

SLAM 过程中的机器人位姿估计优化算法研究^①

禹鑫燚^② 朱熠琛 詹益安 欧林林^③

(浙江工业大学信息工程学院 杭州 310000)

摘 要 为了在图像位姿估计过程中充分利用更多空间点信息及减少噪声的影响,本文提出了一种基于 EPnP 算法的相机位姿估计优化方法。该方法先基于 EPnP 算法估计相机位姿,再把得到的相机位姿和空间点位置都作为优化变量,与相机的观测数据一起构建关于相机位姿的最小二乘优化问题,进行 BA 优化,来提高位姿的精确度。同时在 SLAM 后端对不同时刻视觉里程计的位姿和回环检测的信息进行优化,再把优化后的相机位姿作为观测数据和激光雷达信息进行融合,从而进一步提高 SLAM 过程的可靠性。实验结果证明了本文所提出的算法在机器人室内自定位和建图方面的准确性和可靠性。

关键词 位姿估计, EPnP, 最小二乘, SLAM 后端, 回环检测

0 引 言

近年来,基于多传感器融合的 SLAM(simulation localization and mapping)是移动机器人领域的重点研究方向。为了在未知环境中进行导航,移动机器人需要构建环境地图并且同时定位自身在地图中的位置,这就需要传感器不断地去获取周围信息并对这些信息进行处理。然而,在整个 SLAM 过程中受限于传感器本身的噪声影响,会不断地累积误差导致地图误匹配以及自定位不精确等问题。因此,多传感器融合的 SLAM 研究具有非常重要的意义。

在 SLAM 过程中,基于相机的位姿估计问题是假定已知世界坐标系中 n 个控制点及其在图像坐标系中对应的二维投影后,求解相机坐标系相对于世界坐标系的位姿问题。该问题也是模式识别、计算机视觉、机器人自主避障和导航等领域中的关键问题之一^[1,2]。由于空间 3D 点和二维图像点间的映射关系已知,因此位姿估计问题也经常被称为 n 点透视(perspective- n -point, PnP)问题。自 1981 年第

一次被提出以来^[3], PnP 问题受到了广泛的关注。解析算法和迭代算法是解决 PnP 问题最常用的两种算法。当 $n > 5$ 时, PnP 问题被认为是一个直接线性变换问题(direct linear transformation, DLT),其能够用此线性求解^[4],因大部分针对解析算法的研究都被聚焦到 P3P、P4P 和 P5P 问题的解决上^[5]。在文献[6,7]中提出了有 n 个任意控制点时 PnP 问题的计算方法,但其计算复杂度大且效率低。此外,解析算法只有在控制点像素坐标能被精确计算的前提下,才能得到可靠的位姿估计结果,而当控制点像素坐标在特征提取过程中受到噪声干扰时,位姿估计结果将变得不准确。

和解析算法相比,迭代算法不仅在计算过程上更加简便,而且对于图像噪声的鲁棒性也更好。典型的迭代算法将位姿估计问题构建为非线性最小二乘问题,并设定约束条件为旋转矩阵正交^[8,9],从而可以用非线性化算法解决,如 Levenberg-Marquadt 算法或 Gaussion-Newton 算法^[10,11]等。但是这些算法完全从数学的角度出发,没有充分利用成像模型的结构特点,影响了算法的性能。

① 国家重点研发计划“智能机器人”重点专项(2017YFB1303900),企业合作项目(KYY-HX-20180162)和宁波重点(2014B10017)资助项目。

② 男,1979 年生,博士,讲师;研究方向:工业机器人轨迹规划,工业机器人运动控制等;E-mail: yuxinyinet@163.com

③ 通讯作者,E-mail: linlinou@zjut.edu.cn

(收稿日期:2018-02-09)

考虑到相机的深度图存在噪声,而且有可能存在数据丢失的情况,使用 ICP 算法^[12]会导致位姿匹配时候的误差越来越大,并且如果特征点丢失的太多,可能引起由于特征点太少,无法进行运动估计的情况。近年来,EPnP(efficient perspective-n-point)算法受到了学者们的广泛关注^[13],由于其计算复杂度为 $O(n)$,该算法被认为是最高效的解析算法。然而,作为一种解析算法^[14],EPnP 算法对于图像噪声的鲁棒性不强。文献[15]中提出了一种改进的 EPnP 算法,将 EPnP 算法的估计值和空间点位置都作为优化变量,但没有将前一个关键帧的相机位姿和观测数据考虑进去。

本文研究了 SLAM 过程中的位姿估计问题,针对目前视觉里程计中相机无法充分利用之前时刻获得的图像信息从而导致视觉里程计误差累积的问题,提出一种新的优化算法,提高了机器人在移动过程中位姿估计的准确性与可靠性。本文首先将 EPnP 算法的位姿估计值作为初始值,把估计的相机位姿和空间点位置作为优化变量,与相机的观测数据一起构建关于相机位姿的最小二乘优化问题,然后进行 BA(bundle adjustment)优化^[16],以增强算法的鲁棒性。同时在 SLAM 后端对不同时刻视觉里程计的位姿和回环检测的信息进行优化,再把优化后的相机位姿作为观测数据和激光雷达信息进行融合,从而进一步提高了移动机器人在三维建图中的可靠性及准确性,易于工程实现。本文的研究安排如下:首先在第 1 节阐述机器人位姿优化过程,第 2 节阐述了多传感器融合 SLAM 过程,第 3 节对实验结果进行分析,最后在第 4 节对本文进行总结。

1 机器人位姿优化过程

本文先基于 EPnP 算法估计相机位姿,把估计的相机位姿和空间点位置都作为优化变量,与相机的观测数据一起构建最小二乘优化问题,然后进行 BA 优化,从而获得更为精确的机器人位姿。

1.1 BA 优化原理

假设有一个 3D 空间中的点,被位于不同位置的多个摄像机看到,光束平差法 BA 就是能够从这

些多视角信息中提取出 3D 点的坐标以及各个摄像机的相对位置和光学信息的过程。而 BA 问题,是一个最小化重投影误差的问题。考虑 n 个三维空间点 P 及其投影 p ,为了计算相机的位姿 R, t , 它的李代数表示为 ξ , 假设某空间点坐标为 $P_i = [X_i, Y_i, Z_i]^T$, 其投影的像素坐标为 $U_i = [u_i, v_i]^T$ 。像素位置与空间点位置的关系如下:

$$s_i U_i = K \exp(\xi^\wedge) P_i \quad (1)$$

其中, s_i 表示深度图的尺度因子, K 代表相机的内参矩阵。由于相机位姿未知及观测点的噪声影响,把误差进行求和,构建最小二乘问题,然后寻找最优的相机位姿,使它最小化:

$$\xi^* = \arg \min_{\xi} \frac{1}{2} \sum_{i=1}^n \left\| U_i - \frac{1}{s_i} K \exp(\xi^\wedge) P_i \right\| \quad (2)$$

上述问题的误差项,是将像素坐标与 3D 点按照当前估计的位姿进行投影得到的位置相比较得到的误差,称为重投影误差。使用齐次坐标时,误差有 3 维。由于 U 最后一维为 1,该维度的误差一直为零,因此更多时候使用非齐次坐标,误差就变成 2 维。如图 1 所示,通过特征匹配得知 p_1 和 p_2 是同一个空间点 P 的投影,在初始值中 P 的投影 $\hat{p}_2$ 与实际 p_2 之间有一定的距离,因而需要不断地调整相机的位姿,使 $\hat{p}_2$ 和 p_2 的距离变短。由于调整需要考虑

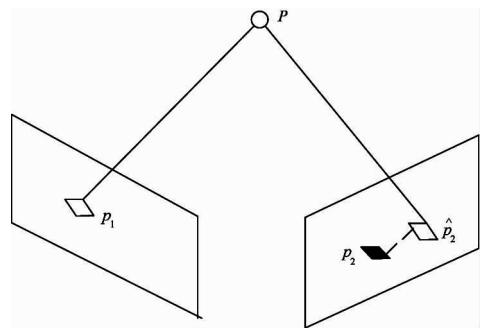


图 1 重投影误差示意图

很多点,并且每个点的误差通常都不会精确为零,最后使用李代数,可以构建无约束的优化问题。在求解之前,需要知道每个误差项关于优化变量的导数,即线性化:

$$e(x + \Delta x) \approx e(x) + J \Delta x \quad (3)$$

其中 x 为相机位姿, e 为像素坐标误差, J 为一个 2×6 的矩阵。接着使用扰动模型来求解观测方程关于相机位姿与特征点的两个导数矩阵 $\frac{\partial e}{\partial \delta \xi}$ 和 $\frac{\partial e}{\partial P}$, 最后通过 Guassion-Newton、Levenberg-Marquadt 方法等优化算法对误差项进行求解。

1.2 相机位姿优化过程

首先,为了解决图像尺度不变性,将摄像头得到的图片通过梯次向下采样构建图像高斯金字塔,总共分为 N 层,接着对每层进行 FAST 角点检测,并对每个特征点按照如下公式计算 Harris^[17] 响应值,保留每个节点中响应值 R 最大的那个特征点:

$$R = \det(\mathbf{M}) - k \times (\text{trace}(\mathbf{M})^2) \tag{4}$$

$$\mathbf{M} = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \tag{5}$$

其中 R 为 Harris 响应值, $\mathbf{M}$ 为一个 2×2 矩阵,可由图像的倒数求得, $k = 0.04 \sim 0.06$, $w(x,y)$ 为图像窗口函数, I_x 为特征点在水平方向上的变化量, I_y 为特征点在竖直方向上的变化量。为了使得到的特征点保持旋转不变性,使用灰度质心法 (intensity centroid) 来计算特征点方向。首先在一个小的图像块 B 中,定义图像块的矩为

$$m_{pg} = \sum_{x,y \in B} x^p y^q I(x,y) \tag{6}$$

其中 $p,q = \{0,1\}$, $I(x,y)$ 为对应像素的灰度值, x,y 为对应像素坐标;然后通过矩可以找到图像块的质心:

$$c = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \tag{7}$$

其中 0 阶矩 (m_{00}) 为图像块的质量, 1 阶矩 (m_{10} , m_{01}) 表示图像块的质心;接着连接图像块的几何中心 O 和质心 C , 得到一个方向向量 $\overrightarrow{OC}$, 于是特征点的方向可定义为

$$\theta = \arctan\left(\frac{m_{01}}{m_{10}}\right) \tag{8}$$

通过以上方法,FAST 角点便具有了尺度与旋转的描述,从而进一步提升了其不同图像之间表述的健壮性。

然后使用汉明距离 (Hamming distance) 对两幅图像进行描述子匹配,去除其中的误匹配信息。接

着根据匹配的点采用 EPnP 算法来估计相机的运动。EPnP 算法通常选取世界坐标系下的 4 个控制点坐标,通过 n 个 3D 点在相机平面的投影关系,以及与这 4 个控制点的权重关系,构建一个 12×12 方阵,求得其零空间特征向量,可以得到虚拟控制点的相机平面坐标,再使用 POSIT 算法^[18] 求解相机位姿。实验将 EPnP 算法求出的解作为初始值,并运用 BA 算法来进行优化,把问题建模成一个最小二乘的图优化问题。在图 2 的图优化中,节点为第 1 张和第 2 张图像的位姿节点 $\xi_1, \xi_2 \in se(3)$, 所有特征点的空间位置 $P_j \in R^3$, 边为每个 3D 点在第 1 和第 2 个相机中的投影,分别为 z_1, z_2 , 以观测方程来描述:

$$z_1 + z_2 = h(\xi_1, P_j) + h(\xi_2, P_j) \tag{9}$$

为了将其它时刻的观测量也考虑进来,在整个 SLAM 的后端实验将多幅图像匹配到的位姿和空间点进行迭代优化,设 $z_{i,j}$ 为在位姿 ξ_i 处观察路标 P_j 产生的数据,那么整体的代价函数为

$$\frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n \|e_{ij}\|^2 = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n \|z_{i,j} - h(\xi_i, P_j)\|^2 \tag{10}$$

对式 (10) 进行求解,相当于对位姿和路标同时做了调整。

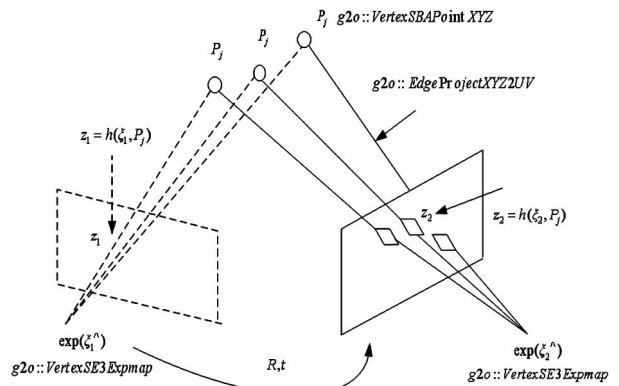


图 2 EPnP 的图优化示意图

实验采用高斯牛顿法来求解代价函数,相应的增量线性方程为

$$\mathbf{H} \Delta x = g \tag{11}$$

其中 $\mathbf{H}$ 矩阵是一个稀疏结构,可以显式地用图优化来表示^[19,20]。于是对应的线性方程组可以变为如

下形式:

$$\begin{bmatrix} \mathbf{B} & \mathbf{E} \\ \mathbf{E}^T & \mathbf{C} \end{bmatrix} \begin{bmatrix} \Delta x_c \\ \Delta x_p \end{bmatrix} = \begin{bmatrix} v \\ w \end{bmatrix} \quad (12)$$

其中 $\mathbf{B}$ 是对角块矩阵,每个对角块的维度和相机参数的维度相同,三维空间中每个路标点为三维,于是 $\mathbf{C}$ 矩阵为对角块矩阵,每个块为 3×3 矩阵。然后对式(12)进行高斯消元,得到关于位姿部分的增量方程:

$$[\mathbf{B} - \mathbf{E}\mathbf{C}^{-1}\mathbf{E}^T]\Delta x_c = v - \mathbf{E}\mathbf{C}^{-1}w \quad (13)$$

最后对上述方程进行 Schur 消元^[21],求出对应的 Δx_c , Δx_p 。

2 多传感器融合 SLAM 过程

由于机器人自身的里程计容易造成误差累积导致机器人位姿不精确,而应用回环检测的视觉里程计可以大幅度地减小累积误差^[22]。因此,实验中将视觉里程计测量的相机位姿及回环检测的信息进行优化,得到更为精确的机器人位姿,把得到的机器人位姿作为观测数据不断地去修正激光雷达的建图过程,来提高 SLAM 过程的准确性与稳定性。

在机器人定位中,机器人位置的后验概率密度可表述为

$$B(x_t) = p(x_t | z_t, u_{t-1}, z_{t-1}, u_{t-2}, \dots, u_0, z_0) \quad (14)$$

其中, x_t 表示机器人在 t 时刻的位置, u_{t-1} 表示相机在时间段 $[t-1, t]$ 内对机器人位移的计算值, z_t 表示激光雷达在时刻 t 的观测数据。然后根据 Bayes 准则进行迭代更新:

$$B(x_t) = \propto p(z_t | x_t) Q \quad (15)$$

$$Q = \int p(x_t | x_{t-1}, u_{t-1}) B(x_{t-1}) dx_{t-1} \quad (16)$$

粒子滤波算法用赋权的粒子的集合 $S_t^N = \{\langle x_t^i, w_t^i \rangle | i = 0, 1, 2, \dots, N\}$ (其中 x 为位置估计, w 为权重, $\langle x, w \rangle$ 记为粒子)来表示 $B(x_t)$, 即对机器人位置的后验估计;并通过序贯权重采样法实现 Bayes 更新。机器人运动过程中位姿估计具体步骤如下:

步骤 1:对图像进行 ORB 特征提取与匹配,并采用 EPnP 算法估计相机的位姿;

步骤 2:根据式(10)构建 BA 优化的代价函数,再运用高斯牛顿法求解增量线性方程,并进行高斯消元;

步骤 3:对式(13)进行 Schur 消元,求解相机的位姿;

步骤 4:将视觉里程计测量的相机位姿及回环检测的信息进行优化,得到更精确的机器人位姿;

步骤 5:权重抽取,从 S_{t-1}^N 中按照权重对粒子进行抽取,所得样点集合 X' 符合 $B(x_{t-1})$, 即对前一时刻位置估计的后验概率分布;

步骤 6:运动更新。对 X' 中的每个样点按照运动模型 $p(x' | x, u)$, 根据相机的测量数据 u_{t-1} 进行更新,从而得到新的样点集合 X ;

步骤 7:粒子赋权。根据 Bayes 准则,根据传感器的测量数据 z_t 对 X 中的每个样点 x_t^i 赋权值 $w_t^i = p(z_t | x_t^i)$, 从而得到对当前时刻位置估计的粒子表示集合 $S_t^N = \{\langle x_t^i, w_t^i \rangle\}$ 。

迭代计算上述的步骤,可实现对机器人位置的动态估计。

3 实验结果

3.1 位姿估计优化算法验证

图 3 为实验中使用的移动机器人,底座是 Irobot create2,使用的传感器为华硕的 RGB-D 摄像头以及思岚的 RPLIDAR-A2 激光雷达,使用的处理器是联想笔记本电脑。图 4 为搭建的算法验证模拟环境。



图 3 移动机器人



图 4 算法模拟平台

实验中让机器人先按照 x 正轴行走 60cm,再沿 y 正轴行走 50cm,确保机器人原点与此刻拍摄到的图片有足够的匹配点,然后采用 EPnP 算法初步计算机器人位姿,接着将得到的相机位姿和空间点位置都作为优化变量,与相机的观测数据一起进行 BA 优化,最后得到位姿变换矩阵,对比文献[15]中提出的方法的输出结果,如表 1 所示。由于在二维平面上运行, z 轴的位移为 0,实验只考虑 x 轴和 y 轴的位移量。根据表 1 可知,本文所提出的算法比文献[15]中提出的算法在输出的移动机器人位姿精度上有了提高,但是本文只验证了相邻图像之间的算法对比,在之后机器人的大规模建图定位过程中,可以将更多幅图像的数据进行迭代化,甚至把整个 SLAM 过程都放进去,来减少机器人输出位姿的累积误差。接着让机器人分别使用上述两种算法从起始点移动到指点的目标点,由图 5 和图 6 对比可知,机器人转移一定角度时采用本文所提出的算法能收敛得更快,更加平稳精确地移动到目标位置。

表 1 两种算法对比结果

实验次数		本文算法 (cm)	文献[15]算法 (cm)
t_1	x	60.4753	60.8874
	y	49.7263	49.2576
t_2	x	60.5681	61.0265
	y	49.5289	49.0964
t_3	x	60.6518	60.8155
	y	49.8703	50.4069

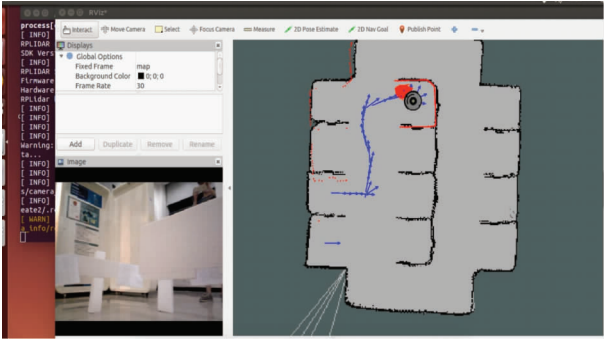


图 5 文献[15]中算法实验结果

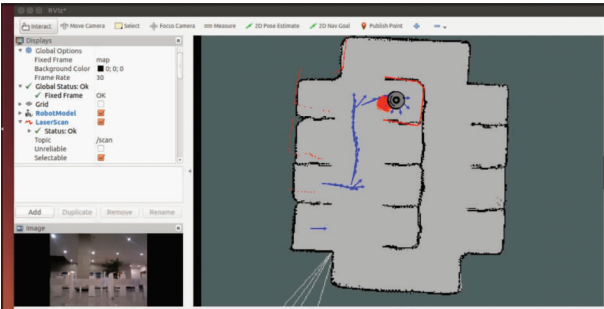


图 6 本文提出的算法实验结果

3.2 多传感器融合建图

首先在模拟环境中让机器人以较慢的速度在室内建立三维地图,并分别对比了有视觉和无视觉观测数据的激光雷达建图效果,由图 7 和图 8 对比可知,在慢速情况下有视觉观测数据的激光雷达建图后地图的匹配度更高,跟模拟环境基本吻合。

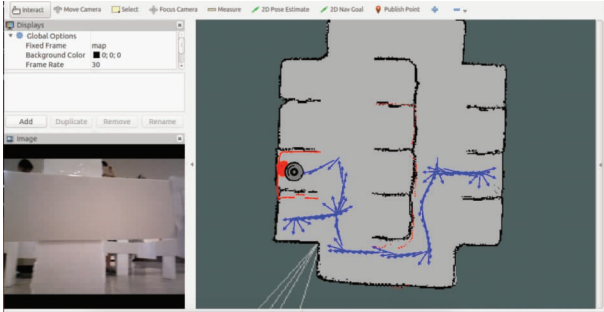


图 7 低速下激光雷达建图

考虑到机器人在低速下建图效果对比不明显,实验中让机器人以相对较快的速度在室内环境建立三维地图,由图 9 和图 10 对比可知,仅依靠激光雷达去建立三维地图,会不断地累积误差,导致地图误匹配,建图效果很差,而加入视觉观测数据后,通过不断地去修正机器人的位姿,减少了机器人的累计

误差和地图的误匹配,优化了输出结果。

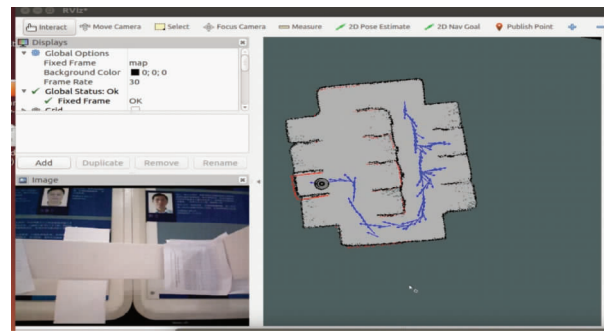


图 8 低速下视觉和激光雷达融合建图

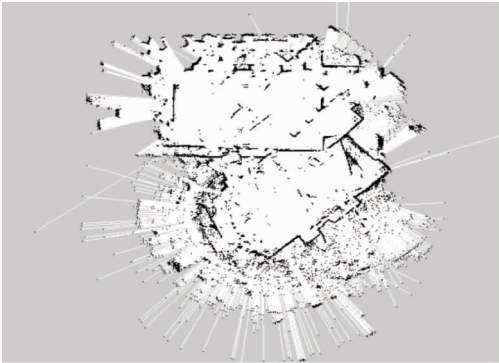


图 9 高速下激光雷达建图

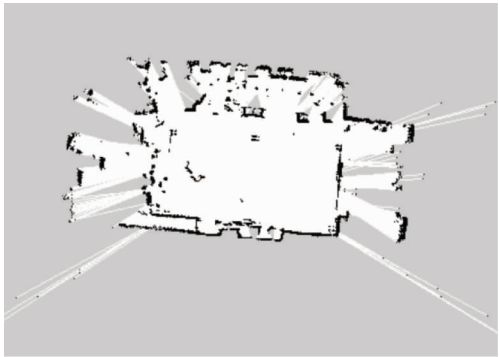


图 10 高速下视觉和激光雷达融合建图

为了验证实验中移动机器人在多传感器融合建图过程中位姿的精确度,加入了 Vicon 三维运动捕捉系统来给整个地图环境建立一个全局定位信息,如图 11 所示。Vicon 的采样频率很快,每秒可达到 300 帧,输出的位姿精确到毫米级别。在机器人建图过程中多次通过 Vicon 采样并对比移动机器人自身的输出位姿。根据表 2 对比数据可知,机器人自身的输出位姿与 Vicon 输出的位姿在误差精度上相差无几,验证了本文算法的精确性和可靠性。



图 11 Vicon 三维运动捕捉系统

表 2 输出位姿对比

实验次数		机器人位姿	Vicon 输出位姿
t_1	x	60.4753	60.1357
	y	49.7263	49.8473
t_2	x	80.5681	80.2457
	y	79.5289	80.1457
t_3	x	70.6518	70.2457
	y	49.6703	50.0469

4 结 论

本文提出了一种基于 EPnP 的相机位姿估计优化算法,同时采用了融合视觉与激光雷达信息的方法来建立室内三维地图。实验过程中采用本文提出的相机位姿优化算法,并分别对比单激光雷达建图和融合视觉与激光雷达建图的实验效果,表明了本文提出的算法有效地减少了机器人移动过程中的误差累积,同时提高了建图过程中地图的匹配度。

本文提出的算法能够作为计算机视觉领域特征点匹配优化以及移动机器人三维建图过程中的有效的方法,对于后期机器人建图和自定位的研究具有重要的意义。

参考文献

[1] 洪洋, 孙秀霞, 王栋, 等. 基于矩形几何特性的小型无人机快速位姿估计方法[J]. 中国激光, 2016, 43 (5): 226-238

[2] 张超, 韩成, 杨华民, 等. 基于位姿约束的大视场双目视觉标定算法[J]. 光学学报, 2016 (1): 197-206

- [3] Fischer J, Eichler M, Bartz D, et al. Model-based Hybrid Tracking for Medical Augmented Reality [C]// EGVE. 2006: 71-80
- [4] 吴福朝, 胡占义. 关于 P5P 问题的研究[J]. 软件学报, 2001, 12(5): 768-775
- [5] 吴福朝, 胡占义. PnP 问题的线性求解算法[J]. 软件学报, 2003, 14(3): 682-688
- [6] Quan L, Lan Z. Linear n-point camera pose determination [J]. *IEEE Transactions on pattern analysis and machine intelligence*, 1999, 21(8): 774-780
- [7] Ansar A, Daniilidis K. Linear pose estimation from points or lines[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(5): 578-589
- [8] 朱仁璋, 林彦, 张磊. 航天器交会计算机视觉系统测距求解新算法[J]. 北京航空航天大学学报, 2006, 32(7): 764-768
- [9] 李鑫, 张跃强, 刘进博, 等. 基于直线段对应的相机位姿估计直接最小二乘法[J]. 光学学报, 2015, 35(6): 203-213
- [10] Liu M L, Wong K H. Pose estimation using four corresponding points[J]. *Pattern Recognition Letters*, 1999, 20(1): 69-74
- [11] 戴东凯, 王省书, 胡春生, 等. 基于天文观测的相机标定及姿态测量技术研究[J]. 光学学报, 2012, 32(3): 100-104
- [12] Besl P J, McKay N D. A method for registration of 3-D shapes[J]. *IEEE Transactions on pattern analysis and machine intelligence*, 1992, 14(2): 239-256
- [13] Lepetit V, Moreno-Noguer F, Fua P. Epnp: An accurate o (n) solution to the pnp problem[J]. *International journal of computer vision*, 2009, 81(2): 155-166
- [14] Hmam H, Kim J. Optimal non-iterative pose estimation via convex relaxation[J]. *Image and Vision Computing*, 2010, 28(11): 1515-1523
- [15] 高翔等. 视觉 SLAM 十四讲: 从理论到实践[M]. 北京: 电子工业出版社, 2017. 157-170
- [16] Triggs B, McLauchlan P F, Hartley R I, et al. Bundle adjustment—a modern synthesis[C]. In: *Proceedings of the International Workshop on Vision Algorithms*, Corfu, Greece, 1999. 298-372
- [17] Harris C, Stephens M. A combined corner and edge detector[C]. In: *Proceedings of the 4th Alvey Vision Conference*, Manchester, UK, 1988. 147-151
- [18] DeMenthon D F, Davis L S. Model-based object pose in 25 lines of code[C]. In: *Proceedings of the European Conference on Computer Vision*, Santa Margherita Ligure, Italy, 1992. 335-343
- [19] Kümmerle R, Grisetti G, Strasdat H, et al. g 2 o: A general framework for graph optimization[C]. In: *Proceedings of the 2011 IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China, 2011. 3607-3613
- [20] Polok L, Ila V, Solony M, et al. Incremental block cholecky factorization for nonlinear least squares in robotics [C]. In: *Proceedings of the Robotics: Science and Systems*, 2013. doi: 10.15607/RSS.2013.IX.042
- [21] Sibley G, Matthies L, Sukhatme G. Sliding window filter with application to planetary landing[J]. *Journal of Field Robotics*, 2010, 27(5): 587-608
- [22] 王超磊, 王田苗, 梁建宏, 等. 小型无人直升机单目视觉 FastSLAM 研究[J]. 高技术通讯, 2013, 23(10): 1061-1067

Research on pose estimation optimization algorithm in mobile robot SLAM process

Yu Xinyi, Zhu Yichen, Zhan Yian, Ou Linlin

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310000)

Abstract

A camera pose estimation optimization algorithm based on EPnP (efficient perspective-n-point) is proposed to make full use of spatial points and reduce the effect of noise on camera pose estimation process. Firstly, the camera pose is estimated based on EPnP. And then the least squares optimization problem for the camera pose is constructed by observing data, the camera pose and positions of spatial points which are seen as optimization variables, which can improve the accuracy of camera pose. Meanwhile camera poses and the data of loop-closure detection at different moments are optimized at SLAM back-end, and the camera pose is fused with laser radar data to improve the effectiveness in SLAM process. Experimental results show the effectiveness and accuracy of the proposed method on 3D localization and mapping.

Key words: pose estimation, EPnP(efficient perspective-n-point), least squares, SLAM back-end, loop-closure detection