

文章编号 1004-924X(2017)增-0160-07

RGB-D 视觉 SLAM 算法的改进及应用

成 怡^{1,2}, 白佳奇^{1,2*}, 修春波^{1,2}

- (1. 天津工业大学 电气工程与自动化学院, 天津 300387;
2. 天津市电工电能新技术重点实验室, 天津 300387)

摘要:针对传统视觉 SLAM 算法位置误差积累及计算量大上的突出问题,提出了一种改进的 RGB-D 的视觉 SLAM 算法。在提出的算法中采用 SURF 算法提取特征,利用随机采样一致性和迭代最近点算法根据匹配结果估算并优化运动过程;在后端图优化的过程中,引入半随机的回环检测,通过改变关键帧的提取方法,克服当前帧最近检测与历史帧回环检测的缺点,提高匹配的准确性,同时减少了计算量。通过开源机器人操作系统的实验结果表明,提出的改进 RGB-D 的视觉 SLAM 算法处理一帧数据的速度为 0.030 s 左右,在满足了所得 3D 点云地图的准确性和完整性的基础上,加快了处理速度,节省了计算量,增强了算法的实时性。

关键词:RGB-D;同时定位与地图创建;特征点匹配;回环检测

中图分类号:TP391.41 **文献标识码:**A **doi:**10.3788/OPE.20172513.0160

Improvement and application of RGB-D vision SLAM algorithm

CHENG Yi^{1,2}, BAI Jia-qi^{1,2*}, XIU Chun-bo^{1,2}

- (1. Tianjin Polytechnic University, Electronic Engineering and Automation, Tianjin 300387, China;
2. Key Laboratory of Advanced Electrical Engineering and Energy Technology, Tianjin 300387, China)

* Corresponding author, E-mail:bjiaqi711@163.com

Abstract: An improved visual SLAM algorithm of RGB-D was proposed aimed at prominent problem of accumulative position error and large calculated amount of traditional visual SLAM algorithm under the application background of wheeled mobile robot in the thesis. SURF algorithm was used to extract features, and random sample consensus (Ransac) and iterative closest point (ICP) algorithm were used to estimate and optimize movement process according to matching result in proposed algorithm; semi-random loop detection was introduced in the optimization process of back-end graph. Defects of recent detection of current frame and loop detection of historical frames were overcome by changing extraction method of key frames, which improved accuracy of matching and reduced calculated amount at the same time. Experimental result of open-source robot operating system (ROS) shows that speed of processing one frame data by proposed improved visual SLAM algorithm of RGB-D is 0.030 s approximately, which accelerates processing speed, saves calculated amount and strengthens real-time performance of the algorithm on the basis

收稿日期:2017-04-28;修订日期:2017-06-10.

基金项目:国家自然科学基金资助项目(No. 61203302);天津市应用基础与前沿技术研究计划资助项目(No. 14JCYBJC18900)

of satisfying accuracy and completeness of obtained 3D point cloud map.

Key words: RGB-D; Simultaneous Localization and Mapping(SLAM); feature point matching; loop closures detection

1 引 言

同时定位与地图构建(Simultaneous Localization And Mapping, SLAM)一直是移动机器人最具挑战性的问题之一。早期的 SLAM 研究侧重于使用滤波器理论来最小化运动体位姿和地图的路标点的噪声。随着 SLAM 问题的研究和各种视觉传感器的发展,基于各种视觉传感器的视觉 SLAM 算法得到了广泛的关注。特别是最近几年推出的微软的 Kinect 深度摄像头或华硕的 Xtion Pro Live,因其价格低、大小和重量合适^[1]等优点被应用于各种地图创建^[2]和 3D 环境重建^[3]中。

近年来,学者们开始借鉴运动恢复结构(Structure from Motion, SfM)中的方式,以优化理论为基础求解 SLAM 问题。该方法在三维环境重建^[4]和机器人避障导航^[5]等方面取得了一定的成就,并在视觉 SLAM 领域中取得了主导地位。例如 Henry 等^[6]人提出了使用 FAST(Features from Accelerated Segment Test)算法和 ICP(Iterative Closest Point)算法相结合,利用 Kinect 相机实现了构建密集 3D 室内环境地图的 SLAM 算法。Dryanovski 等^[7]人提出的卡尔曼滤波与 ICP 算法相结合的 RGB-SLAM 算法和 Endres 等^[8]人提出的基于图像特征匹配的 RGB-D SLAM 算法。目前较成熟的算法是 Endres 等人提出的 RGBD-SLAM-v2 算法,该算法整合了 SLAM 领域里的图像特征、点云、图优化、回环检测等技术^[9],不仅 UI 界面做的美观,也基本可以完成同时定位与建图的过程。但是由于该算法的代码庞大,提取 SIFT 特征、渲染点云等很费时,导致算法的处理速度慢,实时性不好,并且程序采集关键帧的频率很高,不适合做长时间的 SLAM 等。

在执行视觉 SLAM 的过程中,机器人环境地图的构建是一个递推的过程,即根据以前时刻的图像位置推算下一时刻的图像位置,并且不断地更新,在这个过程中,会产生不可避免的误差累

积。所以,为了保证构建地图的准确性,通常会在后端图优化过程中加入回环检测,即把当前帧与离当前帧最近的历史帧中的几帧进行匹配,判断该帧是否为关键帧,进而判断机器人是否来过这个地方。但是,此方法的准确性不高,所得到的地图有过多的重复错误。为了改进这一问题,本文提出了一种半随机回环检测方法,即在回环检测的过程中,改变了历史帧的选取准则,由半随机提取方式代替最普通的就近选取方式,从而增加匹配的准确性,减小累积误差。

2 视觉 SLAM 算法基本原理

移动机器人视觉 SLAM 包含定位和建图两个过程。机器人在未知的环境中运动,利用其自身配置的传感器计算其位置信息及周围环境信息,图 1 描述了移动机器人视觉 SLAM 的原理,其中 X_p^i 表示 $i(i=1, \dots, n)$ 时刻机器人的位置, X_l^j 表示第 $j(j=1, \dots, m)$ 个路标,即在图像处理过程中由特征提取得到的关键点, u 表示机器人的运动信息, z 表示观测信息。

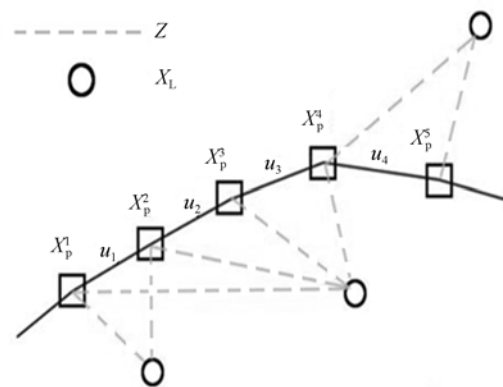


图 1 移动机器人 SLAM 系统示意图

Fig. 1 Schematic diagram of mobile robot SLAM system

视觉 SLAM 算法描述载体运动的方程为:

$$X_p^{i+1} = f(X_p^i, u_i) + \omega_i, \quad (1)$$

其中 ω_i 是系统噪声。系统的观测方程为:

$$z_{i,j} = h(X_p^i, X_l^i) + v_{i,j}, \quad (2)$$

其中 $v_{i,j}$ 是观测噪声,公式(2)表示在第 i 帧中看到了第 j 个路标。视觉 SLAM 算法利用两帧间的位移差,估计移动机器人的运动,然后将运动累加起来,最后得到最终的运动轨迹和地图。

3 改进 RGB-D 视觉 SLAM 算法

RGB-D 视觉 SLAM 算法的工作流程,如图 2 所示。算法分为图像前端处理和后端位姿优化两部分^[10]。前端主要进行数据处理,将 2D 图像数据转成 3D 点云,对环境特征进行提取和匹配,利用随机采样一致性(Ransac)和 ICP 算法得到当前帧与历史帧的平移、旋转向量,进行移动机器人的位姿估计,并对初始估计结果进行修正;后端位姿优化的核心是利用图优化软件 G2O 对位姿进行优化,最后得到一个 3D 点云地图和移动机器人运动的轨迹。

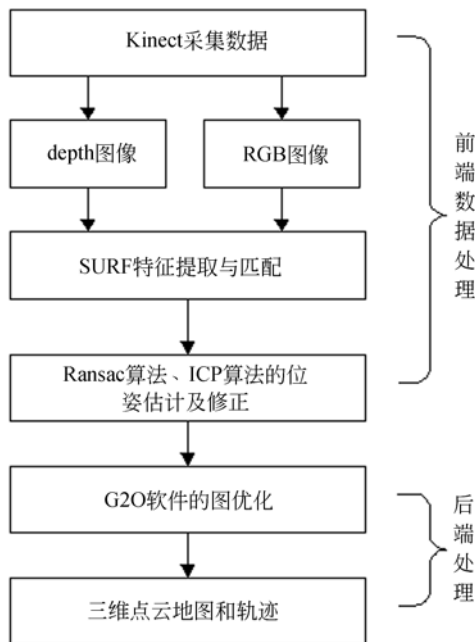


图 2 算法流程图

Fig. 2 Flow-sheet of algorithm

3.1 前端数据处理

3.1.1 特征点提取

Kinect 传感器采集到新的图像数据后,首先要进行图像特征的提取与匹配,特征点匹配的精确度直接影响算法累积误差的大小,匹配越精确,累积误差越小,最终得到的机器人路径越接近真

实路径,3D 环境地图越接近真实的环境。

常用的特征提取算法有 SIFT 算法^[11-12]、SURF 算法^[13]和 ORB(Oriented FAST and Rotated BRIEF)算法^[14]等,并且各种算法的优缺点不同。SIFT 算法提取的特征点多,误差小,但速度慢;SURF 算法提取特征点相对减少,但是速度快,鲁棒性能好;ORB 算法虽然速度很快但是鲁棒性能差。本文选用 SURF 算法来提取特征,得到特征点的二维坐标 (u, v) 。

3.1.2 点云配准与姿态估计

每一个点云都由 r, g, b, x, y, z 这 6 个分量分别表示它们的颜色 (r, g, b) 和空间位置 (x, y, z) 。

假设 RGB 图像中某个特征点的像素坐标为 $[u, v, d]$,其中 d 是在相应的 DEPTH 图像中得到的深度值,则该特征点相应的三维空间坐标为 $[x, y, z]$,两坐标之间的对应关系如式(3)所示:

$$\begin{cases} x = (u - c_x) \cdot \frac{z}{f_x} \\ y = (v - c_y) \cdot \frac{z}{f_y} \\ z = \frac{d}{s} \end{cases}, \quad (3)$$

其中, f_x, f_y, c_x, c_y 为相机的内参, s 指深度图的缩放因子。

得到二维特征匹配点集合后,对于集合中的每一对特征匹配点,利用公式(3)得到特征点在空间中的三维坐标,构成三维坐标匹配点集。根据该集合中匹配点之间位置变化关系,利用 Ransac 算法进行位姿估算,得到一个初始的运动变换,然后利用 ICP 算法进行优化^[15],以此进行运动变化修正,得到更精确的运动变换。

假设已知有两个帧 F_1 和 F_2 ,并且其对应的特征点集为:

$$\begin{cases} P = \{p_1, p_2, \dots, p_n\} \in F_1 \\ Q = \{q_1, q_2, \dots, q_n\} \in F_2 \end{cases}, P, Q \in \mathbf{R}^3. \quad (4)$$

利用 Ransac 算法求解旋转矩阵 \mathbf{R} 和位移矢量 \mathbf{t} ,得到位姿变换矩阵 \mathbf{T} ,进行姿态估计。该过程得到的运动变换模型是用变换矩阵来描述的:

$$\mathbf{T} = \begin{bmatrix} \mathbf{R}_{3 \times 3} & \mathbf{t}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \in \mathbf{R}^{4 \times 4}, \quad (5)$$

其中, \mathbf{R} 是 3×3 的旋转矩阵, \mathbf{t} 是 3×1 的位移矢量。

在移动机器人的实际运动过程中,存在着不

可避免的误差,则利用 ICP 算法进行运动变换修正。即通过最小化一个误差来求解 \mathbf{R} 和 \mathbf{t} :

$$\min_{\mathbf{R}, \mathbf{t}} \sum_{i=1}^n \| p_i - (\mathbf{R}q_i + \mathbf{t}) \|_2. \quad (6)$$

3.2 后端地图创建与优化

机器人的位姿估计由 SLAM 算法前端计算完成,算法的后端包括回环检测、图优化、地图创建和生成轨迹。

3.2.1 半随机回环检测

回环检测的本质就是识别曾经到过的地方。最简单的回环检测方法,就是把新的一帧数据与之前所有的关键帧进行比较,但是随着时间变长,需要比较的帧就会变得越多,这样会严重影响程序的运行速度。目前回环检测算法有:Random Ferns 算法^[16]、Bag-of-Words (BoW) 算法^[17]等。其中比较常用的是 BoW 算法。但是 BoW 算法要求事先构造视觉词典,所以该方法前期工作较为复杂。因此,SLAM 算法的研究学者们一直在探索高效的回环检测的方法。

本文提出了一种半随机回环检测方法,即将新得到的当前帧数据与按照一定的间隔在历史帧中提取的关键帧进行比较。具体步骤为:当得到新的一帧数据时,首先需要在历史帧序列中每隔 t (随帧数的变化而变化)帧数据提取一个关键帧,共取 m 个关键帧,同时在序列的末尾提取 n 个关键帧,然后将该帧数据和之前采取的 $m+n$ 个关键帧进行匹配。如果匹配结果符合保留条件,则将此帧数据放入关键帧序列的末尾;如果匹配结果符合丢弃条件,则将此帧数据丢弃。其流程如图 3 所示。

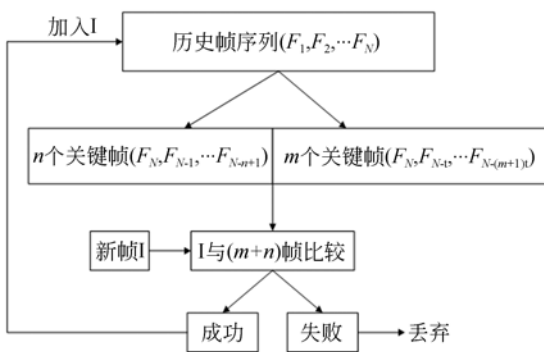


图 3 半随机回环检测算法流程图

Fig. 3 Flow-sheet of semi-random loop closures detection method

其中 t 可根据公式(7)取值:

$$t = \text{INT}\left(\frac{F}{4 \times i}\right) \quad i = 1 \sim m, \quad (7)$$

其中, F 为数据帧的总数; m, n 的取值受算法运行速度及所得地图精度直接影响,通常取 $10 \sim 20$ 。

3.2.2 图优化

为了得到全局最优的运动变换约束和最优机器人位姿,需要对姿态图进行图优化处理。常用的图优化模型框架有 GraphSLAM、TORO 和 G2O (General Graph Optimization)。其中 G2O 带有各种各样的求解器,提供了精确修正误差函数并使其最小化的方法,最具通用性。因此,在本文中使用了 G2O 进行处理,优化姿态图。

姿态图由节点和边构成: $G = \{V, E\}$, 机器人的各个姿态为姿态图的节点:

$$V_i = [x, y, z, q_x, q_y, q_z, q_w] = T_i = \begin{bmatrix} \mathbf{R}_{3 \times 3} & \mathbf{t}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}_i. \quad (8)$$

边是指两个节点之间的变换:

$$E_{i,j} = T_{i,j} = \begin{bmatrix} \mathbf{R}_{3 \times 3} & \mathbf{t}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}_{i,j}. \quad (9)$$

优化图大体图形如图 4 所示(具体图形由帧间匹配的约束决定):

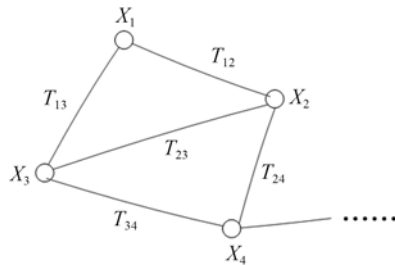


图 4 姿态图

Fig. 4 Poses graph

其中, $X_i (i = 1, 2, 3, \dots)$ 为姿态图的节点, $T_{i,j} (i, j = 1, 2, 3, \dots)$ 为姿态图的边。

图优化过程结束后,即得到机器人运动的轨迹和重建的三维点云地图。

4 实验结果及分析

本实验基于移动机器人 ROS 操作系统,实验硬件平台为一台配置为四核 i3 处理器、内存为 4G 的联想笔记本电脑。

本实验视觉图形来源于标准测试数据集 FR1/room 数据包,该数据集包含 1 300 帧 RGB 和 DEPTH 图像、与其对应的 groundtruth 数据、和标准的运行轨迹。实验场景为包含桌子、椅子、柜子和人等多种信息的实验室,由算法所得图像为全方位的图像。

首先,将本算法的半随机回环检测与近距离回环检测算法进行实验比较。实验所得结果如图 5 和图 6 所示:

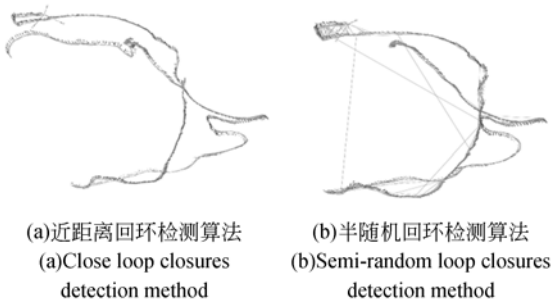


图 5 G2O 位姿优化结果
Fig. 5 G2O pose optimization results

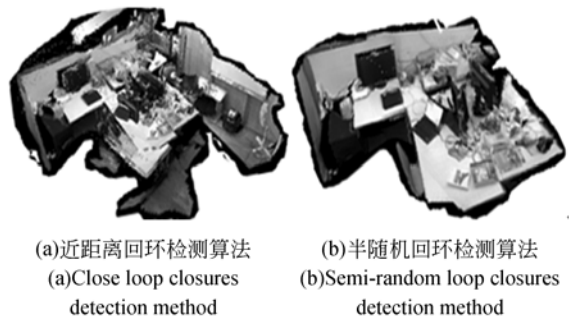


图 6 部分 3D 点云地图
Fig. 6 3D point cloud map

由图 5 可以看出,由于近距离回环检测算法位姿之间的约束力度不足,程序运行中存在很大的误差。但本文提出的半随机回环检测算法很好地解决了误差较大的问题。由图 6 也能直观地看到,因为存在误差的原因,左边图中桌子出现了明显的不重合,而右边的地图准确性得到了明显地提高。同样在图 7 的位姿误差分析图中可以看出本文提出的半随机回环检测算法更接近真实值,减小了运动过程中的累积误差。

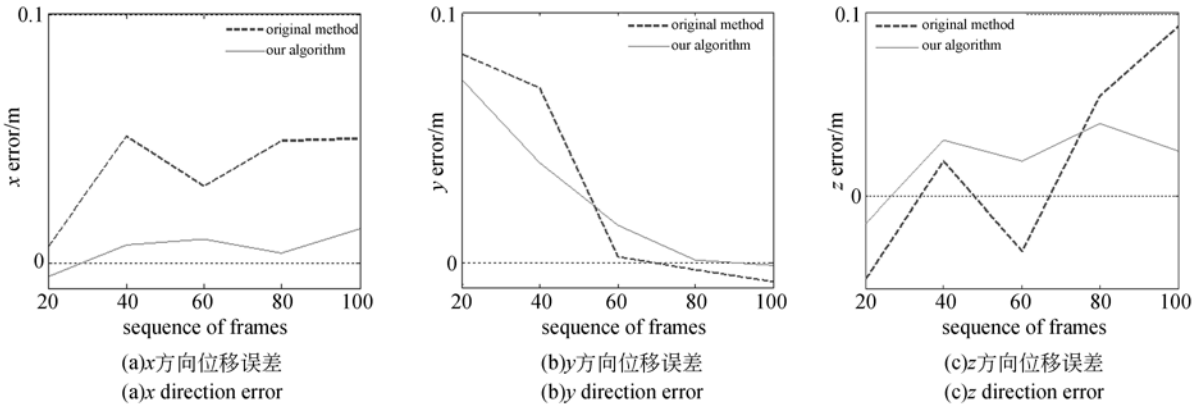


图 7 位姿误差分析图
Fig. 7 Position error analysis

算法执行时间对比结果如表 1 所示,从表 1 可以看出,传统的 RGBD-SLAM 算法处理一帧数据平均用时 0.318 s,不适用于长时间实时 SLAM 应用。本文算法处理一帧数据需要 0.030 s 左右,基本满足实时性的要求。本文提出的改进方法运行速度快,优于传统的 RGBD-SLAM 算法。

表 1 改进算法与传统算法运行时间对比

Tab. 1 Comparison of processing time between improved and conventional algorithms

数据包/room	长度/mm	运行时间/s	
		传统 RGBD	本文方法
FR1	15.99	0.318	0.030

改进算法得到的实验 3D 环境地图如图 8 所示,实验轨迹如图 9(b)所示:



图 8 3D 环境点云地图

Fig. 8 3D environment point cloud map

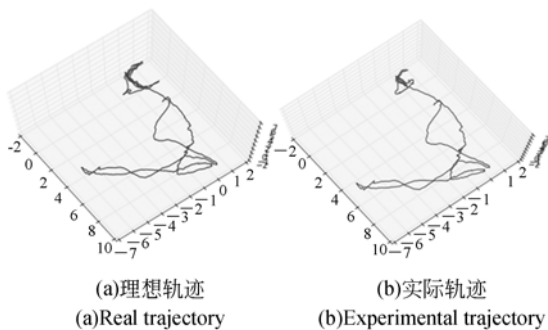


图 9 运动轨迹图

Fig. 9 Trajectory results

由图 8 所示的 3D 点云环境地图可以发现:从地图的任何角度观测,都可以清晰地辨认出电

脑、桌子、柜子、椅子、人及其他物体,与现实相符度很高,说明改进算法使得现实场景得到了较好地重建。同时,将实验获得的运动轨迹与真实的运动轨迹进行对比,对比的结果如图 9 所示,从图中可以看出:改进算法估计的运动轨迹与真实轨迹相差无几。

因此,本文提出的后端的半随机回环检测方法减小了运动过程中的累积误差,满足了构建地图和生成轨迹的准确性,从而很好地完成了机器人的同时定位与地图构建。改进算法的计算量较小,可以在很大程度上提高算法的处理速度,并且位姿误差满足实验误差范围,具有高度的准确性和良好的实时性。

5 结 论

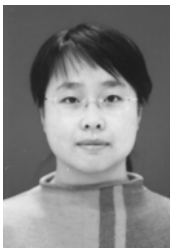
本文提出了一种改进的 RGB-D 视觉 SLAM 算法,在算法后端利用一个半随机回环检测的方法,克服传统算法中计算量大和计算时间长的问题。并通过机器人开源操作系统 ROS 进行实验,验证了提出的改进算法提高了构建的三维地图的准确性,减少了误差的积累,使数据处理过程中,每帧处理速度的平均用时缩短为 0.030 s,在很大程度上提升了算法的运算速度,增强了本算法的应用实时性。

参考文献:

- [1] ENDRES F, HESS J, STURM J, et al.. 3D mapping with an RGB-D camera[C]. *Robotics, IEEE Transactions on*, 2014, 30(1):1-11.
- [2] ALMANSA-VALVERDE S, CASTILLO JC, CABALLER OA. Mobile robot map building from time-of-flight camera [J]. *Expert Systems with Applications*, 2012, 39: 8835-8843.
- [3] JO S J, JO H G, JUN S W, et al.. 3D Environment reconstruction using a ToF camera[C]. *Proceedings of the 9th Korea Robotics Society Annual Conference*, Buyeo, Korea, 2014.
- [4] 王欣,袁坤,于晓,等.基于运动恢复的双目视觉三维重建系统设计[J]. *光学精密工程*, 2014, 22(5): 1379-1387.
WANG X, YUAN K, YU X, et al.. Design of binocular vision 3D reconstruction system based on motion recovery [J]. *Opt. Precision Eng.*, 2014, 22(5):1379-1387. (in Chinese)
- [5] 林义阔,吕乃光,姜小平,等.用于弱纹理场景三维重建的机器人视觉系统[J]. *光学精密工程*, 2015, 23(2):540-549.
LIN Y M, LV N G, LOU X P, et al.. Robot vision system for 3D reconstruction in low texture environment[J]. *Opt. Precision Eng.*, 2015, 23(2): 540-549. (in Chinese)
- [6] HENRY P, KRAININ M, HERBST E, et al.. RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments[C]. *12th International Symposium on Experimental Robotics, Berlin, Germany; Springer*, 2014: 477-491.
- [7] DRYANOVSKI I, VALENTI R G, XIAO J Z. Fast visual odometry and mapping from RGB-D data[C].

- IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2013: 2305-2310.
- [8] ENDRES F, HESS J, ENGELHARD N, *et al.*. An evaluation of the RGB-D SLAM system[C]. *IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2012:1691-1696.*
- [9] 王忠立, 赵杰, 蔡鹤皋. 大规模环境下基于图优化 SLAM 的图构建方法[J]. *哈尔滨工业大学学报*, 2015, 1(1):75-85.
WANG ZH L, ZHANG J. Method for constructing graph based optimization SLAM in large scale environment [J]. *Journal of Harbin Institute of Technology*, 2015, 1(1):75-85. (in Chinese)
- [10] 付梦印, 吕宪伟, 刘彤, 等. 基于 RGB-D 数据的实时 SLAM 算法[J]. *机器人*, 2015, 37(6):683-692.
FU M Y, LIU T. Real time SLAM algorithm based on RGB-D data[J]. *Robot*, 2015, 37(6): 683-692. (in Chinese)
- [11] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. *International journal of computer vision*, 2004, 60(2):91-110.
- [12] 曾峦, 王元钦, 谭久彬. 改进的 SIFT 特征提取和匹配算法[J]. *光学精密工程*, 2011, 19(6): 1391-1397.
ZENG L, WANG Y Q, TAN J B, *et al.*. Improved SIFT feature extraction and matching algorithm[J]. *Opt. Precision Eng.*, 2011, 19(6): 1391-1397. (in Chinese)
- [13] BAY H, ESS A, TUYTELAARS T, *et al.*. Speeded-up robust features (SURF) [J]. *Computer vision and image understanding*, 2008, 110(3):346-359.
- [14] RUBLEE E, RABAU D V, KONOLIGE K, *et al.*. ORB: An efficient alternative to SIFT or SURF [C]. *IEEE International Conference on Computer Vision. Piscataway, USA: IEEE, 2011: 2564-2571.*
- [15] 肖军浩[译], 机器人操作系统(ROS)浅析[M]. 2013.
JASON M. O'Kane. A Gentle Introduction to ROS [M]. 2013. (in Chinese)
- [16] KERL C, STURM J, CREMERS D. Dense Visual SLAM for RGB-D Cameras[C]. *In Proc. of the Int. Conf. on Intelligent Robot Systems (IROS)*, 2013.
- [17] SCHERER S A, DUBE D, ZELLI A. Using depth in visual simultaneous localization and mapping [C]. *2012 IEEE International Conference on Robotics and Automation (ICRA)*, 2012: 5216-5221.

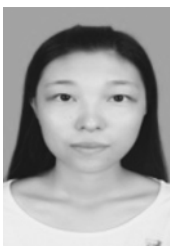
导师简介:



成 怡(1979—), 女, 黑龙江富裕, 博士, 副教授, 硕士生导师, 2004 年、2008 年于哈尔滨工程大学分别获得硕士、博士学位, 主要研究方向为机器人被动导航和视觉导航。E-mail: chengstu@163.com



修春波(1978—), 男, 黑龙江大庆, 博士, 教授, 硕士生导师, 2005 年于北京理工大学获得博士学位, 主要研究方向为神经网络和混沌控制等。E-mail: xiuchunbo@tjpu.edu.cn



白佳奇(1991—), 女, 河北唐山, 硕士研究生, 主要研究方向为机器人视觉 SLAM 算法的研究。E-mail: bjiaqi711@163.com