

# 基于特征点改进的视觉 SLAM 定位研究

王 伟<sup>1</sup>, 汤琴琴<sup>2</sup>, 汪先伟<sup>1</sup>

(1. 南京信息工程大学 自动化学院, 南京 210000; 2. 无锡学院 轨道交通学院, 江苏 无锡 214000)

**摘要:** 为改善视觉 SLAM 系统在低纹理环境下定位精度较低的现象, 提出一种改进的 ORB 特征点提取策略和一种关键帧选择机制; 首先采用多尺度分析和基于局部灰度的特征检测方法克服一般 ORB 算法缺乏尺度和旋转描述的缺点; 其次提出一种基于高斯模糊的图像信息增强方法解决传统 ORB 特征点提取方法在纹理信息不被突出环境下容易失效的问题, 并对图像进行象限分割使特征点均匀分布; 最后为剔除劣质关键帧, 设计了一种综合时间因素与特征点数量因素的关键帧选择机制; 将提出的方法移植到 ORB\_SLAM2 上, 并在 TUM 数据集上测试, 实验结果表明, 视觉 SLAM 系统的定位误差平均降低 14.688%, 证实了该方法的有效性。

**关键词:** 视觉 SLAM; 低纹理; 特征点; 关键帧; 定位

## Research on Improved Visual SLAM Localization Based on Feature Points

WANG Wei<sup>1</sup>, TANG Qinqin<sup>2</sup>, WANG Xianwei<sup>1</sup>

(1. School of Automation, Nanjing University of Information Engineering, Nanjing 210000, China;

2. School of Rail Transit, Wuxi University, Wuxi 214000, China)

**Abstract:** In order to improve the low localization accuracy of simultaneous localization and mapping (SLAM) system in low texture environment, this paper proposes an improved oriented FAST and rotated BRIEF (ORB) feature point extraction strategy and key frame selection mechanism. Firstly, multi-scale analysis and feature detection method based on local gray level are used to overcome the shortcomings of lack scale and rotation description in general ORB algorithm. Secondly, an image information enhancement method based on Gaussian blur is proposed to solve the problem that the texture information of the traditional ORB feature point extraction method is easy to fail in special environment, and the image is segmented to evenly distribute the feature points. Finally, in order to eliminate the inferior key frames, a key frame selection mechanism combining time factor and feature point number factor is designed. The proposed method is transplanted to the ORB\_SLAM2, and testing on the TUM dataset, the experimental results show that the localization error of the visual SLAM system is reduced by 14.688% on average, which proves the effectiveness of the proposed method.

**Keywords:** visual SLAM; low texture; feature points; key frames; localization

## 0 引言

视觉 SLAM (simultaneous localization and mapping) 是指通过视觉传感器对环境进行感知, 同时实现机器人或无人驾驶汽车在未知环境下的自主定位和建图<sup>[1]</sup>。在解决机器人和自动驾驶等领域的应用问题中, 视觉 SLAM 成为最受重视和广泛应用的技术之一<sup>[2-4]</sup>。通过研究 SLAM 技术, 可以为实现更加智能的、精准的机器人定位导航提供支持<sup>[5]</sup>, 同时也可构建数字地图和城市规划等领域提供数据支撑。

目前, ORB (oriented FAST and rotated BRIEF) 算法作为一种关键点检测与描述子生成算法<sup>[6]</sup>, 在 SLAM 领域被广泛使用。然而, ORB 特征点的数量在一定程度上受限

于图像纹理信息, 即图像分辨率较低、场景光照黑暗或者曝光等极端条件、结构信息较少导致图像平整度较高等因素使得环境的纹理信息不被突出时, 难以提取足额数量的 ORB 特征点。因此, 如何提高 ORB 特征点的数量和质量以提升 SLAM 定位精度成为当前 SLAM 领域需要解决的核心问题之一<sup>[7]</sup>。国内外学者们以往主要通过设立特征点检测阈值、图像预处理、图像处理与阈值结合等方式来解决在低纹理场景下 ORB 特征点数量不足的问题, 使得 ORB 算法更加鲁棒。

为通过设置阈值增强 ORB 算法在低纹理环境下的鲁棒性以便提取更多的特征点, Murartal 等人<sup>[8]</sup>在 ORB\_SLAM2 中设置了固定阈值, 即当最大固定灰度阈值搜索不到特征点时, 系统自动启用最小固定灰度阈值进行 ORB 特

收稿日期: 2023-04-25; 修回日期: 2023-05-07。

基金项目: 第十六批次江苏省“六大人才高峰”高层次人才项目(XYDXX-045); 江苏省自然科学基金面上项目(BK20211037); 南京信息工程大学无锡校区研究生创新实践项目(WXCX202121)。

作者简介: 王 伟(1999-), 男, 硕士研究生。

通讯作者: 汤琴琴(1985-), 女, 高级工程师, 硕士生导师。

引用格式: 王 伟, 汤琴琴, 汪先伟. 基于特征点改进的视觉 SLAM 定位研究[J]. 计算机测量与控制, 2024, 32(2): 219-226.

征点检测。Wu 等人<sup>[9]</sup>采用动态阈值的思想,去除最高亮度和最低亮度的像素后计算平均亮度,通过自适应阈值选择准则为每个点都设置专属阈值,解决 ORB 对光变化的敏感性。焦嵩鸣等<sup>[10]</sup>则针对固定阈值提出了一种全局和局部的自适应阈值结合的思路,针对整体与局部的亮度变化两种情况各自采取全局或半全局自适应阈值策略,并结合于局部自适应阈值,以此增加特征点提取的数量。另外,支持向量机(SVM, support vector machine)<sup>[11]</sup>和卷积神经网络(CNN, convolutional neural networks)<sup>[12]</sup>也被提出用以预测自适应阈值,基于机器学习的自适应阈值方法需要大量训练数据和计算资源的数量,以及对数据分布和功能有很高的要求,应用于现实有很大的局限性。这类方法的缺点:1) 如果采用固定的全局或局部阈值,无法兼顾整个图像的各个部分,难以达到理想的效果;2) 两次阈值可能会造成图像某些区域块的 ORB 特征点被重复提取,有些区域被遗忘检测,不仅使特征点冗余,又影响大规模特征匹配的准确性;3) 阈值的方法依赖图像本身的灰度动态范围,在低纹理环境下,灰度动态范围往往较低,此类方法容易失效,适应性差,而且阈值受环境噪声的干扰,阈值的取值并不准确。

所以,学者们继而研究对图像进行预处理,增强图像的边缘、角落等细节以改变低纹理场景下的灰度范围。Cheng 等人<sup>[13]</sup>通过对图像进行低光增强处理,增强了系统在低光环境的鲁棒性。陆佳嘉等人<sup>[14]</sup>使用二维高斯对图像金字塔进行降噪,并对像素进行划分,提高部分区域提取到特征点的概率。李宪华等人<sup>[15]</sup>通过自适应阈值伽马变换,对低照度、低纹理的图像块进行增强,提高在这些低光环境下提取到的特征点的数量。Dai 等人<sup>[16]</sup>提出一种图像增强与截断自适应阈值相结合的方法改进 ORB 特征提取算法,将原始图像转换为灰度后通过应用高斯滤波、截断自适应伽马亮度调节和非锐化掩蔽增强图像,将图像分成若干子区域,并利用改进的大津法获得相应阈值,以此解决了在光照过多或不足等低纹理环境中无法提取足够特征点的问题。现有方法:1) 低纹理场景变化是复杂的,包括分辨率的高低变化、场景结构信息的差异导致平整度的不同,难以保证 ORB 算法在这些因素下的适应性与稳定性;2) 结合额外的阈值后算法不准确性增加。

此外,关键帧选择在视觉 SLAM 算法中也扮演重要角色,关键帧选择策略的质量直接影响着算法的定位精度和运行效率<sup>[17-19]</sup>。传统的关键帧选择策略通常只考虑了时间因素,根据距离阈值、角度阈值或关键帧密度等指标进行筛选<sup>[20]</sup>,将一些劣质的关键帧纳入队列,难以针对不同类型场景进行调整导致定位精度下降、模型计算量大。

为解决上述问题,本文利用多尺度分析与局部灰度的特征检测为一般 ORB 特征点添加尺度与旋转的描述,利用基于高斯模糊的图像信息增强方法加强 ORB 算法在低纹理环境下的稳定性,象限分割图像以使特征点均匀分布,并在只考虑时间因素的传统关键帧选择机制中兼容了特征点

数量因素,降低定位误差。

## 1 视觉 SLAM 定位方法

基于特征点的视觉 SLAM 的定位过程如图 1 所示。

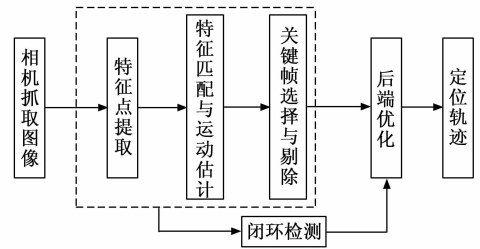


图 1 视觉 SLAM 定位方法

1) 相机调整参数之后利用 ORB 特征提取器检测图像特征点,在当前帧与上一帧之间通过随机采样一致性算法(RANSAC)<sup>[21]</sup>寻找匹配的特征点。

2) 通过特征点坐标差异计算相机运动。

3) 但是为了提升系统响应速度,降低系统定位误差,通过选择合适的关键帧,使系统只保留最必要的信息进行位姿估计。

4) 基于词袋模型构成系统的闭环,用于减少定位过程中的累计误差。闭环检测<sup>[22]</sup>比较当前帧与之前帧,找到或代表同一区域的特征,计算相似度,并在相似度高于一定阈值时确立闭环。

5) 根据运动估计获得的相对位置信息,进行后端的全局优化,使误差最小化。

本文将特征点看作视觉 SLAM 系统正式进行定位的输入数据,对图像中的特征点加以描述,并克服传统 ORB 算法在低纹理的环境下的不稳定性,过滤劣质关键帧,降低视觉 SLAM 的定位误差。

## 2 特征点提取过程以及改进

特征点对视觉 SLAM 定位的重要性是显著的。在视觉 SLAM 中,特征点是用于进行相邻帧间匹配的关键点,它们被用来计算相机的运动和场景的深度信息。特征点数量越充足时,视觉 SLAM 算法在定位中就有越多可用的信息。在视觉 SLAM 定位中,需要依靠匹配当前帧与先前帧之间的特征点来获取相机的运动估计。在低纹理、低分辨的环境中,传统 ORB 算法容易失效导致特征点数量不足,如果特征点数量太少,匹配误差就会增加,从而导致相机姿态估计不准确或者无法成功匹配,进而导致整个 SLAM 系统定位失效。

### 2.1 传统 ORB 特征点提取

传统 ORB 特征点提取方法使用 FAST 角点检测器寻找图像中明显的突变点,并将这些点组成候选关键点,原理如图 2 所示,比较圆心点  $O$  且半径等于 3 的圆上 16 个连续像素点与  $O$  的灰度差来提取 FAST 点。若圆周上连续着  $M$  个 ( $M$  一般取 12) 像素点的灰度值与  $O$  的差已超规定阈值  $T$ ,则判定被选中的圆心  $O$  作为特征点。详细过程如下:

- 1) 在图像上选取一点  $O$ , 灰度值假设为  $I_0$ , 以  $O$  为圆心且 3 为半径作圆, 共有 16 个点。
- 2) 设定灰度阈值  $T$ , 比如  $T=I_0 \times 0.2$ 。
- 3) 若圆上有连续 12 个点的灰度值大于  $T+I_0$  或小于  $T-I_0$ , 判断  $O$  为关键点。

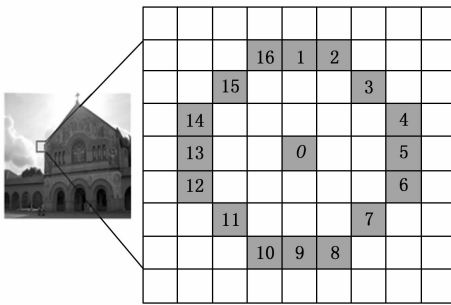


图 2 FAST 角点提取示意图

接着计算每个关键点的描述子, 描述子是用来描述关键点周围区域的一系列二进制的向量, 在 ORB 算法当中, 使用 BRIEF 算法生成每个关键点的描述子, 即在公式 (1) 中:  $P(x)$ 、 $P(y)$  分别为点  $x$ 、 $y$  处的灰度值大小。

$$f(P; x, y) = \begin{cases} 1, & P(y) < P(x) \\ 0, & \text{else} \end{cases} \quad (1)$$

其具体思想是在特征点  $P$  周围以特定模式选取  $N$  对点, 比较灰度值的大小 (比如  $x, y$ )。若  $x$  比  $y$  大, 则取 1, 反之就取 0。

故 BRIEF 描述子可以表示为:

$$F_n(p) = \sum_{1 \leq n \leq N} 2^{n-1} f(P; x_n; y_n) \quad (2)$$

## 2.2 改进的 ORB 特征点提取过程

实际上, 图像中的物体可能会以不同尺度或角度出现, 故提出一种多尺度分析方法, 在不同尺度下检测物体的特征, 即通过对原始图像进行重复缩减或扩大操作, 生成不同分辨率的图像, 每缩放一次, 图像的分辨率就会下降, 并在这些不同尺度下进行特征提取, 使得 ORB 特征点具备尺度不变性。如图 3 所示, 本文将原始图像作为待缩图, 长为  $W_0$ , 宽为  $H_0$ , 以 1.2 为尺度系数进行缩放, 第  $n$  次得到的图像长为  $W_n$ , 宽为  $H_n$ , 面积为  $S_n$ 。

$W_n$  与  $W_0$  的关系为:

$$W_n = \frac{W_{n-1}}{1.2} = \frac{W_{n-2}}{1.2^2} = \dots \frac{W_0}{1.2^n} \quad n = 1, 2, 3 \dots \quad (3)$$

$H_n$  与  $H_0$  的关系为:

$$H_n = \frac{H_{n-1}}{1.2} = \frac{H_{n-2}}{1.2^2} = \dots \frac{H_0}{1.2^n} \quad n = 1, 2, 3 \dots \quad (4)$$

故  $S_n$  可表示为:

$$S_n = W_n H_n = \frac{W_0 H_0}{(1.2^n)^2} \quad (5)$$

由于本文以 1.2 为尺度系数, 所以每缩放一次后的图像面积是上一次的 69.44%。

ORB 算法添加了通过平移模拟旋转的过程来应对

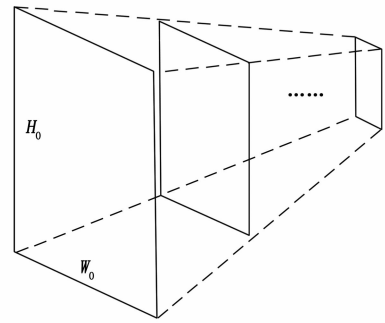


图 3 多尺度缩放示意图

BRIEF 描述子对旋转不太敏感的问题, 本文引入基于局部灰度分布的特征点检测方法以提高 ORB 算法对旋转的稳定性, 即以特征点为原点构建局部坐标系, 在构建好的坐标系中, 通过对邻域内每个像素的灰度加权平均, 计算出邻域内灰度值的中心位置, 最后用特征点指向灰度质心的特征向量代表此特征点的方向。具体步骤如下:

- 1) 定义某图像块  $A$  的矩定义为:

$$m_{pq} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I(i, j) \cdot (i - C_x)^p \cdot (j - C_y)^q \quad (6)$$

其中:  $I(i, j)$  表示在位置  $(i, j)$  处的像素灰度值,  $M$  与  $N$  分别表示图像块  $A$  的宽度与高度,  $p, q$  表示矩阵阶数,  $C_x$  与  $C_y$  分别表示图像块  $A$  灰度质心坐标, 即图像块中所有像素的坐标的加权平均值。

- 2) 图像块  $A$  灰度质心  $C$  的坐标公式为:

$$(C_x, C_y) = \left( \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I(i, j) \cdot i}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I(i, j)}, \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I(i, j) \cdot j}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I(i, j)} \right) \quad (7)$$

- 3) 由于  $p, q = \{0, 1\}$ , 所以  $C$  的坐标可以进一步表示为:

$$(C_x, C_y) = \left( \frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (8)$$

- 4) 将图像块  $A$  几何中心  $O$  与质心相连得到方向向量, 将特征点方向定义为:

$$\theta = \arctan\left(\frac{m_{01}}{m_{10}}\right) \quad (9)$$

在描述子的基础上, 经过多尺度缩放和灰度质心方向描述, ORB 特征点更加稳健, 同时增强算法的性能和稳定性。从角点提取过程可以看出, ORB 的提取依赖角点周围像素的差异程度, 高纹理的图像像素之间的变化明显大于低纹理的图像, 故而更容易提取到特征点。因此, 基于扩大图像灰度动态范围的思想, 在角点检测前提出一种基于高斯模糊的图像信息增强方法, 提升图像灰度动态范围, 突出图像的纹理信息, 从而增加可提取的 ORB 特征点数量。

图像信息增强方法首先对原图像进行 3 个高斯模糊, 得到 3 个不同尺度的模糊图像。因为高斯模糊会平滑掉图像中的细节信息, 所以在每个模糊图像与原图之间做差, 获得这些细节信息, 并通过一定的组合公式将其融合到原

图像中, 进而增强原图像的信息表达能力, 形成更高动态范围的灰度图像。具体步骤如下:

1) 定义待处理图像为  $I^*(x, y)$ , 选择标准差分别为  $\sigma_1, \sigma_2, \sigma_3$  的高斯核函数  $G_1(x, y), G_2(x, y), G_3(x, y)$ , 函数表达式如 (10) 所示, 其中  $\sigma_1 = 1.0, \sigma_2 = 2.0, \sigma_3 = 4.0$  表示模糊程度,  $x$  和  $y$  表示在高斯核函数的中心点的偏移量。

$$G_i(x, y) = \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{(x^2 + y^2)}{2\sigma_i^2}\right), i = 1, 2, 3 \quad (10)$$

2) 利用 3 个不同尺度的高斯核函数分别平滑处理  $I^*(x, y)$ , 得到处理后的图像称为  $B_1(x, y), B_2(x, y), B_3(x, y)$ , 公式如下:

$$\begin{cases} B_1(x, y) = G_1(x, y) * I^*(x, y) \\ B_2(x, y) = G_2(x, y) * I^*(x, y) \\ B_3(x, y) = G_3(x, y) * I^*(x, y) \end{cases} \quad (11)$$

3) 如公式 (12) 所示, 将所得的  $B_1(x, y), B_2(x, y), B_3(x, y)$  分别与原图  $I^*(x, y)$  做差, 获得 3 个不同层次的细节信息, 记作  $D_1(x, y), D_2(x, y), D_3(x, y)$ 。

$$\begin{cases} D_1(x, y) = I^*(x, y) - B_1(x, y) \\ D_2(x, y) = I^*(x, y) - B_2(x, y) \\ D_3(x, y) = I^*(x, y) - B_3(x, y) \end{cases} \quad (12)$$

4) 选择  $w_1, w_2, w_3$  三个加权系数, 通过组合函数得到信息增强后的图像  $D^*(x, y)$ 。组合函数的公式如下:

$$D^*(x, y) = (1 - w_1 \times \text{sgn}(D_1(x, y))) \times D_1(x, y) + w_2 \times D_2(x, y) + w_3 \times D_3(x, y) \quad (13)$$

多次实验对比后得到  $w_1 = w_2 = 0.25, w_3 = 0.5$  时效果最佳。

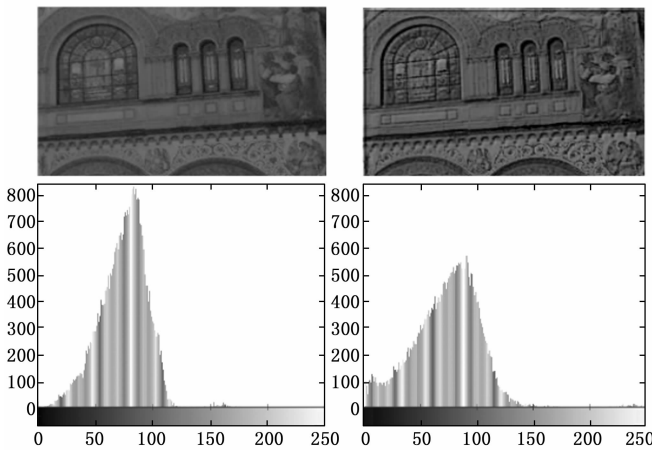
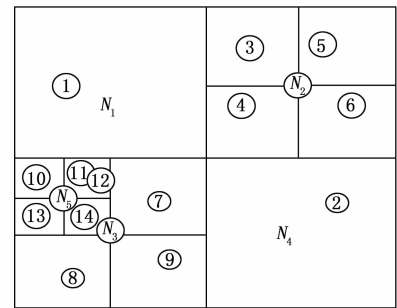


图 4 图像信息增强前后灰度直方图对比

ORB 算法对角点的响应较为敏感, 所以 ORB 算法提取的特征点往往集中在图像角点等边缘区域。此外, 由于采用了在不同尺度空间下检测和描述特征点, 使得特征点分布较为集中, 因为在不同尺度下相邻像素的信息是高度相关的, 因此在不同尺度下提取到的特征点往往存在冗余。本文通过对图像进行象限划分结合非极大值抑制的方法过

滤部分冗余的特征点, 使特征点分布均匀化, 减小计算量, 提高匹配效率, 进而提升视觉 SLAM 的定位精度。

如图 5 所示, 将图像划分为 4 个面积相等的矩形象限, 并作为 4 个原始象限  $N_1, N_2, N_3$  和  $N_4$ , 统计每个象限中的特征点的总数。对于每一个象限, 若包含的特征点数量大于 1, 则作进一步细分, 即将该象限分成 4 个子象限, 继续对子象限进行特征点数量的统计与划分, 直到每个象限中只包含 1 或 0 个特征点, 停止划分。对于已经划分出来的象限进行后处理, 将每个象限中的所有特征点按照响应强度进行排序, 并保留最强的一个特征点, 其余的特征点将被舍弃。在实际应用中, 由于特征点可能会堆积在一起, 导致检测结果出现重叠, 因此需要设定一个参数  $k$ , 限制象限的最大划分次数。当划分次数已达  $k$  时, 对于仍然包含多个特征点的象限, 只保留其中响应强度最高的一个特征点。



(a) 象限划分

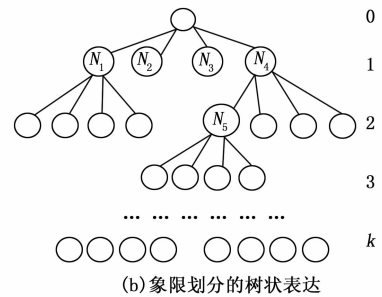


图 5 象限划分原理图

总之, 通过多尺度分析方法及局部灰度特征检测的方法使得 FAST 角点具备尺度和旋转的描述, 利用基于高斯模糊的信息增强预处理视觉传感器抓取的图像, 解决 ORB 算法提取特征点数量存在不足的缺陷, 并对图像进行象限分割融合非极大值抑制将冗余特征点剔除, 使特征点分布均匀化。具体的改进后 ORB 特征点提取流程如图 6 所示。

### 3 关键帧选择机制

视觉 SLAM 系统将前后两帧图像里提取到的特征点进行匹配, 得到它们在两幅图像中的对应关系, 通过对这些点进行几何变换, 估计前一帧到当前帧之间的相对运动, 包括平移与旋转。如图 7 所示,  $O_1$  和  $O_2$  是两个视觉相机中心, 后一帧  $KF_2$  的特征点  $P_2$  是前一帧  $KF_1$  的特征点  $P_1$  通过匹配得到的对应点,  $O_1P_1$  与  $O_2P_2$  交于点  $P$ , 并构成一个

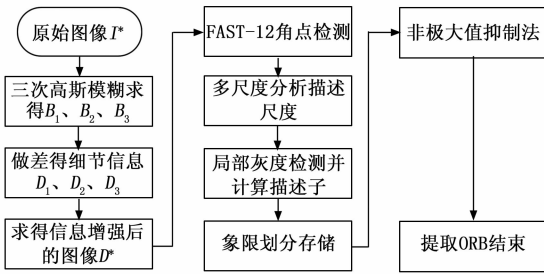


图 6 改进的 ORB 特征点提取过程

极平面,  $O_1O_2$  为基线, 线与  $KF_1$ 、 $KF_2$  交于点  $e_1$ 、 $e_2$ , 极平面与  $KF_1$ 、 $KF_2$  交于极线  $l_1$ 、 $l_2$ 。

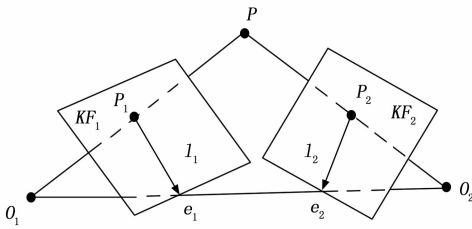


图 7 相对运动估计示意图

假设  $KF_1$  帧与  $KF_2$  帧对应的旋转和平移向量分别是  $\mathbf{R}$ 、 $\mathbf{t}$ , 点  $P$ 、 $P_1$  和  $P_2$  三者之间的关系如式 (14) 所示:

$$\begin{cases} Zp_1 = \mathbf{K}P \\ Zp_2 = \mathbf{K}(\mathbf{R}P + \mathbf{t}) \end{cases} \quad (14)$$

取  $Z=1$  归一化平面, 联解 (14) 中的两个式子, 可以得到式 (15):

$$P_2^T \mathbf{K}^{-T} \mathbf{t} \mathbf{R} \mathbf{K}^{-1} P_1 = 0 \quad (15)$$

其中, 记  $\mathbf{E} = \mathbf{t}^T \mathbf{R}$ ,  $\mathbf{E}$  为本质矩阵 (Essential Matrix), 记  $\mathbf{F} = \mathbf{K}^{-T} \mathbf{E} \mathbf{K}^{-1}$ ,  $\mathbf{F}$  为基础矩阵 (Fundamental Matrix),  $\mathbf{E}$  和  $\mathbf{F}$  之间相差了相机内参  $\mathbf{K}$ 。

将式 (15) 写成本质矩阵  $\mathbf{E}$  的形式, 可以进一步改写为式 (16)。 $\mathbf{E}$  为一个三行三列的矩阵,  $(u_1, v_1, 1)$  和  $(u_2, v_2, 1)$  分别是  $P_1$  和  $P_2$  归一化平面上的投影。

$$[u_1, v_1, 1] \begin{bmatrix} e_1 & e_2 & e_3 \\ e_4 & e_5 & e_6 \\ e_7 & e_8 & e_9 \end{bmatrix} \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = 0 \quad (16)$$

通过八点法, 可以求解本质矩阵  $\mathbf{E}$ , 通过本质矩阵进行奇异值分解求解旋转矩阵  $\mathbf{R}$  和平移向量  $\mathbf{t}$ 。即通过  $P_1$  和  $P_2$  的匹配关系计算本质矩阵, 奇异值分解得到旋转矩阵和平移向量, 这些矩阵和向量描述了相机之间的运动。

利用相对运动信息更新位姿估计, 生成最终的轨迹。但如果所有的帧都用来做匹配和运动估计, 会导致计算量巨大。同时, 由于两帧之间或存在很多相似之处, 使用所有帧进行匹配可能造成冗余信息, 降低 SLAM 系统的定位精度。故需要选择能够提高独特信息的帧作为关键帧, 减少系统计算量和提升系统的效率。关键帧要选取能够代表整个场景的帧, 以保证位置更加精确, 选取描述信息能力强的帧以提高定位精度, 同时时间间隔要适当。一般的关键帧选择机制没有着重考虑图像中的冗余信息, 即便提取

出的特征不同, 但由于图像内容相似, 所选帧之间信息重叠, 导致匹配优化时间过长。为了剔除冗余的劣质关键帧, 本文提出一种综合时间因素和特征点数量因素的关键帧选择机制。

只考虑时间因素时, 首先要保证当前帧的内点数大于最小阈值且重叠度不高, 以确保关键帧跟踪指令的准确性。一旦条件满足, 就会对以下条件进行判断, 满足条件之一, 即可插入关键帧。

1) 为了避免追踪失效, 当与上一个关键帧的时间间隔  $T$  超过预设的最大值  $MAX$  时, 需要插入新的关键帧。

2) 为了避免图像重叠度过高, 当与上一个插入关键帧的时间间隔  $T$  达到预设的最小值  $MIN$  时, 且局部建图线程空余, 插入新的关键帧。

3) 由于局部建图线程处理能力限制, 需要在该线程中保证关键帧队列的帧数量不超过 3 个。

只考虑时间因素, 在视觉相机转弯或移动速度过慢时, 帧与帧之间的信息重叠过大导致定位精度下降, 针对这一问题, 在考虑时间因素的基础上, 本文将特征点因素作为考量, 设计下述劣质关键帧判断机制, 减少计算开销, 提升视觉 SLAM 系统的定位精度。

1) 3 个相邻的关键帧利用随机采样一致性方法匹配 ORB 特征点。

2) 分别将相邻的两组关键帧成功匹配点的数量记为  $Num_1$  和  $Num_2$ , 相隔的两个关键帧成功匹配的特征点数则记为  $Num_3$ 。

3) 如果满足  $Num_3 > Num_1$  和  $Num_3 > Num_2$ , 则中间帧被认为是劣质关键帧。

4) 剔除被选的劣质关键帧, 则进行下一次判断。

本文关键帧选择策略伪代码如下:

```

If  $T > MAX$  Then
    插入关键帧 or
If  $T \geq MIN$  and 局部建图线程空闲 Then
    插入关键帧 or
If 局部建图线程关键帧队列帧数  $< 3$  Then
    插入关键帧
If 关键帧库内关键帧总数超过 3 Then
If  $Num_3 > Num_1$  Then
If  $Num_3 > Num_2$  Then
    剔除劣质的关键帧
EndIf
EndIf
EndIf
EndIf
EndIf
EndIf
    
```

### 4 实验测试与分析

选用 CPU 为 i7-11800H, 内存 16 GB 的计算机, 软件平台为基于 Ubuntu18.04 的 ROS 系统, 选用 Kinect V2 视觉相机拍摄真实图像和 TUM 数据集作为测试数据。

### 4.1 提取特征点的数量实验

为验证本文改进的 ORB 特征点提取方法在不同分辨率、不同光照、不同平整度等因素导致环境纹理信息不被突出时的稳定性,分别选取 3 组各 5 张图像作为实验对象, A1~A5 图像的分辨率分别为 100、150、200、250、300 dpi, B1~B5 图像的光照程度分别为黑暗、稍暗、正常、明亮、高光, C1~C5 图像因为环境结构的差异平整度各不相同。提取效果如图 8 所示, 每组的第一行图像采用传统 ORB 算法提取特征点, 第二行图像利用本文提出改进 ORB 特征点提取策略提取特征点。

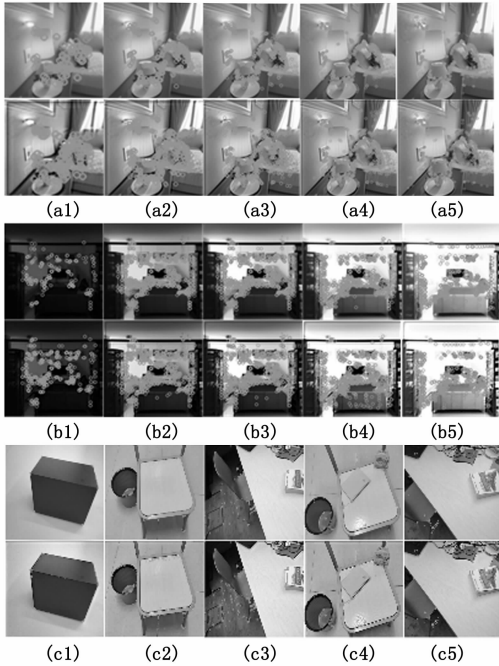


图 8 改进前后提取特征点对比

将传统 ORB 算法提取的特征点总数记为  $Sum_1$ , 本文方法提取的特征点总数记为  $Sum_2$ 。如公式 (17) 所示, 引入 ORB 特征点增长系数  $Q$ , 用于对改进后的 ORB 特征点提取方法在不同因素下的性能表现进行评价,  $Q$  值越大, 说明算法的适应性与稳定性越好。

$$Q = Sum_2 / Sum_1 \quad (17)$$

首先, 通过表 1 数据可以看出, a1 至 c5 所有图像在使用本文改进的方法提取特征点后, 特征点数量均有明显增加, 表明本文方法在不同分辨率、不同光照、不同平整度的情况下均有良好效果。

其次, 经计算, 分辨率因素下的  $Q$  平均值为 1.44; 光照因素下的  $Q$  平均值为 1.31; 平整度因素下  $Q$  平均值为 1.57。对于分辨率因素而言,  $Q$  与分辨率大致成反比; 就光照而言, 越是黑暗和高光这种极端条件下,  $Q$  值越是突出; 就平整度而言, 结构信息越少, 平整度越高,  $Q$  值也越大。

综上所述, 本文改进后的 ORB 特征点能够在纹理信息不被突出的环境下有良好的性能表现, 尤其是在分辨率较低、光照极端、结构信息越少的条件下, 具有良好的适应

性和稳定性。

表 1 图像 ORB 特征点数量对比

类型	序号	$Sum_1$ /个	$Sum_2$ /个	$Q$
分辨率	A1	116	180	1.55
	A2	303	560	1.85
	A3	536	802	1.50
	A4	760	919	1.21
	A5	914	979	1.07
光照	B1	228	465	2.04
	B2	807	871	1.08
	B3	820	892	1.09
	B4	775	848	1.09
	B5	581	724	1.25
平整度	C1	63	149	2.37
	C2	561	980	1.75
	C3	628	925	1.47
	C4	835	941	1.13
	C5	818	911	1.11

### 4.2 均匀度实验

通过均匀度量, 对比在真实场景下传统 ORB 特征点提取方法、ORB\_SLAM2 和本文方法在提取特征点上均匀化的表现。

实验采用的均匀化评价方法如下:

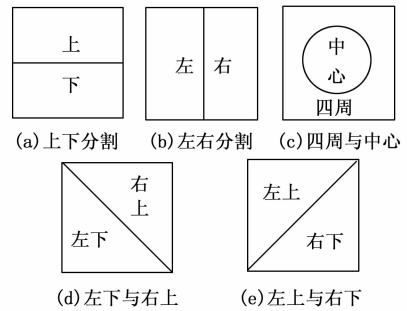


图 9 图像分割示意图

先对图像进行水平、垂直、中心和四周 (中心与四周面积相等)、左上至右下和左下至右上 5 种方式分割得到 10 个区域。分别计算每个区域内的特征点数量, 并构建一个区域统计分布向量。再计算特征点区域统计分布的标准差, 以此表示均匀度, 具体如公式 (18) 所示。标准差越小, 特征点分布越是均匀, 反之亦然。

$$S = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_{10} - \bar{x})^2}{10}} \quad (18)$$

其中:  $S$  是特征点分布均匀度,  $x_1 \dots x_{10}$  是划分的 10 个区域内各自的特征点总数。 $\bar{x}$  是图像特征点总数的平均数。为确保均匀度计算的正确性, 取每张在 3 种方法下都能提取到特征点总数为 1000 个特征点的图像, 测试结果如图 10 所示。

第一行是传统 ORB 算法测试结果, 第二行是本文提出的改进 ORB 特征点提取方法的测试结果。根据式 (18) 量化后的结果如表 2 所示。

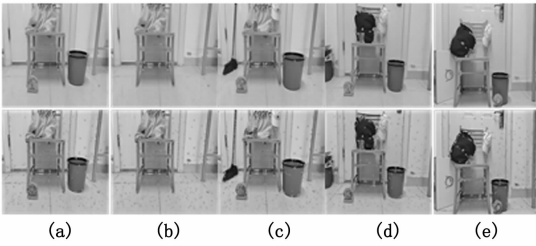


图 10 均匀度测试结果

表 2 均匀度量化结果图像

图像	传统	ORB_SLAM2	本文
1	199.82	174.65	160.75
2	209.56	177.17	158.37
3	229.86	183.13	172.21
4	194.78	172.96	165.47
5	180.08	157.28	146.41
平均	202.82	173.04	160.68

通过表 2 计算, 本文方法在均匀度性能上相较于 ORB\_SLAM2 与传统 ORB 特征点提取方法, 分别提升了 7.14% 和 20.78%, 表明了改进过后的算法可使特征点分布更加均匀。

### 4.3 关键帧选择机制效果测试

为验证在只考虑时间因素的传统关键帧选择机制上融合特征点数量因素的方法更加有效, 将本文提出的关键帧选择机制移植于 ORB\_SLAM2, 与只使用传统关键帧选择机制的 ORB\_SLAM2 共同在 TUM 数据集上进行测试, 在考虑了特征点数量因素之后, 视觉 SLAM 需要处理的关键帧数量减少, 结果如表 3 所示。

表 3 关键帧数量 个

数据集	时间因素	时间+数量因素	剔除
fr1_desk2	134	126	8
fr1_desk	120	103	17
fr3_nstf	48	39	9
fr1_teddy	246	233	13
fr3_cabinet	176	166	10

将移植前后视觉 SLAM 系统求解的轨迹与真实轨迹进行均方根误差 (RMSE) 误差对比, RMSE 越小说明系统定位精度越高, 轨迹越拟合于真实轨迹, 定位精度结果如表 4 所示, RMSE 的减少说明了剔除劣质关键帧可以降低视觉 SLAM 的定位误差, 从而证明本文的关键帧选择机制可以优化关键帧质量。

表 4 剔除劣质帧前后定位精度对比

数据集	移植前	移植后	提升/%
fr1_desk2	0.031	0.029	6.452
fr1_desk	0.029	0.027	6.896
fr3_nstf	0.039	0.037	5.128
fr1_teddy	0.054	0.051	5.556
fr3_cabinet	0.056	0.053	5.357

### 4.4 轨迹精度测试

为验证本文提出的改进 ORB 特征点提取和设计的关键帧选择机制结合使用的有效性与可行性, 首先, 将本文提出的改进 ORB 特征点提取策略与 ORB\_SLAM2 结合为 ORB\_V1 系统, 精度提升结果如表 5 所示。

表 5 ORB\_V1 提升效果

数据集	ORB_SLAM2	ORB_V1	提升/%
fr1_desk2	0.031	0.027	12.903
fr1_desk	0.029	0.026	10.345
fr3_nstf	0.039	0.036	7.692
fr1_teddy	0.054	0.049	9.259
fr3_cabinet	0.056	0.050	10.714

在 ORB\_V1 系统的基础上结合本文设计的关键帧选择机制, 组成 ORB\_V2 系统, 精度测试结果如表 6 所示。

表 6 ORB\_V2 提升效果

数据集	ORB_SLAM2	ORB_V2	提升/%
fr1_desk2	0.031	0.026	16.129
fr1_desk	0.029	0.024	17.241
fr3_nstf	0.039	0.034	12.821
fr1_teddy	0.054	0.047	12.963
fr3_cabinet	0.056	0.048	14.286

如图 11 所示, 以 fr3\_cabinet 为例, ORB\_SLAM2 求解出的轨迹 traj1 的 RMSE 为 0.056, 在未添加本文提出的关键帧选择机制时, ORB\_V1 求解轨迹 traj2 的 RMSE 为 0.050, ORB\_V2 求得轨迹 traj3 的 RMSE 是 0.048。相比于 traj2 和 traj3, traj1 因为环境纹理太少, 一般 ORB 算法几乎失效, 没有足够的特征点用来计算相机的运动, 使得估算轨迹初始阶段很大程度偏离真实轨迹, 且导致中途定位跟踪失败。traj3 在 traj2 的基础上, 再加入了综合时间因素和特征点因素的关键帧选取机制, 降低了劣质关键帧对定位精度的影响, 定位精度从 10.714% 进一步提升至 14.286%, 更贴合真实轨迹。

## 5 结束语

本文利用多尺度分析与局部灰度的特征检测为一般 ORB 特征点添加尺度与旋转的描述, 提升 ORB 算法的稳定性, 利用基于高斯模糊的信息增强方法克服传统 ORB 算法在纹理信息不被突出的环境下的不稳定性, 通过对图像进行象限分割去除多余的特征点, 使特征点均匀分布, 提升视觉 SLAM 系统的定位精度。在只考虑时间因素的传统关键帧选择机制中结合了特征点数量因素, 过滤劣质关键帧, 进一步降低定位误差。

经计算, 整个系统的定位精度误差平均减少了 14.688%。但是由于提取了更多的特征点, 造成运动估计计算需要更长的运行时间, 因此可考虑将深度学习与视觉 SLAM 结合, 提升 ORB 算法识别的速度, 或是研究基于 SLAM 方法的传感

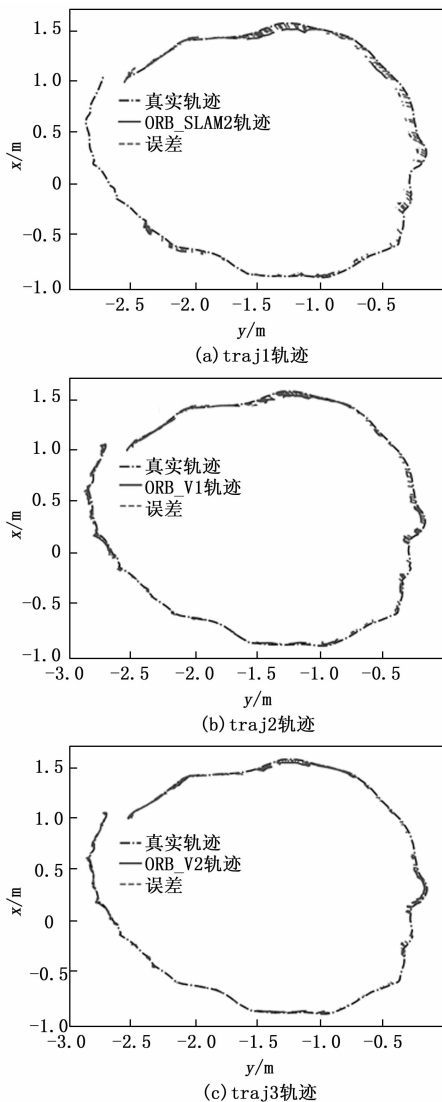


图 11 以 fr3\_cabinet 为例的轨迹对比

器(例如激光雷达、惯性测量器件等)数据融合技术,以提高定位精度。

参考文献:

[1] 王 霞, 左一凡. 视觉 SLAM 研究进展 [J]. 智能系统学报, 2020, 15 (5): 825-834.

[2] 阴贺生, 裴 硕, 徐 磊, 等. 多机器人视觉同时定位与建图技术研究综述 [J]. 机械工程学报, 2022, 58 (11): 11-36.

[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] 吴建清, 宋修广. 同步定位与建图技术发展综述 [J]. 山东大学学报(工学版), 2021, 51 (5): 16-31.

[5] 陈 常, 朱 华, 由韶泽. 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 (3): 641-647.

[6] 崔学荣, 周伟帅, 李 娟, 等. 同步定位与建图特征提取和匹

配算法研究 [J]. 导航定位学报, 2022, 10 (5): 121-127.

[7] 李国竣, 徐延海, 段杰文, 等. 利用局部自适应阈值方法提取 ORB-SLAM 特征点 [J]. 测绘通报, 2021 (9): 32-36.

[8] MURARTAL R, TARDOS 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.

[9] WU G Y, ZHANG Z P. An Improved ORB feature extraction and matching algorithm [C] // Chinese Control and Decision Conference (CCDC), Kunming, China, 2021: 7289-7292.

[10] 焦嵩鸣, 钟宇飞, 姚 鑫, 等. 一种改进自适应阈值的 ORB 算法研究 [J]. 激光杂志, 2022, 43 (6): 85-90.

[11] FAISA D A, AGUS H. Object searching on video using ORB descriptor and support vector machine [C] // Advances in Computational Collective Intelligence: 12th International Conference (ICCCI), DaNang, Vietnam, 2020: 239-251.

[12] GORLA B, PINJARI A K. Elephant herding with whale optimization enabled orb features and CNN for iris recognition [J]. Multimedia Tools and Applications, 2022, 81 (4): 5761-5794.

[13] CHENG W, ZHANG Y, QI Y. A fast global adaptive solution to low-light images enhancement in visual SLAM [C] // IEEE 6th International Conference on Computer and Communications (ICCC), Chengdu, China, 2020: 1874-1879.

[14] 陆佳嘉, 柯福阳, 余晓栋, 等. 机器人 SLAM 的特征点实时提取改进 [J]. 计算机工程与设计, 2022, 43 (6): 1768-1776.

[15] 李宪华, 田硕宇, 宋 韬. 局部自适应图像增强的 ORB-SLAM2 算法 [J]. 安徽理工大学学报(自然科学版), 2022, 42 (5): 15-23.

[16] DAI Y, WU J. An improved ORB feature extraction algorithm based on enhanced image and truncated adaptive threshold [J]. IEEE Access, 2023, 11: 32073-32081.

[17] 徐 武, 高 寒, 王欣达, 等. 改进 ORB-SLAM2 算法的关键帧选取及地图构建研究 [J]. 电子测量技术, 2022, 45 (20): 143-150.

[18] XIE P, SU W, LI B, et al. Modified keyframe selection algorithm and map visualization based on ORB-SLAM2 [C] // International Conference on Robotics and Automation Sciences (ICRAS), 2020: 142-147.

[19] YOUNES G, ASMAR D, SHAMMAS E, et al. Keyframe-based monocular SLAM: Design, survey, and future directions [J]. Robotics and Autonomous Systems, 2017 (98): 67-88.

[20] 白乐乐. 室内动态场景下的视觉 SLAM 算法研究 [D]. 西安: 西安理工大学, 2021.

[21] ZHU W, SUN W, WANG Y, et al. An improved RANSAC algorithm based on similar structure constraints [C] // Singapore: International Conference on Robots & Intelligent System (ICRIS), 2016: 94-98.

[22] 杨雪梅, 李帅永. 移动机器人视觉 SLAM 回环检测原理现状及趋势 [J]. 电子测量与仪器学报, 2022, 36 (8): 1-12.