

DOI:10.19651/j.cnki.emt.2314728

# 改进 ORB 提取匹配算法的 SLAM 应用研究 \*

张钧程<sup>1</sup> 柯福阳<sup>2,3</sup> 王旭<sup>1</sup>

(1. 南京信息工程大学自动化学院 南京 210044; 2. 南京信息工程大学软件学院 南京 210044;

3. 南京信息工程大学无锡研究院 无锡 214000)

**摘要:** 由于传统的 ORB 特征点提取匹配方法在图像纹理信息不丰富或者光照变化剧烈时极易产生特征点丢失、分布不均等问题,不利于 SLAM 系统的定位与建图。为此本文提出了一套较为鲁棒、精度较高的提取匹配算法。首先基于 ORB 特征点对其提取算法进行改进,计算自适应阈值并基于网格模型提取特征点,可提高特征点提取的鲁棒性并使其分布均匀。此外还提出了 G-R 图像匹配算法,基于网格特征计算邻域支持估计量来区分正误匹配点,再结合引入评价函数的 RANSAC 算法进一步剔除误匹配点,相比 ORB-SLAM2 原始匹配算法提高匹配精度 9.36%,并减少时间消耗约 13.6%。最后将本文提出的特征点提取匹配算法加入到 ORB-SLAM2 算法框架,经数据集与实际场景验证本文方法能有效提高 ORB-SLAM2 系统定位精度 36.6% 以上,使系统更具鲁棒性。

**关键词:** 数字图像处理;ORB 特征点;视觉 SLAM;四叉树;GMS 匹配

中图分类号: TP391.9 文献标识码: A 国家标准学科分类代码: 510.99

## Research on SLAM application with improved ORB extraction and matching algorithms

Zhang Juncheng<sup>1</sup> Ke Fuyang<sup>2,3</sup> Wang Xu<sup>1</sup>

(1. College of Automation, Nanjing University of Information Science and Technology, Nanjing 210044, China;

2. College of Software, Nanjing University of Information Science and Technology, Nanjing 210044, China;

3. Wuxi Research Institute of Nanjing University of Information Science and Technology, Wuxi 214000, China)

**Abstract:** As the traditional ORB feature point extraction and matching method is not rich in image texture information or when the lighting changes drastically, it is very easy to produce feature point loss, uneven distribution and other problems, which is not conducive to the location and construction of the SLAM system. In this paper, a set of more robust and higher accuracy extraction matching algorithm is proposed. Firstly, the extraction algorithm is improved based on the ORB feature points, the adaptive threshold is calculated and the feature points are extracted based on the grid model, which can improve the robustness of feature point extraction and make its distribution uniform. In addition, the G-R image matching algorithm is also proposed, which calculates the neighborhood support estimator based on grid features to distinguish between positive and incorrect matches, and then combines with the RANSAC algorithm that introduces the evaluation function to further eliminate incorrect matches, which improves the matching accuracy by 9.36% compared with the original matching algorithm of ORB-SLAM2, and reduces the time consumption by about 13.6%. Finally, the feature point extraction matching algorithm proposed in this paper is added to the ORB-SLAM2 algorithm framework, which is verified by the dataset and the actual scene that the method in this paper can effectively improve the positioning accuracy of the ORB-SLAM2 system by more than 36.6% and make the system more robust.

**Keywords:** digital image processing; ORB feature points; visual SLAM; quadtree; GMS matching

## 0 引言

mapping, SLAM)一直是移动机器人在未知环境中实现自主导航的关键,而图像特征提取匹配又是视觉 SLAM 中的重要问题,目前来看,一方面,传统的基于 ORB<sup>[2]</sup> (oriented

即时定位与地图构建<sup>[1]</sup> (simultaneous localization and

fast and rotated brief) 的特征点总是倾向于在强纹理的区域提取, 而一旦面对一些纹理信息不够丰富的场景, 提取到的特征点难以保证足够多且均匀的分布于图像, 也难以准确充分地反映画面信息。另一方面, 传统的图像匹配算法中基于汉明距离做暴力匹配的方法面对较复杂的图像时很容易出现误匹配的情况。特征点提取和匹配一旦变得不可靠都会极大的影响系统正确地估计位姿, 进而导致系统跟踪失败, 难以实现较好的定位和建图效果。

基于特征去做图像配准是目前的主流匹配方法。最早由 Moravec 提出了特征点的概念和基于角点的检测算法, 但该算法在图像经过旋转和尺度变化后无法实现图像配准。随后, Harris 等提出并构建了图像尺度空间, 主要使用差分运算和自相关矩阵进行角点检测, 该方法使图像在不同尺度下具有更好的匹配效果。尺度特征不变算法 (scale invariant feature transform, SIFT)<sup>[3-4]</sup> 由 Lowe 等提出, 由于其 128 维的特征描述向量, 它具有很强的鲁棒性和尺度不变性, 但也因此在数据处理中消耗了大量时间。随后加速稳健特征 (speeded up robust features, SURF)<sup>[5]</sup> 算法由 Bay 提出, 该算法基于 SIFT 对特征向量进行了降维处理, 缩短了匹配时间, 但降低了匹配精度。Rublee 等<sup>[6]</sup> 提出了加快特征点提取和描述的 ORB 算法, 不同于前两者基于浮点类型作描述符, ORB 是基于二进制描述符做位运算, 加快了运算速度, 并且将 Fast (features from accelerated segment test) 角点检测和 Brief (binary robust independent elementary features) 特征描述子相结合, 具有提取质量高, 速度快等优点。2017 年, MurArta 等<sup>[7-8]</sup> 基于 ORB 提出了 ORB-SLAM2 系统, 使用四叉树方法改善了特征点提取的均匀性, 但运行时间明显增加, 运行效率有待提高。2018 年秦晓飞等<sup>[9]</sup> 提出通过八点改进法减少随机采样一致算法 (random sample consensus, RANSAC) 迭代次数, 采用极线约束来剔除误匹配, 但极限约束的方法可能存在没有可行解的情况, 导致该方法适用性并不强。Bian 等<sup>[10]</sup> 提出了一种基于网格运动统计 (grid-based motion statistics, GMS) 的匹配算法, 该方法处理容量大, 计算速度快, 但仍然存在较多误匹配点。2022 年 Han 等<sup>[11]</sup> 采用改进的近似最近邻快速库 (fast Library for approximate nearest neighbors, FLANN) 算法进行自适应特征匹配并使用 RANSAC 消除选定待匹配点后的假匹配点对, 然而该方法匹配正确率不够高。崔建国等<sup>[12]</sup> 提出了 BEBLID 二进制描述子与改进 SURF 算法相结合的方法来提高匹配正确率, 但仍然无法较好地满足视觉 SLAM 实时性的要求。

目前视觉 SLAM 系统在光照变化剧烈、高速移动和弱纹理环境下特征点跟踪易丢失, 且图像匹配时间长、匹配准确度低, 导致系统计算得到的轨迹误差较大、鲁棒性不够强。针对以上问题本文在 ORB-SLAM2<sup>[8]</sup> 的基础上结合现有技术的优点, 提出了基于 ORB 特征点改进的提取匹配算法, 本文具体贡献如下:

1) 为了增强特征点提取的稳定性, 改善 ORB 提取器对光照变化敏感、弱纹理场景无法有效提取特征点的问题, 采用自适应阈值保证特征点提取性能更加鲁棒, 同时提出了一种基于网格划分的 ORB 提取算法, 保证提取的均匀性。

2) 为改善匹配精度, 本文提出 G-R 图像匹配算法, 首先基于网格划分计算邻域支持估计量, 区分正确匹配点与错误匹配点, 在筛选结果的基础上引入评价函数来评估点对间距离相似性, 进一步增加局内点占比, 最后通过 RANSAC 算法迭代模型实现较高的匹配准确率。

3) 为检验本文算法的实际工程应用效果, 将改进的特征点提取匹配算法加入到 ORB-SLAM2 算法框架, 经数据集与实际场景验证本文方法能有效提高 ORB-SLAM2 系统定位精度, 减小轨迹误差, 增强系统的鲁棒性。

## 1 基本原理

ORB 算法<sup>[6]</sup> 由检测 FAST 角点和计算 BRIEF 描述子两部分组成。

FAST 算法是一种快速的二维图像的特征检测算法, 它通过所选元素与周围元素的像素灰度之间的差异来检测角点。如图 1 所示, 以像素  $P$  为例, 在半径 3 的圆上总共发现了 16 个像素点。如果这 16 个像素中的 9 个连续像素与  $P$  之间的灰度差大于设置的阈值  $T$ , 则判断  $P$  为 FAST 的角点。可用计算公式表达为:

$$N = \sum_{x=1}^n |I(x) - I(p)| \geq T \quad (1)$$

其中,  $n$  为  $p$  点圆周上总共的像素点数,  $I(x)$  是圆周上任意像素点的灰度值,  $I(p)$  是要检测的像素的灰度值,  $T$  是角点检测阈值, 本文中使用 FAST-9 算法, 即此处  $N=9$ 。

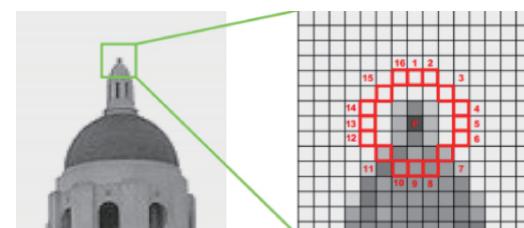


图 1 FAST 角点检测原理

然而由于传统的 FAST 角点<sup>[13]</sup> 不具有尺度和方向性, 仅仅通过改善提取稳定性获得的 ORB 特征点仍然是无法直接使用的。本文采用图像金字塔和灰度质心法来满足尺度不变性与旋转不变性。所谓尺度不变性就是需要算法能在不同距离的图像帧中稳定地提取特征点, 所谓旋转不变性就是需要算法能在不同角度的图像帧中计算特征点的主要方向, 从而稳定地提取特征点。图像金字塔是同一幅图像在不同尺度下的集合, 包括图像的上采样和下采样。示意图如图 2 所示, 金字塔底层是原始图像, 逐层往上是将图像进行固定倍率缩放的结果(根据工程经验一般把缩放因子

$scare$  取为 1.2), 这样就得到了不同分辨率的图像, 金字塔层数越高则等同于相机离物体越远, 对于第  $i$  层图像尺寸可表示为:

$$Size_i = \left( H \times \frac{1}{scare^i} \right) \times \left( W \times \frac{1}{scare^i} \right) \quad (2)$$

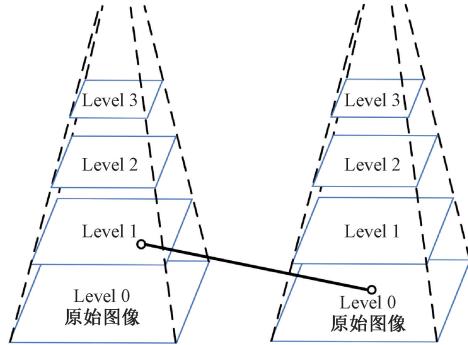


图 2 图像金字塔示意图

由于 FAST 关键点缺少角度信息, 这样当图像发生旋转后, BRIEF 描述子也会发生变化, 使得特征点对旋转不鲁棒, 所以本文采用灰度质心法——连接图像块的几何中心与质心得到方向向量, 来满足旋转不变性。所谓质心就是指以图像块灰度值作为权重的中心, 在后续计算描述子时, 就可以将目标特征点周围像素旋转到主方向上来计算。如示意如图 3 所示, 其旋转角度可定义为:

$$\theta = \arctan\left(\frac{m_{01}}{m_{10}}\right) \quad (3)$$

其中,  $m_{01}$  与  $m_{10}$  为沿两个坐标轴  $X, Y$  方向的图像矩:

$$m_{10} = \sum_{x=-R}^R \sum_{y=-R}^R x I(x, y) \quad (4)$$

$$m_{01} = \sum_{x=-R}^R \sum_{y=-R}^R y I(x, y) \quad (5)$$

特征点的描述使用二进制的 BFIEF 描述子, 其核心思想是在关键点  $P$  的周围以一定模式选取  $N$  对点对, 将  $N$

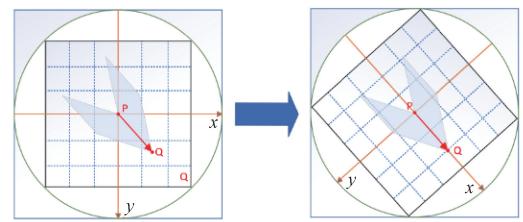


图 3 灰度质心法示意图

对点对的比较结果组合起来作为描述子。其描述向量由多个 0 和 1 组成。具体计算公式为:

$$\tau(p; x, y) = \begin{cases} 1, & p(x) < p(y) \\ 0, & p(x) \geq p(y) \end{cases} \quad (6)$$

其中,  $p(x)$  和  $p(y)$  分别是点  $x$  和点  $y$  处像素的灰度值,  $\tau$  是描述符的值,  $x$  和  $y$  处的灰度值是上述子窗口的灰度值之和。在窗口中随机选取  $N$  对随机点(一般  $N = 256$ ), 形成一个二进制编码, 这个编码就是特征点描述子。此外, 在式(3)中计算的方向被添加到描述子, 使得描述子具有良好的旋转不变性。

## 2 关键技术研究

### 2.1 算法基本框架

本文基于 ORB 特征点改进的提取匹配算法框架如图 4 所示。算法接收到图像序列之后, 首先通过创建 8 幅不同尺寸的图像, 建立图像金字塔以实现尺度不变性, 然后在各层图像金字塔中基于网格划分的方法计算每个网格需要提取的特征点数<sup>[14]</sup>, 这样提取到的特征点具备较好的均匀性, 可以防止特征点过于密集, 使算法更加鲁棒, 再根据图像灰度信息计算适应阈值完成 ORB 特征点的提取。将两幅图像中处理好的特征点通过邻域支持估计量完成正误匹配点的区分, 再建立评价函数评估各点对间距离的相似性, 利用 RANSAC 方法剔除误匹配点, 最后输出完成匹配的图像。

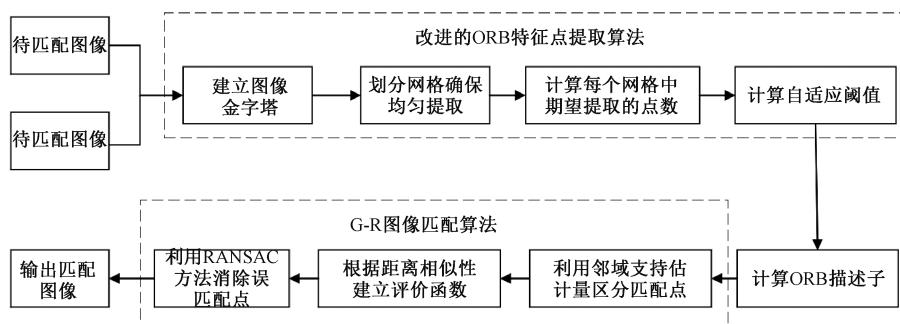


图 4 特征点提取匹配算法框架图

### 2.2 基于 ORB 特征点的提取算法改进

ORB-SLAM2 算法使用固定阈值  $T$  来提取特征点, 并结合四叉树算法来抑制非极大值<sup>[15]</sup>, 在一定程度上避免了特征点过于集中的情况, 然而使用固定阈值不能满足不同

场景下图像特征点提取的要求, 而用于均匀化的四叉树算法有可能会导致一些原本质量较好的点被剔除, 一些质量较差的点被保留。为此我们提出了基于网格与自适应阈值的 ORB 特征提取方法。

为了使每层金字塔都能均匀合理地提取特征点, 每层特征点数量由总的图像金字塔层数和待提取的特征点总数关系确定, 其计算公式如下:

$$DesF_i = N \times \left( \sum_{i=0}^n Size_i \right) \quad (7)$$

其中,  $DesF_i$  为第  $i$  层图像金字塔所需要提取的特征

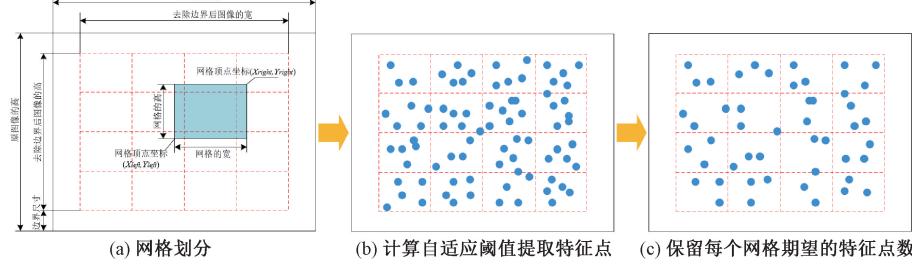


图 5 基于网格划分的 ORB 提取过程

由于图像边缘特征点不太可靠且易丢失, 依赖于边缘特征有可能会导致匹配失败, 因此在金字塔每层图像中设置了边界, 如图 5(a)所示。此时每层金字塔图像实际可用于特征提取的宽高尺寸为。

$$w_{inew} = w_i - 2EdgeTh \quad (8)$$

$$h_{inew} = h_i - 2EdgeTh \quad (9)$$

这里的  $w_{inew}$  和  $h_{inew}$  为去除边界后图像的宽和高,  $w_i$  和  $h_i$  为第  $i$  层图像金字塔的宽和高,  $EdgeTH$  为去除边界的尺寸。

网格的划分应与图像大小保持等比例的对应关系, 以使特征点更加均匀, 且网格数量与每层期望提取的特征点数量正相关, 因此网格的列数和行数按公式确定。

$$col_i = \sqrt{\frac{DesF_i}{ImRat_i}} / s \quad (10)$$

$$row_i = ImRat_i \times col_i \quad (11)$$

这里的  $DesF_i$  为每层金字塔图像需要提取的特征点数, 由式(7)确定,  $s$  为控制网格划分的系数,  $ImRat_i$  为图像宽高比, 可由公式  $ImRat_i = w_{inew}/h_{inew}$  确定。进一步, 各个网格 4 个顶点的位置可表示为:

$$X_{left} = n \times Mw_i - Th \quad (12)$$

$$Y_{left} = m \times Mh_i - Th \quad (13)$$

$$X_{right} = X_{left} + Mw_i + Th \quad (14)$$

$$Y_{right} = Y_{left} + Mh_i + Th \quad (15)$$

其中,  $X_{left}$ 、 $Y_{left}$ 、 $X_{right}$ 、 $Y_{right}$  为图 5 中的 4 个边界点,  $Th$  为网格拓展系数, 可以更好地提取网格边缘特征,  $Mw_i$  为该网格的宽,  $Mh_i$  为该网格的高, 可以以如下公式表示:

$$Mw_i = w_{inew} / col_i \quad (16)$$

$$Mh_i = h_{inew} / row_i \quad (17)$$

然后对每个网格进行划分, 提取 ORB 特征。为使每个网格中的特征点分布尽可能均匀, 确定图像金字塔第  $i$  层的每个网格的期望特征点数目。

点数,  $N$  为期望提取的总的特征点数,  $Size_i$  为第  $i$  层金字塔图像尺寸, 其大小可由式(2)求得,  $\sum_{i=0}^n Size_i$  为  $n$  层金字塔总的尺寸面积。当确定图像金字塔中每一层的特征提取数量时, 利用网格模型的方法均匀地提取特征点, 如图 5 所示。

$$MDesF_i = \frac{DesF_i}{row_i \times col_i} \quad (18)$$

阈值  $T$  的选择决定了提取的结果, 在 ORB-SLAM2 中阈值由工程经验值设置为一个固定的值, 不能较好地满足实际环境。例如, 在一些纹理信息较弱的图像区域, 高阈值不能提取足够的特征。Rublee 等<sup>[6]</sup>提出的自适应阈值仍然包含人为设置的比例因子。因此, 本文在网格化的基础上, 基于图像灰度信息的自适应阈值<sup>[16]</sup>进行 FAST 角点检测, 其值能根据图像全局灰度分布的变化进行合理的变化, 更鲁棒地提取特征点, 初始阈值  $T_{ini}$  的计算方法如公式所示:

$$T_{ini} = \left( \frac{1}{n} \sum_{i=1}^n (I(x_i) - \overline{I(x)})^2 \right) / \overline{I(x)} \quad (19)$$

其中,  $I(x_i)$  是图像中每个像素的灰度值,  $\overline{I(x)}$  是图像灰度的平均值。利用上述公式可以根据不同的图像计算出不同的比较阈值, 从而使算法即使在结构稀疏的弱纹理场景下也能具有更强的抗干扰能力。

最后, 对于各网格中提取到的特征点按照 Harris 响应值大小进行排序, 保留前  $MDesF_i$  个特征点以用于后续的匹配。

### 2.3 G-R 图像匹配算法

ORB-SLAM2 中使用基于汉明距离结合 RANSAC 的方法来实现特征点的匹配, 然而对描述子逐对计算汉明距离耗费了大量的时间, 而且传统的 RANSAC 算法需要预设一组“局内点”作初值, 该方法具有随机性, 它要求大部分的错误匹配要被预先剔除, 一旦最近邻匹配集中存在大量的错误匹配的情况, 则需要剔除大部分正确的匹配来减弱错误匹配的影响, 使计算时间指数型增长。

本文基于 RANSAC 提出了一套新的匹配算法, 首先利用网格结构计算邻域支持估计量<sup>[10]</sup>进行初步筛选, 区分正误匹配点, 将得到的正确匹配结果作为先验信息, 再建立评价函数<sup>[17-18]</sup>, 根据匹配点间距离相似性特点进行筛

选, 提高“局内点”的占比, 最后利用 RANSAC 方法不断迭代计算出参数模型, 剔除误匹配点, 使图像匹配实现较高的准确率。为方便后文描述, 将其命名为 G-R 图像匹配算法。

图 6 为正误匹配对分布示意图, 对于同一场景下不同角度的图像, 其对应的特征点匹配集合为  $\{M, N\}$ , 在区域  $a$ 、区域  $b$  内匹配特征点集合为  $\{x_1, x_2, x_3, \dots, x_n\}$ , 其中  $x_i$  表示一对正确的匹配点对, 正确匹配点对附近存在若干邻域支持匹配点对, 而错误匹配点对  $x_j$  邻域内几乎没有可供支持的匹配对。如图 6(a)所示, 正确的邻域支持匹配点对  $s_i = 3$ , 错误的邻域支持匹配点对  $s_j = 0$ 。

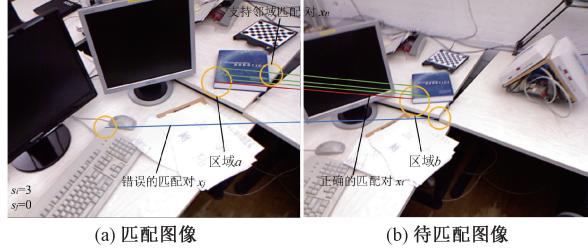


图 6 正误匹配对分布示意图

为了计算区域  $a$  到区域  $b$  匹配正确的概率, 用图 7 表示匹配过程中的事件空间, 图 7(a)表示区域  $a$  到区域  $b$  中正确的匹配, 将该事件记为  $T^{ab}$ , 图 7(b)表示区域  $a$  到区域  $b$  中错误的匹配, 将该事件记为  $F^{ab}$ ,  $f_a^b$  表示特征点从区域  $a$  匹配到区域  $b$ , 特征点从区域  $a$  匹配到区域  $b$  以外的区域则用  $\bar{f}_a^b$  表示,  $f_a^t$  表示由区域  $a$  到区域  $b$  正确的匹配, 由区域  $a$  到区域  $b$  错误的匹配则用  $f_a^f$  表示。那么当一个点对匹配错误发生时, 这样的匹配点落在区域  $b$  的概率为:

$$p(f_a^f | f_a^t) = \beta n / N \quad (20)$$

其中,  $n$  是区域  $b$  中特征点的个数,  $N$  为待匹配图像中特征点的总数量,  $\beta$  是权重值。

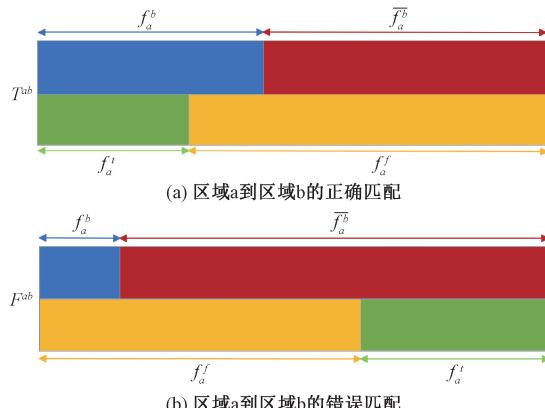


图 7 区域  $a$  到区域  $b$  匹配事件空间

令某特征点由区域  $a$  到区域  $b$  正确匹配发生的概率为  $p(f_a^t) = t$ , 那么可以得到正确匹配发生时被邻域支持的概率  $p_t$  以及错误匹配发生时被邻域支持的概率  $p_f$ :

$$p_t = p(f_a^b | T^{ab}) = t + (1-t)\beta n / N \quad (21)$$

$$p_f = p(f_a^b | F^{ab}) = \beta(1-t)(n/N) \quad (22)$$

由式(21)、(22)可以近似得到邻域支持估计量  $|s_i|$  的二项分布:

$$|s_i| \sim \begin{cases} B(n, p_t), & \text{匹配正确} \\ B(n, p_f), & \text{匹配错误} \end{cases} \quad (23)$$

于是真假匹配的可区分度被表示为如下形式:

$$P = \frac{|E_t - E_f|}{\sqrt{V_t} + \sqrt{V_f}} = \frac{n \times (p_t - p_f)}{\sqrt{n} \times p_t (1-p_t) + \sqrt{n} \times p_f (1-p_f)} \quad (24)$$

其中,  $E_t$  与  $E_f$  分别为正确与错误邻域估计量的期望,  $V_t$  与  $V_f$  分别为正确与错误邻域估计量的方差, 由上式可知  $P \propto \sqrt{n}$ , 即当图像匹配点数越多, 正确与错误的匹配区分度就越强。在实际的匹配过程中, 利用特征点提取时划分好的网格结构, 将邻域支持估计量分配给对应的网格, 如图 8 所示, 算法总体复杂度被降低到  $O(N)$ , 从而起到加速区分正确和错误匹配的效果。通过阈值化  $|s_i|$  来判断区域匹配对  $\{i, j\}$  的真假:

$$cell-pair\{i, j\} \in \begin{cases} T, & \text{if } |s_i| > \tau_i \\ F, & \text{其他} \end{cases} \quad (25)$$

其中,  $i$  表示匹配图像(a)中第  $i$  个网格区域,  $j$  表示待匹配图像图 8(b)中第  $j$  个网格区域,  $T$  和  $F$  分别表示真对应集和假对应集。基于式(24), 将  $\tau_i$  设置为:

$$\tau_i = \alpha \sqrt{n} \quad (26)$$

其中,  $\alpha$  是一个超参数, 根据工程经验, 当  $\alpha$  取到  $4 \sim 6$  之间时, 能达到较好的匹配性能。通过上述公式, 我们将邻域支持估计量  $|s_i| > \tau_i$  的网格区域匹配对保留下来, 保留下来的网格中存在大量正确的匹配对, 为后续进一步的匹配筛选提供了较好的先验条件。

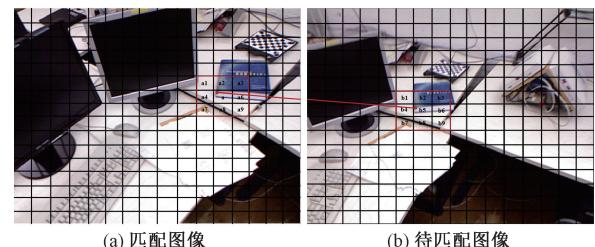


图 8 网格运动示意图

图 9 表示了匹配点对间的距离相似性,  $(A_i, B_i)$  和  $(A_j, B_j)$  为两帧图像间的两对最邻近匹配点对, 那么  $l(A_i, B_i)$  应相似于  $l(A_j, B_j)$ , 该距离  $l$  应满足  $l \in [L_{\min}, \lambda \times L_{\max}]$ , 其中,  $L_{\min}$  为匹配距离最小值,  $L_{\max}$  为匹配距离最大值,  $\lambda \in [0, 1]$  为比例因子, 根据工程经验一般取 0.7, 构建如下评价函数来描述  $(A_i, B_i)$  与  $(A_j, B_j)$  点对间距离的相似性程度。

$$F(i) = \sum_{j=1}^c \frac{R(i, j)}{1 + Y(i, j)} \quad (27)$$

$$R(i, j) = \exp\left(-\frac{|l(A_i, B_i) - l(A_j, B_j)|}{Y(i, j)}\right) \quad (28)$$

$$Y(i, j) = \frac{l(A_i, B_i) + l(A_j, B_j)}{2} \quad (29)$$

其中,  $c$  为内点个数,  $R(i, j)$  表示点对间距离的相对差异,  $Y(i, j)$  表示点对间的平均距离, 式(27)中的值越小则说明点对之间相似程度越高, 目标匹配点越稳定, 计算目标点对间的距离相似性, 若其值超过了各匹配点对评价函数的平均值, 则将目标点对进行剔除。最后, 将筛选得到的匹配点采用 RANSAC 方法计算迭代参数模型, 得到最终的正确匹配点。

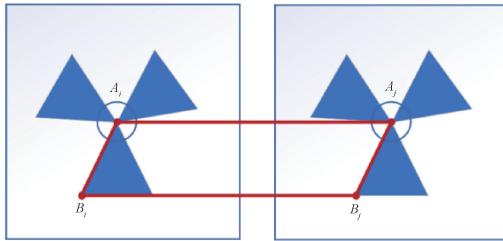


图 9 距离相似性示意图

### 3 实验结果与分析

在不同的数据集下分别做提取匹配性能实验, 经定性与定量分析, 充分验证本文算法在不同环境中的较为鲁棒的提取性能、较准确的图像匹配率以及较高的运算效率。将本文提出的完整的特征点提取匹配算法加入到 ORB-SLAM2 算法框架并分别在数据集和实时环境下进行测试, 与改进前的 ORB-SLAM2 以及 ORB-SLAM3 作轨迹误差对比, 证明本文算法能有效改善视觉 SLAM 的鲁棒性和定位精度, 且具有良好的工程应用性。

用作提取匹配实验的图像共有 4 组, 每组 6 张, 如图 10 所示, 前三组图像来自于牛津 Oxford 标准图像数据集, 分别代表了复杂结构、模仿高速运动产生的模糊以及光照变化剧烈导致的低曝光等场景, 最后一张图像来自于德国慕尼黑工业大学 TUM 数据集, 因含有平滑的桌面、地面等纹理单一的要素, 故可以用来检验本文算法在弱纹理结构下提取匹配效果。

#### 3.1 改进 ORB 特征点提取算法性能测评

为了验证本文提出的改进 ORB 特征点提取算法具备较好的提取性能, 在图 10 中每组图像各选取一张以代表不同的场景, 将提取初始期望值都设置为 2 000。图 11 与表 1 展示了 SIFT 算法、ORB-SLAM 中的提取算法、ORB-SLAM2(AT) 与本文方法实际提取效果的对比情况, 其中 ORB-SLAM2(AT) 表示在 ORB-SLAM2 提取算法的基础上使用了自适应阈值。

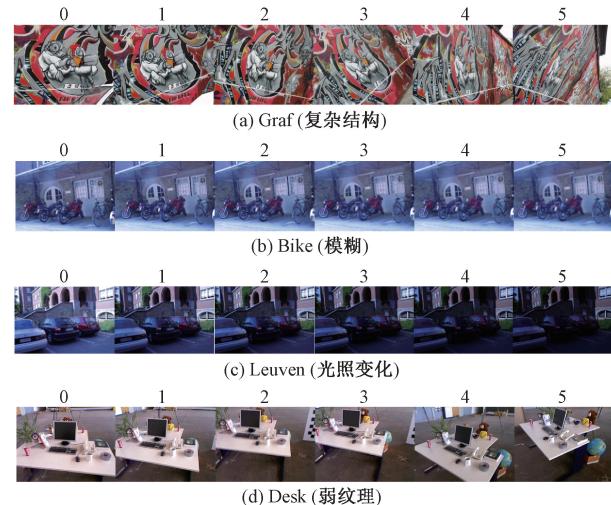
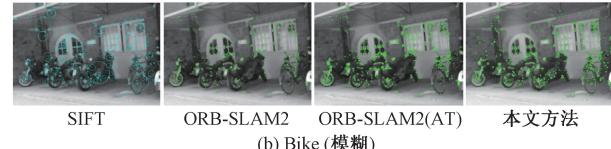


图 10 Oxford/TUM 数据集示例



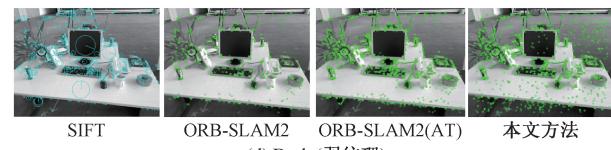
(a) Graf (复杂结构)



(b) Bike (模糊)



(c) Leuven (光照变化)



(d) Desk (弱纹理)

图 11 不同算法提取特征点效果对比

在一般的复杂结构场景(Graf)中, 4 种算法都提取到了较多的点, ORB-SLAM2(AT)由于采用了自适应阈值, 相比于 ORB-SLAM2 提取到了 42.9% 的点, 且与本文算法提取数量接近, 但由于采用的仍然是四叉树的均匀化策略, 在实时性表现上不如本文方法; 在模仿高速运动导致的模糊场景(Bike)中, SIFT 表现明显下滑, 且耗时也最多, ORB-SLAM2(AT)相比 ORB-SLAM2 多提取到了 32.8% 的点, 本文算法相比 ORB-SLAM2 多提取到了 37.8% 的点; 在模仿剧烈光照变化引起的低曝光场景(Leuven)中, ORB-SLAM2(AT)与本文算法比 ORB-SLAM2 明显提取到了更多的点, 但在右下角的暗部区域 ORB-SLAM2(AT)采用的四叉树均匀化策略没有产生很好的表现, 几

表 1 不同算法提取特征点性能对比

图像序列	Contrast parameter	SIFT	ORB-SLAM2	ORB-SLAM2(AT)	本文方法
Graf	Feature number	1 225	1 324	1 892	1 948
	Running time/ms	217.0	132.3	144.2	62.6
Bike	Feature number	814	1 359	1 804	1 872
	Running time/ms	318.8	163.8	158.2	41.1
Leuven	Feature number	1 041	1 481	1 749	1 902
	Running time/ms	208.2	127.4	134.9	54.4
Desk	Feature number	1 159	1 265	1 896	1 942
	Running time/ms	164.8	83.7	87.1	47.9

乎没有保留到可用特征点,相比之下采用网格模型的本文算法更能适应不同的光照环境,获得更好的均匀化效果;而在弱纹理场景(Desk)中,ORB-SLAM2 在左下角和右上角的桌面、地面等弱纹理区域提取表现一般,而本文方法采用了基于网格模型与自适应阈值,在弱纹理区域相比于 ORB-SLAM2 明显提取到了更多的特征点,且取得了更好的匀质化效果。在时间消耗方面,各个场景下 SIFT 实时性最差,而本文方法均以更少的时间代价完成了特征点的均匀化提取,在 Bike 场景中相比于 ORB-SLAM2 提速了 74.9%,在 Desk 场景中相比于 ORB-SLAM2 提速了 42.8%。

### 3.2 G-R 图像匹配算法测评

为验证本文提出的 G-R 图像匹配算法良好的匹配效果,在图 10 所示的图例下基于 ORB 特征点与主流的 FLANN、K 近邻算法(k-nearest neighbor, KNN)以及 ORB-SLAM2 中的图像匹配算法作对比,采用匹配正确率 CMR<sup>[19]</sup>与运行时间等指标来定性定量的全面衡量本文特征点匹配算法性能,CMR 值越大,运行时间越少则说明匹配性能越好。CMR 的值定义为:

$$CMR = \frac{m}{n} \quad (30)$$

其中,  $m$  表示匹配正确的特征点数,  $n$  表示提取到的总的特征点数。

图 12 显示了图 10(a)组中 0 号图像与 2 号图像在各算法下匹配的效果对比,由于 FLANN 和 KNN 本质都是利用近邻去寻找匹配点,当图像中存在大量相似的复杂结构时,精度较差,所以进行匹配时都出现了较多的错误匹配,而在使用 ORB-SLAM2 的匹配算法时由于采用了 RANSAC 进行迭代提出了大量误匹配点,匹配正确率明显提高,但在过程中由于缺乏较为可靠的初值也过度剔除了大量有用的特征点,而在本文 G-R 算法中有更多的可用特征点保留了下来,且匹配正确率高于 ORB-SLAM2 的匹配算法。

在其他各组图像中也按照同样的方法进行匹配实验,图 13 展示了不同算法在不同场景下的匹配正确率。横轴序号 1~5 分别代表了每组编号为 1~5 的图像与编号 0 的

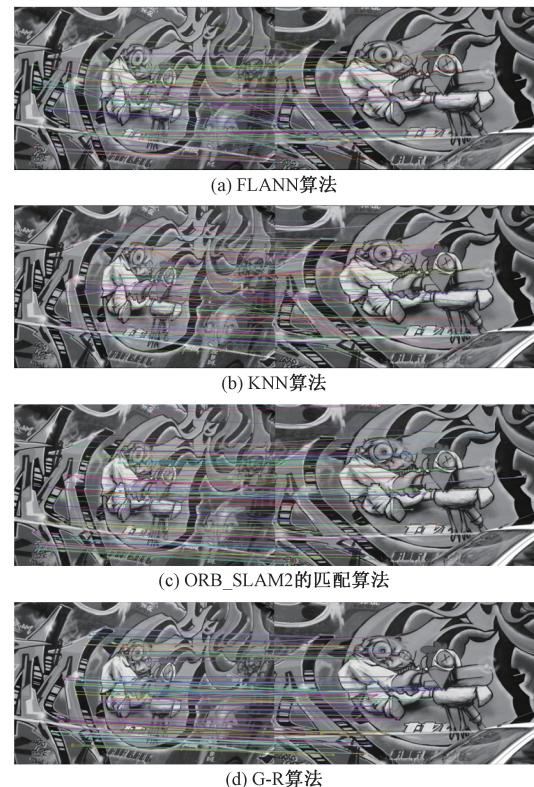


图 12 不同算法匹配性能对比

首张图像依次进行匹配,纵轴代表了各算法匹配正确率 CMR 的值。可以看到对于 Graf(复杂结构),由于编号 4 和编号 5 图像视角变化较大,各算法匹配正确率明显下滑,但本文 G-R 算法相比于同类其他算法仍能保持相对较高的正确率,对于 Bike(模糊)以及 Leuven(光照变化)中,本文算法表现稳定,匹配正确率保持在 95% 以上,这是由于运动网格的划分,正误匹配点得到了较好的区分,在 Desk(弱纹理)场景中,本文算法在不同角度下也均优于其他算法。总体来看,在 20 对图像实验中,本文提出的 G-R 算法匹配正确率相比于 ORB-SLAM2 中的匹配算法平均提高 9.36%,并且也明显优于 KNN 算法与 FLANN 算法,其不同环境下匹配的准确性为后续图像信息的进一步应用提供了可靠保证。

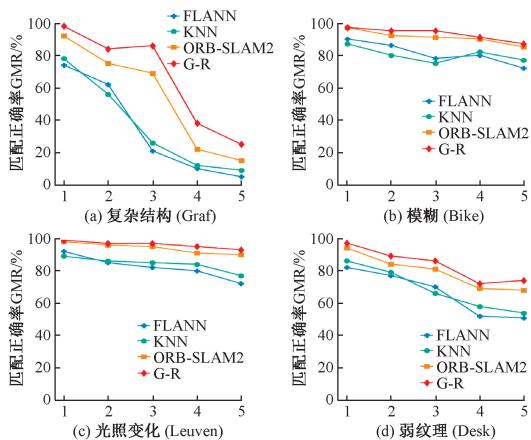


图 13 不同算法匹配正确率

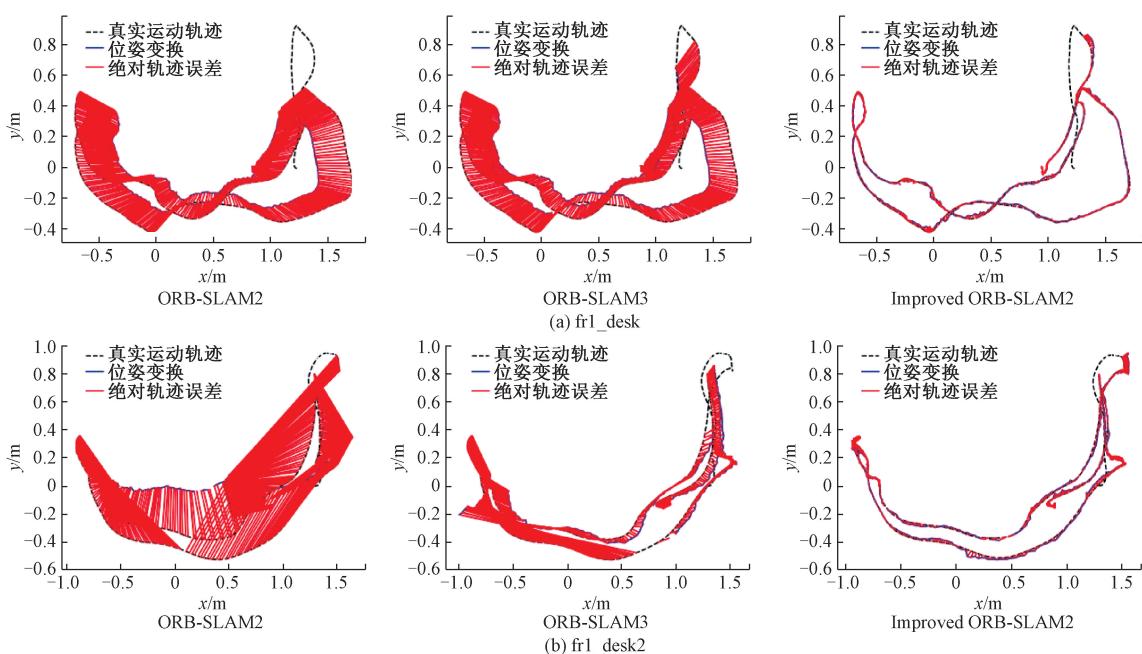
表 2 展示了各组图像场景在不同算法下匹配所消耗的平均时间,在 Graf(复杂结构)场景中,由于提取到较多特征点,各算法普遍耗时较多,本文 G-R 算法相比于最慢的 KNN 算法提速约 58%,相比于 ORB-SLAM2 匹配算法提速 14.8%,在 Bike(模糊)以及 Leuven(光照变化)场景中,G-R 算法相比于 ORB-SLAM2 匹配算法差距不大,但速度优势仍然存在,在 Desk(弱纹理)场景中,G-R 算法运行速度明显优于其他算法,相比于 ORB-SLAM2 匹配算法运行耗时约降低 19.6%。四组图像序列中,G-R 算法相比于 ORB-SLAM2 匹配算法运行耗时平均降低 13.6%。经过分析指出,本文 G-R 算法在保证较高匹配率的同时仍拥有较稳定的速度优势,保障了其应用的实时性。

表 2 不同算法匹配时间 ms

图像序列	FLANN	KNN	ORB-SLAM2	G-R
Graf	9.6	11.2	5.4	4.6
Bike	5.6	7.2	3.6	3.2
Leuven	5.9	8.1	3.4	3.1
Desk	7.2	9.5	4.6	3.7

### 3.3 SLAM 应用测评

为了验证本文算法具备的良好的工程应用性,将该套完整的特征点提取匹配算法加入到 ORB-SLAM2 算法框架,使用 TUM 数据集进行视觉里程计精度测定,利用 evaluate 测评工具对其进行了测评,使用绝对轨迹误差 (absolute trajectory error, APE) 来评判视觉里程计的精度,并加入 ORB-SLAM3<sup>[20]</sup> 进行对比测试,ORB-SLAM3 是西班牙萨拉戈萨大学于 2020 年开源的算法,由于加入了多地图合并,相比 ORB-SLAM2 具有更好的定位效果。图 14 展示了以 fr1\_desk、fr1\_desk2、fr1\_desk\_with\_person 为例,与 ORB-SLAM2 和 ORB-SLAM3 的绝对轨迹误差对比情况,其中黑色虚线代表 ground\_truth 真实运动轨迹,实线代表算法估计的位姿变换,线段组成的阴影区域代表绝对轨迹误差量的大小,阴影区域越大则代表误差量越大,定位越不精准。由于本文算法较好的特征点提取性能和较高的匹配正确率,可以看出改进后的 ORB-SLAM2 由于采用了本文的特征点提取匹配算法,前端特征点配准性能的提升使 SLAM 在各数据集的 ground\_truth 真值相



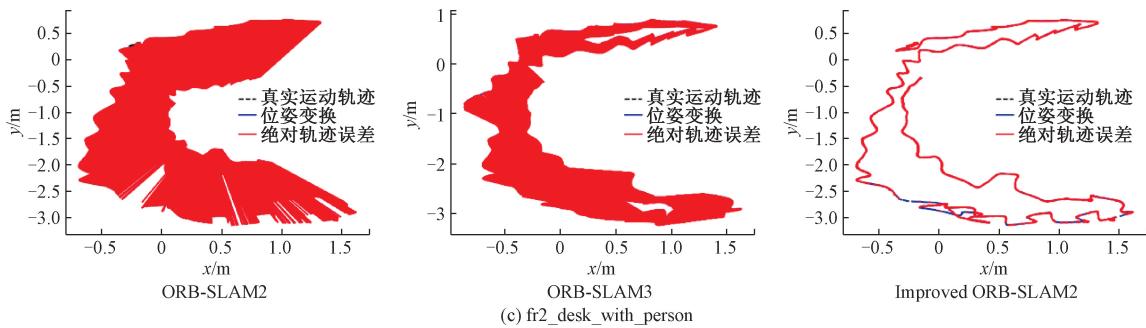


图 14 绝对轨迹误差对比图

比有着较小的 APE 误差, 其定位效果明显优于原始 ORB-SLAM2 和 ORB-SLAM3。

为了更加量化地评测本文算法对于 ORB-SLAM2 的精度提升效果, 在更多的数据集下进行测试, 参数采用均方根误差 (root mean square error, RMSE) 和标准差 (standard deviation, STD) 进行评价, RMSE 与 STD 的值突出显示为易受大误差或偶尔误差影响, 因此, RMSE 和 STD 可以比平均值和中值更好地表示系统的鲁棒性和稳定性。并且为了减少偶然误差, 提高数据的真实性, 表中

每组数据都取重复运行 3 次后的平均值。具体的实验结果如表 3 所示。在数据集 fr2\_desk\_with\_person 中, 改进后的 ORB-SLAM2 效果提升最为明显, 相比原始的 ORB-SLAM2, RMSE 减小了 63.2%, STD 减小了 58.8%。改进后的 ORB-SLAM2 在六组数据集中平均 RMSE 减小了 48.8%, STD 减小了 46.2%, 且相比于 ORB-SLAM3 也有着更好的表现。综合图 14 和表 3 说明了本文通过增强前端特征点可靠性的方法能有效保证 SLAM 系统在初始化完成后稳定地跟踪特征点, 使位姿估计结果更加精准。

表 3 数据集精度对比

图像序列	ORB-SLAM2		ORB-SLAM3		Improve ORB-SLAM2		Improvement	
	RMSE/mm	STD/mm	RMSE/mm	STD/mm	RMSE/mm	STD/mm	RMSE/%	STD/%
fr1_desk	41.4	22.7	36.0	18.5	26.4	12.5	36.6	45.0
fr1_desk2	47.9	24.8	28.6	16.3	19.7	13.1	58.9	47.1
fr1_room	75.2	39.3	68.7	25.5	44.4	23.0	41.0	41.5
fr1_xyz	62.6	32.4	57.2	26.3	38.5	19.2	38.5	40.8
fr2_desk_with_person	94.9	49.7	79.0	35.4	34.9	20.5	63.2	58.8
fr2_large_with_loop	12.0	8.8	10.6	5.9	5.48	4.92	54.3	44.1

为了验证经本文方法改进的 ORB-SLAM2 系统在实际室内环境下的定位准确性, 基于 APE 利用无人机采集平台对校园教学楼内场景进行了室内轨迹测试。用于数据采集无人机实验平台及实际环境如图 15 所示, 小车由实验团队自行组建。控制底盘采用松灵机器人公司出产的 MR2000, 在此基础上设计安装了 Intel D435i 深度相机、维特 HWT605-6 轴 IMU、厦门四信 F-V111-RTK 和处理器为 CPU-Intel i7, 运行内存 16 G 的工控电脑。工控机操作系统为 Ubuntu18.04, 对应的 ROS 机器人操作系统为 Melodic。进行测试的楼层环境复杂, 包含桌椅、门窗、行人、展架、标语牌以及纹理特征较少的回廊, 为实验创造了合适的室内环境条件。在该场景下按照有无回环和移动范围的大小分别进行实验。

图 16(a)为无回环小场景下无人机的各算法得到的运动轨迹运动图, 小车从左上角出发以 1.5 m/s 的平均速度绕大厅前台行驶一圈后达到左上角停止运动, 图 16(b)为



图 15 实验小车及实际环境

带回环大场景中无人机的各算法得到的运动轨迹图, 小车从大厅出发经过各教室回廊绕整个楼层行驶一圈后回到大厅, 完成回环。图 16 中黑色虚线代表 GNSS 真值轨迹, 线段分别代表原始 ORB-SLAM2、ORB-SLAM3 和经本文算法改进的 ORB-SLAM2 方案的轨迹。可以看出, 原始 ORB-SLAM2 表现最差, ORB-SLAM3 表现略好, 改进后的 ORB-SLAM2 误差最小, 在两种数据集上都和真值轨迹最为贴近。

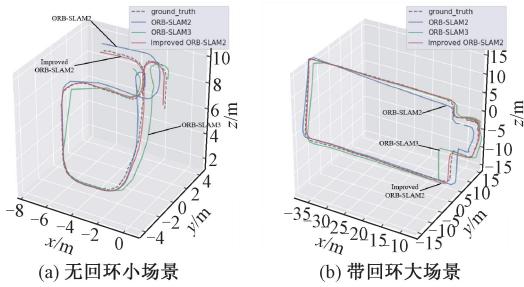


图 16 实际环境轨迹对比图

表 4 更加量化地展示了不同算法在实际环境下的定位效果,仍然采用 RMSE 和 STD 作为评价参数。在无回

环的小场景下,原始的 ORB-SLAM2 由于采集到的特征点较少且存在较多误匹配,影响力系统定位准确性,而改进后的 ORB-SLAM2 算法精度提升明显,均方根误差减小 58.0%,标准差减小 33.7%,并且表现也优于 ORB-SLAM3,在带回环的小场景下由于经过弱纹理的回廊结构,特征点极易丢失,且不利于重定位,但经过回环矫正后三者算法误差均得到减小,但对比来看仍然是改进后的 ORB-SLAM2 算法精度较高,在大场景下利用纯视觉已经达到了分米级,而原始 ORB-SLAM2 和 ORB-SLAM3 均方根误差分别为 2.01 m 和 1.61 m。结合图 16 和表 4 整体上看采用本文方法改进后的 ORB-SLAM2 算法具有更好的鲁棒表现和定位精度。

表 4 实际环境 ATE 误差对比

图像序列	ORB-SLAM2		ORB-SLAM3		Improved ORB-SLAM2		Improvement	
	RMSE/m	STD/m	RMSE/m	STD/m	RMSE/m	STD/m	RMSE/%	STD/%
无回环小场景	3.05	1.01	2.65	1.22	1.28	0.67	58.0	33.7
带回环大场景	2.01	0.89	1.61	0.66	0.87	0.43	56.7	51.7

## 4 结论

传统特征点提取匹配算法在弱纹理、高速移动及光照变化剧烈等环境下容易引起特征点丢失和图像配准失败等问题,从而导致视觉 SLAM 系统定位效果不佳。针对这样的问题,本文提出了一套基于 ORB 的鲁棒性强、精度高的特征点提取匹配算法。首先在特征点提取方面,本文提出了基于网格模型的特征点提取策略,并根据图像灰度信息计算自适应阈值提取特征点,保证了提取的稳定性和均匀性。此外本文还提出了 G-R 图像匹配算法:首先基于网格特征计算邻域支持估计量来进行正误点区分,再建立评价函数进一步确定内点,最后通过 RANSAC 算法迭代模型得到最终的正确匹配点。在各类数据集中通过实验证明了本文提取匹配算法的稳定性和较高的正确率。为了检测其在 SLAM 工程应用所带来的实际效果,我们将其加入到经典的 ORB-SLAM2 算法框架,在 TUM 数据集与自采集环境下进行了大量实验测试,并基于绝对轨迹误差,采用均方根误差与标准差作为测评参数,与主流的视觉 SLAM 算法进行对比,证明了本文提出的特征点提取匹配算法能切实提高视觉 SLAM 系统的定位精度,减小误差,且在不同场景下具备较好的稳定性。后期更多实验中发现当在室外移动物体如人、车等占据画面较大比例时,动态特征点也被用于帧间匹配与位姿估计,会在一定程度上影响系统的精度。因此还考虑加上目标检测算法,结合语义信息对画面中的动态对象特征点实现精准的剔除,从而在动态环境下也能构建全局一致的地图。

## 参考文献

[1] TAKETOMI T, UCHIYAMA H, IKEDA S. Visual

SLAM algorithms: A survey from 2010 to 2016 [J]. IPSJ Transactions on Computer Vision and Applications, 2017, 9(1): 1-11.

- [2] TAREEN S A K, SALEEM Z. A comparative analysis of sift, surf, kaze, akaze, orb, and brisk [C]. International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), IEEE, 2018: 1-10.
- [3] KARAMI E, PRASAD S, SHEHATA M. Image matching using SIFT, SURF, BRIEF and ORB: Performance comparison for distorted images [J]. ArXiv Preprint, 2017, ArXiv:1710.02726, 2017.
- [4] 朱翔宇, 汪丙南, 向茂生. 基于非线性尺度空间 SIFT 的 SAR 图像特征点提取 [J]. 国外电子测量技术, 2016, 35(9): 96-100.
- [5] OYALLON E, RABIN J. An analysis of the SURF method [J]. Image Processing On Line, 2015, 5: 176-218.
- [6] RUBLEE E, RABAUD V, KONOLIGE K, et al. ORB: An efficient alternative to SIFT or SURF [C]. IEEE International Conference on Computer Vision, 2012: 2564-2571.
- [7] MUR-ARTAL R, MONTIEL J M M, TARDOS J D. ORB-SLAM: A versatile and accurate monocular SLAM system [J]. IEEE Transactions on Robotics, 2015, 31(5): 1147-1163.
- [8] MUR-ARTAL R, JUAN D. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras [C]. IEEE Transactions on

- Robotics, 2017.
- [9] 秦晓飞, 皮军强, 李峰. 基于极线约束的 ORB 特征匹配算法 [J]. 计算机应用研究, 2018, 35 (8): 2865-2868.
- [10] BIAN J, LIN W Y, MATSUSHITA Y, et al. GMS: Grid-based motion statistics for fast, ultra-robust feature correspondence [C]. IEEE Conference on Computer Vision & Pattern Recognition, 2017: 2828-2837.
- [11] HAN L, WANG J, ZHANG Y, et al. Research on adaptive ORB-SURF image matching algorithm based on fusion of edge features [J]. IEEE Access, 2022(10): 109488-109497.
- [12] 崔建国, 孙长库, 李玉鹏, 等. 基于 SURF 的快速图像匹配改进算法 [J]. 仪器仪表学报, 2022, 43(8): 47-53.
- [13] 张绍荣, 张闻宇, 李云, 等. 基于 FAST 角点和 FREAK 描述符改进的无人机景象匹配算法 [J]. 电子测量与仪器学报, 2020, 34(4): 102-110.
- [14] GUPTA S, KUMAR M, GARG A. Improved object recognition results using SIFT and ORB feature detector [J]. Multimedia Tools and Applications, 2019, 78: 34157-34171.
- [15] SARKHEL R, DAS N, DAS A, et al. A multi-scale deep quad tree based feature extraction method for the recognition of isolated handwritten characters of popular indic scripts [J]. Pattern Recognition, 2017, 71: 78-93.
- [16] YAO J, ZHANG P, WANG Y, et al. An adaptive uniform distribution ORB based on improved quadtree [C]. IEEE Access, 2019.
- [17] 杜媛. 基于图像质量分析与分割算子的图像特征点匹配算法 [J]. 电子测量技术, 2017, 40(8): 144-149.
- [18] 朱成德, 李志伟, 王凯, 等. 基于改进网格运动统计特征的图像匹配算法 [J]. 计算机应用, 2019, 39(8): 2396-2401.
- [19] YANG S, FAN G, BAI L, et al. MGC-VSLAM: A meshing-based and geometric constraint VSLAM for dynamic indoor environments [J]. IEEE Access, 2020, 8: 81007-81021.
- [20] CAMPOS C, ELVIRA R, RODRIGUEZ J G, et al. ORB-SLAM3: An accurate open-source library for visual, visual-inertial, and multimap SLAM [J]. IEEE Transactions on Robotics: A publication of the IEEE Robotics and Automation Society, 2021(6):37.

## 作者简介

张钧程, 硕士研究生, 主要研究方向为视觉 SLAM 与自主导航。

E-mail: 1316254766@qq.com

柯福阳(通信作者), 博士, 教授, 博士生导师, 主要研究方向为测绘科学与技术、SLAM 无人驾驶特定场景应用。

E-mail: ke.fuyang@qq.com

王旭, 硕士研究生, 主要研究方向为激光即时定位与建图(SLAM)。

E-mail: 980883829@qq.com