



计算机应用研究
Application Research of Computers
ISSN 1001-3695, CN 51-1196/TP

《计算机应用研究》网络首发论文

题目: 基于 ORB-SLAM3 的改进型特征匹配与稠密建图算法
作者: 刘畅, 党淑雯, 陈丽
DOI: 10.19734/j.issn.1001-3695.2023.03.0098
收稿日期: 2023-03-06
网络首发日期: 2023-05-28
引用格式: 刘畅, 党淑雯, 陈丽. 基于 ORB-SLAM3 的改进型特征匹配与稠密建图算法 [J/OL]. 计算机应用研究. <https://doi.org/10.19734/j.issn.1001-3695.2023.03.0098>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式 (包括网络呈现版式) 排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊 (光盘版)》电子杂志社有限公司签约, 在《中国学术期刊 (网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊 (网络版)》是国家新闻出版广电总局批准的网络连续型出版物 (ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

基于 ORB-SLAM3 的改进型 特征匹配与稠密建图算法

刘 畅, 党淑雯[†], 陈 丽

(上海工程技术大学 航空运输学院, 上海 201600)

摘 要: 针对传统 ORB 算法存在提取的特征点极易堆积在纹理丰富的区域及误匹配率高等而导致无法满足高精度定位要求, 以及 ORB-SLAM3 系统无法构建稠密地图的问题, 提出一种基于 ORB-SLAM3 的改进型 ORB-GMS 特征匹配方法, 并增加稠密建图线程来实现稠密地图的构建。首先, 在特征点提取过程中引入四叉树策略, 将图像帧分为若干个网格, 并在每个网格中提取最优特征点; 然后, 在特征匹配过程中将运动平滑约束转换为剔除错误匹配的统计量, 通过比较匹配对邻域内的匹配对数量和阈值来快速筛选正确匹配; 最后, 完成位姿估计并根据关键帧与相应位姿完成稠密点云地图的构建。采用 TUM 的 RGB-D 数据集进行实验, 改进算法提取的特征点相较于传统 ORB 算法分布更加均匀, 匹配数比 ORB-SLAM3 增加 64.5%, 比 GMS 算法增加 4.7%, 匹配耗时比 ORB-SLAM3 减少 20.4%, 比 GMS 算法减少 94.6%, 从而验证了改进算法在特征点提取与匹配方面的优越性, 并且相较于 ORB-SLAM3, 改进算法的定位精度提高了 3.75%, 从而验证了其在总体上提高定位精度进而实现稠密建图的可行性和有效性。

关键词: 特征点; 特征匹配; 四叉树原理; 基于网格的运动统计; 稠密建图

中图分类号: TP242.6 doi: 10.19734/j.issn.1001-3695.2023.03.0098

Improved feature matching and dense-mapping algorithm based on ORB-SLAM3

Liu Chang, Dang Shuwen[†], Chen Li

(School of Air Transport, Shanghai University of Engineering Science, Shanghai 201600, China)

Abstract: In view of the problems that the extracted feature points of the traditional ORB algorithm tend to accumulate in the texture-rich area and the high false matching rate, which cannot meet the requirements of high-precision positioning, and that the ORB-SLAM3 system cannot build a dense map, this paper proposes an improved ORB-GMS feature matching method based on ORB-SLAM3, and adds dense-mapping thread to realize the construction of dense maps. Firstly, the feature points extraction process adopts the quadtree strategy to divide the image frame into several meshes, and extracts the best feature points in each mesh. Then, replace the motion smoothing constraint with a statistic that rejects incorrect matches during feature matching, and use a comparison of the number of matching pairs and thresholds within the neighborhood of matching pairs to quickly filter correct matches. Finally, complete the positional estimation and construct a dense point cloud map by keyframes and corresponding poses. Testing by RGB-D dataset from TUM, the improved algorithm can extract uniform feature points, the number of matches increases by 64.5% than ORB-SLAM3, increases by 4.7% than the GMS algorithm, and the matching elapsed time decreases by 20.4% than ORB-SLAM3, and decreased by 94.6% than the GMS algorithm. This proves that the improved algorithm is superior in feature point extraction and matching. And compared to the ORB-SLAM3, the accuracy of the improved algorithm increases by 3.75%, thus, demonstrating its feasibility and effectiveness in improving the localization accuracy and building dense maps.

Key words: feature point; feature matching; quad tree principle; grid-based motion statistics; dense mapping

0 引言

同时定位与地图构建 (simultaneous localization and mapping, SLAM)^[1-3]是指在未知环境下机器人根据传感器采集的数据实时构建目标环境地图的同时进行定位的新技术^[4]。

视觉 SLAM (Visual SALM, VSLAM) 采用视觉传感器作为数据源, 成本相对激光雷达等其他 SLAM 技术更加低廉, 是当前 SLAM 研究领域热点方向之一^[5]。

虽然传统视觉 SLAM 技术不断在发展, 但仍无法满足实际环境应用的需求, 所以前人从不同方面对其进行了改进。

收稿日期: 2023-03-06; 修回日期: 2023-04-23 基金项目: 国家自然科学基金资助项目 (52175103)

作者简介: 刘畅(1996-), 男, 安徽宿州人, 硕士研究生, 主要研究方向为视觉 SLAM; 党淑雯(1979-), 女(通信作者), 内蒙古乌兰察布人, 副教授, 硕士, 主要研究方向为 SLAM、多传感器数据融合 (swdang@126.com); 陈丽(1975-), 女, 黑龙江牡丹江人, 教授, 硕士, 博士, 主要研究方向为飞行器和机器人动力学与控制。

有学者围绕以相机作为单一传感器的 V-SLAM 系统开展研究, Davison^[6]率先设计出了使用单目相机的实时单目 VSLAM 系统; Klein 和 Murray^[7]在 2007 年提出了并行追踪与建图(Parallel Tracking and Mapping, PTAM)方法, 其实现了并行化跟踪与建图过程和非线性优化, 对传统滤波器, 如卡尔曼滤波^[8]、粒子滤波^[9]和组合滤波^[10]等作为后端的算法方案进行了改善。文献[11]研究了以 PTAM 作为基础的单目 RGB-D-SLAM 系统, 而 ORB-SLAM2^[12]是一种使用双目 RGB-D 相机, 并以稀疏特征为基础的轻量级 SLAM 定位方法, 但由于特征追踪存在积累误差, 使后面帧和实际位姿的距离增加, 进而降低系统整体精度。有人通过融合其他传感器(如 GPS、IMU 等)设计出多传感器 V-SLAM 系统, 提高了算法在大多数环境的鲁棒性。Mourikis 等^[13]在 2007 年提出基于 EKF 的 MSCKF 系统, 该方法利用 IMU 数据实现预测阶段系统状态的传递, 并且通过将线性关联代替原来三维点的高阶关联来减少计算量。OKVIS^[14]是 Leutenegger 等在 2014 年提出的典型紧耦合视觉-惯性框架, 该算法为确保实时性, 通过优化关键帧选择策略实现静止时维持固定姿态, 并且边缘化旧状态使有界大小的滑动窗口保持非线性优化。David^[15]等于 2019 年设计的自适应因子图优化的多传感器融合系统在应对复杂的真实场景时自动选择状态最优的传感器信息。ORB-SLAM3^[16]是通过在 ORB-SLAM2 基础上增加 IMU 融合算法和多地图系统而提出的基于特征点法的紧耦合 VIO 系统, 该算法能充分利用数据关联来完成系统漂移重置和回环校正, 具有更强的鲁棒性。2022 年, 张昊宇等^[17]抛弃 ORB-SLAM3 框架的原特征点提取算法, 采用 Geometric Correspondence Network v2 网络^[18]提取特征点并生成二值描述子, 然后利用最近邻匹配算法和基于运动统计的快速鲁棒特征匹配过滤算法进行特征匹配及剔除误匹配。薛秀杰等^[19]提出一种基于点线特征的立体视觉惯性 SLAM 算法。该算法在 ORB-SLAM3 上增添线特征, 利用改进型 LSD 检测线特征, 利用 IMU 或等速模型提高线特征投影匹配效率, 通过构建点线重投影误差, 同时计算 IMU 的位置和位姿。胡章芳等^[20]研究出使用实时 STDC 网络的语义线程对动态对象进行分段的短期密集级联网络 SLAM(STDC-SLAM), 该算法基于 ORB-SLAM3, 一方面设计分割细化模块, 利用图像深度信息优化语义分割图, 另一方面通过增加接受线程中低质量特征点的迭代来改进 Qtree-ORB 算法, STDC-SLAM 在定位精度和运行速度方面比 Dyna SLAM^[21]和 PSP Net-SLAM^[22]有明显提高。吕克峰等^[23]于 2023 年设计出基于顺序图像分割的 SIIS-SLAM, 该算法在 ORB-SLAM3 基础上集成顺序图像分割实例和光流动态检测模块, 其中顺序图像分割模块能够消除序列帧间进行相对位姿估计时动态对象的有效性, 从而提高实例分割的速度和精度。通过融合实例分割结果和光流模块有效检测动态物体并获得过滤动态对象后的特征点, 进而提高算法的准确性和鲁棒性并在动态环境中实现更好的精度。

为解决传统 ORB 算法特征点提取分布不均匀且误匹配率高, 以及无法实现稠密建图等问题, 本文从提高特征匹配精度和生成易进行二次加工的地图这两个角度出发, 提出一种基于 ORB-SLAM3 算法的改进型 ORB-GMS 特征匹配策略, 并增加稠密建图线程以实现精准定位及稠密建图。该方法首先按照 ORB-SLAM3 算法利用四叉树策略将图像分为若干个网格, 再在每个网格中提取最佳特征点; 然后在匹配过程中引入 GMS 匹配方法^[24], 利用基于网格的剔除误匹配的统计量来筛选正确匹配, 以得到分布均匀且正确的匹配对;

最后进行位姿估计, 并根据关键帧与相应位姿构建稠密点云地图。此外, 本文利用 TUM 数据集集中的 RGB-D 标准数据开展仿真实验, 通过与原始算法进行对比来验证所提方法的有效性及其优越性。

1 系统结构框架

ORB-SLAM 作为一种适用于单目、双目、RGB-D 等多种相机模型的视觉 SLAM 系统, 该系统从 ORB 特征点提取出发, 再结合追踪、局部建图和回环检测三个线程来构成整个 SLAM 系统。本文在 ORB-SLAM3 的结构基础上, 增加了稠密建图线程, 整体框架如图 1 所示。

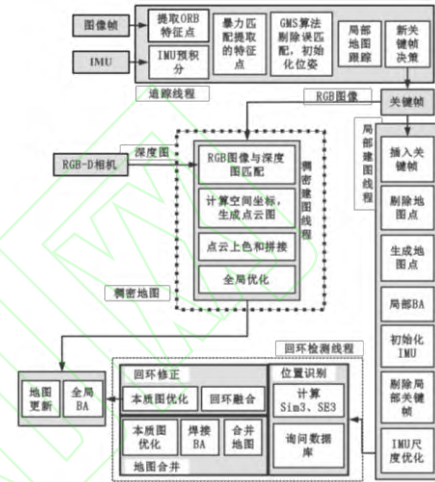


图 1 算法流程图

Fig. 1 Algorithm flowchart

1.1 追踪线程

在追踪线程, 根据传感器的不同需要, 选择不同的初始化方式。首先, 对图像进行 ORB 特征提取与匹配, 由于传统 ORB 算法提取的特征点分布不均匀且误匹配率过高, 因此本文基于 ORB-SLAM3 算法对特征匹配环节进行改进。首先在特征点提取过程利用四叉树策略将图像分为若干个网格, 分别在每个网格中提取最佳特征点, 然后在特征匹配阶段引入 GMS 匹配方法, 使用基于网格剔除误匹配的统计量筛选正确匹配, 从而得到分布比较均匀且正确的匹配对; 其次, 利用恒速运动跟踪模型和参考帧跟踪模型计算初始位姿, 两个模型分别用于正常跟踪模式与跟踪失败后的重定位; 然后, 跟踪局部地图以确定更多局部地图点与当前帧图像的特征匹配关系, 进而基于图优化模型进一步优化相机位姿; 最后, 根据关键帧选取条件确定新关键帧。此外, IMU 模式中的跟踪线程还需要计算 IMU 预积分, 从而使局部建图线程在视觉与 IMU 结合后进行局部 BA 优化。

1.2 局部建图线程

局部建图线程的流程主要包含两方面, 首先是对局部地图进行维护, 即将新关键帧和地图点插入活动地图, 同时剔除不满足条件的关键帧和地图点, 然后进行局部 BA 进一步优化关键帧位姿和地图点空间坐标。因为 ORB-SLAM3 的传感器类型包含 IMU, 所以该算法的局部地图构建线程还包括 IMU 的初始化, 该过程的目的是为了给局部 BA 和全局 BA 提供一个更好的初始值以减少 IMU 噪声积累。

1.3 回环检测线程

回环检测线程主要包括回环检测和回环校正。首先在关键帧数据库中确认候选闭环帧, 并与当前关键帧进行特征匹配, 然后通过 Sim3 变换计算相似变换, 其次当前帧的共视

帧和候选闭环帧进行投影匹配, 当匹配数满足相应条件时则可检测到回环, 最后校正关键帧位姿和局部地图点三维坐标, 并融合地图点和优化本质图。回环校正完成后 ORB-SLAM3 系统通过在独立线程中启动完整的 BA 细化地图映射, 并且可保障实时性不受影响。

2 改进型 ORB-GMS-SLAM 算法

2.1 基于四叉树策略的特征点提取

作为目前特征点提取匹配中常用的算法之一, ORB 特征包括关键点和描述子两部分, 其中关键点称为 Oriented FAST, 是一种具有不变尺度和旋转的改进 FAST 角点^[25], 描述子称为 BRIEF^[26], 其使用随机选点比较, 并且利用二进制表达, 所以两点的特征比较速度非常快, 存储也方便。

传统 ORB 算法提取的特征点过于集中, 经常出现扎堆现象, 本文通过对图片划分区域提取特征点, 引入四叉树策略进行筛选, 最大程度使用图片信息的同时也增加了算法的鲁棒性。具体步骤如下:

a) 网格划分: 建立图像金字塔, 各层图像划分成若干网格。

b) FAST 角点提取: 对每层金字塔图像使用 FAST 角点检测器进行角点提取, 对无法提取出特征点的区域, 通过降低阈值以确保获取足量角点。

当周围 16 个像素点与 p 点的灰度差值之和大于预设阈值 T , 并符合角点响应函数, 则确定像素 P 为 FAST 角点。角点响应函数如下:

$$N = \sum_{i=1}^{16} |I(i) - I(p)| > T \quad (1)$$

c) 特征点筛选。利用四叉树的思想, 对每个网格中提取的特征点进行筛选, 如果该网格中不存在特征点, 则删除该网格; 如果该网格中存在特征点, 则继续将该网格划分成四等份, 再次判断每个网格中是否存在特征点, 最后按照一定的策略挑选出分布均匀的 FAST 角点。四叉树策略筛选特征点过程如下:

(a) 首先大概分析初始图可分裂出几个原始节点, 然后将完成特征点提取的图像划分。一般按照四叉树的根节点以分辨率为 640×480 的图像为一个节点;

(b) 1 个根节点 N 分裂成 4 个节点, 计算各节点的特征点数量, 数量为零的节点直接删除, 数量为 1 的节点终止分裂并保留, 当前的节点总数若低于设定值, 特征点数量大于 1 的节点继续分裂;

(c) 特征点数量大于 1 的节点继续分裂, 如(b)。如此往复, 直到节点数量大于所需的特征点数量或者各节点的特征点数量都不大于 1 时终止; 当面对同时存在多个特征点数量大于 1 的节点, 优先分裂特征点数量最多的节点, 然后分裂特征点数量次高的节点, 依此类推, 从而保证特征密集区域也能够细分。面对特征点比较少的节点, 若可能提前达到节点阈值则不再分裂, 保证充分考虑所有节点的分裂, 提高提取的准确性。

(d) 当达到节点阈值终止分裂, 利用非极大值抑制思想将角点响应值最高的特征点作为该节点的代表特征点, 其他特征点删除, 最终得到分布均匀且足量的特征点。

利用四叉树策略筛选特征点的过程如图 2 所示。

2.2 基于 GMS 算法的误匹配剔除

ORB-SLAM3 进行特征匹配通常使用暴力匹配算法, 该方法通过求解描述子之间的汉明距离寻找距离最近的匹配点。ORB-SLAM3 算法首先通过接收最近特征点距离与第二近距离的商小于预设阈值的特征点, 即采用 ratio test 筛选相对显著的匹配, 然后在以该特征点为中心的一定区域内选择候选

匹配特征点, 将旋转角度置于长度为 30 的直方图, 最后依据角度一致性剔除角度差大于 12° 的外点, 从而确定相对“正确”的匹配。然而正由于角度差的限制, ORB-SLAM3 算法无法剔除角度差低于 12° 的错误匹配或者应用于相机视角变化较大的场景。因此为了提升 ORB 特征点匹配效率, 本文采用 GMS 算法剔除误匹配。

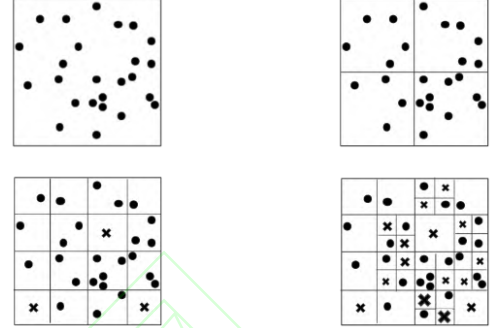


图 2 四叉树筛选特征点流程

Fig. 2 Quad tree screening feature point process

网格运动统计算法通过统计正确和错误匹配邻域的可区分度来筛选正确匹配和剔除误匹配。该方法主要基于假设匹配图像的所有像素点都具有运动平滑性, 正确匹配周围的匹配大多正确, 因此会集中在三维空间的同一区域, 同样错误匹配周围的匹配会出现在三维空间的不同位置, 不会集中在同一个区域, 即正确匹配周围会有较多的匹配去支持它, 而错误的匹配周围支持它的匹配很少。

本文改进算法通过在 ORB-SLAM3 的特征匹配环节引入 GMS 算法来实现同时降低匹配的错误率和耗时, 该算法利用运动平滑性约束原理, 通过统计特征点邻域内的匹配点数量筛选正确匹配。此外, 网格化图片使用 3×3 网格代替原本圆形的中心点邻域, 从而满足 SLAM 算法的高实时性要求, 由于 GMS 算法忽视了置信度会随着距离变化而变化, 改进算法选择根据网格与中心网格的距离对其进行权重赋值来实现公平计数。

假设存在两张待匹配图像 $\{I_a, I_b\}$, 其特征匹配数分别为 N 和 M , 图 I_a 到图 I_b 的最近邻匹配定义为 $C = \{c_1, c_2, \dots, c_i\}$, 其中 c_i 表示像素点 p_i 与 q_i 的匹配; 定义 c_i 的邻域 N_i 和相似邻域 S_i 为

$$N_i = \{c_j | c_j \in C, c_j \neq c_i, d(p_i, p_j)\} \quad (2)$$

$$S_i = \{c_j | c_j \in N_i, d(q_i, q_j) < r_2\} \quad (3)$$

其中: d 是两点之间的欧式距离。

如图 3 所示, 按照统计学, 设 ORB 正确匹配概率为 t , 匹配点对应邻域中的特征点数量为 n 和 m , 针对空间中同一区域, 正确匹配被该邻域窗口匹配支持的概率为 p_t , 错误匹配被该邻域窗口匹配支持的概率为 p_f 。

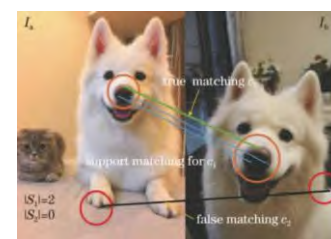


图 3 GMS 算法示意图

Fig. 3 Schematic representation of GMS algorithm

$$p_t = t + (1-t)\beta m / M \quad (4)$$

$$p_f = \beta(1-t)m / M \quad (5)$$

其中: β 是一个自适应参数。假设特征点独立分布, 根据式 (4) 和 (5), 可用二项分布估计 S_i 中匹配对数 $|S_i|$ 。

$$|S_i| = \begin{cases} B(n, p_t), & \text{if } c_i \text{ is true} \\ B(n, p_f), & \text{if } c_i \text{ is false} \end{cases} \quad (6)$$

通过分析得到 $|S_i|$ 期望值和方差分别为

$$E_{|S_i|} = \begin{cases} E_t = n \cdot p_t, & \text{if } c_i \text{ is true} \\ E_f = n \cdot p_f, & \text{if } c_i \text{ is false} \end{cases} \quad (7)$$

$$V_{|S_i|} = \begin{cases} V_t = n \cdot p_t \cdot (1-p_t), & \text{if } c_i \text{ is true} \\ V_f = n \cdot p_f \cdot (1-p_f), & \text{if } c_i \text{ is false} \end{cases} \quad (8)$$

通过设定阈值, 可以借助 $|S_i|$ 的期望和方差来衡量正确和错误匹配的可区分度, 如式 (9) 所示。

$$P = \frac{|E_t - E_f|}{\sqrt{V_t} + \sqrt{V_f}} = \frac{n \times (p_t - p_f)}{\sqrt{n \cdot p_t \cdot (1-p_t)} + \sqrt{n \cdot p_f \cdot (1-p_f)}} \quad (9)$$

其中: $P \propto \sqrt{n}$ 。利用不同的 P 值可以筛选粗匹配特征点, 根据 $n \rightarrow \infty$ 时, $P \rightarrow \infty$, 因此通过增加特征点检测数量可以获得图像中的正确匹配对。GMS 算法利用对 $|S_i|$ 设置阈值 τ_i 判断 c_i 是正确或错误匹配, 判断函数为

$$c_i \in \begin{cases} T, & \text{if } |S_i| > \tau_i \\ F, & \text{otherwise} \end{cases} \quad (10)$$

其中: T 表示正确匹配, F 表示错误匹配。

$$\tau_i = \delta \sqrt{n} \quad (11)$$

其中: δ 为一个超参数。通常 δ 值较小时会降低正确和错误匹配的可区分度; δ 值较大时则会剔除大量正确匹配, 减少最终匹配数。

利用基于划分网格的算法可以减少以上算法的耗时, 将两幅待匹配图像划分成若干小网格, c_i 表示网格 a 和 b 中的匹配对, 如图 4 所示使用一条实线表示。则 c_i 的邻域 (网格 a 中的匹配) N_i 和相似邻域 S_i 可以重新定义为

$$N_i = \{c_j | c_j \in C_a, c_i \neq c_j\} \quad (12)$$

$$S_i = \{c_j | c_j \in C_{ab}, c_i \neq c_j\} \quad (13)$$

其中: C_a 表示网格 a 中的匹配对, C_{ab} 表示同时在网格 a 和 b 中的匹配对。

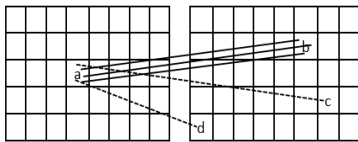


图 4 基于网格的框架

Fig. 4 Framework based on Grid

GMS 算法定义的邻域是指在同一网格中的匹配, 相似邻域则指同时在两个网格中的匹配, 即区对。网格的大小对该算法的性能具有一定的影响, 网格若过小容易忽略邻域信息, 过大则无法确保只包含正确匹配对, 合适的网格大小既能保证不忽略过多邻域信息, 又能保证所包含的匹配对尽可能正确。此外, 可以采用运动内核来解决无法包含更多邻域信息的问题, 并且为满足算法实时性, 设置 3×3 网格, 将原本邻域内匹配点问题进行简化, 转换为网格的对应匹配问题, 如图 5 所示。

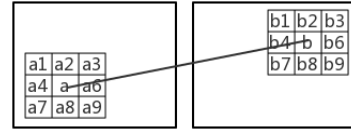


图 5 运动核

Fig. 5 Movement nuclei

设置网格并采用运动内核思路, 定义网格的邻域为它周围的 8 个网格, 在其基础上可以构建一个更大的网格 ($C_{a1b1}, \dots, C_{a9b9}$), 由此重新定义邻域和相似邻域为

$$N_i = \{c_j | c_j \in C_a, c_i \neq c_j\} \quad (14)$$

$$S_i = \{c_j | c_j \in C_{AB}, c_i \neq c_j\} \quad (15)$$

其中: $C_{AB} = C_{a1b1} \cup C_{a2b2} \cup \dots \cup C_{a9b9}$ 。

面对匹配过程中对旋转不变性与尺度不变性的限制以及重复统计网格分界线上的特征点, 本文提出的改进型特征匹配方法采用以下方法弥补缺陷, 一是在 x 、 y 轴上同时迁移划分网格的尺寸单位, 重复计算并在实际统计中忽略分界线上的特征点, 从而不重复统计分界线上的特征点; 二是针对旋转问题, 利用旋转运动核模拟不同方向的旋转, 如图 6 所示, 固定 C_a, C_b 顺时针旋转后可以得到 8 个运动核, 然后在所有运动核上使用 GMS 算法并作出最佳选择, 即选择匹配数量最多的; 三是通过比例系数改变被匹配图像的网格大小, 重复 GMS 算法对特征点进行筛选, 以解决尺度不变性的问题。

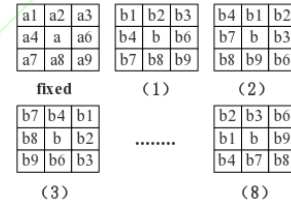


图 6 旋转运动核

Fig. 6 Rotating moving nucleus

综上, 改进型 ORB-SLAM3-GMS 算法的总体流程如图 7 所示。

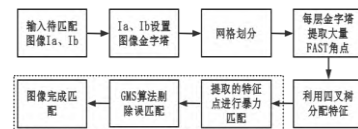


图 7 改进型 ORB 算法总流程

Fig. 7 Procedure of improved ORB algorithm

3 改进算法的稠密点云地图构建

ORB-SLAM3 作为常用视觉 SLAM 系统之一, 虽然具有强鲁棒性和较强性能, 但只能构建三维稀疏特征地图, 若使机器人能够完美完成导航、避障以及路径规划等任务, 必须改进 ORB-SLAM3 系统使其能够实现三维稠密建图。本文在上述改进型 ORB-SLAM3-GMS 算法中新增稠密建图线程 (如图 1 虚线框内所示), 结构如图 8 所示。

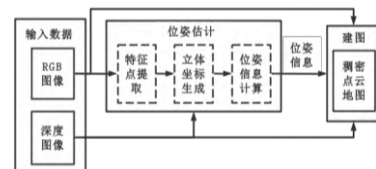


图 8 稠密建图线程结构图

Fig. 8 Dense mapping thread structure diagram

稠密建图线程的工作过程包括输入数据、位姿估计和建图。首先输入建图必需的环境源数据, 包括 RGB 图像、深度图和 ORB 特征点; 然后利用 ORB-SLAM3 为每帧符合建图需求的关键帧提供位姿估计; 最后接收图像数据和位姿信息, 并完成三维点云拼接以及点云数据管理。

生成三维点云的原理是针孔相机成像, 假设完成匹配的 RGB 图像、深度图提取出像素坐标 (u, v) 和深度值 d , W 表示深度值和实际空间距离的比例关系因子, 则数学模型为

$$W \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = T(RQ_Z + k) = \begin{bmatrix} f_x & 0 & d_x \\ 0 & f_y & d_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} + k \quad (16)$$

由于已完成相机标定, 则相机内参矩阵 T 为已知量, 此外追踪线程进行位姿估计可以确定相机的旋转矩阵 R 和平移矩阵 k , 因此最终可以计算出 RGB 图像中每个像素的空间坐标。利用 PCL 库(Point Cloud Library)对点云数据进行处理, 点云 p_i 坐标计算公式如下:

$$\begin{bmatrix} p_{i,z} \\ p_{i,x} \\ p_{i,y} \end{bmatrix} = \begin{bmatrix} d/s \\ (u-d_x)*p_{i,z}/f_x \\ (u-d_y)*p_{i,x}/f_y \end{bmatrix} \quad (17)$$

其中, s 是深度尺度, 一般取 1000。得到空间坐标后利用 RGB 图像色彩信息对生成的点云图上色, 然后读取新增的关键帧, 与生成的地图进行点云拼接和全局优化, 最终实现稠密地图的构建。

4 仿真实验及分析

4.1 实验环境

为验证改进算法的有效性和优越性, 通过与不同传统方案开展仿真对比实验。本文仿真实验平台是搭载 i3-10100 处理器, 16G 内存的台式电脑, 安装有 Ubuntu18.04 版本 Linux 操作系统, 并使用 TUM 标准 RGBD 数据集开展大量性能测试

和三维建图仿真实验。

4.2 改进算法图像提取与匹配仿真实验

对比特征点提取性能的优劣主要围绕耗时与分布情况, 而对比特征匹配性能的优劣主要参照匹配速度、匹配对数与匹配正确率, 本文采用相同图像帧进行特征点提取的对比实验, 图 9 为 ORB 算法与 ORB-SLAM3 算法提取数量为 1000 的特征点的效果。



(a)ORB 算法

(b)ORB-SLAM3 算法

图 9 不同算法特征点提取效果

Fig. 9 Points extraction effect of different algorithms

如图 9(a)所示, 传统 ORB 算法所提取到的特征点大多数堆积在纹理相对丰富的桌面上的物体上, 而在纹理相对较弱的地板和桌面却无法提取足量的特征点, 甚至部分区域一个特征点也没有提取出, 从而浪费了图片大部分区域的信息, 而 ORB-SLAM3 算法在特征点提取过程中引入四叉树策略, 在提取的特征点数量相同的条件下, 如图 9(b)所示, 所有的特征点相较于 ORB 算法不仅分布均匀, 地板和桌面等纹理比较弱的区域也提取出比较多的特征点, 使图片的有用信息被最大程度地保留。

特征匹配的对比实验需要提取两帧相近的图像, 为提高实验真实性和准确性, 从 7 组 Handheld SLAM 类别的 RGB-D 数据集中提取图像帧并从匹配对数、匹配耗时和匹配率三方面进行不同算法的对比。不同算法特征匹配效果对比如表 1 所示。

表 1 不同算法特征匹配效果对比

Tab. 1 Comparison of the feature matching effects of different algorithms

	匹配对数			匹配耗时/ms			匹配率/%		
	ORB-SLAM3	GMS	OURS	ORB-SLAM3	GMS	OURS	ORB-SLAM3	GMS	OURS
fr1/desk	169	354	371	8.220	73.816	5.782	16.9	35.4	37.1
f1/roo	369	554	586	2.676	115.809	1.368	36.9	55.4	58.6
f1/flo	294	408	419	3.157	66.851	1.982	29.4	40.8	41.9
f2/kidnap	368	524	530	7.070	84.549	4.946	36.8	52.4	53
f2/dsek	445	554	586	6.657	94.301	6.363	44.5	55.4	58.6
f2/l-no-l	365	644	665	6.833	88.014	6.066	36.5	64.4	66.5
f3/l-office-h	86	254	291	5.996	73.749	5.796	8.6	25.4	29.1
平均	299.429	470.286	492.571	5.801	85.298	4.615	29.943	47.029	49.257

由表 1 可知, 改进算法的匹配数比 ORB-SLAM3 增加 64.5%, 比 GMS 算法增加 4.7%, 匹配耗时比 ORB-SLAM3 减少 20.4%, 比 GMS 算法减少 94.6%, 明显降低了大幅度增加匹配对数情况下的匹配时间成本, 从而验证了其在特征点提取与匹配方面的优越性。

此外, 在本文 3.2 小节提到 ORB-SLAM3 算法无法剔除角度差低于 12° 的错误匹配或者应用于相机视角变化较大的场景, 因此从 f3/l-office-h 数据集中选取相机分别在办公椅的侧后方和正后方进行拍摄的图像帧, 如图 10 所示。其中侧后方可以明显看到垃圾桶证明相机视角变化大, 经过特征匹配

实验, 其中 ORB-SLAM3 只有 86 对匹配点, 表现并不佳, 而改进算法增加了 2.4 倍, 说明改进算法可以有效解决 ORB-SLAM3 在相机视角变化较大的场景鲁棒性低的问题。

4.3 改进算法精度仿真实验

定位精度指估计轨迹与真实轨迹接近程度, SLAM 算法常用性能评估指标包括绝对轨迹误差(Absolute Trajectory Error, ATE)和相对轨迹误差(Relative Pose Error, RPE), 其中绝对轨迹误差指相机真实位姿和估计位姿的差值, 其与定位精度存在反比关系, 可以评估定位精度, 第 n 帧 ATE 模型如下:

$$F_n = Q_n^{-1} SP_n \quad (18)$$

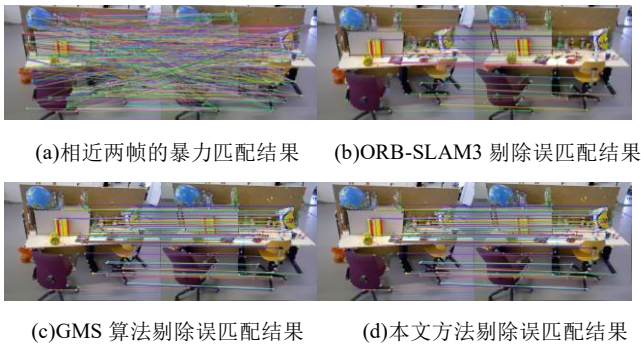


图 10 不同算法筛选暴力匹配误匹配效果

Fig. 10 Different algorithms screen BF match

ATE 的数值大小可以使用均方根误差(RMSE)统计:

$$RMSE(F_{tr}, \Delta) = \left(\frac{1}{m} \sum_{n=1}^m \|\text{trans}(F_n)\|^2 \right)^{\frac{1}{2}} \quad (19)$$

表 2 改进算法与 ORB-SLAM3 的 ATE 对比

Tab. 2 Comparison of absolute trajectory error between OURS and ORB-SLAM3

	长度/m	平均误差(mean)		均方根误差(rmse)		误差平方和(sse)		标准差(std)	
		ORB-SLAM3	OURS	ORB-SLAM3	OURS	ORB-SLAM3	OURS	ORB-SLAM3	OURS
fr1/360	5.818	0.2112	0.1525	0.2541	0.1769	48.0308	23.2905	0.1412	0.0896
fr1/desk	9.263	0.0156	0.0140	0.0186	0.0179	0.1983	0.1838	0.0102	0.0111
f1/room	15.989	0.0638	0.0578	0.0779	0.0686	8.1958	6.3788	0.0446	0.0370
f1/floor	12.569	0.0589	0.0454	0.1063	0.0620	13.5814	4.6169	0.0885	0.0422
f2/kidnap	14.286	0.2080	0.2162	0.3882	0.3943	198.9469	205.2974	0.3278	0.3297
f2/dsek	18.880	0.0187	0.0175	0.0196	0.0196	0.8392	0.8354	0.0060	0.0086
f2/l-no-l	26.086	0.6428	0.5506	0.6742	0.5773	295.0412	216.3068	0.2036	0.1734
fr2/l-with-l	39.111	0.2463	0.1377	0.2517	0.1517	76.7289	27.8896	0.0518	0.0637
f3/l-office-h	21.455	0.0093	0.0106	0.0101	0.0121	0.2575	0.3697	0.0042	0.0059
平均	-	0.1638	0.1335	0.2001	0.1644	71.3133	53.9076	0.0975	0.0845

为了进一步验证改进算法的改进型特征匹配对提高定位精度的有效性, 在 5 组不同 TUM 数据集下与同样通过采用提高匹配准确度来提高精度的 GCN v2^[27,28]、FAST ORB-SLAM^[29]以及 PL-SLAM^[30]算法进行对比, 表 3 为四种算法的 ATE 的均方根误差(RMSE)对比结果, 其中 GCN v2、FAST ORB-SLAM 和 PL-SLAM 算法的相关数据分别取自文献[28~30], 根据表 3 可以进一步确定改进算法在总体上实现了更高的定位精度。

表 3 不同方案在 TUM 数据集下的 RMSE(m)对比

Tab. 3 RMSE comparison between different algorithms on the TUM				
	GCN v2	FAST ORB-SLAM	PL-SLAM	OURS
fr1/xyz	---	0.010	0.012	0.010
fr1/desk	0.031	0.025	---	0.017
f2/dsek	---	0.009	---	0.019
fr2/l-with-l	---	0.181	---	0.151
f3/l-office-h	0.021	0.018	0.019	0.012

注: --表示文献中没有相应值, 斜体加粗表示误差最小。

4.4 改进算法稠密建图仿真实验

ORB-SLAM3 生成的稀疏地图不便于人工观测, 生成地图所保存的路标点难于被理解, 而且无法满足机器人进行自动导航和避障的需要, 故此, 改进算法通过在 ORB-SLAM3 系统上增加稠密点云构建线程, 利用 RGB-D 相机获取的深度信息来生成三维的彩色空间点云以改善上述现象。根据估计轨迹与真实轨迹重合度的不同, 数据来源选择 TUM 数据

ATE 包含旋转误差和平移误差, 式中 $\text{trans}(F_n)$ 为 APE 中的平移部分, 旋转误差求取公式同理, 但只利用平移误差对比即可。本文采用 9 组 RGB-D 标准数据集作为模拟环境, 对改进算法及 ORB-SLAM3 算法进行评估, 表 2 为两种方案在不同 TUM 数据集下的 ATE 对比。

根据表 2 改进算法与 ORB-SLAM3 在不同数据集下的 ATE 对比可以得知, fr1/desk、f1/room、f2/kidnap 和 f3/l-office-h 四组数据集下两种方法的绝对轨迹误差数值变化量都小于 0.01m, 表明改进算法与 ORB-SLAM3 的定位精度相同, 没有发挥其优势, 其他数据集中改进算法的绝对轨迹误差全面降低, 表明改进算法的定位精度得到提高。按平均表现看, 改进算法的绝对轨迹误差平均减少 0.0375m, 说明其相较于 ORB-SLAM3 的定位精度提高 3.75%, 从而验证了本文改进算法在总体上提高定位精度的可行性。

集中的 fr1/room、fr2/l-with-l 和 fr3/l-office-h 图像序列。表 4 为 3 组 RGB-D 数据集构建稠密地图的参数, 图 11 是 3 组数据集下的稠密建图效果及相机估计的轨迹, 从图中可以看到稠密点云地图可以清晰可见的显示实际环境信息, 能够较真实的恢复场景的三维结构。构建稠密地图为接下来转换为八叉树地图并应用于移动机器人自动导航和避障奠定了基础。

表 4 稠密建图计算数据

Tab. 4 Dense construction map calculation data

	Keyframe	Length/m	Nums
fr1/room	203	20.04	6431351
fr2/l-with	915	109.104	28573613
fr3/l-office	198	23.752	7651540

5 结束语

本文针对传统 ORB 算法提取的特征点分布不均匀、误匹配率高的问题, 提出一种基于 ORB-SLAM3 的改进型特征匹配策略, 并增加稠密建图线程来实现稠密地图的构建, 具体内容和结论如下:

a)提出了一种基于四叉树的特征点提取方法, 与传统 ORB 算法对比, 提取的特征点分布更加均匀, 使图像的有用信息得到充分利用。

b)提出了一种基于 GMS 和 ORB-SLAM3 的误匹配剔除方法, 不仅使匹配对数相较 ORB-SLAM3 和 GMS 算法大幅增加, 匹配耗时也明显低于两种原始算法, 在大幅增加匹配

对数的同时也明显降低了匹配时间, 验证了改进算法在特征点提取与匹配方面的优越性。

c)通过 TUM 数据集进行仿真实验验证了算法的优异性。实验结果表明, 改进算法相较于 ORB-SLAM3 的定位精度提高了 3.75%, 证明了其在总体上提高定位精度的可行性。

d)改进算法在 ORB-SLAM3 系统添加基于 PCL 库的稠密构建线程来构建稠密地图, 建立的稠密点云地图既清晰显示了实际环境信息, 也较真实的恢复了场景的三维结构, 为稠密地图接下来转换为八叉树地图并应用于机器人自动导航和避障奠定了基础。

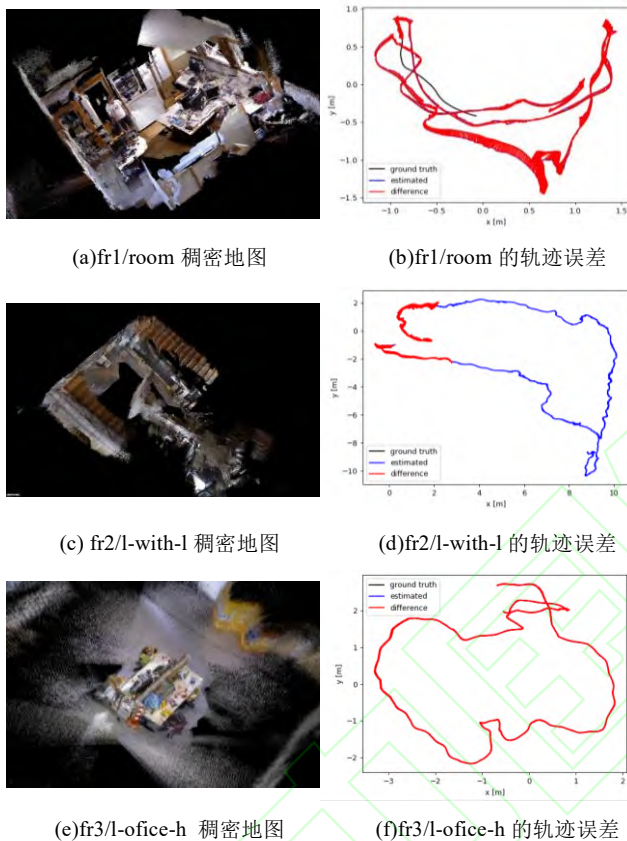


图 11 三组数据集下的稠密建图效果及相机估计的误差

Fig. 11 Effects of dense mapping and camera estimation error in the three datasets

参考文献:

- [1] Hugh D W, Tim B. Simultaneous localization and mapping: part I [C]// Proc of IEEE Robotics & Automation Magazine, 2006, 13 (2): 99-110.
- [2] Tim B, Hugh D W. Simultaneous localization and mapping (SLAM): part II [C]// Proc of IEEE Robotics & Automation Magazine, 2006, 13 (3): 108-117.
- [3] 何俊学, 李战明. 基于视觉的同时定位与地图构建方法综述 [J]. 计算机应用研究, 2010, 27 (8): 2839-2844. (He Junxue, Li Zhanming. Overview of vision-based simultaneous localization and map construction methods [J]. Computer Application Research, 2010, 27 (8): 2839-2844.)
- [4] 陈常, 朱华, 由绍泽. 基于视觉的同时定位与地图构建的研究进展 [J]. 计算机应用研究, 2018, 35 (03): 641-647. (Chen Chang, Zhu Hua, You ShaoZe. Progress in simultaneous localization and map construction base on vision [J]. Computer Application Research, 2018, 35 (03): 641-647.)
- [5] 席志红, 王洪旭, 韩双全. 基于 ORB-SLAM2 系统的快速误匹配剔除算法与地图构建 [J]. 计算机应用, 2020, 40 (11): 3289-3294. (Xi Zhihong, Wang Hongxu, Han Shuangquan. Fast mismatch exclusion algorithm and map construction based on the ORB-SLAM2 system [J]. Computer Applications, 2020, 40 (11): 3289-3294.)
- [6] Davison. Real-time simultaneous localisation and mapping with a single camera [C]// Proc of the 9th IEEE International Conference on Computer Vision, 2003, 1403-1410.
- [7] Georg K, David M. Parallel Tracking and Mapping for Small AR Workspaces [C]// Proc of the 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, 2007, 225-234.
- [8] Quan Meixiang, Piao Songhao, Tan Minglang, *et al.* Accurate monocular visual-inertial SLAM using a map-assisted EKF approach [J]. IEEE Access, 2017, 7: 34289-34300.
- [9] Robert S, Pantelis E, Matt G, *et al.* Design and analysis of a framework for real-time vision-based SLAM using Rao-Blackwellised particle filters [C]// Proc of the 3rd Canadian Conference on Computer and Robot Vision, Quebec, Canada, 2006: 21-21.
- [10] Luo Jingwen, Qin Shiyin. A Fast Algorithm of SLAM Based on Combinatorial Interval Filters [J]. IEEE Access, 2018, 6: 28174-28192.
- [11] 陈世浪, 吴俊君. 基于 RGB-D 相机的 SLAM 技术研究综述 [J]. 计算机工程与应用, 2019, 55 (7): 34-44+131. (Chen Shilang, Wu Junjun. Review of SLAM technology research based on RGB-D camera [J]. Computer Engineering and Application, 2019, 55 (7): 34-44+131.)
- [12] Raul M A, Juan D T. ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras [J]. IEEE, 2017 (5): 1255-1262.
- [13] Anastasios I M, Stergios I R. A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation [C]// Proc of 2007 IEEE International Conference on Robotics and Automation, Rome, Italy, 2007: 3565-3572.
- [14] Leutenegger S, Lynen S, Bosse M, *et al.* Keyframe-based visual-inertial odometry using nonlinear optimization [J]. International Journal of Robotics Research, 2014, 34 (3): 314-334.
- [15] Wisth D, Camurri M, Das S, *et al.* Unified Multi-Modal Landmark Tracking for Tightly Coupled Lidar-Visual-Inertial Odometry [J]. IEEE Robotics and Automation Letters, 2021, (99): 1004-1011.
- [16] Carlos C, Richard E, Juan J G R, *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.
- [17] 张昊宇, 柳祥乐, 王思山. 基于深度学习特征提取的改进 ORB-SLAM3 算法 [J]. 湖北汽车工业学院学报, 2022, 36 (02): 55-59. (Zhang Haoyu, Liu Xiangle, Wang Shishan. Improved ORB-SLAM3 algorithm based on deep learning feature extraction [J]. Journal of Hubei Institute of Automotive Industry, 2022, 36 (02): 55-59.)
- [18] Tang Jiexiong, John F, Patric J. Geometric Correspondence Network for Camera Motion Estimation [J]. IEEE Robotics and Automation Letters, 2018, 3 (2): 1010-1017.
- [19] Xue Xiujie, Lv Xiafu. Stereo Visual Inertial SLAM Algorithm Fusing Point and Line Features [C]// Proc of 2022 China Automation Congress (CAC), Xiamen, China, 2022: 1687-1692.
- [20] Hu Zhangfang, Chen Jian, Luo Yuan, *et al.* STDC-SLAM: A Real-Time Semantic SLAM Detect Object by Short-Term Dense Concatenate Network [J]. IEEE Access, 2022 (10): 129419-129428, 2022.
- [21] Berta B, Jose M F, Javier C, *et al.* DynaSLAM: Tracking, Mapping and Inpainting in Dynamic Scenes [J]. IEEE Robotics and Automation Letters, 2018, 3 (4), 4076-4083.
- [22] Long Xudong, Zhang Weiwei, Zhao Bo. PSPNet-SLAM: A Semantic

- SLAM Detect Dynamic Object by Pyramid Scene Parsing Network [J]. IEEE Access, 2020 (8): 214685-214695.
- [23] Lv Kefeng, Zhang Yongsheng, Yu Ying, *et al.* SIIS-SLAM: A Vision SLAM Based on Sequential Image Instance Segmentation [J]. IEEE Access, 2023 (11): 17430-17440.
- [24] Bian Jiawang, Lin Wenyan, Yasuyuki M, *et al.* Gms: Grid-based motion statistics for fast, ultra-robust feature correspondence [C]// Proc of IEEE CVPR, 2017: 4181-4190.
- [25] Rosten E, Drummond T. Machine learning for high-speed corner detection [C]// Proc of ECCV. Springer, Berlin, Heidelberg, 2006: 430-443.
- [26] Calonder M, Lepetit V, Strecha C, *et al.* Brief: Binary robust independent elementary features [C]// Proc of ECCV. Springer, Berlin, Heidelberg, 2010: 778-792.
- [27] Tang Jiexiong, Ludvig E, John F, *et al.* GCNv2: Efficient Correspondence Prediction for Real-Time SLAM [J]. IEEE Robotics and Automation Letters, 2019, 10 (4): 3505-3512.
- [28] Fu Qiang, Yu Hongshan, Wang Xiaolong, *et al.* Fast ORB-SLAM without keypoint descriptors [J]. IEEE Trans on Image Processing, 2021, 31: 1433-1446.
- [29] Ruben G O, Francisco A M, David Z N, *et al.* PL-SLAM: A stereo SLAM system through the combination of points and line segments [J]. IEEE Trans on Robotics, 2019, 35 (3): 734-746.
- [30] 伍晓东, 张松柏, 汤连荣, 等. 基于改进关键帧选择的 ORB-SLAM3 算法 [J/OL]. 计算机应用研究: 1-7 [2023-03-06]. (Wu Xiaodong, Zhang Songbai, Tang Shirong, *et al.* ORB-SLAM3 algorithm based on improved keyframe selection [J/OL]. Computer Application Research: 1-7 [2023-03-06].)
- [31] 潘峰, 沈建新, 秦顺, 等. 融合 GMS 的 ORB 特征点提取与匹配算法 [J]. 计算机工程与设计, 2022, 43 (08): 2244-2251. (Pan Feng, Shen Jianxin, Qin Shun, *et al.* The ORB feature point extraction was fused to GMS with the matching algorithm [J]. Computer Engineering and Design, 2022, 43 (08): 2244-2251.)
- [32] 袁俊鹏. 基于 ORB-GMS 的视觉惯导 SLAM 研究 [D]. 西南交通大学, 2021. (Yuan Junpeng. SLAM based on ORB-GMS [D]. Southwest Jiaotong University, 2021.)
- [33] 刘洋, 陈俊, 胡诗佳, 等. 改进特征匹配环节的 ORB-SLAM 稠密建图算法 [J/OL]. 计算机工程: 1-11 [2023-02-10]. (Liu Yang, Chen Jun, Hu Shijia, *et al.* ORB-SLAM dense map algorithm [J/OL]. Computer Engineering: 1-11 [2023-02-10].)
- [34] 席志红, 王洪旭, 韩双全. 基于 ORB-SLAM2 系统的快速误匹配剔除算法与地图构建 [J]. 计算机应用, 2020, 40 (11): 3289-3294. (Xi Zhihong, Wang Hongxu, Han Shuangquan. The map algorithm based on the ORB-SLAM 2 system [J]. Computer Applications, 2020, 40 (11): 3289-3294.)
- [35] 马靖煊, 王红雨, 曹彦, 等. 面向嵌入式平台的单目 ORB-SLAM 稠密化建图实现 [J]. 计算机工程与应用, 2022, 58 (16): 213-218. (Ma Jingxuan, Wang Hongyu, Cao Yan, *et al.* Implementation of Singular ORB-SLAM for Embedded Platform [J]. Computer Engineering and Application, 2022, 58 (16): 213-218.)
- [36] 洪运志, 吴怀宇, 陈志环. 室内环境下深度相机 V-SLAM 的稠密建图 [J]. 计算机工程与设计, 2020, 41 (9): 2552-2558. (Hong Yunzhi, Wu Huaiyu, Chen Zhihuan. Dense mapping of deep camera V-SLAM in indoor environment [J]. Computer Engineering and Design, 2020, 41 (9): 2552-2558.)