

DOI: 10.11992/tis.201612004

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20170702.0425.012.html>

基于图优化的移动机器人视觉 SLAM

张毅, 沙建松

(重庆邮电大学 智能系统及机器人实验室, 重庆 400065)

摘 要: 针对目前移动机器人视觉 SLAM(simultaneous localization and mapping) 研究中存在的实时性差、精确度不高、无法稠密化建图等问题, 提出了一种基于 RGB-D 数据的实时 SLAM 算法。在本算法前端处理中, 采用了鲁棒性与实时性更好的 ORB 特征检测。利用 RANSAC 算法对可能存在的误匹配点进行剔除完成初始匹配, 对所得内点进行 PNP 求解, 用于机器人相邻位姿的增量估计。在后端优化中, 设计了一种遵循图优化思想的非线性优化方法对移动机器人位姿进行优化。同时结合闭环检测机制, 提出了一种点云优化算法, 用于抑制系统的累积误差, 进一步提升位姿与点云的精确性。实验验证了本文所提方法能够迅速、准确地重构出稠密化的三维环境模型。

关键词: RGB-D; 移动机器人; 图优化; 同时定位与地图构建; 位姿与点云优化

中图分类号: TP24 **文献标志码:** A **文章编号:** 1673-4785(2018)02-0290-06

中文引用格式: 张毅, 沙建松. 基于图优化的移动机器人视觉 SLAM[J]. 智能系统学报, 2018, 13(2): 290-295.

英文引用格式: ZHANG Yi, SHA Jiansong. Visual-SLAM for mobile robot based on graph optimization[J]. CAAI transactions on intelligent systems, 2018, 13(2): 290-295.

Visual-SLAM for mobile robot based on graph optimization

ZHANG Yi, SHA Jiansong

(Laboratory of Intelligent System and Robotics, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: In the present research on the visual SLAM (Simultaneous Localization and Mapping) of mobile robot, some defects as inferior real-timeness, low accuracy and hard to densified mapping exist, therefore, the paper proposed a real-time SLAM algorithm based on RGB-D data. In the front-end processing of the algorithm, the ORB feature detection with better robustness and real-timeness was adopted. RANSAC algorithm was utilized to get rid of the possible mismatch points and complete the initial match. For the obtained inner point, PNP solution was carried out for using as the increment estimate of the adjacent pose of robot. In the rear-end optimization, a nonlinear optimization method obeying image optimization thought was used for optimizing the pose of a moving robot. In addition, in combination with the closed-loop detection mechanism, a point cloud optimization algorithm was proposed to suppress the cumulative error of the system and further improve the accuracy of pose and point cloud. The experimental results show that the proposed method can reconstruct the dense 3D environment model quickly and accurately.

Keywords: RGB-D; mobile robot; graph optimization; SLAM; pose and point cloud optimization

同时定位与地图创建 (SLAM) 技术已经成为目前机器人研究领域的热点, 随着 SLAM 问题研究的不断深入, 对机器人位姿监测和环境地图创建的精确度、实时性也都有了更高的要求。因此, 如何在精确度与实时性之间寻求一种平衡是目前众多

SLAM 算法中亟待解决的一个问题。而新型 RGB-D 摄像机 Kinect^[1], 具有丰富 3 维空间信息与颜色纹理信息, 为上述问题的解决提供了的可能。文献 [2] 提出的基于 RGB-D 相机的室内 SLAM 方法可实现对室内环境的稠密化建图, 但论文中的 SIFT 特征提取算法较为费时, 会占用较多的运算资源, 因此其实时性较差。ORB-SLAM^[3] 采用了基于单目摄

收稿日期: 2016-12-05. 网络出版日期: 2017-07-02.

基金项目: 重庆市科学技术委员会项目 (CSTC2015jcyjBX0066).

通信作者: 沙建松. E-mail: 277829172@qq.com.

像机的 SLAM 方案, 以高效的 ORB 特征贯穿于系统中的特征处理、匹配、回环检测、重定位等方面, 更侧重于系统处理的实时性, 但因为单目相机的限制, ORB-SLAM 算法在稠密化建图方面还有所欠缺。针对目前移动机器人视觉 SLAM 研究中存在的实时性差、精确度不高、无法稠密化建图等问题, 本文提出了一种基于 Kinect 深度摄像机的移动机器人三维视觉 SLAM 算法, 以实现对环境进行实时的、精确的稠密化建图。在本算法前端处理中, 通过分析多种特征点检测算法的优劣, 采用了鲁棒性与实时性更好的 ORB 特征检测。利用 RANSAC 算法^[4]对可能存在的误匹配点进行剔除完成初始匹配, 以 PNP(perspective-N-point)^[5]问题求解的方式实现了图像的快速配准。在后端优化中, 通过对

SLAM 问题建模分析, 遵循图优化思想, 设计了一种非线性优化方法, 实现对机器人位姿的优化。并与闭环检测^[6]机制相结合, 提出了一种基于局部闭环与全局闭环约束的点云优化算法, 同时完成对位姿与稠密化点云的优化, 可有效地抑制系统的累积误差, 进一步提升整个系统的精确性。最后通过实验, 验证了本文 SLAM 方法的可行性和有效性。

1 系统结构

本文提出的室内环境下移动机器人三维视觉 SLAM 算法的总体框架如图 1 所示, 主要分为前端的关键帧筛选、帧间配准、闭环检测和后端的图优化、点云融合、点云优化几个部分。

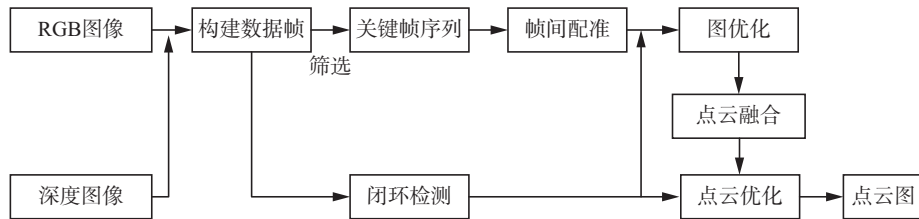


图 1 基于图优化的视觉 SLAM 系统框架

Fig. 1 Visual-SLAM system framework based on graph optimization

2 特征点的检测与配准

2.1 特征点的检测

特征点的检测与匹配是视觉 SLAM 问题中“图像前端”的部分, 也是视觉 SLAM 中占用计算资源最多, 最影响系统实时性的部分。基于 FAST 特征检测改进的 ORB 特征^[7]具有旋转不变和尺度不变特性的二进制特征, 它是一种高效的具有良好针对视图不变的特征, 比 SIFT 特征检测快两个数量级, 同时特征点的质量较高、鲁棒性好。考虑到 SLAM 问题对鲁棒性、实时性的要求, 本文采用了 ORB 特征提取算法。

2.2 特征点的配准

图像产生的本质是光线通过相机透镜投影到图像传感器上, 这个过程是 3-D 场景到 2-D 平面的投影, 因此场景与它的成像之间存在重要的联系, 相同场景的不同成像之间也有联系。而通过这些联系, 可以实现特征点的配准。

特征点配准算法可以分为以下 5 个步骤:

1) 提取特征点: Opencv 提供了通用接口 `cv::FeatureDetector`, 这里将其设定为 ORB 特征。

2) 提取描述子: 对于检测出的角点, 用一些数学上的特征对其进行描述, 这里通过 `cv::FeatureDe-`

`tectorExtractor` 接口设置特征描述子提取器。

3) 匹配: 通过各个角点的描述子来判断它们在两张图像中的对应关系。这里采用的方法为快速最近邻算法 FLANN。

4) 去外点: 利用 RANSAC 算法去除错误匹配的外点, 保留正确的内点。

5) 位姿求解: 调用 Opencv 的 `PnP solver()` 函数, 求解相机的位姿运动 (R, t) 。

3 位姿优化

针对传统的利用滤波方法估计位姿存在线性化误差、更新效率低下等问题, 本文依据 SLAM 研究中的图优化^[8]思想, 设计了一种基于非线性优化的位姿优化方法, 同时实现局部位姿优化和全局位姿优化。

3.1 模型的建立

SLAM 的经典数学模型通常分为 3 部分: 状态变量、运动方程、观测方程。

$$\mathbf{X} = [x_1 \ x_2 \ \cdots \ x_N \ y_1 \ y_2 \ \cdots \ y_M] \quad (1)$$

$$x_k = f(x_{k-1}, u_k) + u_k \quad (2)$$

$$z_{k,j} = g(x_k, y_j) + n_{k,j} \quad (3)$$

本文主要以 Kinect 作为主要的传感器输入, 通过对视觉信号的处理, 构建视觉里程计, 用于位姿

估计。所以这里只考虑状态变量和观测方程, 构建出目标函数:

$$e(x_k, y_j, z_{k,j}) = z_{k,j} - g(x_k, y_j) \quad (4)$$

$$E(X) = \sum_{(k,j) \in C} e(x_k, y_j, z_{k,j})^T \Omega_{k,j}^{-1} e(x_k, y_j, z_{k,j}) \quad (5)$$

$$X^* = \arg \min_X E(X) \quad (6)$$

式中: x_k 表示第 k 时刻的机器人位姿, 也称为图优化的一个 pose 节点; y_j 表示第 k 时刻机器人所能观测到的路标, 也称为图优化的一个 point 节点; $z_{k,j}$ 表示 pose 节点与 point 节点的约束; 特别地, 在闭环检测优化过程中, $z_{k,j}$ 代表 pose 节点与 pose 节点的约束; $e(x_k, y_j, z_{k,j})$ 为由两顶点组成的边所产生的向量误差函数, 代表两顶点满足约束条件 $z_{k,j}$ 的程度, 当它的值为 0 时, 表示两顶点完全满足匹配约束。

目标函数的物理意义: 这里假设了误差函数 $e(\cdot)$ 符合高斯分布 $(0, \Omega)$, 则 SLAM 问题的贝叶斯概率描述为, 求解适合的 (x, y) , 使系统最有可能产生当前的观测数据 Z , 即最大似然 (MLE):

$$(x, y)^* = \arg \max_{x, y} p(z|x, y) \quad (7)$$

最终所求的 $e(x_k, y_j, z_{k,j})$ 的二次型 (式 (6)), 实质上是在求解最大似然。即在假设噪声是高斯分布之后, 将整个 SLAM 问题等效为一个最小二乘问题。进一步, 利用李代数^[9]的扰动模型, 求得目标函数关于 pose 和 point 的导数 J_x 和 J_y , 可对此最小二乘问题求解。以此建立最终的求解方程:

$$e(x + \delta_x, y + \delta_y) \approx e(x, y) + J_x \delta_x + J_y \delta_y \quad (8)$$

$$H = \sum_i J_i^T \Omega^{-1} J_i \quad (9)$$

$$b = \sum_i J_i^T \Omega^{-1} e_i \quad (10)$$

$$H \delta_x = -b \quad (11)$$

式中: J_x 代表相机微小运动后, 像素的变化; J_y 代表空间点微小运动后, 像素的变化。式 (11) 代表了考虑所有的变量与边的 Full SLAM 问题的求解。结合 SLAM 问题的稀疏性, 选用合适的梯度下降方法^[10], 用求得的 δ_x 用于状态变量 X 的更新, 通过多次迭代, 最终求得满足目标方程的状态变量 X 的收敛值。可借助通用图优化工具 g2o^[11] 来解决上述非线性最小二乘问题, 求解过程如图 2 所示。

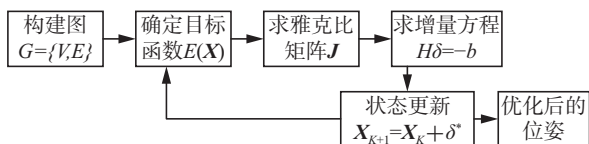
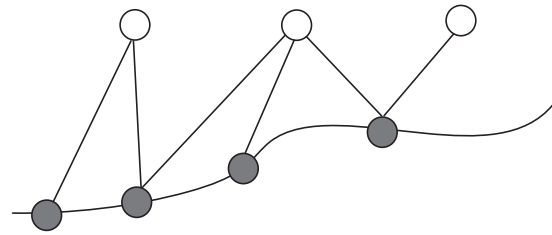


图2 优化求解流程

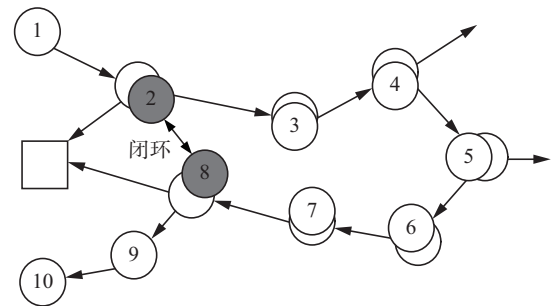
Fig. 2 Process of optimizing solution

3.2 图的建立与优化

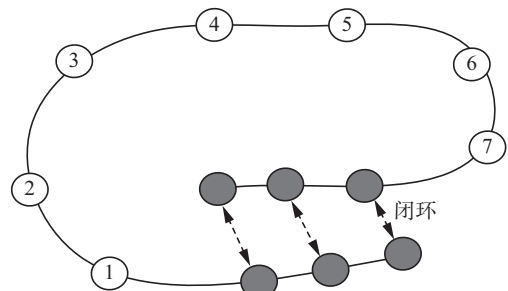
上述 SLAM 问题全部抽象成节点与边, 节点代表了优化变量, 包括 pose 节点和 point 节点。边代表了误差项约束条件, 包括 pose-point 约束和 pose-pose 约束。pose-point 约束依赖于相机观测产生的约束, 如图 3(a) 所示。pose-pose 约束则依赖于系统中闭环检测产生的约束, 如图 3(b)、(c) 所示, 可有效地抑制系统的累积误差, 从而得到信息一致的地图。



(a) 相机观测约束下的位姿优化



(b) 局部闭环约束下的位姿优化



(c) 全局闭环约束下的位姿优化

图3 图的建立与优化

Fig. 3 Establishment and optimization of graph

4 闭环检测下的点云优化

在基于位姿增量估计的 SLAM 建图中, 会产生误差的累积, 则在长时间的大范围 SLAM 建图中, 就需要闭环检测对位姿和生成的点云进行优化。RGBD-SLAM 算法是在局部回环检测后, 只使用局部回环优化后的位姿作为三维地图点云的重建。ORB-SLAM 算法则是在运动跟踪过程中, 时刻检测并维护关键帧之间的点云共视关系, 同时实现对点云优化。但这种方法只适用于规模有限的稀疏点云, 并不适用于稠密化点云的优化。基于此, 提出

了一种基于局部闭环与全局闭环约束的点云优化算法,同时完成对位姿与稠密化点云的优化。算法具体步骤如下:

1) 封装构造数据帧 Frame, 根据帧间的关联程度, 筛选出关键帧序列。

2) 执行帧间配准。完成相机位姿初始估计。

3) 在相机观测约束下, 直接应用式 (6), 构建如图 3(a) 所示的图优化。求得优化后的相机的位姿变化(R, t), 将原始点云进行坐标系变换, 实现点云的拼接融合。

4) 执行闭环检测。通过动态计算闭环帧筛选阈值, 得到匹配质量较高的闭环帧候选序列。

5) 闭环检测结束后, 遍历闭环帧候选序列, 在当前帧与闭环候选帧之间, 进行 Sim3(相似变换矩阵) 求解与优化。当某一候选帧通过了 Sim3 求解与优化, 则认为闭环检测成功, 跳出对闭环帧候选序列的遍历。求得对应相似变换矩阵为

$$S = \begin{bmatrix} sR & t \\ 0 & 1 \end{bmatrix}, R \in SO(3), s \in \mathbb{R}, t \in \mathbb{R}^3 \quad (12)$$

6) 位姿优化的传播。当前帧坐标系与世界坐标系之间的变换在 5) 中已经确定并优化。通过一级相连关系的传播, 可以确定这些相连的关键帧坐标系与世界坐标系之间的相似变换矩阵。

7) 点云优化。利用 6) 中的相似变换矩阵 S , 将原始点云的世界坐标系转换到当前闭环帧处的坐标系, 进行点云的拼接融合, 实现点云的优化。

5 实验结果及分析

5.1 实验平台

为了验证算法的有效性, 并保证本文实验的一致可对比性, 故采用了知名的 TUM 标准数据集^[12], 并选用了基于 Kinect 视觉传感器采集的室内基准数据包。该基准数据包中包含 Kinect 产生的彩色图像和深度图像序列, 以及机器人的真实运动位姿。本文实验数据处理平台为一台 Intel 双核 2.4 GHz 主频的 PC 机, 运行 Ubuntu 14.04 操作系统。程序均设定以 30 帧/秒的速度同时读取数据包中的 RGB 数据和深度数据。

5.2 室内环境下的三维视觉 SLAM

5.2.1 特征提取与特征点配准

对两帧图像进行 ORB 特征提取后, 在利用特征匹配算法 (FLANN) 获得相邻帧间对应关系的基础上, 利用 RANSAC 算法对点云进行匹配, 匹配结果如图 4(c) 所示。在图 4(a) 中, 通过特征匹配算

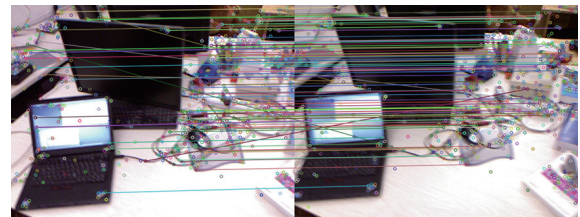
法, 一共产生了 929 个匹配特征点, 但误匹配点过多。因此, 通过人工制定规则的筛选, 去除了部分误匹配点, 此时的匹配个数为 250。为进一步提升匹配精度, 通过 RANSAC 算法对图 4(b) 进行进一步的匹配筛选。最终, 产生了 97 个经筛选后的匹配点, 实验中整个匹配所耗时间为 0.003 5 s。



(a) 特征匹配算法后



(b) 人工规则筛选后



(c) RANSAC 算法匹配

图 4 特征提取与配准实验结果

Fig. 4 The experimental results of feature extraction and matching

5.2.2 回环检测约束

通过回环检测, 可以增加帧间的约束, 进一步提升系统的鲁棒性, 进而提升整个 SLAM 的轨迹与地图的质量。在本文算法中, 通过筛选关键帧之后进行回环检测, 并以 PNP 求解的结果作为对边的估计, 增加到图优化中, 如图 5 所示。

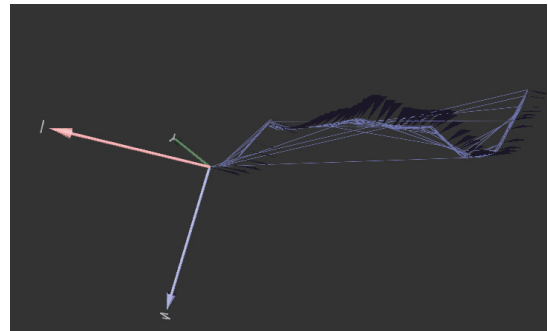


图 5 回环检测产生的约束

Fig. 5 Loop closure of constraint

5.2.3 三维环境重构效果对比

图 6(a) 为采用了 ORB-SLAM 方法得到的三维环境重构结果, 所构建的三维地图为稀疏点云。图 6(b) 为采用 RGBD-SLAM 方法得到的三维环境重构结果, 物体轮廓可以辨别, 但冗余点相对较多。图 6(c) 为采用本文的 SLAM 方法得到的三维环境重构结果, 物体轮廓清晰可辨, 冗余点明显减少。

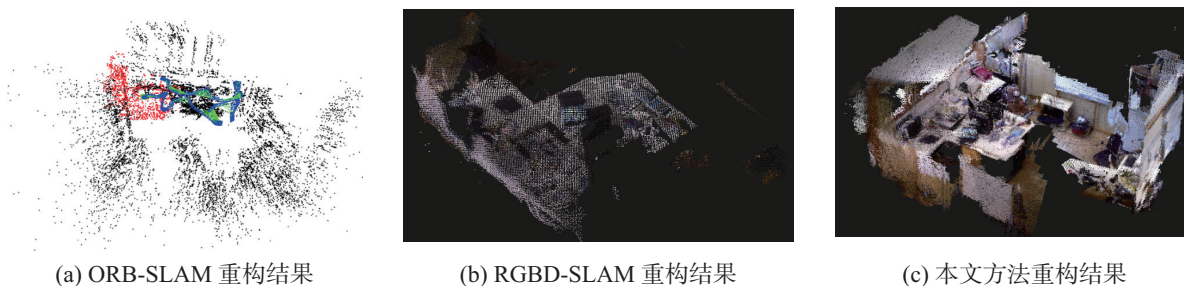


图 6 室内环境三维重构结果

Fig. 6 3D reconstruction result of the indoor environment

5.2.4 机器人运动轨迹跟踪

为了获得机器人位姿的变化, 本文以 3D-2D 的特征点约束, 采用 PNP 求解的方法对机器人相邻位姿进行估计, 并借助图优化方法对估计位姿进行优化。为了使运动轨迹更为直观, 实验给出了本文算法在不同数据集下, 机器人运动轨迹在 X/Y 平面上

本算法实时性也远优于 RGBD-SLAM, 与 ORB-SLAM 的实时性相当。RGBD-SLAM 处理一帧数据平均用时约为 350 ms, 每秒钟只能处理 3 帧数据, 因此不适于实时 SLAM。而本算法前端处理一帧数据平均约为 31 ms, 后端经过位姿优化与点云优化, 以约 5 帧/秒的频率更新点云地图。可以满足 SLAM 系统对实时性要求。

的投影, 如图 7 所示。表 1 则记录了不同算法在不同数据集下, 由计算所得的机器人在三维空间中的运动轨迹的真实位姿之间的均方根误差 (RMSE)。通过均方根误差的对比, 可有效地评估算法对机器人位姿跟踪的误差大小。可以看出本文方法能较好地实现机器人的运动轨迹跟踪。

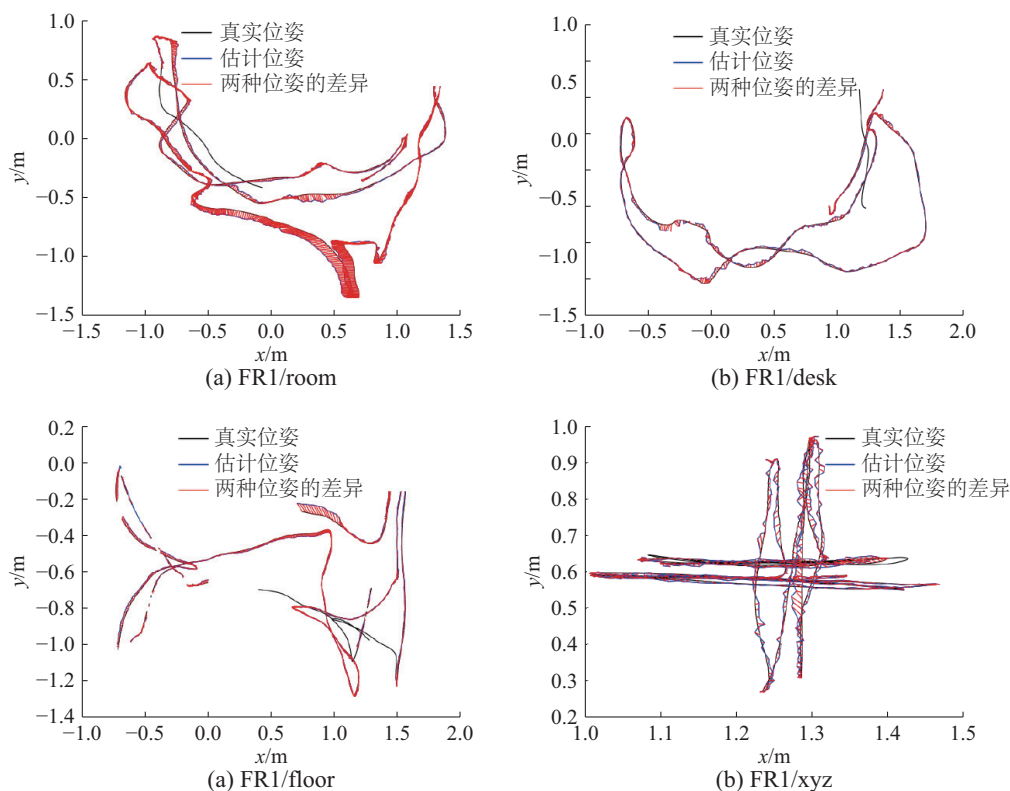


图 7 真实位姿与估计位姿对比图

Fig. 7 Comparison between ground truth and estimated trajectories

表1 不同算法位姿估计的均方根误差对比
Table 1 Pose estimation comparison of different algorithms about RMSE /m

数据来源	ORB-SLAM	RGBD-SLAM	本文方法
FR1/room	0.056	0.087	0.051
FR1/desk	0.014	0.026	0.016
FR1/floor	0.013	0.023	0.015
FR1/xyz	0.004	0.018	0.009

6 结束语

在本文基于图优化的视觉SLAM算法前端框架中,采用的ORB特征检测、RANSAC算法等方法,较好地解决了帧间配准过程中匹配误差较大、效率低下的问题。在三维地图生成部分,以图优化后的位姿为基础,进行点云构建,并引入闭环检测的机制,实现了对位姿与稠密化点云的双重优化。实验验证了本文所提方法能够准确、迅速地重构出稠密化的三维环境模型,并能精确估计出移动机器人的运动轨迹。下一步计划对系统进行两方面的改进:1)改进前端的帧间配准方法,考虑采用直接法进行位姿跟踪,完成位姿估计;2)研究环境三维地图与语义信息的结合,实现更好的人机交互与导航。

参考文献:

- [1] ZHANG Zhengyou. Microsoft kinect sensor and its effect[J]. IEEE multimedia, 2012, 19(2): 4–10.
- [2] ENDRES F, HESS J, STURM J, et al. 3-D mapping with an RGB-D camera[J]. IEEE transactions on robotics, 2014, 30(1): 177–187.
- [3] MUR-ARTAL R, MONTIEL J M M, TARDÓS J D. ORB-SLAM: a versatile and accurate monocular slam system[J]. IEEE transactions on robotics, 2015, 31(5): 1147–1163.
- [4] RUSU R B, COUSINS S. 3D is here: point cloud library (PCL)[C]//Proceedings of 2011 IEEE International Conference on Robotics and Automation. Shanghai, China: IEEE, 2011: 1–4.
- [5] HARTLEY R, ZISSERMAN A. 计算机视觉中的多视图几何[M]. 韦穗, 杨尚骏, 章权兵, 等, 译. 合肥: 安徽大学出版社, 2002: 158–256.
HARTLEY R, ZISSERMAN A. Multiple view geometry in computer vision[M]. WEI Sui, YANG Shangjun, ZHANG Quanbing, et al, Trans. Hefei: Anhui University Press, 2002: 158–256.
- [6] LABBÉ M, MICHAUD F. Online global loop closure detec-

tion for large-scale multi-session graph-based slam[C]//Proceedings of 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems. Chicago, IL, USA: IEEE, 2014: 2661–2666.

- [7] RUBLEE E, RABAU V, KONOLIGE K, et al. ORB: an efficient alternative to SIFT or SURF[C]//Proceedings of 2011 International conference on computer vision. Barcelona, Spain: IEEE, 2011: 2564–2571.
- [8] 梁明杰, 闵华清, 罗荣华. 基于图优化的同时定位与地图创建综述[J]. 机器人, 2013, 35(4): 500–512.
LIANG Mingjie, MIN Huaqing, LUO Ronghua. Graph-based SLAM: a survey[J]. Robot, 2013, 35(4): 500–512.
- [9] FADDEEV L D, RESHETIKHIN N Y, TAKHTAJAN L A. Quantization of lie groups and lie algebras[M]//KASHIWARA M, KAWAI T. Algebraic analysis: papers dedicated to professor mikio Sato on the occasion of his sixtieth birthday. Boston, MA: Academic Press, 2014: 129–139.
- [10] SHAWASH J, SELVIAH D R. Real-time nonlinear parameter estimation using the Levenberg–Marquardt algorithm on field programmable gate arrays[J]. IEEE transactions on industrial electronics, 2013, 60(1): 170–176.
- [11] KÜMMERLE R, GRISETTI G, STRASDAT H, et al. G2o: a general framework for graph optimization[C]//Proceedings of 2011 IEEE International Conference on Robotics and Automation. Shanghai, China: IEEE, 2011: 3607–3613.
- [12] STURM J, MAGNENAT S, ENGELHARD N, et al. Towards a benchmark for RGB-D SLAM evaluation[C]//RGB-D Workshop on Advanced Reasoning with Depth Cameras at Robotics: Science and Systems Conference. Los Angeles, United States: RSS, 2011: 1–2.

作者简介:



张毅,男,1966年生,教授,博士生导师,中国人工智能学会理事,国家信息无障碍研发中心主任,主要研究方向为智能系统与移动机器人、机器视觉与模式识别、多传感器信息融合。主持完成国家级和省部基金项目10余项,发表学术论文100余篇,被SCI/EI和ISTP收录30余篇次,出版著作5部,获国家发明专利30余项。



沙建松,男,1991年生,硕士研究生,主要研究方向为机器人同时定位与地图创建(SLAM)、基于语义信息的机器人三维视觉导航。