

基于改进双目 ORB-SLAM3 的特征匹配算法

伞红军^{1,2} 冯金祥¹ 陈久朋^{1,2} 彭真¹ 赵龙云¹

(1. 昆明理工大学机电工程学院, 昆明 650500; 2. 云南省先进装备智能制造技术重点实验室, 昆明 650500)

摘要: 针对传统 ORB 算法在双目特征匹配阶段误匹配率高而导致无法满足高精度定位要求的问题,提出了一种基于改进双目 ORB-SLAM3 的特征匹配算法。在特征点匹配阶段引入最近邻匹配算法 (FLANN),通过设定比率阈值筛选出更为精确的匹配对,在双目 ORB-SLAM3 立体匹配中引入自适应加权 SAD-Census 算法,通过考虑像素之间的几何距离,重新计算 SAD 值并与 Census 算法相融合来提高特征匹配稳定性和精度,同时加入自适应的 SAD 窗口滑动范围进一步扩大搜索距离,进而筛选出正确的匹配来提高系统精度。在 EuRoC 数据集和真实室内场景中进行实验,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 23.32%,真实环境中提高近 50%,从而验证了改进算法可行性和有效性。

关键词: 改进双目 ORB-SLAM3; 特征匹配; 最近邻匹配算法; 自适应加权 SAD-Census 算法

中图分类号: TP242; TH112 **文献标识码:** A **文章编号:** 1000-1298(2025)05-0625-10 **OSID:**



Feature Matching Algorithm Based on Improved Binocular ORB-SLAM3

SAN Hongjun^{1,2} FENG Jinxiang¹ CHEN Jiupeng^{1,2} PENG Zhen¹ ZHAO Longyun¹

(1. Faculty of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming 650500, China

2. Yunnan Provincial Key Laboratory of Advanced Equipment Intelligent Manufacturing Technology, Kunming 650500, China)

Abstract: Aiming at the problem that the traditional ORB algorithm fails to meet the high-precision localization requirements due to the high mis-matching rate in the binocular feature matching stage, a feature matching algorithm based on the improved binocular ORB-SLAM3 is proposed. The nearest neighbor matching algorithm (FLANN) is introduced in the feature point matching stage, and more accurate matching pairs are filtered out by setting the ratio threshold, and the adaptive weighted SAD-Census algorithm is introduced in the binocular ORB-SLAM3 three-dimensional matching, and the geometric distances between the cases are taken into account to recalculate the SAD values and merge them with the Census algorithm to improve the stability and accuracy of feature matching, while the adaptive weighted SAD-Census algorithm is introduced. At the same time, the adaptive SAD window sliding range is added to further expand the search distance, so as to filter out the correct matches to improve the accuracy of the system. Experiments are carried out in the EuRoC dataset and real indoor scenes, and the results show that compared with the pre-improved ORB-SLAM3 algorithm, the localization accuracy of the improved algorithm is improved by 23.32% in the dataset, and nearly 50% in the real environment, thus verifying the feasibility and effectiveness of the improved algorithm.

Key words: improved binocular ORB-SLAM3; feature matching; nearest neighbor matching algorithm; adaptive weighted SAD-Census algorithm

0 引言

同时定位与地图构建 (Simultaneous localization and mapping, SLAM) 能在未知环境下估计地图并能

同时定位机器人位置,因此在野外搜索、太空探测和定位侦察等中应用广泛^[1-3]。根据使用传感器类型,SLAM 系统可以分为激光 SLAM 和视觉 SLAM 两大类^[4-6]。其中视觉 SLAM 因其硬件成本低、目

收稿日期: 2024-03-08 修回日期: 2024-04-14

基金项目: 云南省科技厅基础研发计划-青年基金项目 (202301AU070059) 和昆明理工大学人才培养项目 (KKZ320230104)

作者简介: 伞红军 (1976—),男,副教授,博士,主要从事机器人技术及应用研究,E-mail: sanhjun@163.com

通信作者: 陈久朋 (1993—),男,讲师,博士,主要从事机器人技术及应用研究,E-mail: 18314490225@163.com

标检测直观且提供丰富的视觉信息而备受瞩目^[7]。

尽管传统视觉 SLAM 技术取得了一定发展,但在实际环境应用中的性能仍存在一定局限性,对其算法改进成为众多学者探寻方向。MonoSLAM 具有较好的实时性,是第 1 个运用单相机的实时 VSLAM 系统^[8]。算法通过使用扩展卡尔曼滤波来估计相机运动轨迹和三维物体位置,在小规模运动中能够把误差控制在小范围内,但采用稀疏特征点容易导致特征点丢失,且缺乏全局优化和回环检测模块。为了解决 MonoSLAM 计算复杂问题,KLEIN 等^[9]提出了 PTAM,该算法对视觉里程计部分进行了优化,增加了并行化跟踪与建图步骤,通过非线性优化算法来增加定位精度。虽然该方法稳定性高于 MonoSLAM 方法,但其设计过程较为复杂且必须通过手动设计。FORSTER 等^[10]提出了 SVO 算法,该算法融合了特征点法和直接法,速度快且能更好地计算特征点位置,但无法进行回环检测和全局优化。MUR-ARTAL 等^[11-12]提出的 ORB-SLAM 和 ORB-SLAM2 具有良好纹理的连续图像序列上实现定位与地图构建。其中,ORB-SLAM 可以利用关键帧来确定相机位姿及周围环境结构;而 ORB-SLAM2 则是对 ORB-SLAM 的进一步优化,其兼容双目相机和 RGB-D 相机,并且具备 3 条独立执行线程,因为特征追踪存在积累误差,所以系统整体精度可能会降低。FERRERA 等^[13]提出了一种定位精度优于 ORB-SLAM2 的 VO2-SLAM 算法,该算法在回环检测阶段采用一种 iBowLCD 在线词袋,使得 OV2-SLAM 可以适应更多的场景,但该算法使用 looselyBA 进行优化,无法优化全部关键帧位姿和 3D 点。LI 等^[14]在 ORB-SLAM2 的框架上提出了一种基于目标跟踪和静态点检测的语义 SLAM,通过将跟踪算法与目标检测相结合以及引入深度估计来剔除动态点,该算法在动态环境中定位精度有所提升,但算法实时性有所下降。董蕊芳等^[15]提出了一种基于改进 ORB-SLAM2 的视觉导航方法,该算法通过先验多图约束进行位姿估计,并通过运动约束进行位姿优化,使得定位精度有所提升。刘庆运等^[16]提出了一种基于激光雷达与深度相机融合的 SLAM 算法,将激光雷达和深度相机提取的环境信息进行互补,能够构建出精确的栅格地图,同时提高了定位精度。丛佩超等^[17]针对果园机器人定位与建图问题对 ORB-SLAM2 算法进行了改进,通过对算法中的 FAST 角点和阈值进行优化,并结合图像金字塔法和高斯滤波算法来剔除劣质特征点,以提升特征配精度和定位精度。WANG 等^[18]提出了 Stereo-NEC 算法,该算法将陀螺仪零偏估计单独建

模作为优化问题,通过 Refine 和 3DOF-BA 等操作提高初始化精度和鲁棒性。CAMPOS 等^[19]在 ORB-SLAM2 的基础上提出了 ORB-SLAM3,增加了多地图系统和 IMU 融合算法,使得 ORB-SLAM3 能够更有效地利用数据关联进行系统漂移重置和回环校正,从而显著提高了算法鲁棒性。虽然 ORB-SLAM3 通过融合 IMU 提高了定位精度使其算法性能优于 ORB-SLAM2,但传统 ORB 算法误匹配率高问题依然存在,从而影响系统定位精度。

本文以双目 ORB-SLAM3 系统框架为基础,在粗匹配阶段加入最近邻匹配算法 (FLANN)^[20],以减少误匹配。在精匹配阶段对原有的 SAD 算法进行改进,加入基于自适应权重 SAD 算法和 Census 算法相融合的立体匹配算法^[21],通过在匹配过程中自适应地调整每个像素点权重,使位于中心附近像素点具有更高的参考价值,同时使用自适应的 SAD 滑动窗口范围代替固定值滑动窗口范围,以降低固定滑动窗口范围所带来的误匹配,提高双目 ORB-SLAM3 的特征匹配精度。

1 传统 ORB-SLAM3 双目立体匹配原理

ORB-SLAM3 系统由 3 个并行运行的线程构成,其中跟踪线程负责实时追踪相机运动并生成关键帧。局部建图线程主要负责优化局部地图中的关键帧位姿和地图点,并通过预处理新加入的关键帧和剔除冗余信息,确保地图准确性和高效性。回环检测与地图融合线程,主要负责检测机器人回环运动并将地图进行全局优化,从而提升整个系统稳定性和精度。3 个线程相互协作,共同实现 ORB-SLAM3 系统功能。整体结构框架如图 1 所示。

在 ORB-SLAM3 跟踪线程中,特征提取和匹配是 SLAM 系统中至关重要的步骤,其中特征匹配准确率将直接影响系统定位精度。而在 ORB-SLAM3 中不同相机模式下的特征匹配过程会有所不同,双目模式下特征匹配^[22]过程为:

(1)行特征点统计。对于图 2 左图中极线上投影像素点,图 2 右图中搜索纵坐标范围为 $r_{\min} \sim r_{\max}$ 。为了后续描述方便,对图 2 左半部分称为左图,图 2 右半部分称为右图。以图像金字塔尺度作为偏移量,在当前点上下正负偏移量 r 内纵坐标都认为是匹配点可能存在的行数。意味着左图中一个特征点在右图中可能对应多行中候选特征点,而不是仅仅局限于一行。原因为避免由于极线校正后可能存在的误差而导致漏匹配情况。

(2)粗匹配。在立体校正和行特征点统计后,左图每个特征点都对应右图中相同行上的一组候选

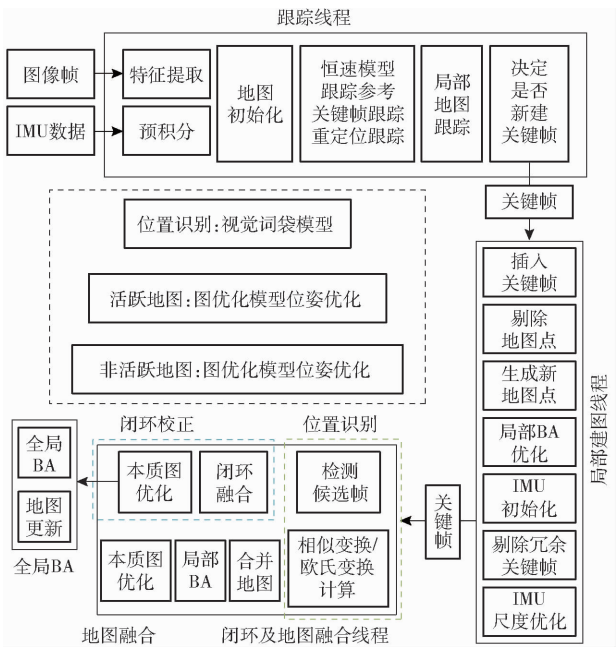


图 1 ORB-SLAM3 算法流程图

Fig. 1 Flowchart of ORB-SLAM3 algorithm

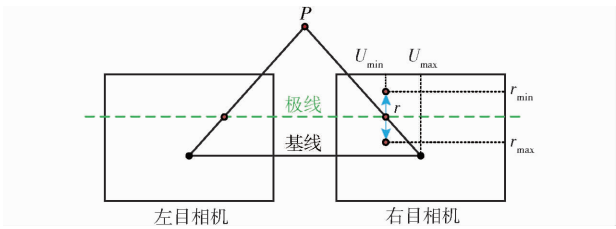


图 2 匹配点的上下偏移量

Fig. 2 Upper and lower offsets of matching points

匹配点。为了找到最佳粗匹配点，只需将左图上的每个特征点与右图中对应行上的候选匹配点进行逐个比较计算描述子之间的距离，选择距离最小的点作为最佳粗匹配点。搜索时不用搜索整行像素的横坐标，根据三维点距离范围可以将横坐标搜索范围限制在 $U_{\min} \sim U_{\max}$ 。其中 $U_{\max}$ 对应的是三维点位于无穷远处，视差为 0 时的横坐标；而 $U_{\min}$ 对应的是三维点位于最近距离（假设是相机基线距离）时的横坐标。

(3)精匹配。精匹配是在粗匹配的基础上，利用 SAD 算法对已得到的匹配对进行进一步验证和筛选，以实现更精确的匹配结果。SAD 是立体匹配中常用的块匹配算法，其基本原理是计算两幅图像中对应像素值的绝对差值之和。其原理如图 3 所示。

以图 3a 为参考图，图 3b 为待匹配图像。设定 SAD 窗口尺寸和移动距离 d ，在图 3a 中选择一个待匹配特征点位置 (x, y) ，以该点为中心构建 SAD 窗口区域 $region_l$ 。在图 3b 中，选取匹配的起始点，位置为 (m, n) ，并以该起始点位置为中心构建相同尺

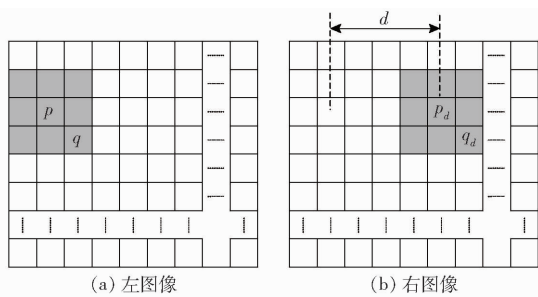


图 3 SAD 算法原理图

Fig. 3 Schematics of SAD algorithm

寸的 SAD 窗口区域 $region_r$ 。然后计算 $region_r$ 和 $region_l$ 内像素值之差的绝对值之和。沿着图 3b 行方向移动 SAD 窗口，并更新 $region_r$ 。在每次移动后，将更新后的 $region_r$ 与 $region_l$ 进行比较，并得出新的计算值。重复这个过程，直到移动了指定距离 d ，在得到的所有差值之和中，找到最小值所对应的位置，即为图 3a 中特征点的最佳匹配位置，此位置所对应的 d 即为视差值。对每个特征点重复以上步骤，即可完成两幅图像的精确匹配。

SAD 算法计算式为

$$C_{SAD} = \sum_{i,j \in W} |I_L(x+i, y+j) - I_R(m+i+d, n+j)| \quad (1)$$

式中 C_{SAD} ——像素点差绝对值和

$I_L(x, y)$ ——左图像中位于坐标 (x, y) 的特征点像素灰度

$I_R(m, n)$ ——在粗匹配阶段找到的与左图中 (x, y) 点相匹配的特征点在右图像中位置 (m, n) 的像素灰度

W ——SAD 窗口尺寸

(4)亚像素插值法。左、右图像在经过粗匹配和精匹配后，得到的匹配像素坐标仍然是整数坐标。整数坐标在对精度要求较高的图像配准和 SLAM 等领域会产生较大误差，这时需要使用浮点类型坐标实现更精确的位置表示。在 ORB-SLAM3 代码中通过亚像素插值法即二次多项式插值得到亚像素坐标。其中亚像素坐标是指坐标误差小于 1 个像素的浮点像素坐标。

图 4 为用亚像素插值得到右图中最佳的匹配点坐标示意图。已知 3 个点 $P_1(x_2 - 1, y_1)$ ， $P_2(x_2, y_2)$ ， $P_3(x_2 + 1, y_3)$ ，其中 P_2 横坐标 x_2 为精匹配得到的匹配结果，纵坐标 y_2 为用 SAD 匹配时误差。为了找到点 P_2 附近更准确的亚像素坐标，需要借助其相邻整数坐标 $x_2 - 1, x_2 + 1$ 及其对应的 SAD 误差 y_1, y_3 来进行二次多项拟合。利用 3 个已知的点 $P_1(x_2 - 1, y_1)$ 、 $P_2(x_2, y_2)$ 、 $P_3(x_2 + 1, y_3)$ 来拟合一个开口向上的抛物线，然后计算抛物线上纵坐标值

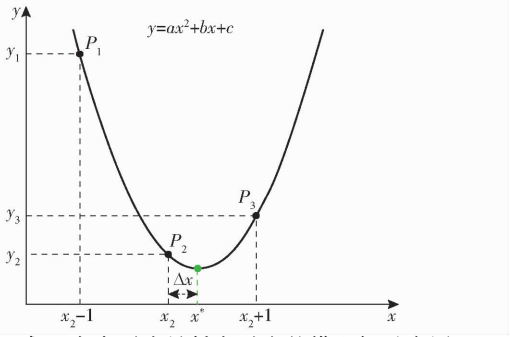


图 4 求二次多项式最低点对应的横坐标示意图
Fig. 4 Graph of horizontal coordinates corresponding to the lowest point of quadratic polynomial

最小点对应的横坐标 x^* , 其过程如下:

假设二次多项式为 $y = ax^2 + bx + c$, 将点 P_1 、 P_2 、 P_3 坐标代入多项式, 则有

$$\begin{cases} y_1 = a(x_2 - 1)^2 + b(x_2 - 1) + c \\ y_2 = ax_2^2 + bx_2 + c \\ y_3 = a(x_2 + 1)^2 + b(x_2 + 1) + c \end{cases} \quad (2)$$

展开后, 使用 $y_1 - y_2$ 和 $y_3 - y_2$ 消去 c 得

$$\begin{cases} y_1 - y_2 = a(-2x_2 + 1) - b \\ y_3 - y_2 = a(2x_2 + 1) + b \end{cases} \quad (3)$$

对式(3)中两式相加消去 b , 可得

$$y_1 + y_3 - 2y_2 = 2a \quad (4)$$

对式(3)中两式相减可得

$$y_1 - y_3 = -4ax_2 - 2b \quad (5)$$

通过式(4)、(5)可得亚像素偏移量 Δx 为

$$\Delta x = x^* - x_2 = \frac{y_1 - y_3}{2(y_1 + y_3 - 2y_2)} \quad (6)$$

根据式(6)可求得亚像素级的匹配坐标 x^* , 得到的匹配结果更加准确。

2 双目立体匹配算法改进

在双目 ORB-SLAM3 中的特征匹配阶段有粗匹配、精匹配以及亚像素法和剔除外点等步骤来提高特征匹配精度, 但提取出的匹配对中仍不可避免地存在大量误匹配。为了降低误匹配率并进一步提升匹配精度, 在 ORB-SALM3 框架上对双目立体匹配阶段进行改进, 在粗匹配阶段加入 FLANN 算法, 同时对精匹配阶段的 SAD 算法进行改进并与 Census 算法进行融合, 从而提高 ORB-SLAM3 双目特征匹配精度, 算法流程如图 5 所示。

2.1 快速近似最近邻算法

快速近似最近邻算法 (FLANN) 是一种用于高维数据的近似最近邻搜索算法。它的目标是在保持搜索质量的同时提高搜索速度。因此, 在 ORB-SLAM3 中引入 FLANN 算法可以有效提高粗匹配阶

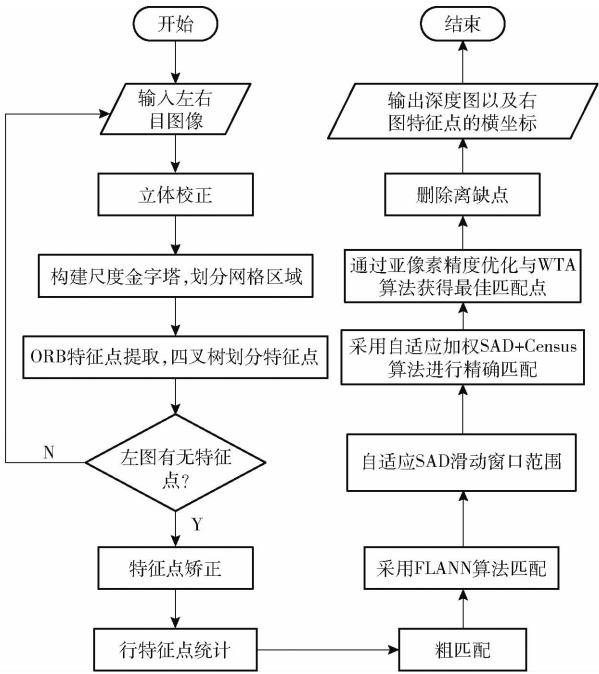


图 5 改进的立体匹配流程图
Fig. 5 Flowchart of improved stereo matching

段效率和准确性, 操作过程为: 将左图中的特征点与右图中的候选匹配点逐一比较, 并计算它们的描述子距离。从这些距离中找到最小两个距离, 记为 d_1 、 d_2 。将最小距离 d_1 除以次最小距离 d_2 得到比率 r 。如果比率 r 低于预先定义的阈值 T , 则认为这对特征点之间匹配正确。在特征空间中, 由于高维性即使两个特征点之间的距离相似, 也可能存在大量的错误匹配, 这会出现较高的 r 。通过调整比值阈值 T 可以控制匹配对数量和匹配精度, 其降低阈值会减少匹配对数量但提高稳定性, 反之亦然。通过对大量图像进行匹配, 得到最优匹配的概率密度函数 (Probability density function, PDF) 与最近邻次近邻距离比关系曲线, 如图 6 所示。

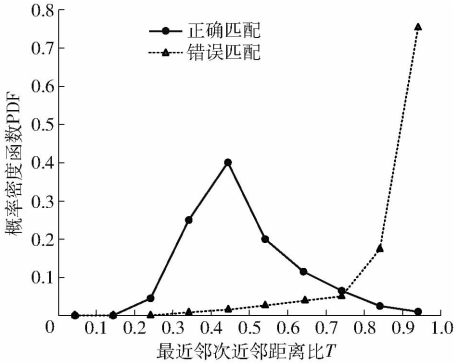


图 6 PDF 与最近邻次近邻距离比关系
Fig. 6 Relationship between PDF and nearest-neighbor next-nearest-neighbor distance ratio

由图 6 可以观察到, 当 T 为 $0.3 \sim 0.6$ 时, 匹配效果最佳。当 T 小于 0.3 时, 匹配点对数量较少; 而当 T 大于 0.6 时, 错误匹配点会明显增多。所以在

对准确率要求较高的匹配中 T 可选择为 0.4, 一般情况下可选择 T 为 0.5, 对于匹配点数目要求比较多的情况下可选择 T 为 0.6。为了保证粗匹配阶段提取的特征点对数量不会太少, 同时也保证匹配对准确率得到提升, 本文选取 T 为 0.6。

2.2 基于自适应加权的 SAD 算法

在传统 SAD 算法中, 直接累加窗口中所有像素灰度差的绝对值计算 SAD 值。表明在窗口中的每个位置, 不论其离中心点距离远近, 其贡献都被视为相等。在实际情况中, 远离中心点像素对于匹配相关性较小。为了更准确反映像素之间的相关性, 把 ORB-SLAM3 中原本的 SAD 算法替换成了基于自适应加权的 SAD 算法。此算法通过引入几何距离作为权重因子, 来调节每个像素点对中心像素点的贡献度。具体来说, 距离中心点越近的像素点将被赋予更高的权重, 因为它们对中心像素点的影响更大; 而距离中心点较远的像素点则会被赋予较低的权重, 因为它们对中心像素点的影响较小。通过考虑像素之间的几何距离, 重新计算的 SAD 更能反映出图像区域内像素之间的实际相关性来提高了匹配准确性, 进而降低噪声影响。其中权重表示为

$$G = e^{-kf^2} \quad (7)$$

$$\text{其中 } k = 200\,000e^{-\left(\frac{c+150}{40}\right)^2} \quad (8)$$

式中 f ——邻域点到中心点的间距
 k ——与 SAD 窗口尺寸相关的常数
 c ——SAD 窗口尺寸

可以看出, 权重 G 不仅受到像素间距离的影响, 而且与 SAD 窗口尺寸有直接联系。从式(7)可得, 靠近中心点像素具有更大权重 G 。

2.3 基于自适应的 SAD 滑动窗口范围

由图 3 可知, 在传统的 ORB-SLAM3 双目立体匹配中, SAD 窗口移动距离 d 固定。虽然这种方法简单易行, 但可能导致未能找到最佳匹配对。因为在右图中通过粗匹配得到的匹配点周围像素灰度比较平滑即像素灰度变化较小时, 会使得通过 SAD 算法得出的匹配对并不是最佳的, 原因为最佳匹配对可能不在固定滑动范围内, 这会导致匹配结果不准确, 进而影响后续定位精度。基于相邻且颜色相近的像素点往往具有相似视差值^[23]这一假设, 提出了一种自适应的 SAD 滑动窗口范围来替代原来的固定滑动距离。通过分析灰度变化, 可以精确地计算出像素点距离图像边缘左侧和右侧距离, 使得右图中 SAD 窗口移动距离 d 不再固定, 而是通过右图候选特征点周围的灰度变化来自适应地确定出移动距离 d , 确保在匹配点周围像素灰度变化较小的情况

下, 能够包含更多周围像素信息, 进而找到最佳匹配对使得匹配精度有所提高。

以求粗匹配中得到与左图特征点相匹配的右图特征点 $p(x, y)$ 的左侧滑动距离为例, 从点 p 开始在同行中向左逐个像素点 $p_i(x_i, y)$ 计算, 检查每一个像素点是否满足滑动距离计算公式

$$\begin{cases} |I(x, y) - I(x_i, y)| < \tau_1, |I(x_i, y) - I(x_i + 1, y)| < \tau_1 \\ D(p, p_i) < l_1 \\ |I(x, y) - I(x_i, y)| < \tau_2, l_2 < D(p, p_i) < l_1 \end{cases} \quad (9)$$

式中 $I(x, y)$ ——像素点 $p(x, y)$ 灰度
 $I(x_i, y)$ ——像素点 $p_i(x_i, y)$ 灰度
 $D(p, p_i)$ ——像素点 $p(x, y)$ 、 $p_i(x_i, y)$ 之间距离
 直到不再有任何像素点满足这些条件为止。而最后一个仍然满足式(9)的像素点与起始点 p 之间的距离即为左侧滑动距离, 求取右侧滑动距离同理。

式(9)第 1 行公式中, 在向左遍历时不仅限制像素点 $p(x, y)$ 和 $p_i(x_i, y)$ 灰度差异小于阈值 τ_1 , 而且限制像素点 $p_i(x_i, y)$ 与其相邻前 1 个像素点 $p(x_i + 1, y)$ 灰度变化也小于阈值 τ_1 , 这样滑动距离延伸就不会穿过边缘像素。而第 2 行公式和第 3 行公式则是为了使弱纹理区域包含更多的像素, 所以采用较大的滑动距离阈值 l_1 。第 3 行公式中的 $l_2 (l_2 < l_1)$ 与 $\tau_2 (\tau_2 < \tau_1)$ 的设定是为了防止颜色差异较大时滑动距离也会延长至 l_1 而设置的约束, 通过设置阈值 l_2 、 τ_2 , 从而确保只有当颜色非常相近时, 才会让滑动距离延长至 l_1 。通过上述步骤即可得到点 p 左右两侧 SAD 窗口滑动距离。

2.4 基于自适应加权的 SAD 与 Census 算法融合

Census 算法也是一种用于图像匹配和立体视觉的特征描述算法^[24]。它主要用于计算图像中每个像素周围邻域灰度的二进制编码, 以便进行匹配和视差估计。首先 Census 算法选择一个固定的窗口, 用于计算每个像素 Census 编码。对于窗口中每个像素, 将其灰度与窗口的中心灰度进行对比。如果当前点灰度小于或等于中心点灰度, 则将相应的二进制位设为 0, 否则设为 1。将所有比较结果组合成一个二进制编码, 形成 Census 编码如图 7 所示。这个编码可看作是一个二进制数字, 其中每个位表示一个像素比较结果。Census 变换公式为

$$b_{ij} = \begin{cases} 1 & (I(i, j) > I(x, y)) \\ 0 & (I(i, j) \leq I(x, y)) \end{cases} \quad (10)$$

式中 b_{ij} 为一个二进制编码, 表示比较结果; $I(x, y)$ 为中心点 (x, y) 灰度; $I(i, j)$ 为邻域点灰度。在对左右两个窗口执行 Census 变换之后, 将这两个窗口的 Census 编码逐位进行异或操作。异或运算结果, 1

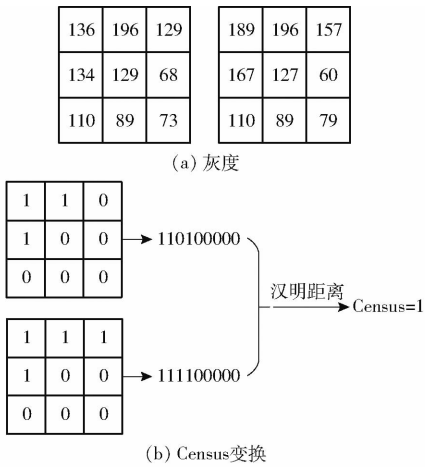


图 7 Census 变换示例

Fig. 7 Example of Census transformation

的个数即为汉明距离。距离越小,表示两个窗口(即对应的像素点)之间的相似度越高。这一过程被逐点地应用于整个图像,即对每个点寻找其相应的匹配点,以完成整体图像匹配任务。

Census 算法具有灰度不变特性,即不会因为像素的具体灰度不同而有太大变化,它更多地反映像素之间的空间结构和分布关系,使得 Census 算法对于灰度变化具有一定的鲁棒性。Census 变换在纠正由光线不均匀引起的灰度误差方面具有显著优势。然而,在图像存在显著灰度变化或具有相似纹理重复区域时,Census 变换计算出匹配代价不能准确体现像素间相似度,进而影响匹配效果,导致误匹配现象较为常见。SAD 变换通过直接计算灰度差绝对值的和来确定匹配代价,因此在图像中灰度变化明显的区域,它能够有效地反映两个窗口之间的相似程度。而基于自适应加权的 SAD 算法,通过为每个像素分配不同权重,进一步提高了匹配准确性,在理论上比传统 SAD 算法具有更好的匹配效果。

为了利用两种算法优点,将 Census 算法与基于自适应加权的 SAD 算法进行融合。先以左图中的待匹配点及右图中通过式(9)确定的左侧滑动距离起点为中心构建 $c \times c$ 的两个子窗口,在两个子窗口内计算经过改进后的 SAD 值和汉明距离,并通过对这两个值加权求和,最终得到一个全新的代价值

$$C = 2 - e^{-\frac{C_{Census}}{100}} - e^{-\frac{C_{SAD}}{500c^2}}$$

(11)

式中 C_{Census} ——汉明距离

C_{SAD} ——改进 SAD 值

在右图中沿行方向向右移动子窗口,每移动一个像素点按照式(11)计算代价值 C ,直到移动范围达到滑动距离 d 为止。在所得代价值中找出最小代价

值,其对应的右图子窗口中心点就是与左图待匹配点相匹配的最佳匹配点,并将该点坐标与左图待匹配点坐标相减即可得出视差值。由于 C_{Census} 、 C_{SAD} 这两个量会受到窗口尺寸影响,因此在公式中包含窗口尺寸,以平衡它们之间的权重。且 C_{Census} 与 C_{SAD} 的取值范围可能存在较大差异,通过取指数方式将它们缩放至 0 ~ 1,可确保计算更加可靠。

3 实验

为验证改进算法在定位精度方面的有效性和优越性,在公开数据集与实际环境中进行了验证,同时在算法评估中还将改进算法与不同算法进行仿真对比实验来验证改进算法优势,实验设备为 Turtlebot2 智能移动平台(图 8a)。此外,实验器材还包括 1 台配置为 HP OMEN 16-wf0032TX 的 PC,该 PC 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,并运行 Ubuntu 18.04 操作系统。其中视觉传感器使用的是 Intel RealSense D455,该相机的分辨率设置为 848 像素 × 480 像素,频率为 30 Hz。采用 USB 3.0 连接 PC 与相机。为了准确评估轨迹数据,采用 EVO 作为轨迹评估工具,采用激光跟踪仪(图 8b)进行数据测量。

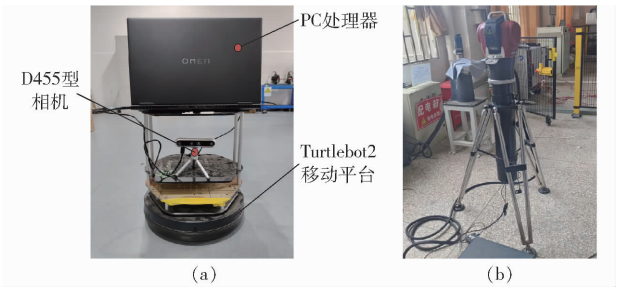


图 8 实验设备

Fig. 8 Experimental equipments

3.1 公开数据集系统实验

为体现改进算法在双目视觉特征匹配阶段优化效果,所有对比实验都是在纯双目模式下进行。由于 EuRoC 数据集^[25]中提供了双目图像序列以及用于评估 SLAM 算法的真实轨迹和位姿信息,同时它还包含在室内和室外不同场景中采集的数据,能够很好地模拟现实中环境多样性。因此在此数据集上进行 ORB-SLAM3 和改进算法评估。

图 9 为 ORB-SLAM3 和本文算法在 MH_01_easy、MH_04_difficult、V1_03_difficult、V2_02_medium 等 4 组数据集下所得到的估计轨迹与真实轨迹对比曲线。由图 9 相机真实轨迹可看出,在图 9a、9b 的大范围转角处相机会出现较大的视角突变以及在图 9c、9d 中相机运动幅度较大情况下,致使相机出现运动模糊现象,而原有的 ORB-SLAM3

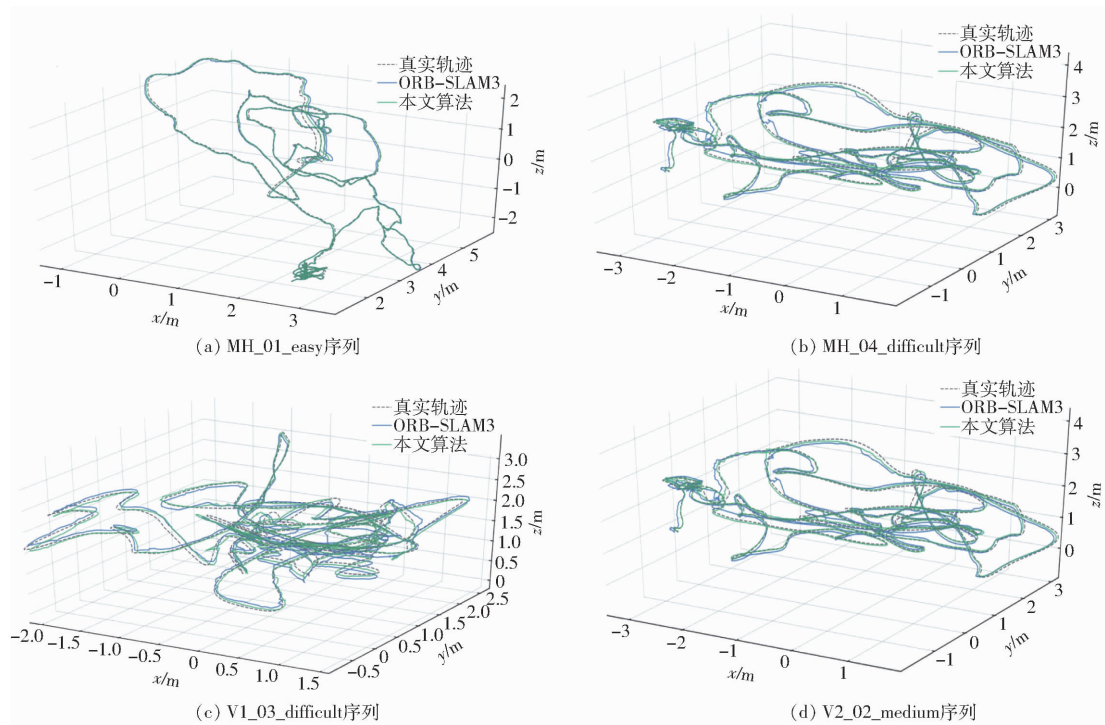


图 9 ORB-SLAM3 与改进算法运动轨迹对比曲线
Fig. 9 Comparison curves of ORB-SLAM3 and improved algorithm motion trajectories

在此处进行特征匹配时会出现大量误匹配现象,使后续估计轨迹明显偏离真实值,而改进算法在这种情况下能够很好地改善在特征匹配时出现的误匹配情况,从而降低估计轨迹与真实轨迹之间的偏差,使得估计轨迹更加接近真实轨迹。

为了更加直观地体现改进算法的特征匹配相比于 ORB-SLAM3 对于定位精度的提高,采用绝对轨迹误差(ATE)^[26]来衡量其优越性。绝对轨迹误差表示真实位姿与估计位姿之间偏差,误差越小,说明相机估计位姿越接近实际位姿,精度越高。第 n 帧 ATE 定义为

$$F_n = Q_n^{-1} S P_n \tag{12}$$

使用均方根误差(RMSE)统计 ATE 为

$$RMSE(F_{1:n}, \Delta) = \left(\frac{1}{m} \sum_{n=1}^m \| \text{trans}(F_n) \|^2 \right)^{\frac{1}{2}} \tag{13}$$

式中 Δ ——真实位置与算法估计位置之间误差
 F_n ——第 n 帧的估计轨迹与真实轨迹之间偏差
 Q_n ——真实位姿信息
 P_n ——算法计算得的估计位姿
 S ——相似旋转矩阵

用于估计位姿转换到真实位姿,以量化两者之间的旋转相似性^[27]。

为了评估估计轨迹准确性,将估计轨迹与真实轨迹进行了对比,如图 10 所示。该曲线直观地展示

了估计轨迹与真实轨迹之间偏差变化趋势。图中为 MH_01_easy、MH_04_difficult、V1_03_difficult、V2_02_medium 等 4 组数据集下绝对轨迹误差曲线,可以看出加入改进后双目立体匹配算法后,估计轨迹与真实轨迹之间的误差显著降低,明显优于 ORB-SLAM3 算法。证明了改进算法在提升定位精度方面的有效性。

同时选取不同数据集下 ATE 的均方根误差、误差平方和和标准差作为评判标准,表 1 为 2 种算法 ATE 对比结果。由表 1 可知,除在 V1_01_easy 和 V2_03_difficult 两组数据集下改进算法和 ORB-SLAM3 的 ATE 变化量小于 10%,表明在此环境下 2 种算法精度相当,改进算法虽未使定位精度明显提升,但与原始 ORB-SLAM3 算法相比改进算法的性能却有所提升。而在其它数据集下改进算法相比于 ORB-SLAM3 的绝对轨迹误差要显著降低,表明其定位精度也是显著提高。从整体数据来看改进算法相较于 ORB-SLAM3 定位精度提高 23.32%,进而验证了本文所提出的改进型双目立体匹配算法对 SLAM 定位精度的提升具有可行性和有效性。同时为了进一步验证改进算法对提高定位精度的有效性,将本文算法与 VINS-Fusion^[28]和 OV2-SLAM 算法在 EuRoC 数据集下进行对比,表 2 为 3 种算法 ATE 均方根误差对比结果,通过表 2 可以进一步确认改进算法确实实现了更高的定位精度,证明了改进算法在定位性能上的优越性。

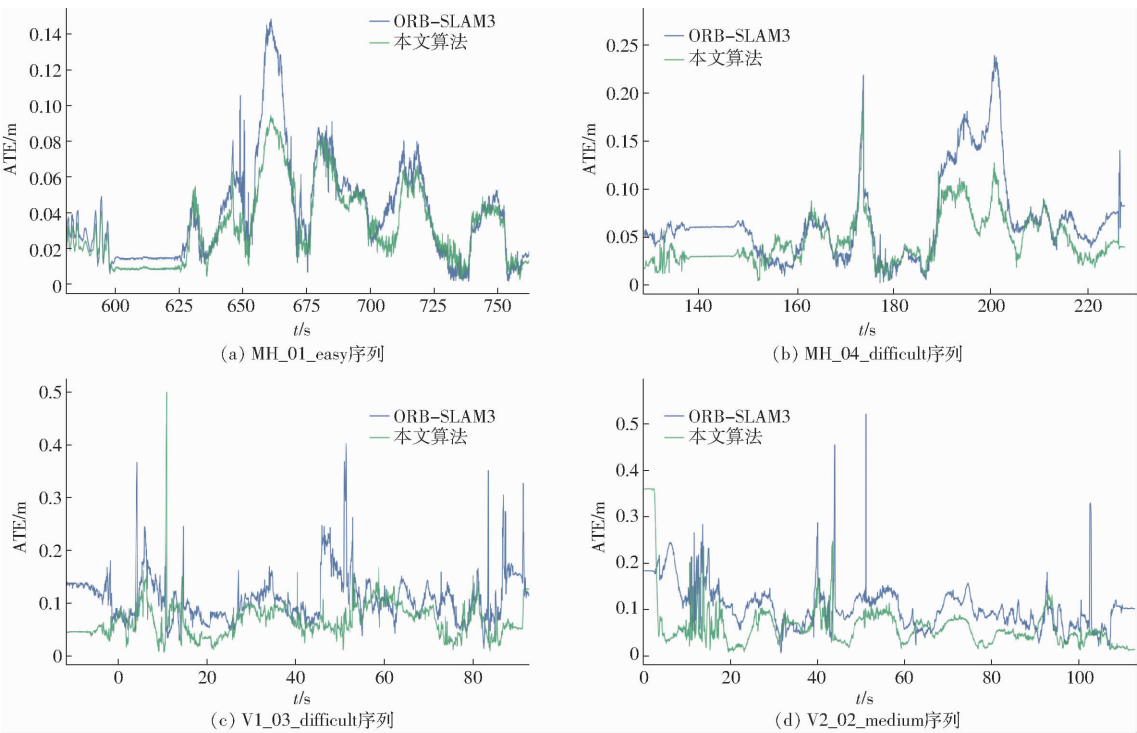


图 10 绝对轨迹误差曲线
Fig. 10 Absolute trajectory error curves

表 1 绝对轨迹误差 (ATE) 结果对比

Tab. 1 Comparison of absolute trajectory error (ATE) results

数据集	均方根误差/m		误差平方和/m ²		标准差/m	
	ORB - SLAM3	本文算法	ORB - SLAM3	本文算法	ORB - SLAM3	本文算法
MH_01_easy	0. 048 8	0. 038 3	8. 668 4	5. 354 5	0. 029 1	0. 021 5
MH_02_easy	0. 045 2	0. 037 3	6. 135 8	4. 183 8	0. 022 5	0. 017 7
MH_03_medium	0. 060 2	0. 040 0	9. 545 3	4. 228 9	0. 023 3	0. 016 2
MH_04_difficult	0. 080 0	0. 052 5	12. 645 0	5. 463 3	0. 044 0	0. 025 1
MH_05_difficult	0. 059 6	0. 041 3	7. 881 0	3. 797 1	0. 032 4	0. 021 7
V1_01_easy	0. 091 3	0. 085 9	23. 951 6	21. 211 8	0. 032 5	0. 029 2
V1_02_medium	0. 073 2	0. 062 3	8. 267 9	6. 018 5	0. 023 9	0. 018 8
V1_03_difficult	0. 119 6	0. 075 3	29. 967 1	11. 881 2	0. 045 2	0. 033 4
V2_01_easy	0. 099 2	0. 075 9	21. 706 4	12. 749 3	0. 063 7	0. 031 1
V2_02_medium	0. 113 4	0. 082 6	29. 690 8	15. 788 4	0. 048 5	0. 055 6
V2_03_difficult	0. 321 6	0. 287 4	173. 074 0	124. 736 0	0. 188 3	0. 163 6

表 2 不同 SLAM 算法绝对轨迹误差 (ATE) 结果对比

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

数据集	VINS - Fusion	OV2 - SLAM	本文算法
MH_01_easy	0. 246 2	0. 059 1	0. 038 3
MH_02_easy	0. 284 9	0. 047 8	0. 037 3
MH_03_medium	0. 428 3	0. 059 7	0. 040 0
MH_04_difficult	0. 545 8	0. 154 4	0. 052 5
MH_05_difficult	0. 480 7	0. 099 4	0. 041 3
V1_01_easy	0. 229 8	0. 092 6	0. 085 9
V1_02_medium	0. 320 5	0. 100 5	0. 062 3
V1_03_difficult	2. 785 9	0. 244 9	0. 075 3
V2_01_easy	0. 264 8	0. 117 2	0. 075 9
V2_02_medium	0. 298 6	0. 101 1	0. 082 6
V2_03_difficult	9. 302 3	2. 050 2	0. 287 4

3. 2 真实环境下评估

为了验证系统有效性和实时性,采用图 8 所示的实验平台进行数据采集和处理,同时考虑到 ORB - SLAM3 是一种基于特征点算法,因此环境中要包含一定的特征信息,并且为了验证本文算法在视角突变和弱纹理环境下特征匹配的效果,选择如图 11 所示环境比较改进算法和 ORB - SLAM3 在实时运行过程中定位精度。

为了验证本文算法和 ORB - SLAM3 算法在 Turtlebot2 机器人直线运动和转弯运动 2 种运动模式下定位精度,通过观察两种算法的轨迹偏移程度和 x 、 y 轴上的偏移,以及比较两种算法下的 ATE 参数来对比两种算法产生误差,进而对比 2 种算法定



图 11 实验环境

Fig. 11 Experimental environment

位精度。在对 Turtlebot2 机器人进行调试以及相机标定等操作后,将机器人按照预设路径进行移动,并使用激光跟踪仪采集其实际运动轨迹数据。同时,利用改进算法和 ORB-SLAM3 的跟踪估计功能得到对应的定位轨迹数据。图 12 为运动轨迹对比。由图 12a 可得,本文算法相比于 ORB-SLAM3 更接近于真实轨迹,尤其在转弯处由于相机出现视角突变以及图像纹理较弱的情况下,本文算法能更好地识别特征点并进行正确的匹配来提高定位精度,从

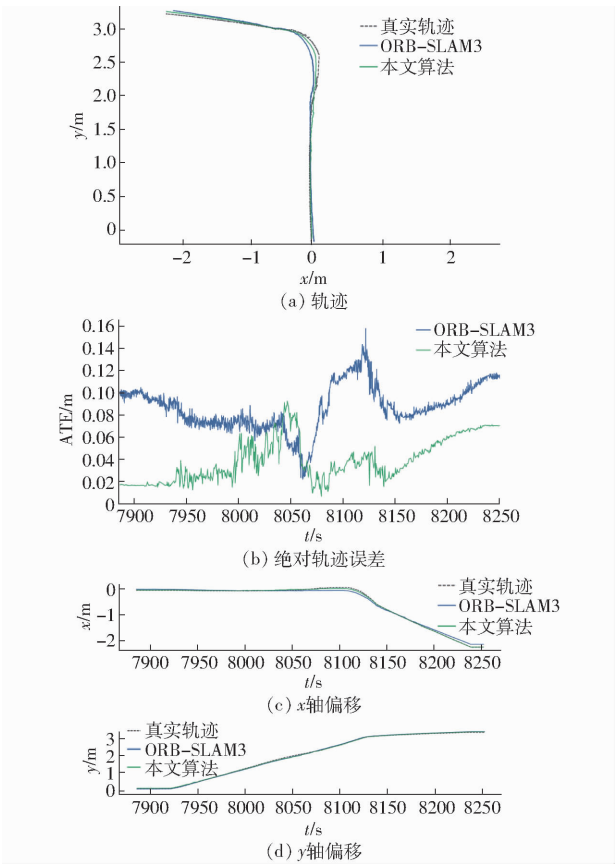


图 12 真实轨迹与估计轨迹对比

Fig. 12 Comparison of real trajectory and estimated trajectory

而使估计轨迹更加接近真实值。从图 12b 可得本文算法误差曲线明显低于 ORB-SLAM3 的误差曲线,并且在图 12c、12d 可以看到估计轨迹在 x 、 y 轴上偏移量小于 ORB-SLAM3 更接近于真实值,因此反映出在同一环境下本文算法在特征提取匹配阶段以及在剔除误匹配的过程中优于 ORB-SLAM3 算法,从而使得本文算法能够提供更高的定位精度。表 3 为本文算法与 ORB-SLAM3 算法在真实环境下得到绝对轨迹误差对比。由表 3 可以看出,本文算法误差均显著低于 ORB-SLAM3 算法误差。本文算法误差显著降低使得改进算法定位精度相较于 ORB-SLAM3 算法提升 50%。因此,本文算法在整体上实现了更高的定位精度。

表 3 真实环境绝对轨迹误差 (ATE) 结果对比		
Tab. 3 Comparison of real environment absolute trajectory error (ATE) results		
参数	ORB-SLAM3	本文算法
最大误差/m	0.158 6	0.092 8
平均误差/m	0.086 5	0.039 6
中值误差/m	0.085 6	0.035 4
最小误差/m	0.021 9	0.006 6
均方根误差/m	0.088 9	0.044 2
误差平方和/m ²	22.872 8	1.452 1
标准差/m	0.020 6	0.019 4

4 结束语

提出了一种基于双目视觉 SLAM 的改进型特征匹配算法,在双目 ORB-SLAM3 算法中立体匹配阶段加入快速近似最近邻算法和改进的 SAD-Census 算法。通过加入快速近似最近邻算法在粗匹配阶段对原有的粗匹配对进行剔除,同时将原有的 SAD 算法替换为改进的 SAD-Census 算法,进一步减少误匹配,提高了匹配对准确率,从而有效提升了 SLAM 系统定位精度。在 EuRoC 数据集和真实室内环境实验结果表明,本文算法在数据集和实时环境中的定位精度分别提升 23.32%、50%,体现了本文算法可行性和优越性。同时与 VINS-Fusion 和 OV2-SLAM 等优秀的视觉 SLAM 算法相比,本文算法也能够实现更高的定位精度。

参 考 文 献

[1] KHAIRUDDIN A R, TALIB M S, HARON H. Review on simultaneous localization and mapping (SLAM)[C]//2015 IEEE International Conference on Control System, Computing and Engineering (ICCSCE). IEEE, 2015: 85-90.

[2] GUPTA A, FERNANDO X. Simultaneous localization and mapping (SLAM) and data fusion in unmanned aerial vehicles: recent advances and challenges[J]. Drones, 2022, 6(4): 85.

[3] CADENA C, CARLONE L, CARRILLO H, et al. Past, present, and future of simultaneous localization and mapping: toward the robust-perception age[J]. IEEE Transactions on Robotics, 2016, 32(6): 1309-1332.

- [4] ZAFFAR M, EHSAN S, STOLKIN R, et al. Sensors, slam and long-term autonomy: a review[C] // 2018 NASA/ESA Conference on Adaptive Hardware and Systems (AHS). IEEE, 2018:285 – 290.
- [5] 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. A review of visual odometry technology for mobile robots[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(3): 1 – 20. (in Chinese)
- [6] TOURANI A, BAVLE H, SANCHEZ-LOPEZ J L, et al. Visual slam: what are the current trends and what to expect? [J]. Sensors, 2022, 22(23): 9297.
- [7] GARCIA-FIDALGO E, ORTIZ A. Vision-based topological mapping and localization methods: a survey[J]. Robotics and Autonomous Systems, 2015, 64: 1 – 20.
- [8] DAVISON A J, REID I D, MOLTON N D, et al. MonoSLAM: real-time single camera SLAM[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(6): 1052 – 1067.
- [9] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces[C] // 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality. IEEE, 2007: 225 – 234.
- [10] FORSTER C, PIZZOLI M, SCARAMUZZA D. SVO: Fast semi-direct monocular visual odometry [C] // 2014 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2014: 15 – 22.
- [11] MUR-ARTAL R, MONTIEL J, TARDÓS J D. ORB-SLAM: a versatile and accurate monocular SLAM system[J]. IEEE Transactions on Robotics, 2015, 31(5): 1147 – 1163.
- [12] MUR-ARTAL R, TARDÓS J D. ORB – SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras [J]. IEEE Transactions on Robotics, 2017, 33(5): 1255 – 1262.
- [13] FERRERA M, EUDES A, MORAS J, et al. Ov2SLAM: a fully online and versatile visual slam for real-time applications[J]. IEEE Robotics and Automation Letters, 2021, 6(2): 1399 – 1406.
- [14] LI G H, CHEN S L. Visual slam in dynamic scenes based on object tracking and static points detection[J]. Journal of Intelligent & Robotic Systems, 2022, 104(2): 33.
- [15] 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. A robot visual navigation method based on improved ORB – SLAM2[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(10): 306 – 317. (in Chinese)
- [16] 刘庆运, 杨华阳, 刘涛, 等. 基于激光雷达与深度相机融合的 SLAM 算法[J]. 农业机械学报, 2023, 54(11): 29 – 38.
LIU Qingyun, YANG Huayang, LIU Tao, et al. SLAM algorithm based on fusion of LiDAR and depth camera [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(11): 29 – 38. (in Chinese)
- [17] 丛佩超, 崔利营, 万现全, 等. 基于改进 ORB – SLAM2 的果园喷药机器人定位与稠密建图算法[J]. 农业机械学报, 2023, 54(7): 45 – 55.
CONG Peichao, CUI Liying, WAN Xianquan, et al. Improved ORB – SLAM2-based localization and dense map building algorithm for orchard spraying robot[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(7): 45 – 55. (in Chinese)
- [18] WANG W, CHOU C, SEVAGAMOORTHY G, et al. Stereo – NEC: enhancing stereo visual-inertial SLAM initialization with normal epipolar constraints[J]. arXiv Preprint arXiv:2403.07225, 2024.
- [19] CAMPOS C, ELVIRA R, RODRÍGUEZ J, 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.
- [20] SUJU D A, JOSE H. FLANN: fast approximate nearest neighbour search algorithm for elucidating human-wildlife conflicts in forest areas[C] // 2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN). IEEE, 2017: 1 – 6.
- [21] 邹进贵, 万茨, 孟丽媛. 一种基于自适应权重 SAD 与 Census 融合的匹配算法[J]. 测绘通报, 2018(11): 11 – 15
ZOU Jingui, WAN Ying, MENG Liyuan. A matching algorithm based on the fusion of adaptive weight SAD and Census[J]. Bulletin of Surveying and Mapping, 2018(11): 11 – 15. (in Chinese)
- [22] 程小六. 视觉惯性 SLAM 理论与源码解析[M]. 北京:电子工业出版社, 2023.
- [23] ZHANG K, LU J, LAFRUIT G. Cross-based local stereo matching using orthogonal integral images[J]. IEEE Transactions on Circuits and Systems For Video Technology, 2009, 19(7): 1073 – 1079.
- [24] LEE J, JUN D, EEM C, et al. Improved census transform for noise robust stereo matching[J]. Optical Engineering, 2016, 55(6): 063107.
- [25] BURRI M, NIKOLIC J, GOHL P, et al. The euroc micro aerial vehicle datasets[J]. The International Journal of Robotics Research, 2016, 35(10): 1157 – 1163.
- [26] STURM J, ENGELHARD N, ENDRES F, et al. A benchmark for the evaluation of RGB – D SLAM systems[C] // 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2012: 573 – 580.
- [27] 张洪, 于源卓, 邱晓天. 基于改进关键帧选择的 ORB – SLAM2 算法[J]. 北京航空航天大学学报, 2023, 49(1): 45 – 52.
ZHANG Hong, YU Yuanzhuo, QIU Xiaotian. ORB – SLAM2 algorithm based on improved keyframe selection[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(1): 45 – 53. (in Chinese)
- [28] QIN T, PAN J, CAO S, et al. A general optimization-based framework for local odometry estimation with multiple sensors[J]. ArXiv, 2019: 1901.03638.