

文章编号: 2095-2163(2019)03-0168-04

中图分类号: TP242

文献标志码: A

基于 RGB-D 相机的移动机器人定位分析与实现

彭蔚枝, 袁锋伟, 周志伟

(南华大学 机械工程学院, 湖南 衡阳 421001)

摘要: 同时定位与地图创建在移动机器人导航系统中至关重要。本文以 RGB-D 深度相机为视觉传感器, 分析移动机器人定位方法, 利用特征提取与匹配实现位姿估计, 实现定位功能, 以 Turtlebot 为载体的实验平台, 运用成熟的 ORB-SLAM 视觉算法实现定位与地图构建, 拟完善中国在核电设施中移动机器人室内定位问题, 自主构建出厂地图, 或者搭载于已有的核装备机器人上完成自身定位与地图构建。

关键词: RGB-D 深度相机; 定位; ORB-SLAM; 视觉

Positioning analysis and implementation of mobile robot based on RGB-D camera

PENG Weizhi, YUAN Fengwei, ZHOU Zhiwei

(School of Mechanical Engineering, University of South China, Hengyang Hunan 421001, China)

[Abstract] Simultaneous localization and mapping are very important in the navigation system of mobile robot. This paper takes RGB-D depth camera as visual sensor, analyses the location method of mobile robot, extracts and matches features to realize position and pose estimation, realizes positioning function. After that, the paper takes Turtlebot as experimental platform, therefore realizes positioning and map construction based on mature ORB-SLAM visual algorithm, which intends to perfect China's nuclear power plant. For the indoor location problem of mobile robots in the facilities, the research could construct the workshop map independently, or carry it on the existing nuclear equipment robots for completing their own location and map construction.

[Key words] RGB-D depth camera; location; ORB-SLAM; vision

0 引言

随着现代科技的迅猛发展,智能导航设备应用正日趋广泛,比较常见的当属基于 GPS/北斗导航系统的移动设备。及至目前备受瞩目的自动驾驶技术,也离不开导航系统的参与配合。众所周知,定位在导航系统中占有举足轻重的地位。本文研究目的即拟在核电站等核设施内完成厂房内地图构建,为此本次研究将选用以室内为运行环境的定位方法,具有代表性的室内定位技术有基于红外传感器、蓝牙、wifi、视觉传感器、电磁轨道等方法。考虑到以视觉传感器的视觉定位建图技术在理论和实践上均取得了重要突破,又适用于未知的室内环境,且运行效果很好。基于此,本文将深度相机为传感器,利用 ORB 特征提取算法对图像进行特征提取与匹配,分析计算出移动机器人的位姿,并用 ORB-SLAM2 为定位建图算法进行分析与实验,实验项目包括:特征提取、特征匹配、位姿估计、室内定位与建图,实验结果表明本文研究最终能取得一个良好的定位与地图构建效果。对此可展开论述如下。

1 背景与选型

1.1 基本概念

同时定位与地图创建(Simultaneous Localization and Mapping, SLAM)是指包括同时定位与地图创建的系统。机器人设备搭载 SLAM 系统能实时获得自身定位,并构建生成地图^[1]。经过几十年研究发展,视觉 SLAM 的整体算法框架已经成熟。视觉 SLAM 即是指搭载了特定视觉传感器的机器人或者移动设备,在没有 GPS 信号或者缺少先知信息的环境下,在运动过程中感知自己位置并构建生成地图。

1.2 选型依据

1.2.1 传感器选型

对于研究者来说,用于视觉的传感器有单目摄像头、双目摄像头、RGB-D 深度摄像头等类别。其中,RGB-D 相机可直接获取空间点深度信息,而无需像单目摄像头一样进行初始化;深度相机的数据初始化则更简单,首帧即为关键帧,而且由于深度信息已知,就可从立体点中直接构建初始化地图。故而,本文将采用 RGB-D 深度相机作为视觉传感器。

基金项目: 湖南省高校重点实验室创新平台开放基金项目(17K079);湖南省研究生科研创新项目(2018KYY072)。

作者简介: 彭蔚枝(1992-),男,硕士研究生,主要研究方向:基于多传感器融合的室内定位导航技术研究。

收稿日期: 2019-02-02

1.2.2 视觉算法选型

经典的视觉 SLAM 系统的设计步骤可阐释分述如下。

(1) 传感器信息读取,在视觉 SLAM 中读取和预处理图像。

(2) 视觉里程计,简称 VO。视觉里程计也称为前端,其定制功能是估算相邻图像之间相机位置,以及局部地图的粗略样子。

(3) 后端优化。后端接收到不同时刻视觉里程计测量出的相机位姿,以及回环检测信息,对其进行优化,得到全局地图。

(4) 回环检测,判断机器人是否到达过先前的位置^[2]。

视觉 SLAM 可分为视觉里程计前端和运动估计的后端。其中,前端是指通过深度相机输出的图像信息大致估计出相机运动,往往深度相机能提供给后端较好的初始值。基于特征点法的前端图像处理,是视觉里程计的主流方法。特征法的前端运行较为稳定,是目前比较成熟的视觉前端的解决方案。本文采用 ORB-SLAM^[3] 算法为整体框架,ORB-SLAM 算法是一个较为完善的 SLAM 系统,现在许多优秀的算法是基于 ORB-SLAM 框架上再做另行深入开发的。ORB 中的特征计算相对简单快速,不像 SIFT 算法或者 SURF 算法那样耗时,可以不使用 GPU 即可完成快速计算;角点选取简捷方便,ORB 增加了尺度和旋转的描述,具有良好的旋转和尺度不变性。ORB 提供的 BRIEF 描述子在大范围运动时可以进行良好的回环检测和重定位。优秀的回环检测算法能有效地防止 ORB-SLAM 中误差的累积,即使丢失帧后也能根据当前帧将其及时找回。

在 PTAM 采用 2 个线程后,ORB-SLAM 采用了 3 个线程同时实现 SLAM 过程。特别地,Tracking 线程是对相机记录图像的特征提取,并与附近图片进行特征匹配,通过特征点的位置估算出相机的位姿。

2 位姿计算

2.1 特征提取

基于 ORB^[4] 算法的特征是由关键点和描述子组成的。关键点为角点,角点选取工作过程可详述如下:

- (1) 在图像中选取像素 p , 假设其亮度为 I_p 。
- (2) 随机设置一个阈值,一般设 $T = 20\%I_p$ 。
- (3) 以像素 p 为中心,选取半径为 3 的圆周上 16 个像素点。如果选取的圆周上有 N 个亮度大于

$I_p + T$ 或者小于 $I_p - T$ 的点,那么 p 可以被认为是关键点, N 可以取 12、6、11,将其分别称为 FAST-12、FAST-6、FAST-11。具体如图 1 所示。

(4) 重复以上步骤,对每一个像素执行相同的操作,即可完成角点的选取。FAST-12 在时效性上做了优化,在 FAST-12 中选取第 1、5、9、13 个像素亮点,只有这 4 个点同时满足亮度要求,才可能是角点。此类优化可以大大加速角点检测,同时也可避免角点集中现象。研究可知,半径固定为 3 的圆,将存在尺度问题。针对这一问题,ORB 在设计中添加了尺度和旋转的描述,步骤详见如下。



(a) 像素点 p (b) 放大后的点 p
(a) Pixel point p (b) Enlarged pixel point p

图 1 Fast 特征点

Fig. 1 Fast characteristic points

(1) 在图像中定义图像块的矩为:

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

通过矩可以找到图像块的质心,可将其写作如下数学形式:

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

连接图像块集合中心 O 与质心 C , 可以得到一个向量,特征点的方向可以定义为:

$$\theta = \arctan(m_{01}/m_{10}). \quad (3)$$

如此将使得 FAST 角点具有旋转不变性,ORB 的尺度不变性可以通过构建图像金字塔解决^[4]。BRIEF 描述子的确定过程可描述为: BRIEF 描述子是由 0 和 1 组成的二进制数字的组合,这里选取关键点附近的 2 个像素进行比较,当点 p 像素大于点 q 像素,记为 1,反之记为 0。

2.2 特征匹配

组合 FAST 和 BRIEF 两者,在 SLAM 中用于图像的匹配。根据匹配,通过 2 幅图中任意 2 个点的描述子比较判断是否为同一个特征,其中相似度可用二进制之间的汉明距离进行表示。研究中给出了实验室内 2 帧图像匹配结果即如图 2 所示。

2.3 位姿估计^[2]

在图像特征提取与匹配后,研究将对相机运动进

行估计。本文采用非线性优化求解方法。将粗略估计的相机位姿设为第一帧,优化匹配点的重投影误差。计算重投影误差的方法过程可解析表述如下。



图2 特征匹配

Fig. 2 Feature matching

三维空间 P 点在第一帧和第二帧中对应的像素点分别为 p_1 和 p_2 , 而在实际的投影中,在第二帧中的投影点为 p_{21} , 需要通过优化的方式使其靠近 p_2 , 尽可能减小与实际投影点之间的误差。假设 2 帧中有 n 对匹配点, 设空间中的一点坐标 $p_i = [x_i, y_i, z_i]^T$, p_i 投影点像素坐标为 $u_i = [u_i, v_i]^T$, 相机内参固定为 K , 变化矩阵李代数为 ξ , 像素点和空间点的位置有如下关系:

$$s_i \begin{bmatrix} \hat{e} u_i \\ \hat{e} v_i \\ \hat{e} 1 \end{bmatrix} = K \exp(\hat{\xi}) \begin{bmatrix} \hat{e} x_i \\ \hat{e} y_i \\ \hat{e} z_i \\ \hat{e} 1 \end{bmatrix}, \quad (4)$$

改写成矩阵形式为:

$$s_i u_i = K \exp(\hat{\xi}) p_i. \quad (5)$$

研究中包括有许多对匹配点, 因此最小二乘误差求和也称作重投影误差。具体计算公式如下:

$$\hat{\xi}^* = \operatorname{argmin} \frac{1}{2} \sum_{i=1}^n \left\| u_i - \frac{1}{s_i} K \exp(\hat{\xi}) p_i \right\|_2^2, \quad (6)$$

记变换到相机坐标系下的空间点坐标为 $P' = (\exp(\hat{\xi}) P)_{1,3} = [X', Y', Z']^T$, 相机投影模型为 $su = KP'$ 代入相机参数可以消去 s , 由此推得:

$$u = f_x \frac{X'}{Z'} + C_x, \quad v = f_y \frac{Y'}{Z'} + C_y, \quad (7)$$

对 $\hat{\xi}$ 左乘扰动变量 $\delta\xi$, 然后考虑误差 e 关于扰动量的导数。研究中可参考运算公式如下:

$$\frac{\partial e}{\partial \delta \xi} = \lim_{\delta \xi \rightarrow 0} \frac{e(\delta \xi \oplus \hat{\xi}) - e(\hat{\xi})}{\delta \xi} = \frac{\partial e}{\partial P'} \frac{\partial P'}{\partial \delta \xi}, \quad (8)$$

其中, \oplus 是李代数上的左乘扰动。左边项是误差关于投影点的导数, 即:

$$\frac{\partial e}{\partial P'} = - \begin{bmatrix} \hat{e} \frac{\partial u}{\partial X'} & \frac{\partial u}{\partial Y'} & \frac{\partial u}{\partial Z'} \\ \hat{e} \frac{\partial v}{\partial X'} & \frac{\partial v}{\partial Y'} & \frac{\partial v}{\partial Z'} \\ \hat{e} 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{e} f_x \\ \hat{e} z' \\ \hat{e} 1 \end{bmatrix} = - \begin{bmatrix} \hat{e} f_x & 0 & -f_x X' \\ \hat{e} z' & 0 & -f_x X' Y' \\ \hat{e} 0 & -f_y & -f_x Y' \end{bmatrix} \begin{bmatrix} \hat{e} u \\ \hat{e} v \\ \hat{e} 1 \end{bmatrix}, \quad (9)$$

右边是变换后的点关于李代数的导数, 即:

$$\frac{\partial (TP)}{\partial \delta \xi} = (TP) \cdot = \begin{bmatrix} \hat{e} I & -\hat{e} P' \\ \hat{e} 0^T & \hat{e} 1 \end{bmatrix}, \quad \frac{\partial P'}{\partial \delta \xi} = [I, -P'], \quad (10)$$

两式相乘可消掉 $\partial P'$, 综合计算后可得数学公式如下:

$$\frac{\partial e}{\partial \delta \xi} = \begin{bmatrix} \hat{e} f_x & 0 & -f_x X' & -f_x X' Y' & f_x + \frac{f_x X^2}{Z^2} & -\frac{f_x Y'}{Z} \\ \hat{e} z' & 0 & -f_x Y' & -f_x Y^2 & \frac{f_x X Y'}{Z^2} & -\frac{f_x Y'}{Z} \\ \hat{e} 0 & \frac{f_y}{Z} & -f_y & -\frac{f_x Y^2}{Z^2} & \frac{f_y X Y'}{Z^2} & \frac{f_y X'}{Z} \end{bmatrix} \begin{bmatrix} \hat{e} u \\ \hat{e} v \\ \hat{e} 1 \end{bmatrix}, \quad (11)$$

这是描述重投影误差关于相机的李代数的一阶变化关系。经计算如下公式:

$$\frac{\partial e}{\partial P'} = - \begin{bmatrix} \hat{e} f_x & 0 & -f_x X' \\ \hat{e} z' & 0 & -f_x Y' \\ \hat{e} 0 & \frac{f_y}{Z} & -f_y \end{bmatrix} \begin{bmatrix} \hat{e} u \\ \hat{e} v \\ \hat{e} 1 \end{bmatrix} R. \quad (12)$$

由此即推导得到了重投影误差关于位姿的导数。根据 2 幅彩色图和 2 幅深度图像, 研究在计算机上求解出来变化矩阵, 至此便完成了位姿估计。最终运算结果如图 3 所示。

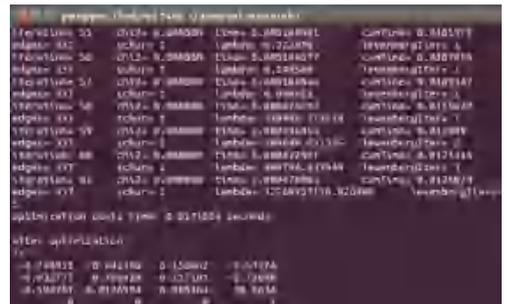


图3 变换矩阵求解值

Fig. 3 Solution value of transform matrix

3 实验

目前, 由于防辐射镜头昂贵, 防辐射电子设备处理仍难于在实验室环境下进行操作。本文将在实验室和实验室楼道里模拟核电站室内环境。本实验的硬件配置为: 课题组购买的实验小车平台 turtlebot2 + ubuntu14.04 操作系统 + ROS indigo + 联想笔记本 T450+XBOX 手柄。运行效果如图 4 所示。

(下转第 175 页)