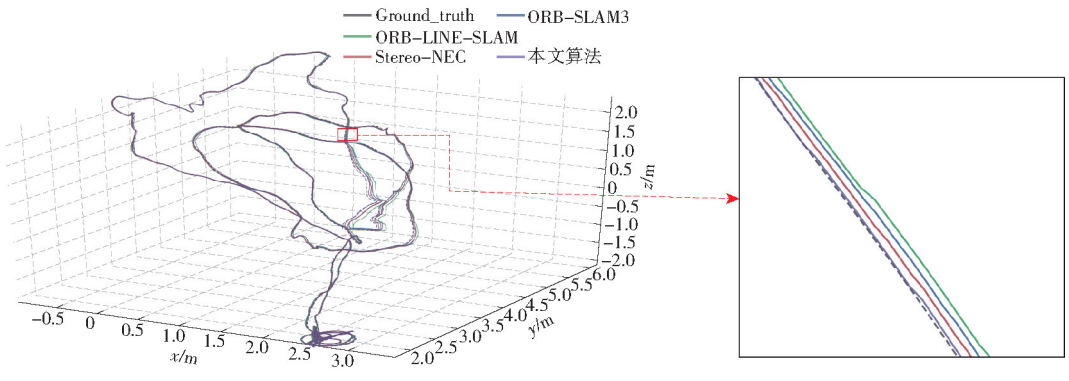
5　实验

5.1　实验平台

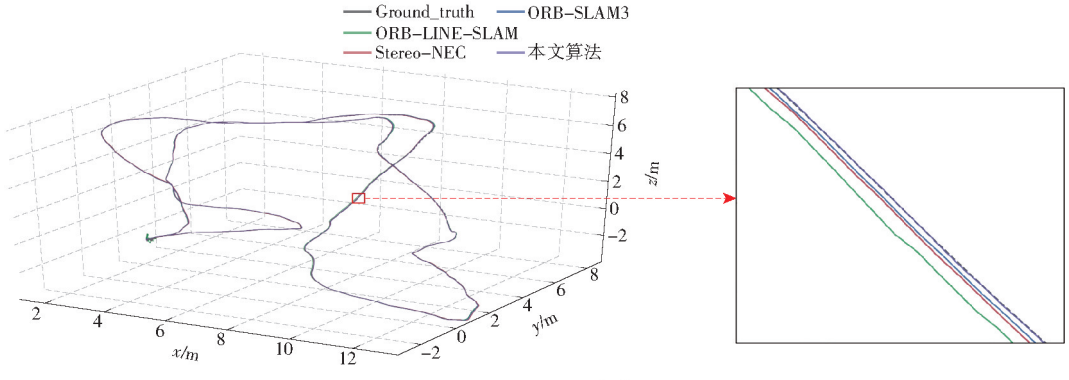
以 Turtlebot2 移动机器人作为硬件平台进行实验，验证所提出的方法，该平台采用了两轮差速的 Kobuki 移动底盘。并在该机器人平台上搭载了 Intel RealSense D455 型双目深度相机。实验平台上使用 HP OMEN 16－wf0032TX PC 机作为上位机，该 PC 机搭载 Intel i9－13900HX 处理器，主频高达 5.40 GHz，配备 16.00 GB 的运行内存，利用串口与移动机器人底盘相连接进行通信。本文以 Linux 开源系统 Ubuntu 18.04 为平台，搭载 Melodic 版本的机器人操作系统 ROS。

5.2　公开数据集实验分析

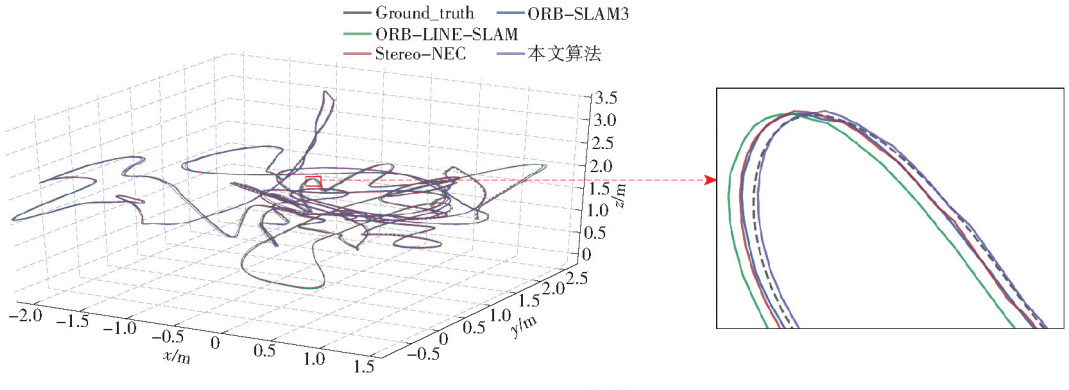
为验证本文提出的双目惯性 SLAM 系统的定位精度，采用公开的数据集对本文系统进行验证。本文采用的公开数据集是 EuRoC 数据集。在测试时为了使实验结果更据有说服力，将本文算法与 ORB－SLAM3、ORB－LINE－SLAM、Stereo－NEC 等优秀算法进行对比，将实验的 4 种算法轨迹与真实值对齐后，采用 EVO 评估工具对定位结果进行系统性量化分析。图 6 展示了 4 种算法在 MH_02_easy、MH_05_difficult、V1_03_difficult 和 V2_03_difficult 4 组数据集下的估计轨迹与真实轨迹的对比曲线。通过分析这些曲线可以发现，尽管 4 种算法在不同数据集中均表现出较高的定位精度，但在处理复杂运动场景时，本文提出的改进算法展现出显著优势。



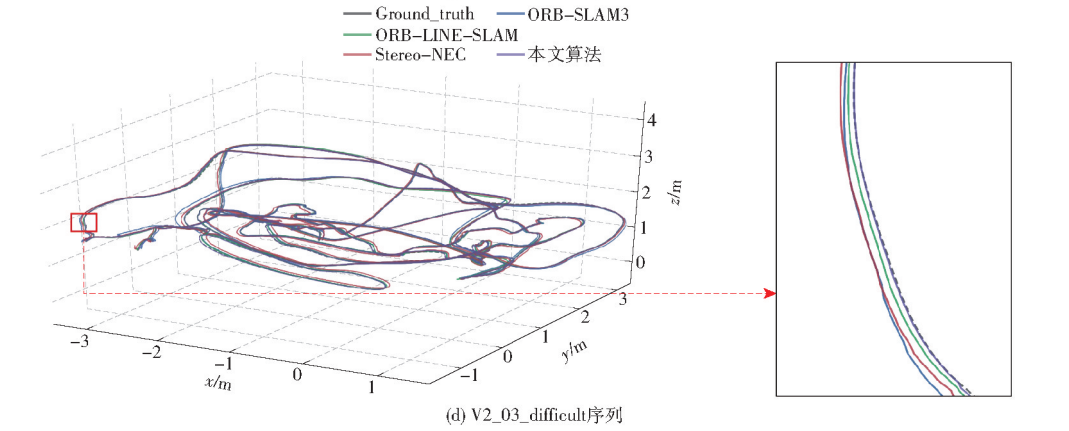
(a) MH_02_easy序列



(b) MH_05_difficult序列



(c) V1_03_difficult序列



(d) V2_03_difficult序列

图 6 各类算法运动轨迹对比曲线

Fig. 6 Comparison curves of motion trajectories of various algorithms

具体而言,当运动速度加快、出现突然抖动或急速转弯等复杂场景时,改进算法的定位精度明显优于其他 3 种算法。例如,在 V1_03_difficult 数据集中,当无人机进行快速转弯时,传统算法出现了明显的轨迹偏移,而改进算法则能够更好地贴合真实轨迹,保持了较高的稳定性。

采用绝对位姿误差(Absolute pose error, APE)作为评估指标,对本文算法在不同序列中的轨迹精

度进行了量化对比。具体实验结果如图 7 所示，图中分别展示了绝对轨迹误差热力图和误差统计图。在图 7 的左图中，热力图直观地反映了算法在各个位置上的轨迹误差分布情况，其中颜色越红表示误差越大，颜色越蓝则表示误差越小。从图 7 可以看出，本文算法的轨迹误差主要集中在蓝色区域，表明其在大多数位置上的误差较小，整体轨迹精度较高。图 7 的右图展示了绝对轨迹误差的统计结果。从图 7 可以观察到，本文算法的中值误差、均方根误差和误差均值均处于较低水平，且标准差范围较小，说明算法的误差分布较为集中，稳定性较好。这一结果进一步验证了本文算法在轨迹估计任务中的优越性。

为了更直观地评估本文算法与现有主流算法在 EuRoC 数据集上的性能表现，在不同时刻轨迹误差方面进行了系统性对比分析，图 8 展示了不同算法在不同序列下的估计轨迹与真实轨迹之间的误差对比。从图 8 可以清晰地观察到，本文算法在整个时

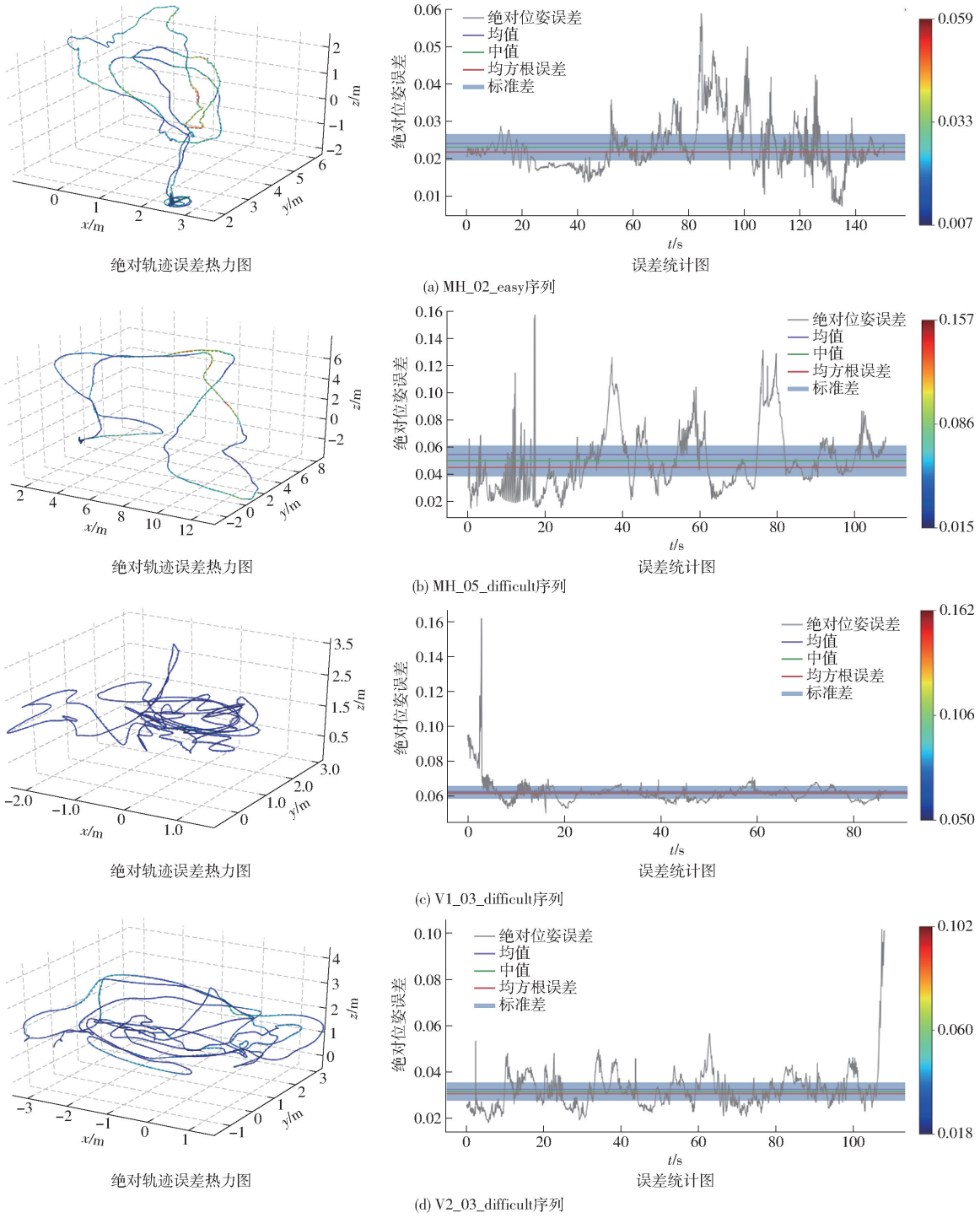


图 7　绝对轨迹误差热力图与误差统计图

Fig. 7　Absolute trajectory error heatmaps and error maps

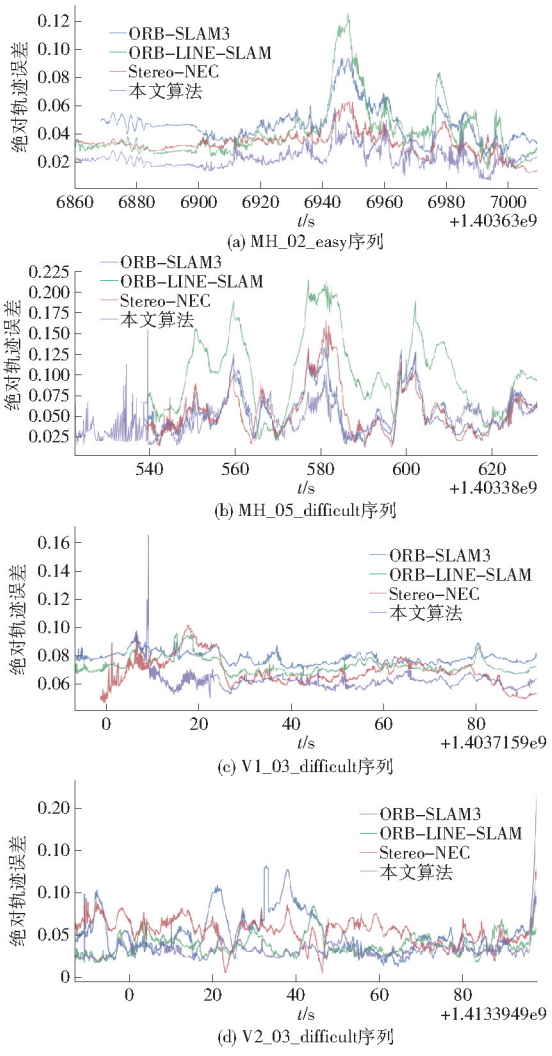


图 8 各类算法绝对轨迹误差对比曲线

Fig. 8 Comparison curves of absolute trajectory errors for various algorithms

间序列上的轨迹估计误差显著降低,其误差曲线明显优于 ORB - SLAM3、ORB - LINE - SLAM 和 Stereo - NEC 等对比算法,这表明本文算法在复杂环境下仍能保持较高的轨迹估计稳定性,有效降低了累计误差。为进一步量化评估算法性能,本文采用 ATE 的均方根误差(RMSE)作为评判指标,对 4 种算法在 EuRoC 数据集不同序列下的误差进行了统计分析,具体结果如表 1 所示。实验数据显示,本文算法在所有测试序列中的 RMSE 均显著低于对比算法。具体而言,相较于 ORB - SLAM3、ORB - LINE - SLAM 和 Stereo - NEC,本文算法的定位精度分别提升了 22.9%、20.6% 和 14.8%。这一显著的性能提升主要得益于本文提出的新型特征匹配策略,以及改进的 IMU 初始化框架,有效提高了系统在复杂场景下的鲁棒性。

5.3 真实环境实验分析

为了评估本文算法的定位性能及其在不同环境下的适应性,选取了室内和走廊2种典型场景进行

表 1 绝对轨迹误差(ATE)结果对比

Tab. 1 Comparison of absolute trajectory error (ATE)

数据集	ORB - SLAM3	ORB - LINE - SLAM	Stereo - NEC	本文算法
MH_01_easy	0.059 293	0.055 326	0.055 222	0.048 380
MH_02_easy	0.048 597	0.045 985	0.034 290	0.024 117
MH_03_medium	0.048 534	0.045 156	0.042 750	0.039 679
MH_04_difficult	0.061 486	0.055 420	0.055 235	0.050 845
MH_05_difficult	0.061 015	0.112 036	0.063 676	0.054 744
V1_01_easy	0.151 451	0.149 138	0.142 616	0.141 017
V1_02_medium	0.062 999	0.061 551	0.055 439	0.053 170
V1_03_difficult	0.077 574	0.072 237	0.068 516	0.062 276
V2_01_easy	0.047 583	0.050 926	0.042 971	0.037 292
V2_02_medium	0.035 447	0.029 267	0.030 741	0.024 410
V2_03_difficult	0.056 466	0.040 341	0.057 809	0.032 757

实验验证。室内环境通常具有复杂的纹理特征,而走廊场景则呈现出明显的线性结构和重复性以及稀疏的纹理特征,这 2 种场景能够充分检验算法的定位精度和环境鲁棒性。

5.3.1 室内环境实验

在室内环境下的定位实验中,选取了如图 9 所示的实验场景进行算法验证。该场景具有丰富的纹理特征,能够为视觉 SLAM 系统提供充足的 ORB 特征点。由于 ORB - SLAM3、ORB - LINE - SLAM 等经典算法以及本文提出的改进算法都需要依赖一定数量的 ORB 特征点来完成位姿估计,因此该场景非常适合用于评估各算法在复杂环境下的定位性能。实验场景中包含了多种视觉元素:密集的纹理区域(虽然特征点丰富,但可能增加计算量并引入误匹配)、重复的图案(容易造成视觉回环检测的歧义)等,这种综合性的实验场景能够全面检验本文算法在面对各种复杂环境时的性能,从而证明了其在实际应用中的鲁棒性。为了进一步验证算法性能,使用 Turtlebot2 移动机器人设计了 U 形路径的移动轨迹。这种轨迹包含大幅度的转向运动,能够有效测试算法在线速度 0.1 ~ 0.5 m/s、角速度最大 0.5 rad/s 的情况下,算法在快速转弯、视角突变等情况下的定位精度和系统稳定性。在轨迹的转折点

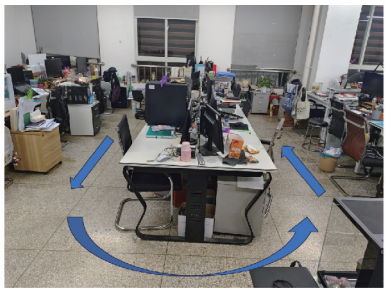


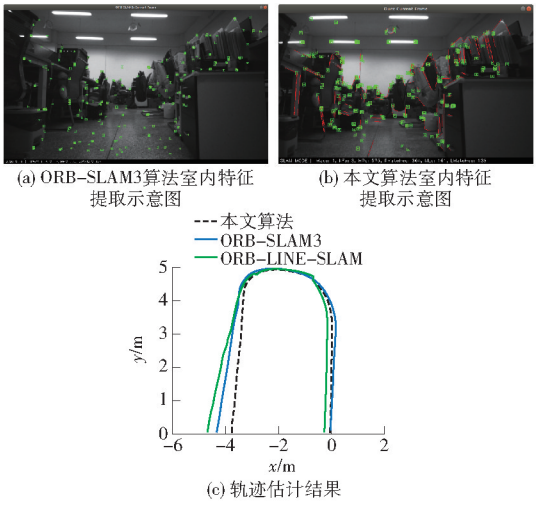
图 9 室内实验环境

Fig. 9 Indoor experimental environment

处，相机的视角会发生剧烈变化，容易导致特征跟踪丢失或位姿估计漂移，这对 SLAM 系统提出了更高的要求。通过对比分析各算法在该轨迹下的定位误差可以更全面地评估本文改进算法在复杂环境中的优势。

为了确保测试场景的一致性和算法评估的可靠性，首先在实际测试场景中控制移动机器人按照预设轨迹运动，并通过传感器采集环境数据、运动状态等信息，形成标准化的测试数据集。随后，在 Ubuntu 系统环境下，基于同一数据集对多种算法进行离线性能验证。这种设计具有以下优势：①通过数据录制与算法验证的分离，有效避免了多次实地测试中可能出现的环境变化（如光照、障碍物分布等）和机器人运动状态差异（如速度波动、定位误差等）对算法评估的干扰。②所有算法均在完全相同的输入数据条件下运行，确保了测试条件的可重复性和结果的可比性。③离线验证模式便于进行大规模参数调优和算法对比实验。

对比了不同算法在室内环境中的特征提取效果。其中，图 10a 展示了基准 ORB－SLAM3 算法的点特征提取效果。从图中可以看到，该算法能够稳定地检测出丰富的特征点，这些特征点为后续的位姿估计提供了基础且鲁棒的视觉信息。图 10b 则展示了本文算法的特征提取效果。与基准版本相比，本文算法不仅保留了丰富的点特征，还成功提取了环境中的关键线特征，从而为后续的位姿估计提供了更加丰富和全面的特征信息。这种点线融合的机制显著增强了算法在纹理稀疏或结构化场景中的鲁棒性，确保了在复杂室内环境中仍能保持较高的环境辨识度。对于真实环境实验，由于缺少可作为真值（Ground truth）的参考轨迹，无法直接计算绝对轨迹误差（ATE）来评估算法精度。因此，通过计算轨迹终点与起点之间的水平偏移量来衡量算法的累积误差和回环校正能力。在图 10 c 中，对比了本文算法、ORB－SLAM3 和 ORB－LINE－SLAM 3 种算法的轨迹估计结果。虽然 3 种算法均能完成基本的运动轨迹估计，但在细节表现上存在明显差异。特别是在转弯区域，本文算法的轨迹曲线相较于 ORB－LINE－SLAM 算法的轨迹曲线更加平滑，表明其在动态运动过程中具有更好的位姿优化能力。通过测试结果可知，本文算法估计的终点与起点之间的水平偏移仅为 0.75 m，而 ORB－SLAM3 和 ORB－LINE－SLAM 算法的偏移量分别达到 1.31 m 和 1.66 m。这一差距进一步验证了本文算法在定位精度上的优势。



(a) ORB-SLAM3算法室内特征提取示意图　(b) 本文算法室内特征提取示意图

(c) 轨迹估计结果

图 10　室内实验结果

Fig. 10　Indoor experimental results

5.3.2　走廊场景实验

为了进一步验证本文算法在重复性纹理和稀疏特征环境下的鲁棒性，又选取了走廊场景作为实验环境，如图 11 所示。该测试场景的纹理特征呈现 3 个显著特点：首先，走廊两侧是大面积的无特征白墙，仅存在少量突兀的门框作为间断性特征点；其次，地面铺设具有高度重复纹理特征的地板，这种重复性图案容易造成特征点误匹配；此外，部分测试路段存在明显的环境亮度不均匀现象，这对视觉算法的光照适应性提出了更高要求。在实验过程中，控制移动平台沿 30 m 的直线走廊行进，重点考察算法在长距离运动、狭窄空间约束下的表现。这是为了评估不同算法对纹理稀疏环境的适应能力，同时检验系统在特征重复区域的轨迹连续性，以及在整个运行过程中的定位精度。



图 11　走廊实验场景

Fig. 11　Corridor experimental scene

图 12 展示了不同 SLAM 算法在走廊环境下的轨迹曲线对比结果。通过分析可以发现，本文提出的算法与 ORB－LINE－SLAM 算法的轨迹曲线明显比 ORB－SLAM3 算法的轨迹更加平滑连续。这种优势主要源于 2 种算法都采用了点线特征融合的方法，在稀疏但几何特征明显的走廊环境中，线特征能够提供更丰富的结构信息，与点特征形成互补，从而显著提升了系统的定位精度和环境适应性。相比之下，仅依赖 ORB 点特征的 ORB－SLAM3 算法在低

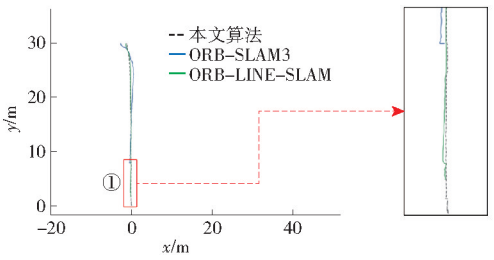


图 12 走廊轨迹估计结果

Fig. 12 Corridor trajectory estimation results

纹理区域容易出现特征匹配不足以及误匹配的情况,导致轨迹出现波动和漂移。特别值得注意的是图中标记①的区域,ORB - SLAM3 和 ORB - LINE - SLAM 算法的轨迹都出现了明显的缺失段,而本文算法则保持了从原点(0,0)开始的完整轨迹。这由于本文算法在初始化阶段通过引入基于法极线约束的视觉惯性初始化方法,不仅缩短了 IMU 参数的初始化时间,同时解决了传统视觉惯性系统在初始化阶段容易出现的尺度不确定性问题,并且显著增强了系统在弱纹理环境下的稳定性。

此外,在本次走廊环境实验中,对移动机器人实际运动的终点与起点所在竖直线之间的垂直距离进行了测量,其值为 0.75 m。值得注意的是,在长达 30 m 的走廊测试中,本文提出的融合算法表现出显著优势,其估计的终点垂直偏移量仅为 0.55 m,而传统 ORB - SLAM3 算法产生的偏移量高达 1.95 m,相比之下 ORB - LINE - SLAM 算法虽然有所提升,但仍存在 0.95 m 的明显偏差。这一结果充分证明了本文算法在具有线性结构和重复性以及稀疏的纹理特征环境中也具有更好的定位精度和鲁棒性。

5.4 消融实验

为了量化本文所提出的各项改进对系统性能的贡献,本文设计了全面的消融实验。“O”表示基准算法,“Line”表示启用了线特征提取与匹配模块,“IMU”表示加入了惯性传感器,“OURS”表示本文完整算法。在 EuRoC 数据集上,对比了 4 种不同的算法配置,并以绝对轨迹误差 ATE 作为主要评估指标。实验结果如表 2 所示。

表 2 不同算法绝对轨迹误差(ATE)结果对比

Tab. 2 Comparison of absolute trajectory error (ATE) results for different algorithms

算法	MH01	MH02	MH03	MH04	MH05	V101	V102	V103	V201	V202	V203
O	0.059	0.049	0.049	0.061	0.061	0.151	0.063	0.078	0.048	0.035	0.056
O + Line	0.054	0.046	0.046	0.056	0.085	0.147	0.054	0.067	0.036	0.032	0.040
O + IMU	0.056	0.038	0.041	0.058	0.055	0.143	0.056	0.076	0.044	0.032	0.053
OURS	0.048	0.024	0.040	0.051	0.055	0.141	0.053	0.062	0.037	0.024	0.033

通过对比基准版本“O”与仅引入线特征的版本“O + Line”,可以观察到在大多数序列中,线特征的引入带来了性能的稳步提升,这印证了线特征作为一种重要的几何约束,能够有效弥补点特征在纹理稀疏或结构化场景中的不足。同时,比较基准版本与仅引入 IMU 的版本“O + IMU”,可以看出 IMU 数据在某些运动模式复杂的序列(如 MH02)中也展现出独特的贡献,为系统提供了更可靠的初始状态估计。本文的完整版本“OURS”在所有测试序列中均展现出卓越的性能,其 ATE 误差始终处于最低水平,强有力地证明了点、线特征与 IMU 残差的紧密融合并非简单的性能叠加,而是通过多传感器信息的互补与冗余,实现了系统鲁棒性与定位精度的提升。综上所述,消融实验结果清晰地表明,本文提出的每一项改进都是系统性能提升的关键因素,并且它们之间的协同作用是实现最终优异性能的核心。

5.5 算法效率实验

为了评估本文所提出的改进对算法实时性的影响,对不同算法版本的平均每帧处理时间进行了对比。实验在相同的硬件环境下进行,以确保对比的公平性。从表 3 所示的实验结果中可以清晰地观察到,相比于基准版本“O”,本文的改进版本在处理时间上均有所增加。具体而言“O + IMU”版本由于额外执行了改进的 IMU 初始化,其处理时间略有增加,但增幅有限。而“O + Line”和“OURS”版本的处理时间则显著高于基准版本,这主要源于线特征提取(LSD 算法)和 LSD 描述子匹配所带来的额外计算开销。

然而,必须将此开销置于整体性能提升的背景下进行考量。虽然完整版本“OURS”的平均处理时间在 75 ms 左右,但这一数值仍远低于 100 ms 的实时性阈值,能够维持 10 f/s 以上的运行帧率,足以满足绝大多数室内移动机器人和增强现实应用的实时性需求。因此,在实现了定位精度显著提升的同时,本文算法仍能够保持良好的实时性。这表明本文所提出的方法在性能和计算开销之间取得了有效且可接受的平衡,证明了其在实际应用中的可行性和价值。

6 结束语

本文针对传统点特征 SLAM 在弱纹理、光照变

表 3　不同算法在 EuRoC 数据集上的平均每帧处理时间

Tab. 3　Frame processing time of different algorithm on the EuRoC dataset　　ms

算法	MH01	MH02	MH03	MH04	MH05	V101	V102	V103	V201	V202	V203
O	22.04	25.01	23.96	24.51	23.93	24.34	24.40	24.45	25.01	24.51	25.33
O + Line	68.77	66.30	68.14	68.32	68.76	70.85	70.10	68.84	74.49	73.93	70.79
O + IMU	25.85	24.69	25.20	25.48	25.02	25.41	25.89	25.19	27.00	26.63	27.07
OURS	79.87	74.45	78.58	74.79	74.06	77.84	77.23	74.02	82.75	82.57	78.98

化及运动模糊场景下的性能退化问题，提出了一种基于点线特征融合的双目视觉 SLAM 系统。该系统在 ORB－SLAM3 框架基础上，融合 ORB 点特征与 LSD 线特征，通过引入改进的线段误差函数与自适应权重因子优化光束平差过程，并利用自适应因子动态调整线特征权重以平衡不同场景下点线特征的贡献。通过在纯运动 BA 与局部 BA 中集成点线约束，显著提升了位姿估计的鲁棒性。同时，改进了法向极线约束方案，扩展构建了双目 MNEC 约束的陀螺仪偏差估计器，采用旋转–平移解耦优化策略，并引入残差评估机制确保初始化可靠性，将 IMU 残差、特征点重投影误差及直线重投影误差共同作为非线性优化的约束条件，从而大幅提高了视觉惯性 SLAM 系统在复杂场景下的初始化精度与鲁棒性。为全面评估性能，在标准 EuRoC 数据集及真实室内、长廊场景中进行了多维度实验。结果表明，本系统在定位精度上显著优于主流算法，相较于 ORB－SLAM3、ORB－LINE－SLAM 和 Stereo－NEC 算法分别提升了 22.9%、20.6% 和 14.8%，并在真实环境中同样展现出更优的定位精度和鲁棒性。

参　考　文　献

[1] 陈明方，黄良恩，王森，等. 移动机器人视觉里程计技术研究综述[J]. 农业机械学报，2024，55(3)：1－20.
CHEN Mingfang, HUANG Liang'en, WANG Sen, et al. Survey of research on visual odometry technology for mobile robots[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(3): 1－20. (in Chinese)

[2] CAMPOS C, ELVIRA R, RODRIGUEZ 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.

[3] YANG G, MENG W D, HOU G D, et al. Real-time visual-inertial odometry based on point-line feature fusion[J]. Gyroscopy and Navigation, 2023, 14(4): 339－352.

[4] LI Q, WANG X, WU T, et al. Point-line feature fusion based field real-time RGB－D SLAM[J]. Computers & Graphics, 2022, 107: 10－19.

[5] GENEVA P, ECKENHOFF K, LEE W, et al. Openvins: a research platform for visual-inertial estimation[C]//2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2020: 4666－4672.

[6] SHAN T, ENGLOT B, RATTI C, et al. LVI－SAM: tightly-coupled LiDAR-visual-inertial odometry via smoothing and mapping[C]//2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2021: 5692－5698.

[7] FU Q, WANG J, YU H, et al. PL－VINS: real-time monocular visual-inertial SLAM with point and line features[J]. arXiv Preprint, arXiv:2009.07462, 2020.

[8] YAN J, ZHENG Y, YANG J, et al. PLPF-VSLAM: an indoor visual SLAM with adaptive fusion of point-line-plane features[J]. Journal of Field Robotics, 2024, 41(1): 50－67.

[9] ZENG D, LIU X, HUANG K, et al. EPL－VINS: efficient point-line fusion visual-inertial SLAM with LK－RG line tracking method and 2－DoF line optimization[J]. IEEE Robotics and Automation Letters, 2024, 9(6): 5911－5918.

[10] YANG H, YUAN J, GAO Y, et al. UPLP－SLAM: unified point-line-plane feature fusion for RGB－D visual SLAM[J]. Information Fusion, 2023, 96: 51－65.

[11] WANG K, YAO X, MA N, et al. PLMOT－SLAM: a point-line features fusion SLAM system with moving object tracking[J]. The Visual Computer, 2025, 41(7): 4547－4565.

[12] KAISER J, MARTINELLI A, FONTANA F, et al. Simultaneous state initialization and gyroscope bias calibration in visual inertial aided navigation[J]. IEEE Robotics and Automation Letters, 2016, 2(1): 18－25.

[13] HE J, LI M, WANG Y, et al. Ple-slam: a visual-inertial slam based on point-line features and efficient imu initialization[J]. IEEE Sensors Journal, 2025, 25: 6801－6811

[14] CAMPOS C, MONTIEL J M M, TARDOS J D. Inertial-only optimization for visual-inertial initialization[C]//2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2020: 51－57.

[15] WANG W, CHOU C, SEVAGAMOORTHY G, et al. Stereo-NEC: enhancing stereo visual-inertial slam initialization with normal epipolar constraints[C]//2024 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2024: 2691－2697.

[16] LEE W, ECKENHOFF K, YANG Y, et al. Visual-inertial-wheel odometry with online calibration[C]//2020 IEEE/RSJ

International Conference on Intelligent Robots and Systems (IROS). IEEE, 2020: 4559 - 4566.
[17] FURGALE P, REHDER J, SIEGWART R. Unified temporal and spatial calibration for multi-sensor systems[C]// 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2013: 1280 - 1286.
[18] CHANG Y, EBADI K, DENNISTON C E, et al. LAMP 2.0: a robust multi-robot SLAM system for operation in challenging large-scale underground environments[J]. IEEE Robotics and Automation Letters, 2022, 7(4): 9175 - 9182.
[19] TEED Z, DENG J. Droid-SLAM: deep visual slam for monocular, stereo, and RGB - D cameras[J]. Advances in neural Information Processing Systems, 2021, 34: 16558 - 16569.
[20] 陈劭, 郭宇翔, 高天啸, 等. 移动机器人 RGB - D 视觉 SLAM 算法[J]. 农业机械学报, 2018, 49(10): 38 - 45.
CHEN Shao, GUO Yuxiang, GAO Tianxiao, et al. RGB - D visual SLAM algorithm for mobile robots[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49 (10): 38 - 45. (in Chinese)
[21] 董蕊芳, 王宇鹏, 阚江明. 基于改进 ORB_SLAM2 的机器人视觉导航方法[J]. 农业机械学报, 2022, 53(10): 306 - 317.
DONG Ruifang, WANG Yupeng, KAN Jiangming. Visual navigation method for robot based on improved ORB_SLAM2[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(10): 306 - 317. (in Chinese)
[22] 齐咏生, 宋继鹏, 刘利强, 等. 基于点线特征融合的延迟边缘化视觉惯性 SLAM 方法[J]. 农业机械学报, 2024, 55(12): 372 - 383.
QI Yongsheng, SONG Jipeng, LIU Liqiang, et al. Delayed marginalization visual inertia SLAM method based on point and line feature fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(12): 373 - 382. (in Chinese)
[23] VON GIOI R G, JAKUBOWICZ J, MOREL J M, et al. LSD: a fast line segment detector with a false detection control[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 32(4): 722 - 732.
[24] ZHANG L, KOCH R. An efficient and robust line segment matching approach based on LBD descriptor and pairwise geometric consistency[J]. Journal of Visual Communication and Image Representation, 2013, 24(7): 794 - 805.
[25] DURRANT-WHYTE H, RYE D, NEBOT E. Localization of autonomous guided vehicles[C]// Robotics Research: the Seventh International Symposium. Springer London, 1996: 613 - 625.
[26] KNEIP L, LYNEN S. Direct optimization of frame-to-frame rotation[C]// Proceedings of the IEEE International Conference on Computer Vision, 2013: 2352 - 2359.
[27] HE Y, XU B, OUYANG Z, et al. A rotation-translation-decoupled solution for robust and efficient visual-inertial initialization [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 739 - 748.

(上接第 329 页)

[14] MENON K, MITTAL R. Quantitative analysis of the kinematics and induced aerodynamic loading of individual vortices in vortex-dominated flows: a computation and data-driven approach[J]. Journal of Computational Physics, 2021, 443: 110515.
[15] HUANG X B, JIN G Q, LIU X D, et al. A method for quantitatively depicting the influence of vortices on the pressure forces acting on a hydrofoil in cavitating flow with tip clearance[J]. International Journal of Multiphase Flow, 2024, 178: 104899.
[16] PEDERSEN N, LARSEN P S, JACOBSEN C B. Flow in a centrifugal pump impeller at design and off-design conditions—part I: particle image velocimetry (PIV) and laser doppler velocimetry (LDV) measurements[J]. Journal of Fluids Engineering, 2003, 125(1): 61 - 72.
[17] HUANG X B, LIU Z Q, LI Y J, et al. Study of the internal characteristics of the stall in a centrifugal pump with a cubic non-linear SGS model[J]. Journal of Hydrodynamics, 2019, 31(4): 788 - 799.
[18] POPE S B. Turbulent flows[M]. Cambridge: Cambridge University Press, 2000.
[19] LIU X D, LI Y J, LIU Z Q, et al. Experimental and numerical investigation of stall mechanism in centrifugal pump impeller [J]. Appl. Fluid Mech., 2022, 15(3): 927 - 941.
[20] LIU B Q, YANG W, CHENG L, et al. Numerical study on jet-wake flow and its evolution in a centrifugal pump with alternating stall[J]. Physics of Fluids, 2023, 35: 115105.