光电技术及应用

DOI：10.16818/j.issn1001－5868.2020.04.019

融合特征法与直接法的 RGB-D 稠密视觉 SLAM 算法

胡章芳*，张 杰，程 亮

（重庆邮电大学 信息无障碍工程研发中心，重庆 400065）

摘 要： 为了保持直接法的快速性与特征法的高精度和闭环能力，提出了一种融合直接法与特征法的 RGB-D 同时定位与地图创建（SLAM）算法。该算法主要包含 3 个并行线程：跟踪线程、局部建图线程和闭环线程。在跟踪线程中对非关键帧进行跟踪，通过最小化光度图像误差来进行相机的初始位姿估计以及像素点的对应关系计算，利用最小化局部地图点重投影误差进一步优化相机位姿，实现快速准确的跟踪与定位；在局部建图线程中对关键帧进行提取并匹配 ORB 特征，执行局部 BA（光束平差法），对局部关键帧位姿和局部地图点的位置进行优化，提高 SLAM 的局部一致性；在闭环线程中执行对关键帧的闭环检测和优化，从而保证 SLAM 全局一致性。另外，根据 RGB-D 图像和相机位姿信息，通过基于 Octomap 的建图框架，构建完整准确的 3D 稠密环境地图。在 TUM 数据集下的实验表明，所提出的方法可以得到与基于特征法相当的精度，且所需时间更少。

关键词： 同步定位与地图创建；直接法；特征法；建图；闭环

中图分类号： TP242 **文章编号：** 1001－5868(2020)04－0548－07

RGB-D Dense Visual SLAM Algorithm Combining Feature Method and Direct Method

HU Zhangfang, ZHANG Jie, CHENG Liang

(Engineering Research Center of Information Accessibility, Chongqing University of Posts & Telecommunications, Chongqing 400065, CHN)

Abstract: In order to maintain the fast performance of the direct method and the high precision and loop closure capability of the feature-based method, a RGB-D simultaneous localization and mapping (SLAM) algorithm combining the direct method and the feature-based method is proposed. The proposed algorithm is composed of three parallel threads: tracking thread, local mapping thread and loop closing thread. In the tracking thread, the non-key frames are tracked, the initial pose estimation and the corresponding relationship calculation of pixel points are carried out by minimizing the photometric image errors, and the camera pose is further optimized by minimizing reprojection errors of the local map points to achieve fast and accurate tracking and positioning. In the local mapping thread, the ORB features are extracted and matched features in the key frames, and the local BA (Bundle Adjustment method) is performed to optimize the position and posture of local key frames and the location of local map points, so as to improve the local consistency of SLAM. In the loop closing thread, the loop detection and the loop optimization for key frames are executed, to enhance the global consistency of SLAM. In addition, according to the RGB-D image and camera pose information, a complete and accurate 3D dense environment map is constructed through Octomap-based mapping framework. Experiments on TUM datasets show that the proposed method

收稿日期：2020－01－01.

基金项目：国家自然科学基金项目(51604056)；重庆市科委项目(csts2017jcyjAX0427).

* **通信作者：**胡章芳 **E-mail：** 495075688@qq.com

achieves the same accuracy as feature-based method with less time.

Key words: simultaneous localization and mapping; direct method; feature-based method; map building; loop closing

0　引言

同时进行定位与地图构建(SLAM)是机器人领域的一个关键挑战和研究重点。未知环境中的自主移动机器人在执行各种任务时需要对环境地图进行构建,例如路径规划、避障和探索等。一方面,移动机器人必须知道其位置才能创建地图;另一方面,移动机器人需要通过地图来获取位置信息。因此,移动机器人必须对环境地图进行创建,同时在地图中定位。

完整的视觉 SLAM 框架由跟踪前端、优化后端、闭环检测以及地图构建等四部分组成。当跟踪丢失时,闭环检测可以实现重定位,对于长时间导航的正确性起着至关重要的作用。视觉 SLAM 中的视觉里程计(Visual Odometry,VO)[1]可分为两类:直接法与特征法。特征法[2]通过提取图像中的特征点和描述子,然后进行特征点匹配,并通过最小化重投影误差计算来获得机器人的位姿。匹配的特征点可用于闭环检测和重定位,这使得 VO 成为完整的 SLAM。然而,特征提取与匹配需要耗费大量的计算时间,使得特征法比直接法耗时更多。直接法[3]通过最小化光度误差来获得机器人的位姿信息,无需进行特征提取。然而基于直接法的 SLAM 闭环检测仍有一个待解决的问题:长时间导航出现的漂移问题。

文献[4]提出了一种基于特征法的 PTAM (Parallel Tracking and Mapping)算法,它属于早期典型的基于非线性优化后端的 SLAM 方法,可以实现跟踪与建图线程的并行化,还可有效依赖基于关键帧的光束平差法(BA)[5]进行实时运行。此后,大多数基于特征的方法都是 PTAM 的改进版本,其中一个是 ORB-SLAM[6]。ORB-SLAM 对于剧烈运动具有鲁棒性,支持宽基线的闭环检测和重定位,以及全自动的初始化,创新性地使用了三线程来完成 SLAM。但它们均采用单目相机,无法恢复真实的位姿信息和场景尺度。文献[7]在 ORB-SLAM 的基础上作了改进,它支持单目、双目以及 RGB-D 三种模式;针对无法建图的问题,提出了一种轻量级的定位模式,能够更加有效地重用地图。虽然 ORB-SLAM 已成为主流特征点 SLAM 方法中的代表,但其也存在一些不足:由于其采用特征法进行计算,需要对每张图像的 ORB 特征进行计算,这会耗费大量时间;另外,ORB 特征采用稀疏特征点进行建图,只能满足定位需求,无法实现导航、避障等诸多功能。

LSD-SLAM (Large-Scale Direct Monocular SLAM)[8]是一种典型的直接 SLAM 算法,其无需对特征点进行计算,还可以对半稠密地图进行构建,能够在没有 GPU 加速的情况下实时运行。但 LSD-SLAM 对于相机的内参和曝光十分敏感,同时在相机快速运动时容易丢失。在 LSD-SLAM 算法的基础上,DSO(Direct Sparse Odometry)[9]针对一个由若干关键帧组成的滑动窗口进行参数优化,并且提出了光度标定,对相机的曝光时间、暗角、伽马响应等参数进行标定后,能够提升直接法的鲁棒性。但是 DSO 不是一个完整的 SLAM,它不具有闭环检测、地图重用以及丢失后的重定位功能,而这些功能在实际应用场景中是不可或缺的。直接法与特征法各具优势,从目前的结果来看,最佳的直接法(即 DSO)在速度与精度方面都比最佳的特征法(即 ORB-SLAM)性能更优。但是 DSO 缺乏闭环检测,不利于长时间工作。因此如何结合这两种方法的优点是本文研究的出发点。

本文提出一种融合特征法与直接法的 RGB-D 稠密视觉 SLAM 算法,以结合直接法的快速性能与直接法的高精度和闭环能力。该算法对关键帧进行提取并匹配 ORB(Oriented FAST and Rotated BRIEF)特征[10],通过最小化光度误差对非关键帧进行跟踪,而不需要特征提取与匹配。关键帧用于局部建图线程与闭环线程,而非关键帧用于跟踪与定位,这使得本文算法耗时更少。此外,本文算法还可根据 RGB-D 图像和相机位姿信息,通过在基于八叉树的建图框架 Octomap[11]上构建 3D 占用网格地图,为移动机器人完成导航和避障等任务提供更加准确的环境地图。

1　算法框架与流程

本文算法在 ORB-SLAM2 的基础上进行研究,

系统构成示意图如图 1 所示，其中红色三角形表示关键帧，白色三角形表示非关键帧。系统基于场景的变化来判断当前帧是否为关键帧，进而提取与匹配每一个关键帧中的 ORB 特征，并且采用不需要特征提取与匹配的直接法来跟踪每个非关键帧。系统选择 ORB 特征，它含有旋转多尺度 FAST 角点和对应的 256 位描述子。此特征可以很快地被提取与匹配，同时具有较好的旋转不变性。关键帧之间的位姿优化通过特征匹配来实现，局部建图线程与闭环检测线程仅与关键帧有关。

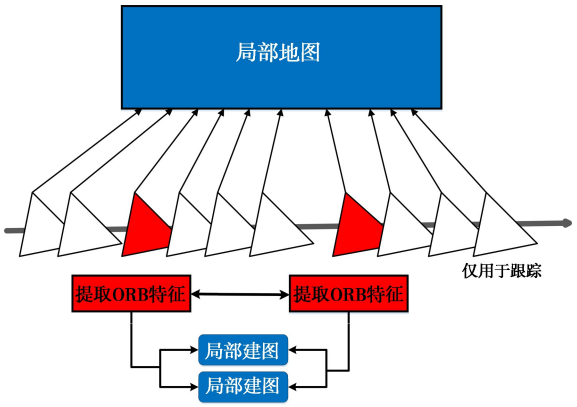


图 1 系统构成示意图

算法线程框图如图 2 所示，其中金色边框模块表示与 ORB-SLAM2 的不同之处，灰色边框模块表示与 ORB-SLAM2 相同。

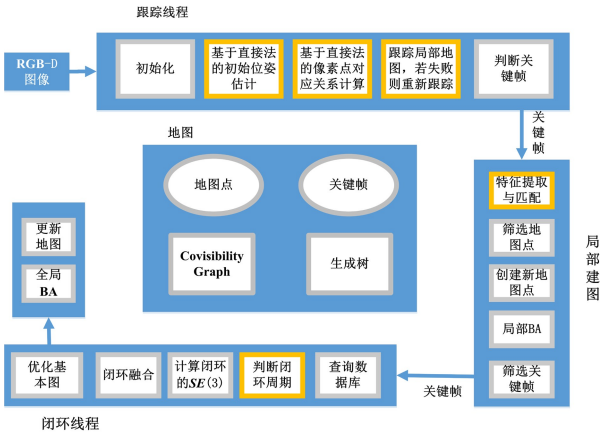


图 2 算法线程框图

该算法有三个并行线程：跟踪、局部建图和闭环检测。在跟踪线程中，由于使用 RGB-D 相机，因此可以直接获取环境的深度信息，不需要像单目相机那样做特定的初始化。在系统启动时，将第一帧作为关键帧，初始化其位姿之后，从所有三维坐标点中创建一个初始化地图。之后通过最小化光度误差来进行相机的初始位姿估计以及像素点的对应关系计算，利用最小化局部地图点的重投影误差再次优化相机位姿，对局部地图进行跟踪，并判断是否生成新的关键帧。若跟踪失败，则触发新的跟踪周期；在局部建图线程中，对当前关键帧图像中的 ORB 特征进行提取与匹配，之后筛选出当前关键帧中冗余的地图点，由特征提取与匹配的结果来创建新的地图点。当处理完所有关键帧之后，执行局部 BA 优化，对周围的关键帧位姿以及地图点位置进行优化，以获得更加准确的定位与建图精度。最后筛选候选关键帧，对冗余的关键帧进行剔除；在闭环检测线程中，通过检索图像识别数据库执行闭环检测[12]，当检测到闭环时，对闭环周期进行判断，若当前关键帧与闭环关键帧不属于同一个周期，则将闭环关键帧的周期与当前关键帧的周期合并为同一周期。采用位姿图优化的方法对运动累积误差进行校正。最后将此周期内的所有关键帧位姿与所有地图点位置作为优化变量执行全局 BA 优化，从而得到全局一致的环境地图和相机运动轨迹。

1.1 基于直接法的跟踪线程

ORB-SLAM2 中跟踪线程的图像帧处理速率由其运行帧率决定。例如在 TUM 数据集 fr3_office 序列中，使用搭载 intel core i7 7567U CPU 的台式机运行 ORB-SLAM2 算法，将每一帧图像提取特征数的阈值设为 1 000，跟踪线程的 4 个主要步骤的平均耗时如表 1 所示。

表 1 ORB-SLAM2 跟踪线程所用时间

步骤	耗时/ms	耗时百分比/%
图像预处理	13.52	59.38
位姿估计	1.36	5.97
局部地图跟踪	6.87	30.17
关键帧判断	1.02	4.48

由表 1 可知，在 ORB-SLAM2 跟踪线程中，对于 RGB-D 相机来说，特征提取与匹配在图像预处理步骤中，如图 3[7]，花费的时间最多，高达 59.38%。所以为了避免特征提取，采用直接法的方式实现像素点的匹配，从而减少跟踪线程所用时间。

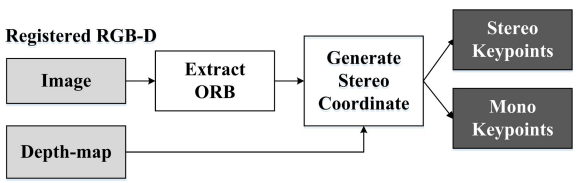


图 3 ORB-SLAM2 跟踪线程的图像预处理

本算法的跟踪线程主要步骤如下：首先通过最小化光度误差获取相机的初始位姿估计；然后通过最小化考虑曝光时间变化的光度误差来获得图像之间的像素点匹配关系信息；最后通过最小化局部地图点的重投影误差来获取更加准确的相机位姿估计，对局部地图进行跟踪，判断新的关键帧，为局部建图线程提供关键帧位姿初值。

1.1.1　初始位姿估计

跟踪线程通过对每一帧图像处理来对当前帧的位姿进行估计，如果上一帧跟踪成功，则利用一个恒速模型来对当前相机的位姿进行预测。根据估计的位姿，将参考帧中跟踪的 3D 地图点投影到当前帧中。当前帧与参考帧的相对位姿变换 $\boldsymbol{T}_{\mathrm{cr}} \in SE(3)$ 则可以通过最小化光度误差计算得到，该误差主要建立在同一个 3D 点对应的图像块之间，即

$$\boldsymbol{T}_{\mathrm{cr}}=\underset{\boldsymbol{T}}{\operatorname{argmin}}\iint_{\mathcal{R}}\rho[\delta I(\boldsymbol{T},\boldsymbol{u})]\mathrm{d}\boldsymbol{u} \tag{1}$$

其中，$\rho[\cdot]=\frac{1}{2}\|\cdot\|^2$，$\boldsymbol{u}$ 是参考帧的像素位置，$\mathcal{R}$ 是图像块区域，光度误差 δI 如下所示：

$$\delta I(\boldsymbol{T},\boldsymbol{u})=I_{\mathrm{c}}(\pi(\boldsymbol{T}\cdot\boldsymbol{x}_{ri}))-I_{\mathrm{r}}(\boldsymbol{u}),\forall\boldsymbol{u}\in\mathcal{R} \tag{2}$$

其中，$\boldsymbol{x}_{ri}$ 是参考帧坐标系中第 i 个被跟踪的 3D 地图点的位置。

1.1.2　基于直接法的像素点对应关系计算

在 ORB-SLAM2 的跟踪线程中，图像帧中的特征点与环境地图点的对应关系是由匹配描述子来获得的，之后再利用最小化重投影误差对参数进行优化。在本算法的跟踪线程中由于没有进行特征提取，所以无法利用描述子匹配的方法获取其对应关系。本文将采用最小化光度误差的方法寻找像素点和地图点的对应。

通过初始位姿估计，可以得到经过优化且与上一帧像素点相关联的地图点集。从该地图点集中可以得到当前帧与局部关键帧可观测到的 Covisibility 地图点，将局部地图点在当前帧中进行投影，从而可通过最小化光度误差的方法获取更多图像像素点和局部地图点的对应关系。另外，本文由初始的投影像素坐标栅格化投影到当前帧中的地图点，采用 10×10 栅格，并根据深度信息对栅格中的地图点进行排列，在计算像素点匹配时，对每个栅格至多选取 6 个不同深度的地图点，从而提高算法的运行效率。

通过上一步的初始位姿估计，可由当前帧和局部关键帧之间的位姿将地图点投影到当前帧上作为当前帧的特征点。而像素匹配点会因相机位姿与地图点位置的不准性产生误差。SVO[13] 在像素点匹配时通过最小化图像小光斑的光度误差来获取亚像素精度的像素点对应关系：

$$\boldsymbol{u}'_i=\underset{\boldsymbol{u}'_i}{\operatorname{argmin}}\|\boldsymbol{I}_k(\boldsymbol{u}'_{i,n})-\boldsymbol{I}_{\mathrm{r}}(\boldsymbol{A}_i\cdot\boldsymbol{u}_{i,n})\|^2 \tag{3}$$

其中，$\boldsymbol{I}_{\mathrm{r}}$ 为局部地图点所在关键帧对应的灰度图，$\boldsymbol{u}_{i,n}$ 为 $\boldsymbol{I}_r$ 对应的图像块中第 n 个像素点的坐标，$\boldsymbol{u}'_{i,n}$ 的定义同理，其属于 $\boldsymbol{I}_k$ 对应的图像块。为了对当前帧和关键帧之间因距离过大引起的图像块变化进行调整[4]，还引入了仿射变换矩阵 $\boldsymbol{A}_i$。

由于普遍使用的相机都包含自动曝光功能，同时 2 个时刻相机的曝光时间也不相同，这会造成环境中同一点在不同图像中灰度误差较大。因此本文通过考虑曝光时间引起的灰度变化，引入光度响应函数[14]对其进行优化：

$$\boldsymbol{u}'_i=\underset{\boldsymbol{u}'_i}{\operatorname{argmin}}\left\|\boldsymbol{I}_k(u'_{i,n})-f\left(\frac{\Delta t_k}{\Delta t_{\mathrm{r}}}f^{-1}(\boldsymbol{I}_{\mathrm{r}}(\boldsymbol{A}_i\cdot\boldsymbol{u}_{i,n}))\right)\right\|^2 \tag{4}$$

式中，Δt_{r} 和 Δt_k 分别是 $\boldsymbol{I}_{\mathrm{r}}$ 和 $\boldsymbol{I}_k$ 的曝光时间，$f(\cdot)$ 为光度响应函数，$f^{-1}(\cdot)$ 为逆光度响应函数。最后利用 G-N(Gauss Newton)算法求解式(4)。

1.1.3　跟踪局部地图

在完成局部地图点与当前帧像素点匹配后，利用 g2o[15] 通过最小化重投影误差再次优化相机的相对位姿 $\boldsymbol{T}_{\mathrm{cw}}$：

$$\boldsymbol{T}_{\mathrm{cw}}=\underset{\boldsymbol{T}_{\mathrm{cw}}}{\operatorname{argmin}}\frac{1}{2}\sum_i\|\boldsymbol{u}_i-\pi(\boldsymbol{T}_{\mathrm{cw}}\cdot\boldsymbol{p}_{\mathrm{w}i})\|^2 \tag{5}$$

其中，$p_{\mathrm{w}i}$ 是世界坐标系中第 i 个 Covisibility 地图点。此过程的优化问题仅对相机位姿执行 BA 优化，由于单个位姿节点的 BA 问题对于地图点的约束太过单一，因此将地图点的优化放在局部 BA 线程中执行。

本文在跟踪线程中虽然没有在当前帧中提取 ORB 特征，但是通过最小化考虑曝光时间变化的光度误差得到了关联地图点的当前帧像素点对应关系，避免了对每一帧图像进行特征提取，从而加快了跟踪线程运行速率。

1.1.4　地图点管理与关键帧处理

如果在上一步中成功跟踪到地图点，那么将成功跟踪的地图点进行缓存，以便在下一帧中对其进行优先处理。最后，确定当前帧是否应作为关键帧生成，关键帧的质量决定了局部地图的好坏，本文选

取关键帧的规则如下：

(1)当前帧至少跟踪到 30 个地图点；

(2)当前帧跟踪到的地图点少于参考关键帧地图点的 90%；

(3)距上一个关键帧插入之后已跟踪超过 20 帧图像；

(4)上一个关键帧与当前帧的位置差值或姿态角差值不小于阈值 $\boldsymbol{T}_t$。

条件(1)可以确保在里程计跟踪丢失后，插入的关键帧对位姿变化不会产生较大影响。条件(2)可以保证关键帧跟踪的质量，同时，避免引入过多的冗余信息。条件(3)保证当局部线程处于繁忙状态时插入关键帧，可以通过发送一个信号暂停局部 BA，以便尽快对新的关键帧进行处理。条件(4)对相机的旋转与平移产生的视差变化进行了考虑。

1.1.5 局部建图与闭环检测

本算法的局部建图和闭环检测线程只与关键帧及其 ORB 特征有关。在局部建图线程中包含关键帧特征提取与匹配、地图点构造、地图点和关键帧筛选以及局部 BA 优化。如果一个图像帧作为关键帧生成，则提取并匹配新的 ORB 特征，同时将成功跟踪的特征保存在此关键帧中。之后通过最小化重投影误差进行局部 BA 来进一步优化关键帧和局部地图点的位姿。闭环检测线程利用 ORB 特征在 DBoW2 上构建的词袋模型来检测闭环。最后通过位姿图优化来纠正累积误差。

2 稠密地图构建

ORB-SLAM2 算法可以实时、精确地估计相机的位姿并构建稀疏地图。然而稀疏地图无法用于导航和避障，因此本文选择基于 Octomap[11](八叉树地图)的概率占据地图构建 3D 稠密地图，而非传统的点云地图。与点云地图相比，八叉树地图采用八叉树分层的方式存储地图，可以节省地图存储空间，能够实时有效地更新重建的地图，同时其分辨率是可调的。此八叉树地图可用于机器人的避障与导航。

八叉树结构的叶子节点以概率的形式来表示空间的占用状态。假设某节点为 n，观测数据为 z，在 t 时刻此节点被占用概率表示为

$$P(n\mid z_{1:t})=\left[1+\frac{1-P(n\mid z_t)}{P(n\mid z_t)}\cdot\frac{1-P(n\mid z_{1:t-1})}{P(n\mid z_{1:t-1})}\cdot\frac{P(n)}{1-P(n)}\right]^{-1} \tag{6}$$

一般情况下假设先验概率 $P(n)=0.5$，为了保证 $P(n)$在[0,1]区间之内，这里引入 $\log it$ 变换：

$$\log it(P)=\ln\left(\frac{P}{1-P}\right) \tag{7}$$

由式(7)可得：

$$\ln\left[\frac{P(n\mid z_{1:t})}{1-P(n\mid z_{1:t})}\right]=-\ln\left[\frac{1-P(n\mid z_t)}{P(n\mid z_t)}\cdot\frac{1-P(n\mid z_{1:t}-1)}{P(n\mid z_{1:t-1})}\cdot\frac{P(n)}{1-P(n)}\right]=\ln\left[\frac{P(n\mid z_{1:t-1})}{1-P(n\mid z_{1:t-1})}\right]+\ln\left[\frac{P(n\mid z_t)}{1-P(n\mid z_t)}\right] \tag{8}$$

其中，

$$L(n\mid z_t)=\ln\left[\frac{P(n\mid z_t)}{1-P(n\mid z_t)}\right] \tag{9}$$

则式(6)可以改写为

$$L(n\mid z_{1:t})=L(n\mid z_{1:t-1})+L(n\mid z_t) \tag{10}$$

从式(10)中可以看出，对于三维空间中的任意点，新的观测值都会直接添加到最后的观测值中，这种更新方式更加灵活且易于实现。在八叉树建图过程中，根节点的占据概率可由子节点的概率计算得出。计算根节点的方法有两种：求平均值法 $\bar{l}(n)=\frac{1}{8}\sum_{i=1}^{8}L(n_i)$ 和求最大值法 $\bar{l}(n)=\max_i L(n_i)$。在创建点云地图的过程中，基于 Octomap 库实时构建 3D 环境的八叉树地图。

3 实验与分析

本算法的运行平台配置如下：CPU 为 intel core i7 7567U，主频为 3.5 GHz，内存为 32 G，不使用 GPU 加速，测试系统为 Ubuntu 16.04。在 TUM RGB-D 数据集[16]下进行算法测试。

3.1 数据集实验

TUM RGB-D 数据集作为德国慕尼黑工业大学公开的一种 RGB-D SLAM 系统标准评价方案，其中包含了由 RGB-D 传感器采集的不同类别的室内图像序列，可用于评估不同纹理、照明以及结构条件下的环境重建和 SLAM 方案。同时，此数据集还提供了用于评估算法定位精度的真实轨迹。为了评估本文的 RGB-D SLAM 系统的性能，本文选择能

够评估大部分 RGB-D SLAM 算法的一些图像序列进行测试，并且与 ORB-SLAM2[7] 和 DVO-SLAM[17] 进行对比。

3.1.1　视觉里程计评估

为了体现本算法的定位效果，首先仅对本算法和 ORB-SLAM2 里程计部分的位姿输出精度进行对比，不包含闭环检测以及闭环优化部分，检测本算法改进后的跟踪线程与局部建图线程对算法定位精度的影响。视觉里程计的评估主要以相对位姿误差(Relative Pose Error，RPE)指标为依据，见表 2。

表 2　里程计输出相对位姿误差对比

图像序列	平移 RMSE/(m·s^{-1})		旋转 RMSE/((°)·s^{-1})	
	本文算法	ORB-SLAM2	本文算法	ORB-SLAM2
fr1/desk	0.029 1	0.028 3	1.504 8	1.503 9
fr1/desk2	0.045 2	0.044 4	2.146 3	2.143 8
fr1/room	0.087 6	0.086 8	2.938 6	2.937 3
fr1/xyz	0.013 2	0.014 3	0.925 8	0.926 7
fr2/desk	0.040 1	0.039 8	1.318 8	1.316 7
fr2/large_no_loop	0.087 5	0.085 7	1.524 3	1.521 9
fr2/large_with_loop	0.228 3	0.227 4	3.678 7	3.677 4
fr2/xyz	0.011 6	0.012 4	0.505 6	0.506 2
fr3/long_office_household	0.027 8	0.026 5	0.845 5	0.843 1
fr3/str_nontex_far	0.015 3	0.017 6	0.532 1	0.541 7
fr3/str_tex_far	0.023 3	0.022 9	0.567 3	0.566 1

由表 2 可知，在大多数情况下，本文算法与 ORB-SLAM2 算法精度相当，甚至在“fr1/xyz”、“fr2/xyz”和“fr3/str_nontex_far”3 个序列中精度都要高于 ORB-SLAM2 算法，由于本文在跟踪线程采用了直接法，因此在无纹理的“fr3/str_nontex_far”序列中，本文方法的旋转与平移 RPE 要明显高于对比方法。

3.1.2　SLAM 系统评估

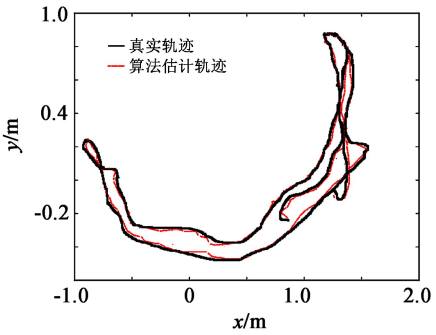
不同于视觉里程计的评估，SLAM 系统在评估方面采用绝对轨迹误差(Absolute Trajectory Error，ATE)指标。该部分实验仍然在 TUM 数据集上进行，并将本文算法与 ORB-SLAM2 和 DVO-SLAM 进行比较，结果见表 3。

表 3　SLAM 系统绝对轨迹误差对比

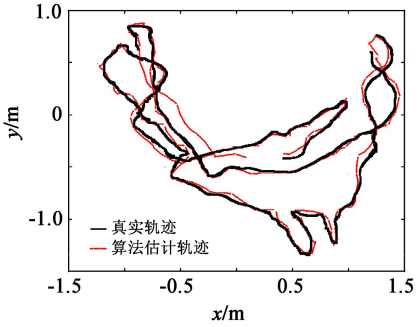
图像序列	本文算法	ORB-SLAM2	DVO-SLAM
fr1/desk	0.016	0.016	0.021
fr1/desk2	0.024	0.022	0.046
fr1/room	0.045	0.047	0.043
fr1/xyz	0.009	0.010	0.011
fr2/desk	0.008	0.009	0.017
fr2/xyz	0.004	0.004	0.018
fr3/long_office_household	0.011	0.010	0.035
fr3/str_nontex_far	0.010	0.012	0.011

在表 3 中，第一列为 TUM 数据集各个图像序列的名称，在包含闭环检测与优化的情况下，将本算法与 ORB-SLAM2 和 DVO-SLAM 进行了对比实验，通过对比 ATE 指标可以发现，本算法取得了与 ORB-SLAM2 算法相当的精度，而在“fr3/str_nontex_far”序列中，在环境结构单一、纹理信息不丰富的情况下，ORB-SLAM2 出现了跟踪失败的情况，本文算法由于可以利用光度信息进行跟踪，没有出现跟踪丢失的情况，鲁棒性较好且取得的结果误差最小。

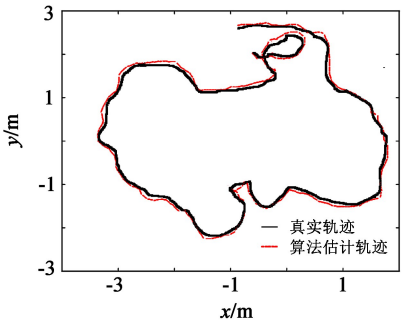
在图 4 中，对 TUM 数据集中 4 个图像序列的 SLAM 算法运行轨迹图与真实轨迹进行对比，均取得了较理想的结果。在“fr3/str_nontex_far”序列中，估计轨迹与真实轨迹几乎达到了一致。这也证明本算法无论面对结构简单、纹理信息缺失的环境还是纹理丰富的环境，都有着较高的准确性及稳定性。



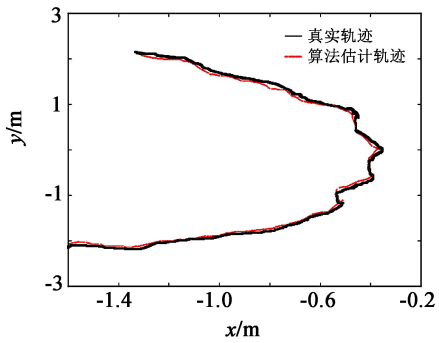
(a)　fr1/desk2



(b)　fr1/room



(c)　fr3/long_office_household



(d) fr3/str_nontex_far

图 4 本文 SLAM 算法对 4 个图像序列估计轨迹对比图

3.1.3 系统实时性评估

通过在 TUM 数据集上进行测试,本算法各个步骤的耗时情况统计如表 4 所示。在 SLAM 系统中,只有跟踪线程需要对 RGB-D 图像进行实时处理,而对于局部建图与闭环检测线程,仅对关键帧进行处理,无需实时运行。因此跟踪线程的运行帧率决定了算法的运行帧率。在表 4 中,本文算法的跟踪线程平均耗时在 11 ms 左右,考虑图像的读取与校正等步骤,其处理的运行帧率在 25～35 Hz。

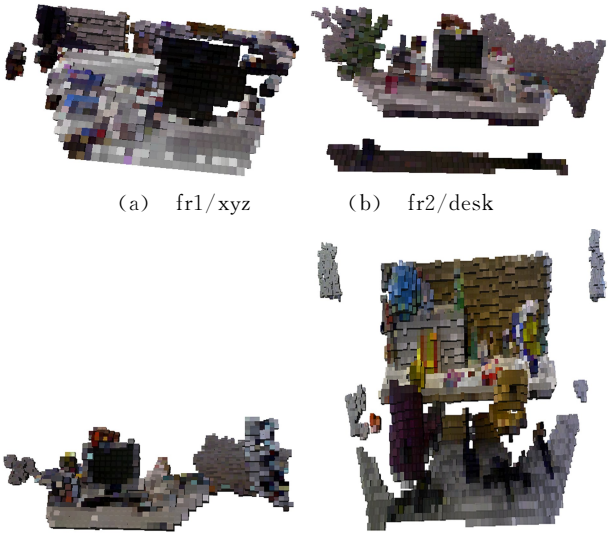
表 4 跟踪线程的耗时统计(单位:ms)

步骤	ORB-SLAM2	本文算法
特征提取	11.42	—
初始位姿估计	1.36	0.68
特征匹配	13.44	7.51
局部地图跟踪	6.87	2.93
关键帧判断	1.02	0.22

表 4 将本文算法与 ORB-SLAM2 算法的跟踪线程完整耗时进行了对比,由于 ORB-SLAM2 算法在对每一帧图像进行处理时都要提取 1 000 个 ORB 特征,且在跟踪局部地图时需进行图像特征点与局部地图点的匹配,而本文算法采用直接法无需进行特征提取,且在跟踪局部地图时可由特征匹配直接完成,因此在跟踪线程中节省了很大一部分时间,提升了算法运行的效率。

3.2 稠密地图构建实验

ORB-SLAM2 算法主要针对相机位姿估计的实时性、准确性以及鲁棒性进行改进,在地图构建部分只能构建稀疏地图,无法实现机器人进行导航和避障等任务。本文算法利用 RGB-D 传感器构建了适用于移动机器人的稠密八叉树地图,解决了单目相机的尺度不确定性问题。将本文提出的方法在 TUM 数据集中的"fr1/xyz"、"fr2/desk"、"fr2/xyz"以及"fr3/long_office_household"4 个图像序列下进行实验,生成了八叉树图,如图 5 所示。四个八叉树图的存储空间仅为 0.38 MB,这表明八叉树图可以节省大量的存储空间,且易于维护。



(a) fr1/xyz (b) fr2/desk

(c) fr2/xyz (d) fr3/long_office_household

图 5 4 个图像序列的八叉树地图

4 结论

本文提出了一种基于 RGB-D 相机的半直接稠密 SLAM 方法。该方法在跟踪线程中,对于非关键帧采用直接法进行初始位姿估计和像素点匹配,无需进行特征提取与匹配,从而提高了跟踪线程的运行速率。而对于关键帧采用特征法进行特征提取与匹配,以保证 SLAM 算法局部优化、全局优化以及闭环检测功能的实现。此外,本文算法基于 RGB-D 相机构建了压缩的且易于维护的八叉树稠密地图,满足了移动机器人导航和避障的需求。在 TUM 数据集上对该算法进行了验证,实验数据表明其能够保持与 ORB-SLAM2 算法相当的定位精度,并且在跟踪线程中的运行帧率得到了显著提升,可实时构建环境信息丰富的 3D 稠密地图。在今后的工作中将进一步对惯性测量单元(IMU)数据进行融合,从而为跟踪线程的初始位姿估计提供初值,并将其作为误差项用于优化局部地图跟踪和局部 BA,进一步提升算法的精度与鲁棒性。

参考文献:

[1] Scaramuzza D, Fraundorfer F. Visual odometry [tutorial] [J]. IEEE Robotics & Automation Magazine, 2011, 18(4): 80-92.

(下转第 559 页)

用光放大补偿的措施来优化附加相位噪声，并明确了放大器的选用原则。结果表明本文的相位噪声模型能够有效应用于微波光子链路，能对附加相位噪声的分析和优化提供理论依据，具有较好的工程指导意义。

参考文献：

[1] Jose C, Dalma N. Microwave photonics combines two worlds [J]. Nature Photonics, 2007, 1: 319-330.

[2] Vilcot A, Cabon B, Chazelas J. Microwave Photonics——From Components to Applications and Systems[M]. Kluwer: Kluwer Academic Publishers, 2003: 306-375.

[3] Seed A J, Williams K J. Microwave photonics [J]. J. Lightwave Technol., 2006, 24(12): 4628-4641.

[4] Williamson R C, Esman R D. RF photonics[J]. J. Lightwave Technol., 2008, 26(9): 1145-1153.

[5] Wang Renyuan, Bhardwaj A, Li Yifei. Efficent RF frequency down conversion using coupling quantum-well optical phase modulator[J]. IEEE Photon. Technol. Lett., 2011, 23(10): 645-647.

[6] Marpaung D, Roeloffzen C, Etten W V. Enhanced dynamic range in a directly modulated analog photonic link[J]. IEEE Photon. Technol. Lett., 2009, 21(24): 1810-1812.

[7] Cox C H, Ackerman E I, Betts G E, et al. Limits on the performance of RF-over-fiber links and their impact on device design[J]. IEEE Trans. Microw. Theory Tech., 2006, 54(2): 906-920.

[8] Urick V J, Bucholtz F, Mckinney J D, et al. Long-haul analog photonics[J]. J. of Lightwave Technol., 2011, 29(8): 1182-1205.

[9] Bibey M B, Deborgies F. Very low phase-noise optical links—experiments and theory [J]. IEEE Trans. on Microwave Theory and Techniques, 1999, 47(12): 2257-2262.

[10] Eliyahu D, Seidel D, Maleki L. RF amplitude and phase-noise reduction of an optical link and an opto-electronic oscillator [J]. IEEE Trans. on Microwave Theory and Techniques, 2008, 56(2): 449-456.

作者简介：

瞿鹏飞(1983—)，男，博士，工程师，研究方向为微波光子学。

(上接第 554 页)

[2] Krombach N, Droeschel D, Houben S, et al. Feature-based visual odometry prior for real-time semi-dense stereo SLAM [J]. Robotics and Autonomous Systems, 2018, 109: 38-58.

[3] Schops T, Sattler T, Pollefeys M. BAD SLAM: Bundle adjusted direct RGB-D SLAM[C]// Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, 2019: 134-144.

[4] Klein G, Murray D. Parallel tracking and mapping for small AR workspaces[C]// Proc. of the 2007 6th IEEE and ACM Inter. Symp. on Mixed and Augmented Reality IEEE Computer Society, 2007: 1-10.

[5] Eudes A, Lhuillier M, Naudet-Collette S, et al. Fast odometry integration in local bundle adjustment-based visual slam [C]// IEEE 2010 20th Inter. Conf. on Pattern Recognition, 2010: 290-293.

[6] Mur-Artal R, Montiel J M M, Tardos J D. ORB-SLAM: a versatile and accurate monocular SLAM system [J]. IEEE Trans. on Robotics, 2015, 31(5): 1147-1163.

[7] Mur-Artal R, Tardos J D. ORB-SLAM2: An open-source slam system for monocular, stereo and RGB-D cameras[J]. IEEE Trans. on Robotics, 2017, 33(5): 1255-1262.

[8] Engel J, Schops T, Cremers D. LSD-SLAM: Large-scale direct monocular SLAM[C]// European Conf. on Computer Vision. Springer, Cham, 2014: 834-849.

[9] Engel J, Koltun V, Cremers D. Direct sparse odometry[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2017, 40(3): 611-625.

[10] Karami E, Prasad S, Shehata M. Image matching using SIFT, SURF, BRIEF and ORB: performance comparison for distorted images[C]// IEEE Newfoundland Electrical and Computer Engin. Conf., 2015.

[11] Hornung A, Wurm K M, Bennewitz M, et al. OctoMap: An efficient probabilistic 3D mapping framework based on octrees [J]. Autonomous Robots, 2013, 34(3): 189-206.

[12] Mo H, Wang K, Wang H, et al. Research on loop closing for SLAM based on RGB-D images [C]// Proc. of 2018 Chinese Intelligent Systems Conf., 2019: 739-748.

[13] Forster C, Pizzoli M, Scaramuzza D. SVO: Fast semi-direct monocular visual odometry[C]//2014 IEEE Inter. Conf. on Robotics and Automation (ICRA), 2014: 15-22.

[14] Shi B, Matsushita Y. Self-calibrating photometric stereo [C]// 2010 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 2010: 1118-1125.

[15] Grisetti G, Kummerle R, Strasdat H, et al. g2o: A general framework for (hyper) graph optimization[C]//Proc. 2011 IEEE Inter. Conf. on Robotics and Automation, 2011.

[16] Sturm J, Engelhard N, Endres F, et al. A benchmark for the evaluation of RGB-D SLAM systems[C]//Proc. 2012 IEEE/RSJ Inter. Conf. on Intelligent Robots and Systems, 2012: 573-580.

[17] Kerl C, Sturm J, Cremers D. Dense visual SLAM for RGB-D cameras [C]// Proc. 2013 IEEE/RSJ Inter. Conf. on Intelligent Robots and Systems, 2013: 2100-2106.

作者简介：

胡章芳(1969—)，女，重庆市人，教授，硕士生导师，主要研究领域为数字图像处理、微机电系统、智能信号处理。