

DOI:10.3969/j.issn.1673-4785.201504003
网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20150630.1541.001.html>

室内环境下移动机器人三维视觉 SLAM

张毅¹, 陈起¹, 罗元²
(1.重庆邮电大学 自动化学院, 重庆 400065; 2.重庆邮电大学 光电工程学院, 重庆 400065)

摘 要:在移动机器人三维视觉 SLAM 中, 针对点云配准过程存在的匹配误差较大、效率低下的问题, 提出一种分层式点云配准策略。利用 RANSAC 算法对可能存在的误匹配点进行剔除完成初始配准, 对所得内点进行最小二乘算法下机器人相邻位姿的估计, 并采用改进的 ICP 算法完成精确配准, 最后采用 g^20 优化方法对移动机器人位姿进行优化。针对点云拼接结果中存在重复点过多以及空洞的问题, 引入基于体积空间融合的匹配算法对空洞进行填补。实验验证了本文所提方法能够准确重构出环境三维模型和精确估计出移动机器人的运动轨迹, 并有效去除拼接结果中的重复点, 填补了空洞。

关键词:移动机器人; Kinect 相机; 同时定位与地图构建; 点云配准; 轨迹跟踪; 机器人导航; 鲁棒性

中图分类号: TP24 **文献标志码:** A **文章编号:** 1673-4785(2015)04-0615-05

中文引用格式: 张毅, 陈起, 罗元. 室内环境下移动机器人三维视觉 SLAM[J]. 智能系统学报, 2015, 10(4): 615-619.
英文引用格式: ZHANG Yi, CHEN Qi, LUO Yuan. Three dimensional visual SLAM for mobile robot in indoor environment[J]. CAAI Transactions on Intelligent Systems, 2015, 10(4): 615-619.

Three dimensional visual SLAM for mobile robots in indoor environments

ZHANG Yi¹, CHEN Qi¹, LUO Yuan²
(1.Automation College, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; 2.College of Electroning Engineering College, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: In this paper, it aiming at the problems of large matching error and inefficiency in the process of point cloud registration in mobile robots 3D visual SLAM, a hierarchical point cloud registration strategy is proposed. It completes the initial registration by eliminating possible false matching points by RANSAC algorithm and then estimates the robot adjacent pose using the least square algorithm of interior points. Meanwhile, accurate registration is completed by the improved ICP algorithm and finally the mobile robot pose is optimized by the g^20 optimization method. Addressing the problem of too much repeated points and hole existing in point cloud splicing results, a matching algorithm based on volume space fusion is introduced to fill the hole. Experimental results obtained demonstrated that the proposed method can accurately reconstruct the 3D model and estimate the trajectory of the mobile robot, and can effectively remove the repeated points and fill the hole.

Keywords: mobile robot; Kinect camera; SLAM; point cloud registration; trajectory tracking; robot navigation; robustness

近年来, 移动机器人同时定位与地图构建已经成为机器人领域研究的热点^[1]。过去, 国内外学者对移动机器人二维 SLAM 进行了大量研究, 然而在二维 SLAM 中丢失了很多有价值的形状和几何信息。三维 SLAM 能够充分利用环境三维属性, 精确描述出局部环境形状以及物体的几何外形, 因而移动机器人三维 SLAM 对于全面实现移动机器人精确导航具有重要意义。在三维环境建模中最常用的方式是利用范围传感器获得的深度点云来表示环境模型^[2]。然而, 这种方法大多利用精度高、价格昂贵的三维激光扫描器。Kinect 具有廉价、方便的特点^[3]。近年来, 将 Kinect 应用于移动机器人三维 SLAM 已成为机器人领域研究的新热点, 并取得了大量研究成果。文献^[4]提出一种室内环境下的三维环境建模方法, 基于 GPU 编程的 SIFT 算法提高了系统的运算效率, 且能实现三维环境的建立, 但效果欠佳且无法进行位姿估计。Henry^[5]等人提出了

收稿日期: 2015-04-07. 网络出版日期: 2015-06-30.
基金项目: 国家自然科学基金资助项目(51075420).
通信作者: 陈起. E-mail: chenqisansan@163.com.

一个基于 Kinect 的交互式的三维模型重建系统,但是该系统依靠关键帧的选取,计算量大、时间复杂度高,难以用于移动机器人三维地图在线创建。针对上述问题,本文提出了一种室内环境下移动机器人三维视觉 SLAM 方法。针对点云配准过程存在的匹配误差较大、效率低下的问题,提出一种分层式点云配准策略。同时引入基于体积融合的匹配算法完成精确配准,有效的去除了拼接结果中的重复点,采用

了 g^20 优化方法对机器人位姿进行优化,实验验证了本文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 系统的总体框架如图 1 所示,主要分为特征点检测与匹配、Kinect 标定与点云生成、分层式点云配准 3 部分。

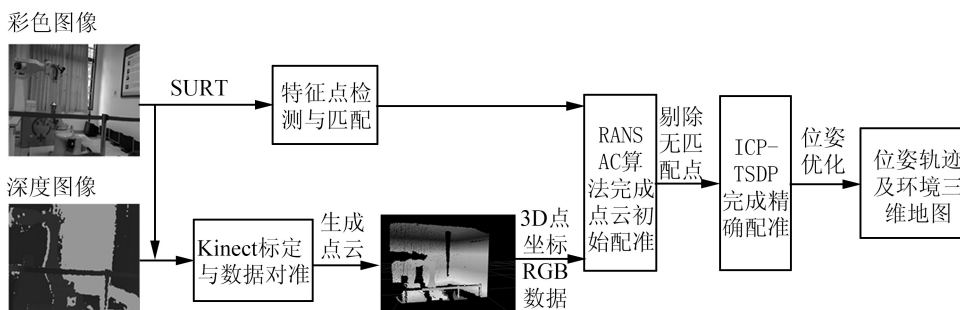


图 1 室内环境下移动机器人三维视觉 SLAM 总体框架图

Fig. 1 Total frame of three dimensional visual SLAM for mobile robot in indoor environment

2 特征点检测与匹配

经典的特征点检测算法包括 Harris^[6]、SIFT^[7]、SURF^[8] 算法等。经综合考虑为获得鲁棒性、稳定性更好的特征点检测结果,本文采用具有旋转和尺度不变性且速度较快的 SURF 算法。SURF 特征点提取过程主要包括特征点检测、特征点主方向选取、生成特征描述子。判断基于描述的特征点是否匹配,可通过计算欧式距离来实现,设匹配图像的特征点集合分别为 X 和 Y 。欧式距离可以描述为

$$D(X, Y) = \|X - Y\| = \sqrt{\sum_{i=1}^d (X_i - Y_i)^w} \quad (1)$$

欧式距离越小表示特征点相似程度越高,本文在匹配过程中采用了最大紧邻向量匹配方法,即通过比较特征点集合中的每一个特征点与特征点集合中各特征点的欧式距离, D_1 和 D_2 分别代表最近和次近的欧式距离,若 $D_1 \leq \alpha D_2$ (α 为设定的最近次近距离比值,本文设定 $\alpha = 0.60$),则认为两个特征点匹配,反之丢弃该点。

3 标定与点云生成

为获取 Kinect 的彩色与深度相机的内参数矩阵以及彩色与深度相机之间的刚体变换,减小深度图像与红外图像之间存在的漂移^[9],需要对 Kinect 进行标定,本文采用文献[10]标定方法对 Kinect 进行标定。在完成 Kinect 对准校正后,可根据图像中任意一点对应的深度值得到该点三维坐标,进而生成三维彩色点云数据。给定任意深度图像像素点

(x_d, y_d) , 投射到三维空间点 $p(x, y, z)$, 其坐标计算如下:

$$x = \frac{(x_d - c_{xd}) \times \text{depth}(x_d, y_d)}{f_{xd}} \quad (2)$$

$$y = \frac{(x_d - c_{yd}) \times \text{depth}(x_d, y_d)}{f_{yd}} \quad (3)$$

$$z = \text{depth}(x_d, y_d) \quad (4)$$

式中: $P_j = P_i T_i^j$ 为当前像素点深度值, c_{xd} 和 c_{yd} 为图像光心坐标, f_{xd} 和 f_{yd} 为深度相机焦距, 其中 f_{xd} f_{yd} 均可由标定得到。

4 分层式点云配准

4.1 RANSAC 算法粗匹配

本文采用 RANSAC 算法进行特征点的粗匹配,以使 2 组点集的公共区域能够大致重合,为下一步精确匹配做准备。算法设定一个容许误差将所有的匹配点对分为内点和外点。本文实验中设定的阈值 $d=3.00$ 舍去误差大于 d 的匹配点对(外点)。为了获得摄像机位姿的变化,本文采用了最小二乘逼近的方法^[11]。利用算法所得内点进行最小二乘算法下机器人相邻位姿的估计。摄像机第 i 时刻的位姿 P_i 相对于第 j 时刻的摄像机位姿 P_j 的转换关系为 $P_j = P_i T_i^j$ 其中 T_i^j 为刚性变换矩阵:

$$T_i^j = \begin{bmatrix} \mathbf{R}_{3 \times 3} & \mathbf{t}_{3 \times 1} \\ 0 & 1 \end{bmatrix} \quad (5)$$

式中: $\mathbf{R}_{3 \times 3}$ 为旋转矩阵; $\mathbf{t}_{3 \times 1}$ 为平移矩阵。对于移动机器人同步定位与地图构建,可通过下式获得各时刻数据相对于初始时刻的变换矩阵:

$$T_0^k = T_0^1 * T_1^2 \cdots T_{k-1}^k \tag{6}$$

4.2 改进的 ICP 算法精确配准

传统的 ICP 算法的基本原理利用最小二乘的优化思想通过计算使函数(7)中的 R 和 T 最小化。

$$f(R,T) = \sum_{i=1}^N \| Q_i - RS_i - T \|^2 \tag{7}$$

式中: N 为对应点对的数量, R 为 3×3 的旋转矩阵, T 为三维的平移向量, S_i 和 Q_i 分别为初始和目标数据点集。ICP 匹配算法是基于两片点云是完全重合这一假设,因此,理论上总是能搜索到最近点。但是在实际应用中,两片点云之间只有部分是重合的,因此,在实际应用中 ICP 匹配算法容易陷入局部最优,甚至不能收敛,要对 ICP 算法的搜索策略进行改进。

本文采用了点到切平面的最近搜索算法。另外,在初值的选取上本文采用了阈值法,利用欧氏距离阈值和方向向量阈值,提高了初值的匹配准确度。欧氏距离阈值法可剔除噪声点,提高初值选择成功率。设有点云 S 和 Q ,对 S 中的一个点 s_i ,搜索 Q 中与 s_i 欧氏距离最近的 3 个点,分别为 q_1 、 q_2 和 q_3 。若满足 $|s_i - q_i| \leq T$,则为正确的匹配点,反之为误匹配点予以剔除。式中 T 为距离超出阈值,经过欧氏距离阈值可基本剔除点云数据的噪声点。然后结合方向向量夹角判断是否为对应点,提高初值选择的正确率。若 s_i 、 q_i 为一对正确的匹配点,在满足式(7)的基础上还应当满足下式:

$$\arccos \frac{n_{s_i} \times n_{q_i}}{|n_{s_i}| \times |n_{q_i}|} \leq \theta \tag{8}$$

式中, n_{s_i} 、 n_{q_i} 分别为 s_i 和 q_i 所在曲面的法向量, θ 为设定的方向向量夹角阈值。经上述分析,本文改进的 ICP 算法步骤如下:

- 1) 在 S 中选择初始点集 s_{i0} 。
- 2) 使用欧氏距离阈值法在目标点集 Q 中求出与 s_{i0} 最近的点集 q_{i0} 、 s_{i0} 和 q_{i0} 构成对应点集。
- 3) 应用方向向量阈值法定点集 s_{i0} 和 q_{i0} ,运算初值点集 s_{i1} 和 q_{i1} 。
- 4) 应用 SVD 法求得点集 s_{i1} 和 q_{i1} 之间的旋转矩阵 R 和平移矢量 T 。
- 5) 根据式 $s_{i2} = R_1 s_{i1} + T_1$,计算数据点集 s_{i1} 经过一次坐标变换后所得到的数据点集 s_{i2} ,然后重复步骤 3)~5),直到满足式(9):

$$\begin{cases} d_k - d_{k+1} < \varepsilon \\ d_k = \frac{1}{N} \sum_{i=1}^N \| q_{ik} - s_{ik} \|^2 \end{cases} \tag{9}$$

式中: ε 表示的是大于零的阈值,判断迭代是否收敛,收敛则迭代结束。

4.3 基于体积空间融合的匹配算法

传统 ICP 算法存在的一个问题是直接融合两帧点云会造成融合后的点云中出现冗余点。因此,本

文引入了基于体积空间融合的匹配算法。算法主要思想是在对点云进行融合时,建立包含全部帧的一个立方体栅格来表示三维空间,以一定的分辨率来划分立方体为空间的像素点。每一个点记录了到物体模型表面的距离如图 2 所示,使用正负来表示表面的被遮挡面和可见面,而过零点就代表是表面上的点。然后依次对每一帧数据进行处理,当有新的数据加入时由式(10)和式(11)进行处理,通过累加将各帧数据融合起来,最后空间体中的零值像素即为最后的曲面上的点。

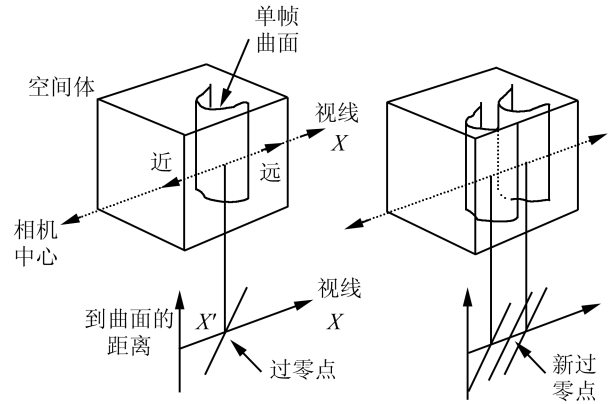


图 2 基于体积空间的点云融合
Fig. 2 Point cloud integration based on space

$$D_{i+1}(x) = \frac{W_i(x)D_i(x) + W_{i+1}(x)D_{i+1}(x)}{W_i(x) + W_{i+1}(x)} \tag{10}$$

$$W_{i+1}(x) = W_i(x) + W_{i+1}(x) \tag{11}$$

$$Err(d) = d^2 \times 0.0035 \tag{12}$$

$$\omega_i(x) = \frac{x - x^*}{Err(x^*)} \tag{13}$$

式中: i 表示栅格距离值, $Err(d)$ 为传感器误差函数, x^* 为单帧曲面上的点到相机中心距离。算法使用权重值进行融合,且具有最小二乘优化性质,因而对传感器的噪声起到了一定的抑制作用。

4.4 位姿优化

在位姿转换计算中通常会存在误差累计,为此采用了 g^20 优化方法,误差函数 $F(x)$ 定义如下:

$$F(X) = \sum_{(i,j) \in C} e(x_i, y_j, z_{ij})^T \Omega_{ij} e(x_i, y_j, z_{ij}) \tag{14}$$

$$X^* = \underset{x}{\operatorname{argmin}} F(x) \tag{15}$$

式中: x_i 表示第 i 时刻的机器人位姿,也称 g^20 为图中的一个点, z_{ij} 表示 x_j 和 x_i 两点之间的约束, $e(x_i, x_j, z_{ij})$ 为 x_i 和 x_j 组成的边所产生的向量误差函数,代表 x_i 和 x_j 满足约束条件 z_{ij} 的程度,当它的值为 0 时, x_i 和 x_j 完全满足匹配约束。

5 实验结果及分析

5.1 实验平台

本文实验平台分为数据采集端和处理端两部

分,其中数据采集端(图 3(b)所示)主要由 Pioneer3-DX 机器人、Kinect 和笔记本电脑组成,Kinect 图像分辨率为 640×480 ,最高帧频为 30 帧/s,水平视场角为 52° 。数据处理端处理器为一台 Intel 双核 2.0 GHz 主频的 PC 机,运行 Ubuntu 12.04 Linux 操作系统。为保证较高的鲁棒性,实验中机器人移动速度 v 为 0.1 m/s、转弯角速度 w 为 0.52 rad/s ($30^\circ/\text{s}$)对每个关键帧完成特征点提取、误匹配剔除和地图更新的时间约为 300 ms。



(a) 实验场景



(b) 实验装置

图 3 实验装置与实验场景

Fig. 3 Experimental apparatus and experimental scene

5.2 真实环境下的三维视觉 SLAM

5.2.1 彩色点云生成与位姿估计

以室内真实环境(图 3(a)所示)作为实验场景,在对 Kinect 进行标定的基础上,由所得内外参数以及对准校正后的像素点深度值将图像点转换为彩色点云数据,如图 4 所示。

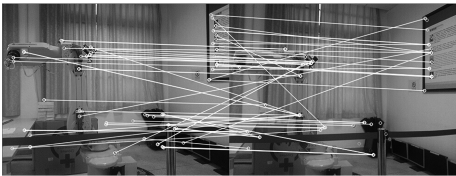


图 4 环境三维彩色点云

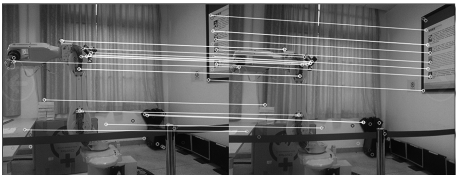
Fig. 4 Environment 3D color point cloud

在利用特征匹配算法获得相邻帧间对应关系的基础上,利用 RANSAC 算法对点云进行粗匹配,粗匹配结果如图 5 所示,由图可以看出,经过 RANSAC 算法进行粗匹配后能够将匹配过程中出现的错误匹

配剔除,提高特征匹配的正确率,实验中完成粗匹配所耗时间为 0.004 25 s。



(a) 未经 RANSAC 算法匹配



(b) 经 RANSAC 算法匹配

图 5 粗匹配实验结果

Fig. 5 The experimental results of coarse matching

5.2.2 三维环境重构效果对比

机器人以 0.1 m/s 的线速度在室内环境下创建地图,实验过程中最大迭代次数为 50,阈值设定为 5 cm,内点数量阈值 $N = 13$;台式机处理采集到的图像帧,当特征点对少于 50 组时,将舍弃该帧数据,不进行初始配准。



(a) 传统方法重构结果



(b) 本文方法重构结果

图 6 室内环境三维重构结果

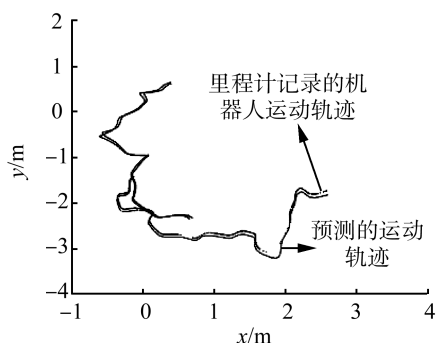
Fig. 6 3D reconstruction result of the indoor environment

图 6(a) 为采用文献[4]方法得到的三维环境重构结果,所构建的三维地图存在 8 冗余即 8 点较多、环境中物体轮廓模糊等问题。图 6(b) 为采用本文方法构建的三维环境重构结果,获得的三维地图较好

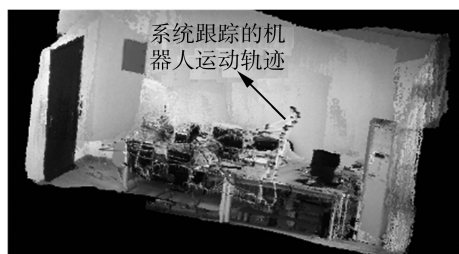
的解决了上述问题,物体轮廓清晰,冗余点减少。

5.2.3 机器人运动轨迹跟踪

为了获得相机位姿的变化,本文采用最小二乘逼近的方法对机器人相邻位姿进行估计,借助 g^20 优化方法对估计位姿进行优化。为了使运动轨迹更为直观,实验给出了机器人运动轨迹在 X/Y 平面上的投影如图 7(a) 所示,以及机器人在三维空间中的运动轨迹如图 7(b) 所示。可以看出本文方法能较好的实现机器人的运动轨迹跟踪。



(a) 运动轨迹在 X/Y 平面上的投影



(b) 轨迹跟踪结果

图 7 机器人运动轨迹跟踪

Fig. 7 Robot trajectory tracking

6 结束语

本文提出的分层式点云配准策略,较好地解决了点云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,引入了基于体积空间的多帧融合方法来减少冗余点。实验结果表明本文所提方法的有效性。下一步计划对系统进行两方面的改进:1)对创建的环境三维地图的全局优化。2)研究采用并行技术,进一步提高系统执行效率,缩短系统的运行时间,使系统具有更好的实时性。

参考文献:

- [1] RYDE J, HU H S. 3D mapping with multi-resolution occupied voxel lists[J]. *Autonomous Robots*, 2010, 28(2): 169-185.
- [2] KOPPILA H S, ANAND A. Semantic labeling of 3D point clouds for indoor scenes[C]//25th Annual Conference on Neural Information Processing Systems. New York, USA: Curran Associates Inc., 2011.

- [3] HERRERA C D, KANNALA J, HEIKKILA J. Joint depth and color camera calibration with distortion correction[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(10): 2058-2064.
- [4] 杨扬, 曹其新, 朱笑笑. 面向机器人手眼协调抓取的 3 维建模方法[J]. *机器人*, 2013, 35(2): 151-155.
- [5] YANG Yang, CAO Qixin, ZHU Xiaoxiao. A 3D modeling method for robot's hand-eye coordinated grasping[J]. *Robot*, 2013, 35(2): 151-155.
- [6] HENRY P, KRAININ M, HERBST E, et al. RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments[J]. *International Journal of Robotics Research*, 2012, 31(5): 647-663.
- [7] HARRIS C, STEPHENS M. A combined corner and edge detector[C]//Proc. *Alvey Vis. Conf.*, 1988: 147-151.
- [8] LOWE D G. Object recognition from local scale-invariant features [C]//International Conference on Computer Vision. Kerkira, Corfu, Greece, 1999: 1150-1157.
- [9] BAY H, ESS A, TUYTELAARS T, et al. Speeded-up robust features (SURF) [J]. *Comput. Vis. Image Understanding*, 2008, 110: 346-359.
- [10] KUMMERLE R, GRISETTI G, STRASDAT H, et al. G2o: A general framework for graph optimization [C]//IEEE International Conference on Robotics and Automation. Piscataway: IEEE, 2011: 3607-3613.
- [11] JAN S, MICHAL J, TOMAS P. 3D with Kinect [C]//2011 IEEE International Conference on Computer Vision Workshops. Barcelona: Institute of Electrical and Electronics Engineers Inc, 2011: 1154-1160.
- [12] TYKKALA T, AUDRAS C, COMPORT A I. Direct iterative closest point for real-time visual odometry [C]//ICCV Workshops. IEEE, Barcelona, Spain, 2011: 2050-2056.

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 160 余篇,SCI/EI 收录 80 余篇,获国家发明专利 10 余项。



陈起,男,1991年生,硕士,主要研究方向为机器人同时定位与地图创建(SLAM)、机器人三维视觉导航。



罗元,女,1972年生,教授,博士,美国 OSA、SPIE、中国仪器仪表学会、中国光学学会高级会员,主要研究方向为人机交互、基于图像视频处理的人机交互、基于图像视频处理的测试。主持和参与国家、省部基金项目 10 余项,发表学术论文 100 余篇,被 SCI/EI 收录 40 余篇,获国家发明专利 6 项。

[责任编辑:陈峰]