

基于均匀 ORB 特征的回环检测算法

陈绵书¹, 于录录¹, 李晓妮¹, 郑宏宇²

(1. 吉林大学 通信工程学院, 长春 130022; 2. 吉林大学 汽车仿真与控制国家重点实验室, 长春 130022)

摘要: 针对传统视觉即时定位与地图构建(SLAM)中 ORB 特征存在的聚集问题, 基于网格划分和关键点分层确定思想, 设计了均匀 FAST 角点提取方法, 进而设计了基于均匀分布的 ORB 特征结合暴力匹配的回环检测方法。与基于词袋(BoW)模型的回环检测算法对比实验表明, 本文算法能显著提高回环检测的准确率。基于机器人操作系统(ROS)平台, 将均匀 ORB 特征回环检测模块与直接稀疏里程计(DSO)相结合, 设计了一种松耦合式的半直接法 SLAM 系统。实验结果表明, 本文系统具有较高的地图构建性能。

关键词: 信息处理技术; 即时定位与地图构建; 回环检测; 均匀 ORB; 暴力匹配; 直接稀疏里程计

中图分类号: TP391.4 文献标志码: A 文章编号: 1671-5497(2023)09-2666-10

DOI: 10.13229/j.cnki.jdxbgxb.20211293

Loop detection based on uniform ORB

CHEN Mian-shu¹, YU Lu-lu¹, LI Xiao-ni¹, ZHENG Hong-yu²

(1. College of Communication Engineering, Jilin University, Changchun 130022, China; 2. The State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun 130022, China)

Abstract: Aiming to solve the aggregation problem of traditional uniform oriented FAST and rotated BRIEF (ORB) features in visual simultaneous localization and mapping (SLAM), a uniform FAST corner extraction method is designed, which is based on grid division and laying to determinate key points. Furthermore, a corresponding loop detection method is designed based on uniform distribution of ORB features combined with brute force matching. Experiment results compared with BoW-based loop detection algorithms show that the proposed algorithm can significantly improve the accuracy of loop detection. Furthermore, a robot operating system (ROS) based loosely coupled semi-direct SLAM system is designed, which combine the uniform ORB feature loop detection module with direct sparse odometry (DSO). The experimental results show that the proposed system has high map construction performance.

Key words: information processing technology; simultaneous localization and mapping (SLAM); loop detection; uniform oriented FAST and rotated BRIEF (ORB); brute force matching; direct sparse odometry (DSO)

收稿日期: 2021-11-29.

基金项目: 国家自然科学基金项目(61771220); 吉林省科技厅重点研发项目(20210203149SF, 20210201079GX).

作者简介: 陈绵书(1973-), 男, 副教授, 博士. 研究方向: 图像处理, 计算机视觉. E-mail: chenms@jlu.edu.cn

0 引言

视觉即时定位和地图构建(Simultaneous localization and mapping, SLAM)^[1,2]是自动驾驶、无人机、虚拟现实和增强现实等新兴技术的基本构建模块,主要由相机传感器、前端、后端、回环检测和地图构建5部分组成。相机传感器采集用于建图的图像序列,它们在时间上彼此相邻。前端根据图像序列计算帧间的位姿和图像中路标点的深度信息。由于这些信息并不准确,因此后端模块对前端计算出的数据进行优化,以获得较为准确的位姿和空间结构信息。地图构建模块利用后端优化后的位姿和空间结构信息建立点云地图。由于内存和计算资源的限制,视觉SLAM系统在完成上述计算任务过程中采用滑动窗口的形式,即位姿的计算和优化被限定在相邻数帧范围内。每次计算和优化都会存在误差,尽管很小,但不能被忽略。误差的长时间积累会导致后续位姿估计和空间结构估计越来越不准确,无法建立全局一致的地图。

回环检测模块通过发现当前场景与历史某一场景高度相似,并验证确定为同一地点,以建立当前帧与历史帧的约束,进而完成位姿矫正。回环检测一直是视觉SLAM系统中较难解决的问题,若回环检测成功,则能够通过位姿矫正消除累积误差,获得全局一致地图。若回环检测失败,则估计的位姿、地图与真实值会有较大误差,使系统无法完成其他复杂任务。因此,回环检测的准确率至关重要。

目前,在主流的视觉SLAM系统框架中,回环检测模块多基于图像间的匹配实现,即将当前帧与历史帧进行比较,计算它们之间的相似度,若相似度超过设定阈值,则判定为发生回环。因此,回环检测问题转变为图像间特征点提取和匹配问题,核心在于特征点的提取、描述子的设计和选取以及相似度的计算。Angeli等^[3]将文本检索领域的词袋(Bag of word, BoW)模型应用于回环检测中,提出采用BoW对图像进行描述并完成图像之间的匹配。Cummins等^[4,5]提出了基于外观的快速映射(Fast appearance based mapping, FAB-MAP)算法,并将其应用于大场景中,其回环检测效果十分优异。梁志伟等^[6]以FAB-MAP为基础,将离线训练字典替换为在线构建字典,进一步提高了系统的实时性。文献^[7]提出了增量式的BoW模型,利用透视不变二值特征描述子完善回

环检测模型,取得了较好的回环检测结果。Oliva等^[8]提出了Gist描述子,它使用Gabor滤波器从频率和方向两个维度提取图像信息,使得Gist成为描述图像的全局描述子。文献^[9]基于Gist描述子提出了一种描述场景信息的全局图像特征,以此优化了BoW模型。

近年来,深度学习技术的兴起促进了回环检测方法的进一步发展。Yang等^[10]提出了一种并行循环搜索和验证方法,将BoW特征与深度神经网络特征相结合,实现了高精度和高速回环检测。Bai等^[11]提出了一种基于卷积神经网络(Convolutional neural network, CNN)特征的回环检测方法,将预训练的CNN中间层输出与传统的基于序列匹配过程输出相结合,从而能够处理视点和条件变化。为了降低全局最优搜索策略的高计算复杂度,Wang等^[12]开发了一种两阶段的回环检测策略:前端序列节点级匹配基于运动过程的局部连续性约束,避免了对全局最优匹配的盲目搜索;后端图像级匹配结合改进的语义模型DeepLab_AE,使用CNN作为特征检测器提取视觉描述符。文献^[13]采用深度反卷积网络将场景表示为一个低维矢量,通过比较矢量实现回环检测。Liu等^[14]借鉴BoW模型思想,将CNN特征聚类生成CNN词,进一步为了编码空间信息,创造了CNN词对,取得了优于直接使用CNN特征的检测效果。

在视觉SLAM系统中用于特征提取的BoW模型,多以ORB(Oriented FAST and rotated BRIEF)特征^[15]为基础。该特征的提取是以图像中的FAST角点为关键点,对于场景中变化平缓的区域不能提取足够数量的特征。本文提出了基于均匀ORB的特征提取方法,克服了经典ORB特征对图像内容描述不充分的问题。进一步,在直接稀疏里程计(Direct sparse odometry, DSO)框架中提出了回环检测模块,该方法显著提高了DSO框架的建图性能。

1 均匀ORB特征

1.1 FAST角点检测原理

ORB特征是图像中关键点处的特征,采用FAST算法^[16]提取图像中的关键点。若某个像素比周围像素过亮或过暗,则该像素可能是一个角点。图1为一幅图像的FAST角点检测示例。

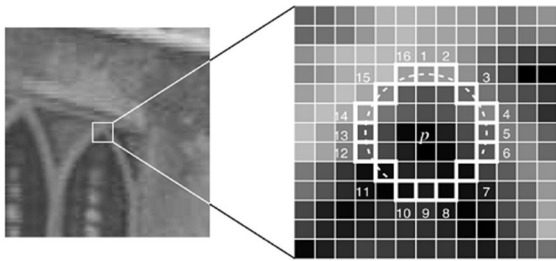


图1 FAST角点

Fig.1 FAST corner

图1中,以点 p 为圆心,以3个像素为半径,若圆周上的像素亮度值比点 p 亮度值高,那么点 p 可能就是一个角点。FAST角点检测步骤如下所示。

- (1)选取一个像素点 p ,设其亮度为 I_p 。
- (2)设定合适的阈值 t 。
- (3)以点 p 为圆心、3个像素为半径作圆,选取该圆周上的点,一共有16个点。
- (4)比较 I_p 与第(3)步中的16个点的亮度值,若这16个点中连续 N 个点亮度值均大于 $I_p + t$,或均小于 $I_p - t$,则点 p 是一个角点。
- (5)循环以上步骤,对每个像素执行同样操作。

1.2 FAST角点的均匀化

图2为OpenCV中FAST角点分布。如图2所示,采用1.1节中阐述的方式提取FAST角点会出现“聚集”现象。“聚集”现象体现在两个方面:一是纹理较丰富的区域,检测出的角点较多,但是平面图像中的角点较多并不意味着空间三维结构的复杂;二是纹理较弱的区域,检测出较少的角点,这意味着对这部分区域三维结构表述不充分。在视觉SLAM中,这种聚集现象或者导致地图信息冗余,或者导致地图信息缺失,造成位姿估计精度较低,不利于视觉SLAM的应用。

针对这种现象,本文提出提取较为均匀的FAST角点的改进方法^[17],均匀FAST角点分布如图3所示。设角点数量为 N ,均匀FAST角点的提取步骤如下所示。

- (1)将每层金字塔图像分割为 30×30 像素的网格。
- (2)以 t_1 为阈值在该层每个网格内提取FAST角点;若提取不到,则以 t_2 为阈值继续提取($t_2 < t_1$)。
- (3)将每层图像划分为 4×3 个等大的正方形格子,将不包含和只包含1个FAST角点的格

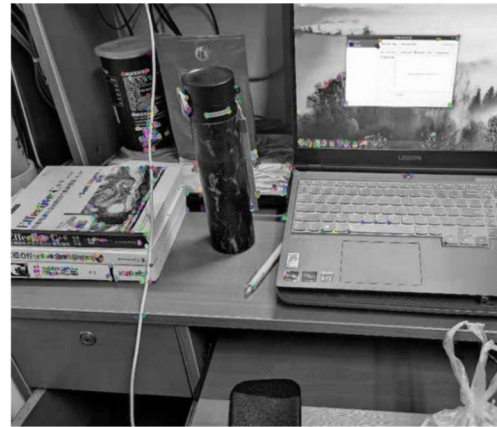


图2 OpenCV中FAST角点分布

Fig.2 FAST corner distribution in OpenCV

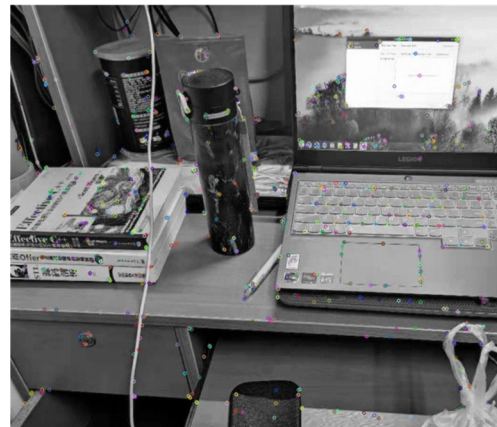


图3 均匀FAST角点分布

Fig.3 Uniform FAST corner distribution

子标记为不可分割,检查包含FAST角点的格子数量,如果大于等于 N ,则跳至第(5)步,否则执行第(4)步。

(4)将每个未被标记的格子分为4个等大的正方形格子,并将不包含和只包含1个FAST角点的格子标记为不可分割,检查包含FAST角点的格子数量,如果大于等于 N ,则执行第(5)步,否则继续执行第(4)步。

(5)选取 N 个包含FAST角点的格子,在每个格子内,选取Harris响应最大的角点作为特征提取的关键点。

1.3 均匀ORB特征点

1.3.1 灰度质心法

FAST角点检测算法在整个检测过程中并没有为角点分配方向。因此,当图像发生旋转后,FAST角点没有旋转不变性,便无法利用其进行图像匹配。ORB特征点提取过程中,利用图像块的灰度质心为FAST角点分配方向信息。图像块的灰度质心通过如下步骤确定。

(1)选取图像块 B ,计算该图像块的矩 m_{pq} ,计算方法为:

$$m_{pq} = \sum_{x,y \in B} x^p y^q f(x,y), \quad p, q \in \{0, 1\} \quad (1)$$

式中: $f(x,y)$ 为图像在 (x,y) 处的取值。

(2)利用图像块的矩计算它的灰度质心 C ,计算方法为:

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (2)$$

(3)设 B 的几何中心为 O ,连接 OC 得到一个方向向量 L_{Oc} ,则角点的方向定义为:

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

上述过程为FAST角点分配了方向,使其具备旋转不变性,提高了角点在不同图像间的鲁棒性。在ORB中,具备方向信息的FAST被称为Oriented FAST。

1.3.2 BRIEF描述子

提取FAST角点并为其分配方向后,生成角点描述子。ORB利用改进的二进制鲁棒独立基本特征(Binary robust independent elementary features, BRIEF)为角点计算描述子。该描述子本质上是二进制的向量,向量中的1和0分别为角点周围某两个像素点(设为 p, q)的亮度值关系:若 p 大于 q ,则取1;若 p 小于 q ,则取0。若选取256个 p, q 点对,则BRIEF描述子就是256维由1和0组成的二值向量。 p, q 按照某种概率分布进行随机选取,速度极快。二进制向量对存储空间的消耗较小,适合实时对图像间的特征点进行匹配。

ORB特征提取算法为了使BRIEF描述子具备旋转不变性,在选取 p, q 点对之前,将像素坐标轴进行旋转,使 x 轴方向与角点的灰度质心方向相同,形成了Rotated BRIEF。

均匀ORB特征点的提取过程如下所示。

(1)建立图像金字塔。

(2)利用1.2节中的方法对每层图像提取均匀FAST角点。

(3)使用灰度质心法计算每个FAST角点的方向。

(4)对每层图像进行高斯模糊。

(5)利用BRIEF计算每个FAST角点的描述子。

(6)将所有FAST角点恢复到原始图像中。

2 特征点的匹配

本文采用暴力匹配结合交叉匹配验证的方式进行两幅图像之间的特征点匹配。设待匹配的两幅图像分别为 f_1 和 f_2 (f_1 为查询图像, f_2 为参考图像), p_1 为 f_1 中的一个特征点。暴力匹配计算 p_1 和 f_2 中所有特征点的汉明距离,选择距离最小的特征点 p_2 作为 p_1 的匹配点。

暴力匹配会产生大量误匹配点。误匹配的消除可以采用随机采样一致性(Random sample consensus, RANSAC)和交叉匹配验证两种方式。RANSAC方法通过计算单一性矩阵验证匹配点,比较适合平面结构。本文采用交叉匹配验证方式消除误匹配点。首先,将 f_1 作为查询图像, f_2 作为参考图像,进行暴力匹配;然后,将 f_2 作为查询图像, f_1 作为参考图像,进行暴力匹配。若两次匹配结果中, p_1 与 f_2 中的特征点 p_2 均为匹配对,则认为 p_1 和 p_2 是正确的匹配,否则丢弃这对匹配点。

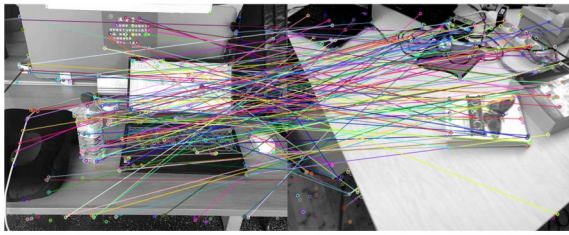
3 相似度计算

首先,分别对查询图像 f_1 和参考图像 f_2 完成均匀ORB特征的提取;然后,利用暴力匹配和交叉匹配验证得到一定数量的特征点对,设共有 N 对。上述过程为查询图像 f_1 中的每一个特征点在参考图像 f_2 中找到最合适的匹配点,但并不意味着每一个匹配对都来自三维空间中的同一点,仍然存在大量的误匹配。图4和图5分别为来自不同地点和相同地点的匹配结果。由图4(a)和图5(a)可以看出:通过暴力匹配和交叉匹配验证,不同地点和相同地点都会得到大量的匹配点对,不能根据匹配结果进行回环检测。

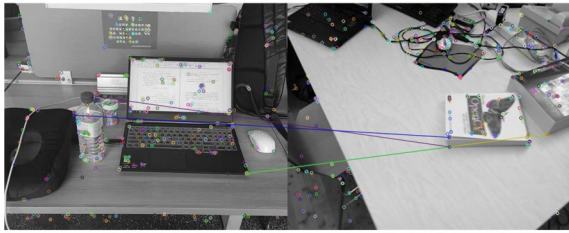
为了较准确地检测出回环,本文设定匹配点对的汉明距离阈值对匹配点对进行筛选,丢弃汉明距离超过阈值的误匹配点对。图4(b)和图5(b)分别为不同地点和相同地点筛选后的结果。由图可以看出:通过阈值筛选,不同地点和相同地点的匹配点对存在较大差异,可以用于进行回环检测。若经过筛选后的匹配点对数记为 M (N 个匹配点对中汉明距离小于阈值的匹配点对数量),则图像间的相似度 $s(f_1, f_2)$ 可定义为:

$$s(f_1, f_2) = M/N \quad (4)$$

式(4)的本质含义是汉明距离满足阈值条件的匹配点数占总匹配点数的比例,利用该公式能



(a)暴力匹配结果



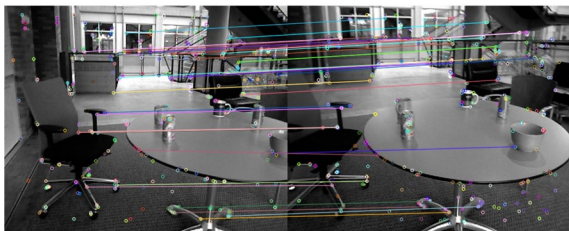
(b)基于汉明距离筛选后结果

图 4 不同地点匹配示意图

Fig. 4 Schematic diagram of different place



(a)暴力匹配结果



(b)基于汉明距离筛选后结果

图 5 同一地点匹配示意图

Fig. 5 Schematic diagram of matching at same place

够得到任意两幅图像间的相似性评分,但对于场景相似的情况,效果可能会很差。比如两间不同卧室有同样的布局,同样的床具,则式(4)的值会很高,但两间卧室并不是同一个地点。鉴于上述情况,使用如下方法进行改进:利用式(4)计算当前帧与前一帧的相似性评分,记为 s_1 ,再计算当前帧与历史帧的相似性评分,记为 s_2 ,若 s_2 与 s_1 的比值超过设定阈值,则认为该历史帧是回环帧,即采用相对相似度 $s_{relative}(f_1, f_2)$ 进行评分:

$$s_{relative}(f_1, f_2) = s_2 / s_1 \quad (5)$$

4 算法流程

设 f_i 为当前帧、 u_i 为 f_i 的均匀 ORB 特征向量、

t 为阈值,则基于均匀 ORB 特征的回环检测算法流程如下所示。

(1)对 f_i 提取均匀 ORB 特征向量 u_i ,并保存于参考向量集合中。

(2)若 f_i 为第 1 帧或第 2 帧,则返回第(1)步,继续处理下一帧;否则,对 f_i 与 f_{i-1} 进行暴力匹配与交叉匹配验证,利用式(5)计算 s_1 。

(3)计算 f_i 与历史帧集合 $F = \{f_1, f_2, \dots, f_{i-2}\}$ 中的每一帧的相似度 $s_2, j = 1, 2, \dots, i - 2$ 。

(4)计算 $s_{relative}$,若其大于 t ,则判定为发生回环。

(5)返回与 f_i 构成回环的帧集合。

5 基于 ROS 的视觉 SLAM 系统设计

机器人操作系统(Robot operating system, ROS)^[18]是一种开源的机器人元操作系统,提供操作系统应有的服务,比如硬件描述、消息通信、驱动管理和功能执行等。其具有点对点设计、多语言支持、精简和集成以及工具包丰富等特点。

本文基于 ROS 设计视觉 SLAM 系统对 DSO^[19]进行改进,融入本文提出的基于均匀 ORB 特征的回环检测模块。DSO 是一种基于直接法的视觉里程计,包括追踪线程和局部优化线程,它选取稀疏的像素点用于追踪和建图。DSO 没有回环检测模块,因此只能建立局部的点云地图,无法保证其全局的一致性,不是一个完整的视觉 SLAM 系统。基于此,将第 4 节设计的回环检测算法应用于 DSO,构成完整的视觉 SLAM 系统,进而达到消除 DSO 生成轨迹累积误差的目的。

DSO 以优化的形式完成初始化和追踪任务,其后端维护一个滑动窗口,窗口内包含 5~7 个关键帧以及一定数量的活跃地图点。当参考帧 f_i 中的点 p 在目标帧 f_j 中被观测到时,两点之间的光度误差 E_{pj} 定义为:

$$E_{pj} = \sum_{p \in N_p} w_p \left\| (f_j(p') - b_j) - \frac{t_j e^{d_j}}{t_i e^{d_i}} (f_i(p) - b_i) \right\|_y \quad (6)$$

$$w_p = c^2 / (c^2 + \|\nabla f_i(p)\|_2^2) \quad (7)$$

$$p' = \prod_c (R \prod_c^{-1}(p, d_p) + \tau) \quad (8)$$

式中: p 对应于图 6 中的黑色点; N_p 为图 6 中灰色

点和黑色点的集合,是由 p 确定的邻域; p' 为点 p 在目标帧 f_j 中的对应点; t_i, t_j 分别为参考帧和目标帧的曝光时间; a_i, b_i 分别为参考帧的光度变换参数; a_j, b_j 分别为目标帧的光度变换参数; $\|\cdot\|_y$ 为 Hubber 范数; Π_c 为利用小孔成像模型将三维空间点投影到像素平面的变换; Π_c^{-1} 为其逆变换; $R \in SO(3)$ 为两帧之间的旋转矩阵; τ 为两帧之间的平移向量; c 为某一常数,用于防止梯度接近于 0 时权重系数过大; d_p 为 p 点的逆深度值。

滑动窗口内整体的待优化目标函数 E_{total} 可以表示为:

$$E_{\text{total}} = \sum_{i \in F} \sum_{p \in P_i} \sum_{j \in \text{obs}(p)} E_{pj} \quad (9)$$

式中: F 为滑动窗口内所有关键帧的集合; P_i 为帧 i 内的所有点; j 为滑动窗口内所有能观测到点 p 的关键帧; $\text{obs}(p)$ 为滑动窗口内所有能观测到点 p 的关键帧。

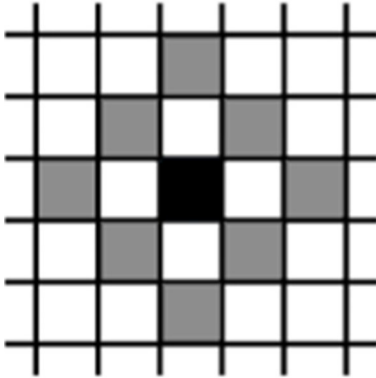


图 6 N_p 集合示意图

Fig. 6 Schematic diagram of N_p collection

当有新的关键帧加入窗口中时,便利用 Gauss-Newton 或 Levenberg-Marquardt 进行优化,待优化变量的更新步长 $\Delta\xi$ 为:

$$\Delta\xi = -(J^T W J)^{-1} J^T W r \quad (10)$$

式中: r 为残差向量; J 为其雅克比矩阵; W 为权值矩阵; ξ 包含滑动窗口内所有待优化的变量,如相机位姿、光度变换参数、点的逆深度和相机内参。

为保持窗口内的关键帧数量在一定范围内,DSO 会利用舒尔补进行关键帧以及点的边缘化。当点的主导帧被边缘化或者点在最新创建的两个关键帧内无法被观测到时,会被边缘化。当某一关键帧内不到 5% 的点能在最新关键帧中被观测到时,该帧会被边缘化;滑动窗口内关键帧数量超过最大值时,最大距离分数 Dis 的关键帧会被边缘化。 Dis 的计算公式为:

$$\text{Dis}(f_i) = \sqrt{d(i, 1)} \sum_{j \in \{3, n\} \setminus \{i\}} (d(i, j) + \epsilon)^{-1} \quad (11)$$

式中: $d(i, j)$ 为 f_i 与 f_j 两帧之间的欧式距离; ϵ 为一个较小的常数。

上述关键帧的边缘化方式能够使关键帧在空间内有较好的分布,即更多的关键帧靠近最新的关键帧。

整个系统利用 ROS 搭建,以 DSO 作为消息的发布者节点,将其边缘化的关键帧和点以消息的形式发布给订阅者节点,这里的订阅者即回环检测模块。由于 DSO 在边缘化关键帧时,并不总是边缘化在时间上最久远的帧,因此回环检测模块需要检查关键帧的时间戳信息,将时间戳超前于当前帧的关键帧直接抛弃。当确定候选关键帧后(设当前帧为 g , 候选帧为 can), 计算候选帧与当前帧的相似变换 $S_{g\text{can}} \in \text{Sim}(3)^{[21, 22]}$, $\text{Sim}(3)$ 为相似变换群:

$$\text{Sim}(3) = \left\{ S = \begin{pmatrix} sR & t \\ 0 & 1 \end{pmatrix} \middle| R \in SO(3) \right\} \quad (12)$$

$$t \in \mathbb{R}^{3 \times 1}, s \in \mathbb{R}$$

若候选关键帧的位姿为 T_{can} , 则当前帧的位姿 T_g 可以表示为 $T_g = T_{\text{can}} S_{g\text{can}}$ 。然后,进行当前帧邻域与回环帧邻域的回环融合工作,参与融合的信息包括位姿和地图点的深度信息。随后,利用位姿图优化进行仅位姿的优化,将累积误差平摊到候选关键帧与当前帧之间的每一个关键帧上。最后,进行全局优化,包括所有关键帧的位姿和地图点信息。如果在视觉 SLAM 系统中确实发生了回环,通过上述步骤,回环检测模块能够生成全局一致的地图,并且得到较为精确的关键帧的位姿信息。

整个视觉 SLAM 系统的追踪和优化基于直接法完成,而回环检测模块是基于特征点法,所以本文设计的系统是半直接法的 SLAM 系统。利用 ROS 将 DSO 封装成消息发布者,将回环检测模块封装成消息订阅者,使得二者之间并没有完成状态的共享。前端与后端能够实时地完成追踪、优化和局部地图构建工作,而回环检测模块由于需要进行均匀 ORB 特征点的提取和逐帧判断相似度操作,其运行速度要慢于 DSO 前端和后端,实时性较差,是一种松耦合式的设计。总体而言,上述设计的是一种松耦合式的半直接法视觉 SLAM 系统。整个系统结构如图 7 所示。

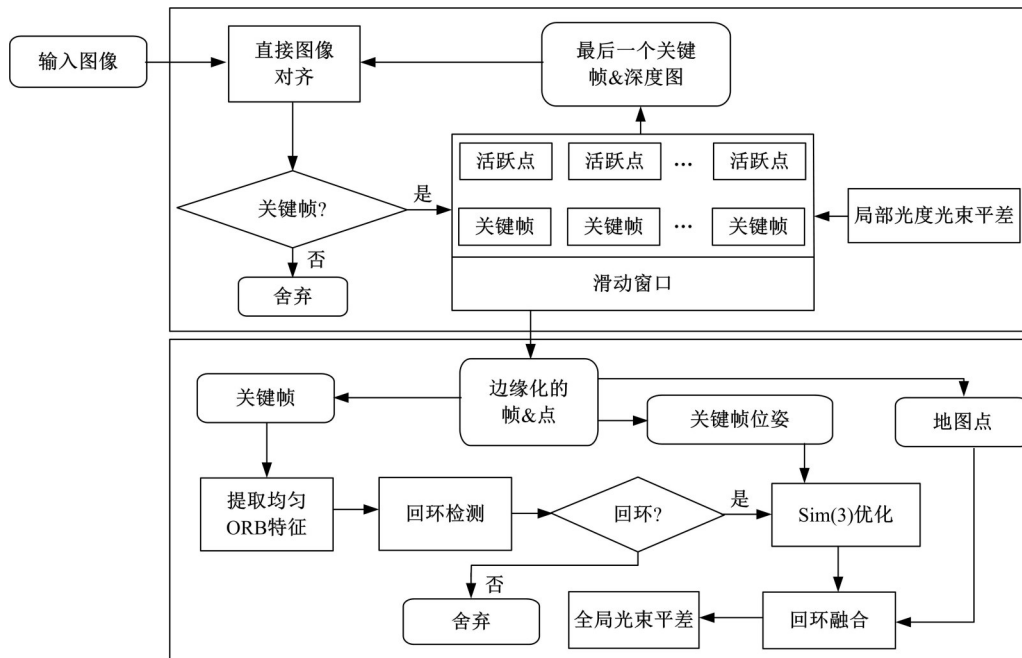


图 7 视觉 SLAM 系统组成框图

Fig. 7 Visual SLAM composition block diagram

6 实验及结果讨论

6.1 实验环境和数据集

实验在笔记本电脑 Lenovo LEGION R7000P 上完成,其具体配置如表 1 所示。实验中使用的第 3 方软件库包括 OpenCV、Eigen3、SuiteSparse、G2O、Pangolin 等。设计视觉 SLAM 系统时使用的 ROS 版本为 ROS Melodic。

表 1 实验环境

Table 1 Experiment environment

处理器	AMD Ryzen7 4800H
显卡	NVIDIA RTX2060
内存	Samsung DDR4 16 GB
固态硬盘	Western Digital 512 GB
操作系统	Ubuntu 18.04 LTS

实验分为两个部分:第 1 部分将基于均匀 ORB 特征的回环检测算法与基于 BoW 的回环检测算法^[20]进行对比分析。使用英国牛津大学 Mobile Robotics Group 的两个数据集 NewCollege 和 CityCentre,分别包含 2146 和 2474 幅图像。图 8(a)(b)分别为两个数据集中的几幅示例图片。图像的分辨率为 640 像素×480 像素。

第 2 部分实验是将本文设计的视觉 SLAM 系统与 DSO 进行对比分析,使用的数据集为 TUM_monoVO^[23],它提供了 50 个图像序列,包含了从室内到室外非常广泛的环境。每一个序列



(a)NewCollege示例图像

(b)CityCentre示例图像

图 8 数据集示例图像

Fig. 8 Datasets example images

的起点和终点均来自同一个地点,可以通过累计漂移评估算法的性能。数据集中的图像分辨率为 1280 像素×1024 像素。图 9 为 TUM_monoVO 数据集的示例图像,每个序列展示一幅图像。



图 9 TUM_Mono 数据集示例图像

Fig. 9 Example images of TUM_Mono dataset

6.2 实验及结果分析

6.2.1 回环检测算法对比分析

本部分实验对比基于均匀 ORB 特征的回环检测算法与基于 BoW 的回环检测算法的匹配性能和性能。两种方法进行相同的参数设置:

特征提取数量为500;金字塔层数为8;构建金字塔的尺度因子为1.2;FAST角点初始阈值为20;BRIEF描述子长度为256 bit。对于基于均匀ORB特征的算法,FAST算法初始阈值为20,最小阈值为15;对于BoW算法^[20]中的ORB特征,FAST角点阈值为20。

(1)匹配性能分析。匹配性能通过准确率-召回率曲线下的面积(Area under curve, AUC)值进行衡量,其含义是回环检测算法的平均准确率。AUC的计算公式为:

$$AUC = \frac{1}{2} \sum_{i=1}^{N-1} (p_i + p_{i+1})(r_{i+1} - r_i) \quad (13)$$

式中: N 为数据集中图像的数量; p_i 、 r_i 分别为点 i 处的准确率和召回率。

图10为两种算法在两种数据集上的AUC值,基于均匀ORB特征的回环检测算法相比于基于BoW的算法,在NewCollege数据集上平均准确率提高了17.4%,在CityCentre数据集上平均准确率提高了18.1%。两个数据集上的实验结果说明,本文算法能够有效提高回环检测的平均准确率。

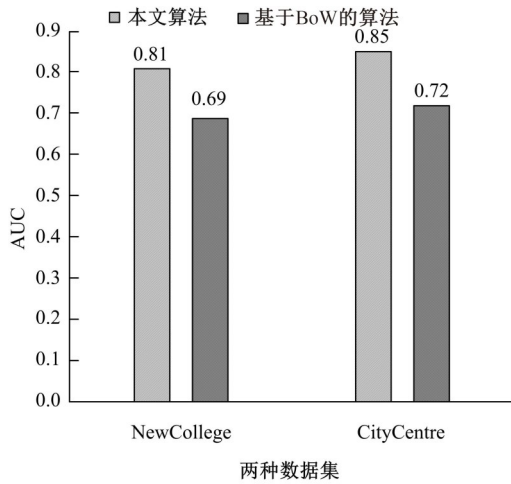


图10 两种算法的AUC对比

Fig. 10 AUC comparison of two algorithms

(2)时间性能分析。表2和表3分别为两种算法在两个数据集上的时间性能表现。对于基于均匀ORB特征的回环检测算法,特征编码时间为平均每幅图像均匀ORB特征点的提取时间,特征匹配时间为两幅图像进行均匀ORB特征点匹配的平均时间。对于基于BoW的回环检测算法,特征编码时间为平均每幅图像ORB特征提取时间与构建单词向量的时间之和,该算法没有特征匹配步骤,故不包含特征匹配时间。在NewCollege

表2 两种算法在NewCollege数据集上的时间性能

Table 2 Time performance on NewCollege of two algorithms

算法	特征编码 时间/ms	特征匹配 时间/ms	相似度计算 时间/ms
均匀ORB	16.21	15.17	0.01
基于BoW	14.09	—	0.05

表3 两种算法在CityCentre数据集上的时间性能

Table 3 Time performance on CityCentre of two algorithms

算法	特征编码 时间/ms	特征匹配 时间/ms	相似度计算 时间/ms
均匀ORB	17.38	16.30	0.01
基于BoW	15.25	—	0.05

数据集上,基于均匀ORB特征的算法完成一次检测需要的总时间为: $t_1=16.21+15.17+0.01=31.39$ ms,基于BoW的算法完成一次检测需要的总时间为: $t_2=14.09+0.05=14.14$ ms,在总时间上,基于均匀ORB特征的算法是基于BoW算法的2.22倍。在CityCentre数据集上,基于均匀ORB特征的算法完成一次检测需要的总时间为: $t_3=17.38+16.30+0.01=33.69$ ms,基于BoW的算法完成一次检测需要的总时间为: $t_4=15.25+0.05=15.30$ ms,在总时间上,基于均匀ORB特征的算法是基于BoW算法的2.20倍。对于时间性能的分析说明,基于均匀ORB特征的回环检测算法是以牺牲时间性能提高匹配性能。

6.2.2 视觉SLAM系统对比分析

将由视觉SLAM系统得到的位姿、轨迹与由DSO得到的位姿、轨迹进行对比分析。为说明不确定因素的影响,对TUM_monoVO数据集的50个图像序列每一个运行10次,共运行500次。图11为本文设计的SLAM系统与DSO在不同对齐误差下的成功运行次数。图11中横坐标为对齐误差,纵坐标为所有500次运行中,对齐误差小于横坐标值的次数。由于对齐误差大于10 m的位姿结果与真实值误差较大,并没有被计入曲线中。对齐误差的定义来自于文献[23]。从图11中可以看出,在相同对齐误差下,本文设计的系统在TUM_monoVO数据集上成功运行次数要高于DSO,意味着失败次数更少;在相同成功运行次数下,本文系统具有较低的对齐误差,验证了本文回环检测算法的有效性。

图12为3组本文设计的SLAM系统与DSO

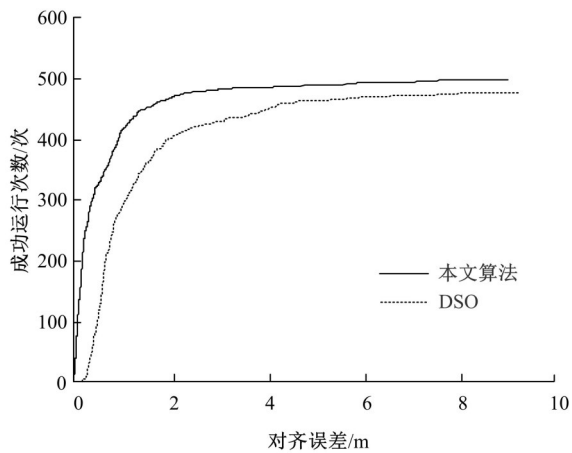
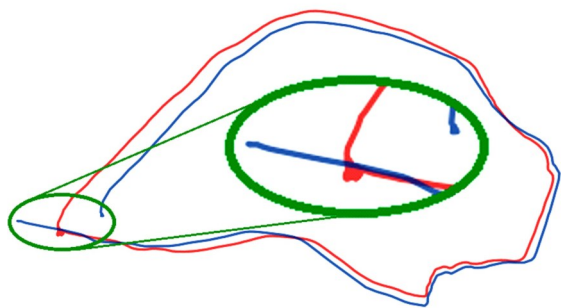
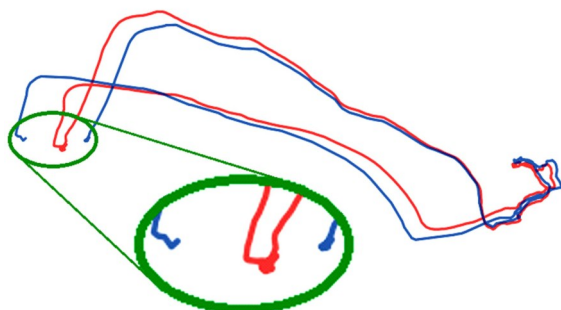


图 11 成功运行次数对比

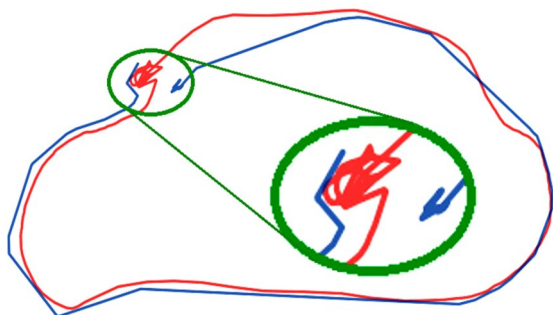
Fig. 11 Comparison of successful runs



(a)序列21轨迹对比



(b)序列22轨迹对比



(c)序列26轨迹对比

图 12 生成轨迹对比

Fig. 12 Comparison of generated trajectories

生成轨迹的对比。真实轨迹为三维轨迹,为了方便对比,将轨迹投影在水平面上。绿色椭圆标出了序列的起点和终点,在没有累积误差的情况下,轨迹的起点与终点应该是重合的。图中蓝色轨迹为 DSO 生成结果,红色轨迹为本文 SLAM 系统生成结果。由图 12 可以看出,DSO 生成的轨迹在起点和终点处无法重合,本文设计系统的生成轨迹可以重合。这是由于 DSO 是一种纯视觉里程计,它只包含位姿的估计和局部优化,没有回环检测模块,故产生较大的累积误差。而本文设计的系统包含回环检测模块,能成功检测到回环并矫正位姿,实验结果验证了本文基于均匀 ORB 特征的回环检测算法的有效性。

7 结束语

本文提出了一种基于均匀 ORB 特征的回环检测算法,在牺牲部分时间性能的情况下,提高了回环检测的平均准确率。将 ROS 平台与 DSO 整合在一起,组建了一种松耦合式的半直接法 SLAM 系统,解决了 DSO 没有回环检测模块导致其生成轨迹累积误差较大的问题。实验结果表明,基于均匀 ORB 特征的回环检测算法的匹配性能优于基于 BoW 的算法,并能有效地完成回环检测任务,减小 DSO 生成轨迹的累积误差。

参考文献:

- [1] Durrant-Whyte H, Bailey T. Simultaneous localization and mapping: Part I[J]. IEEE Robotics & Automation Magazine, 2006, 13(2): 99-108.
- [2] Bailey T, Durrant-Whyte H. Simultaneous localization and mapping (SLAM): part II [J]. IEEE Robotics & Automation Magazine, 2006, 13(3): 108-117.
- [3] Angeli A, Filliat D, Doncieux S, et al. Fast and incremental method for loop-closure detection using bags of visual words[J]. IEEE Transactions on Robotics, 2008, 24(5): 1027-1037.
- [4] Cummins M, Newman P. FAB-MAP: probabilistic localization and mapping in the space of appearance[J]. International Journal of Robotics Research, 2008, 27(6): 647-665.
- [5] Cummins M, Newman P. Appearance-only SLAM at large scale with FAB-MAP 2.0[J]. International Journal of Robotics Research, 2011, 30(9): 1100-1123.
- [6] 梁志伟, 陈燕燕, 朱松豪, 等. 基于视觉词典的单目视觉闭环检测算法[J]. 模式识别与人工智能, 2013,

- 26(6): 561-570.
- Liang Zhi-Wei, Chen Yan-Yan, Zhu Song-Hao, et al. Loop closure detection algorithm based on monocular vision using visual dictionary[J]. *Pattern Recognition and Artificial Intelligence*, 2013, 26(6): 561-570.
- [7] Zhang H, Liu Y L, Tan J D. Loop closing detection in RGB-D SLAM combining appearance and geometric constraints[J]. *Sensors*, 2015, 15(6): 14639-14660.
- [8] Oliva A, Torralba A. Modeling the shape of the scene: a holistic representation of the spatial envelope[J]. *International Journal of Computer Vision*, 2001, 42(3): 145-175.
- [9] Yang L, Hong Z. Visual loop closure detection with a compact image descriptor[C]//2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vilamoura, Portugal, 2012: No. 13195481.
- [10] Yang Z, Pan Y, Deng L, et al. PLSAV: Parallel loop searching and verifying for loop closure detection[J]. *IET Intelligent Transport Systems*, 2021, 15(5): 683-698.
- [11] Bai D D, Wang C Q, Zhang B, et al. CNN feature boosted SeqSLAM for real-time loop closure detection[J]. *Chinese Journal of Electronics*, 2018, 27(3): 488-499.
- [12] Wang Z, Peng Z, Guan Y, et al. Two-stage vSLAM loop closure detection based on sequence node matching and semi-semantic autoencoder[J]. *Journal of Intelligent & Robotic Systems*, 2021, 101(2): 1-21.
- [13] Mukherjee A, Chakraborty S, Saha S K. Detection of loop closure in SLAM: a DeconvNet based approach[J]. *Applied Soft Computing Journal*, 2019, 80: 650-656.
- [14] Liu Qiang, Duan Fu-hai. Loop closure detection using CNN words[J]. *Intelligent Service Robotics*, 2019, 12(4): 303-318.
- [15] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF[C]//Proceedings of the IEEE International Conference on Computer Vision, Barcelona, Spain, 2012: 2564-2571.
- [16] Rosten E, Drummond T. Machine learning for high-speed corner detection[C]//Computer Vision-ECCV 2006: Part I, Berlin, Germany, 2006: 430-443.
- [17] 于录录. 视觉SLAM中回环检测算法的研究[D]. 长春: 吉林大学通信工程学院, 2021.
- Yu Lu-lu. Research on loop detection in visual SLAM[D]. Changchun: College of Communication Engineering, Jilin University, 2021.
- [18] Quigley M, Conley K, Gerkey B, et al. Ros: an open-source robot operating system[J/OL]. [2021-03-22]. <http://www.robotics.stanford.edu/~ang/papers/icraoos09-ROS.pdf>
- [19] Engel J, Koltun V, Cremers D. Direct sparse odometry[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(3): 611-625.
- [20] 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.
- [21] Arun K S, Huang T S, Blostein S D. Least-squares fitting of two 3-D point sets[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1987, 9(5): 698-700.
- [22] Horn B K P. Closed-form solution of absolute orientation using unit quaternions[J]. *Journal of the Optical Society of America A*, 1987, 4(4): 629-642.
- [23] Engel J, Usenko V, Cremers D. A photometrically calibrated benchmark for monocular visual odometry[J/OL]. [2021-03-25]. <https://arxiv.org/abs/1607.02555>