DOI: 10.3901/JME.2019.11.036

基于行人行为学习的机器人同时定位 与可通行区域制图*

邢志伟 朱晓蕊 何超

(哈尔滨工业大学(深圳)机电工程与自动化学院 深圳 518055)

摘要: 在人机共融环境中, 移动机器人自主移动必须解决可通行区域感知与自身定位的问题。因此提出了一种基于行人行为学习的机器人同时定位与可通行区域制图算法。该算法首先基于行人行为学习的方式对非结构化道路环境中的可通行区域进行采样, 并使用区域生长算法得到连通的可通行区域。在定位与制图过程中, 将视觉 SLAM 技术与检测到的行人信息和可通行区域信息结合起来, 在匹配过程中剔除动态行人身上的特征点, 并通过恒速率模型+DLT 算法和贪心搜索+DLT 算法结合的方式提升匹配的速度和鲁棒性, 在 Bundle Adjustment 优化过程中针对不同特征点给予不同的权重以进一步提升精度, 并在构建的环境点云地图中增量式添加所检测到的可通行区域信息。试验表明在复杂的非结构道路环境中, 该算法可以在移动机器人平台上同时实现准确的自定位和包含可通行区域的环境地图。

关键词: 行人行为学习; 非结构道路; 可通行区域; 同时定位与制图

中图分类号: TP242

Simultaneous Localization and Traversable Region Mapping Based on Pedestrian Behavior Learning

XING Zhiwei ZHU Xiaorui HE Chao

(School of Mechanical Engineering and Automation, Harbin Institute of Technology (Shenzhen),
Shenzhen 518055)

Abstract: In human-robot coexist environment, autonomous mobile robots must solve the problem of traversable region perception and self-positioning. Hence proposes an algorithm of simultaneous localization and traversable region mapping based on pedestrian behavior learning. The algorithm samples the traversable region in the unstructured road environment based on pedestrian behavior learning, and uses region growing algorithm to obtain connected traversable region. On simultaneously localization and mapping, visual SLAM technology is combined with the detected pedestrian information and traversable region information. The feature points on dynamic pedestrians are excluded in the matching process, and the constant speed model + DLT algorithm and greedy search + DLT are used to speed up the matching process and enhance the robustness. In the Bundle Adjustment optimization process, different weights are given to different feature points to further improve the accuracy. The detected traversable region information is then incrementally added into the point cloud map. Experiments show that in a complex unstructured road environment, the algorithm can achieve accurately self localization and obtain traversable region mapping on the mobile robot platform.

Key words: pedestrian behavior learning; unstructured road; traversable region; simultaneous localization and mapping

0 前言

近年来, 许多研究致力于使移动机器人完成真正自主的移动。环境中的可通行区域检测、机器人

的自主定位及环境地图的构建, 是完成自主行走任务的两个重要问题。在可通行区域检测方面, 已经有许多基于视觉的方法在结构化道路检测中取得了良好的效果, 而在道路颜色、形状不规则, 没有明显边界与特征的非结构化道路检测中, 由于环境的复杂性, 还很难在运行效率与准确率之间取得令人满意的折中。在机器人自主定位与建图 (Simultaneous localization and mapping, SLAM) 的问题上, 基于视

* 国家自然科学基金资助项目(91648102)。20180402 收到初稿, 20181130 收到修改稿

觉的技术在定位精度与运行效率方面取得了越来越多的进展,但所构建的往往是点云地图,只能用于定位而难以应用于导航规划,如何在所构建的地图中加入对环境的语义理解进而用于指导机器人运动也是当前研究的热点方向之一。

对非结构化道路的检测,许多研究是从图片中提取道路颜色或纹理特征构建外观模型,并用来对道路进行分割。RAMSTROM等^[1]于2005年提出在UV颜色空间中构建高斯混合模型描述道路和背景。2006年,TAN等^[2]在RGB颜色空间中,使用单色直方图描述环境背景,多色直方图表征道路表面。这种方法需要离线训练颜色模型,因此只能局限于应用在外观与训练数据相似的道路区域。另外,有的方法直接从输入图像中进行采样以构建模型。ÁLVAREZ等^[3]和LU等^[4]分别在2011年和2015年使用图像底部部分区域建立不同的模型用于描述道路,这种采样方法对采样的质量依赖很大,如果采样区域不是道路会出现错误的检测结果。除了道路分割法以外,还有方法采用道路边界检测,通过选择指向消失点的边缘或者道路先验形状来寻找道路边界^[5-6],但这些方法也需要有环境中的道路具有较好的消失点或者边缘,同时在进行消失点估计时的运算复杂度比较高。2012年,CHANG等^[7]综合了上述两种方法,首先通过找到两条最重要的指向消失点的边缘来定位道路边界,然后再用从图像底部和中间部分相近区域估计得到的颜色模型来分割道路。ÁLVAREZ等^[8]在2014年也采用了类似的综合方法来检测道路区域。然而上述的所有方法,运算复杂度都比较高,难以在移动机器人平台上实时运行。2017年,DENG等^[9]提出了一种快速算法,以消失点和道路边界为限制区域进行自适应采样,并结合RGB空间,照度不变空间(Illumination invariant space, IIS)空间和局部二值模式(Local binary pattern, LBP)构建多变量高斯混合模型,可以在常规CPU平台上达到实时运行,但此方法仍然对道路结构有一定特征要求。另外,基于深度学习进行可通行区域分割的方法也逐渐成为一个热点研究方向。ÁLVAREZ^[11]2012年使用一个卷积神经网络进行道路分割。2016年,XIAO等^[10]使用基于随机森林的道路检测算法,对非结构道路检测的平均运行时间约70 ms。OLIVEIRA等^[12]也于2016年提出了一种高效的深度网络模型用于道路分割,然而,在NVIDIA Titan X GPU上面的运行时间仍然超过80 ms。对于基于深度学习的算法来讲,目前基本都需要GPU的大量运算资源,无法在移动机器人平台上实际部署应用。因此,本文提出了一种基于行人学

习的环境可通行区域检测方法,可以对多种不同形态的非结构化复杂道路进行准确检测,并在移动机器人平台上实时应用。

在机器人自主定位与建图方面,视觉SLAM技术是近年来研究的热点方向。2015年,MUR-ARTAL等^[13]提出的实时ORB-SLAM框架,结合了PTAM^[14]算法的思想,使用ORB特征完成跟踪、制图、重定位和大尺度环路检测任务,平移漂移误差约为全程的1.15%(KITTI数据集,数百米行程)。此外,为了提高SLAM的准确性和鲁棒性,有研究者在视觉SLAM中融入惯性测量单元IMU的传感器信息,比如MOURIKIS等^[15]在2007年提出一种多状态约束卡尔曼滤波(Multi-state constraint Kalman filter, MSCKF)实时融合单目视觉和IMU信息,在开环条件下行驶3.2 km,误差约7 m。2015年,STEFAN等^[16]采用非线性优化并基于关键帧的方式,通过构建一个结合了视觉路标重映射误差和惯性测量的代价函数,用于实时视觉-惯性里程计,其精度相比较于MSCKF的方法有所提高。但这些基于特征点的方法只能构建环境的稀疏点云地图,无法反映环境的结构特征。近两年国内外的研究者还提出了基于像素的光照不变假设,建立稠密或半稠密环境地图的视觉SLAM方法,相比于上述基于特征的视觉SLAM方法,能够将图像中的所有像素点或者具有较大梯度变化区域的像素点构建到环境地图当中,提供更多机器人周围空间的信息。比较有代表性的是2014年ENGEL等^[17]提出的LSD-SLAM,但其不足之处是定位和重建精度劣于ORB-SLAM。另外,上述的这些方法都假设所处的环境是静态的,并未充分考虑人机共融环境下所存在的动态行人因素对定位与建图的影响,并且所构建的环境地图都只有特征点或像素点的空间几何信息,不具有其他语义信息,不能直接应用于规划移动机器人在环境中的导航运动。因此,本文提出将前述行人检测与可通行区域感知结果与视觉SLAM相融合的算法,提升视觉SLAM算法定位精度的同时,也将可通行区域信息融合到构建的环境点云地图中去。

综上所述,本文的主要贡献在于,提出了一种新的基于行人行为学习的机器人同时定位与可通行区域制图算法。该算法首先基于行人行为学习的方式对非结构化道路环境中的可通行区域进行采样,并使用区域生长算法得到连通的可通行区域。之后将检测到的行人信息与可通行区域信息与视觉SLAM技术结合起来,在匹配过程中剔除动态行人身上的特征点以提升定位精度,并通过恒速率模型+DLT算法和贪心搜索+DLT算法结合的方式提

升匹配的速度和鲁棒性,在捆集调整(Bundle adjustment, BA)优化过程中针对不同特征点给予不同的权重以进一步提升精度,并在构建的环境点云地图中增量式添加所检测到的可通行区域信息。试验表明在复杂的非结构道路环境中,该算法可以在移动机器人平台上同时实现准确的自定位和包含可通行区域的环境地图。

1 基于行人学习的可通行区域感知

在人机共融环境中,行人行走活动的地面状况复杂多样,属于非结构道路环境,其道路颜色、形状不规则,没有明显边界与特征,传统方法很难在计算复杂度、适用范围和准确性等综合性能上取得最佳。本文提出一种基于行人学习的自适应可通行域感知方法,该方法思路如图 1 所示。

图 1 可通行域感知框架图

检测到图像中的行人后,将图像中行人脚部所在的图像区域作为初始高置信度的可通行区域。当检测到多个行人时,对行人脚部位置处的区域建立可通行区域的模板;当检测到的行人较少时,一方面提取行人脚部位置处的区域作为模板,另一方面,结合之前帧中的道路模板综合建立可通行域模型。最后根据可通行区域模板对环境中的可通行域进行分割。因为环境中的道路区域一般为联通的区域,且连通的区域对于运动规划更有意义,所以这里我们希望分割出来的道路区域具有连通的性质。

1.1 快速行人检测

首先,需要快速检测环境中的行人,然后通过行人的位置建立道路区域的采样模板。

机器人在进行可通行区域感知时,首先提取图像的聚合通道特征(Aggregated channel features, ACF),再结合利用 Adaboost 分类算法进行快速的行人检测。为了检测到不同位置的行人,采用滑动窗口法,利用一个固定大小的窗口在待检测图像上面滑动,每次覆盖一个位置,检测所覆盖的图像区域是否是行人。为检测图像中不同尺度的行人,对原图构建特征金字塔,将原图进行不同尺度的缩放,在缩放后的图像上进行行人检测,并将检测结果对应到原图像中的相应像素位置。为了提升检测速度,这里采用了一种快速特征金字塔方法建立特征金字塔,大大提升了行人检测的效率^[18]。将金字塔分组,每

组的方式得到,每组的其他三层通过该组的第一层降采样得到。因此,同样建立八层金字塔的时候,只需要对原图像降采样一次,进行两次特征计算,对特征空间进行六次降采样,这样大大降低了行人检测时间。快速建立特征金字塔的方法如图 2 所示。

图 2 快速建立特征金字塔方法

通过建立特征金字塔,采用滑动窗口的方式进行行人检测时,会在相近位置多次检测到同一行人,产生窗口重叠的问题,如图 3 所示。

图 3 检测窗口重叠示意图

为解决窗口重叠问题,我们提出了一种自适应极大值抑制算法来对重叠的检测框进行优化。从图 3 可以看出,检测框在不同的方向上具有不同的分布特点,在竖直方向上重叠的框有更大的可能属于同一个人,在水平方向上相互重叠的框更有可能属于不同的行人。所以,以两个检测框的相对方向作为参考设定不同的面积重叠阈值,检测框之间的相对位置示意图如图 4 所示,这里所采用的自适应阈值为 $T = T_0(1 - 0.3 \cdot \ln(2 - \cos(\theta)))$, T 表示两个检测框以角度 θ 重合时的阈值, T_0 表示两个检测框以 0° 重合时的阈值。

图 4 检测框的位置示意图

利用自适应极大值抑制算法对行人检测得到的检测框进行优化后,得到的结果如图5所示。从结果可以看到,自适应极大值抑制算法在解决窗口重叠问题的同时,能够准确地将多个相邻的行人分开,得到各个行人的位置。

图5 自适应非极大值抑制效果

1.2 道路采样区域建模

通过快速行人检测算法,可以得到图片中行人位置信息,在行人脚步附近采样一块区域作为可通行区域的模型。当检测到单个行人时,直接在脚步位置附近采样,当检测到多个行人时,根据各个行人在图片中的位置综合得到采样区域,根据采样区域得到采样区域的实例如图6所示。

图6 根据行人位置得到采样区域

得到采样区域后,建立道路采样区域多变量高斯模型来描述道路区域,之后利用区域生长算法得到图像中的所有可通行区域。这里采用RGB颜色空间, IIS 颜色空间, LBP 共七个维度来描述每一个像素。采样区域的均值矢量和协方差矩阵用每一个像素的七个通道计算,计算公式如式(1)、(2)所示

$$\mu_C = \frac{1}{n_s} \sum_{i=1}^{i=n_s} P_{C,i} \quad (1)$$

$$\Sigma_C = \frac{1}{n_s} \sum_{i=1}^{i=n_s} P_{C,i} P_{C,i}^T - \mu_C \mu_C^T \quad (2)$$

式中, $C = \{r, g, b, c_1, c_2, c_3, c_{LBP}\}$ 表示每个像素的七个通道, r, g, b 对应RGB通道, c_1, c_2, c_3 代表 IIS 通道, c_{LBP} 代表 LBP 通道。 n_s 表示样本区域中的像素点个数, $P_{C,i}$ 代表第 i 个像素点对应的七个通

道值构成的矢量。

图像中像素点与道路采样区域模型的相似度可以使用该像素点与多变量高斯模型之间的马氏距离来度量,像素点 $P_{C,i}$ 与模型 (μ_C, Σ_C) 之间的马氏距离记为 $D(P_{C,i})$, 其定义如式(3)所示

$$D(P_{C,i}) = \sqrt{(P_{C,i} - \mu_C)^T \Sigma_C^{-1} (P_{C,i} - \mu_C)} \quad (3)$$

对于道路采样区域中的所有像素,初始的距离均值 μ_{D_0} 和标准差 σ_{D_0} 可以分别按照式(4)、(5)计算

$$\mu_{D_0} = \frac{1}{n_s} \sum_{i=1}^{i=n_s} D(P_{C,i}) \quad (4)$$

$$\sigma_{D_0} = \sqrt{\frac{1}{n_s} \sum_{i=1}^{i=n_s} (D(P_{C,i}) - \mu_{D_0})^2} \quad (5)$$

对于其他位置的像素 $P_{C,j}$, 如果像素 $P_{C,j}$ 满足式6则该像素被判断为道路像素点, 式中 λ 为经验阈值, 阈值过大易将非道路点判断为道路点, 而若阈值过小则易将道路点判断为非道路点, 经实际数据测试, 值取为 1.2~2.0 之间较为合适, 后面的试验中均取为 1.5, μ_{D_k} 为迭代 k 次后的距离均值, σ_{D_k} 为迭代 k 次后的距离标准差, 其定义分别如式(7)、(8)所示

$$|D(P_{C,j}) - \mu_{D_k}| < \lambda \sigma_{D_k} \quad (6)$$

$$\mu_{D_k} = [\mu_{D_{k-1}} \cdot n_{s,k-1} + D(P_j)] / (n_{s,k-1} + 1) \quad (7)$$

$$\sigma_{D_k} = \sqrt{\left[\sigma_{D_{k-1}}^2 \cdot n_{s,k-1} + (D(P_j) - \mu_{D_k})^2 \right] / (n_{s,k-1} + 1)} \quad (8)$$

每次新增加一个像素点为道路区域时,需要更新道路区域距离的均值和标准差,更新后的均值和标准差分别按照式(7)、(8)计算。然后利用更新后的均值和标准差继续判断其他新的像素点是否属于可通行区域,其中 $n_{s,k-1}$ 代表当前可通行区域中像素点个数。

1.3 可通行区域生长算法

在建立道路采样区域的多变量高斯模型后,以式(6)作为准则判断图片中的某一个像素点是否属于道路区域。为了得到图片中所有联通的可通行区域,采用一种区域生长算法来进行可通行域感知。以图像中的某一点为中心,逐渐向四周生长,当周围没有满足判定条件的像素点时算法终止。相比与其他方法,区域生长算法具有效率高,能得到连通区域的优点。

进行区域生长算法时,第一步首先进行算法初始化,将采样区域像素添加到队列。每二步生长过程,将队首像素弹出,判断该像素周围的元素是否

属于道路区域,若属于道路区域,则将该像素标记为道路区域,更新道路区域模型,将像素添加到队列末尾,之后重复执行第二步;若不属于道路区域,则跳到下一步骤。第三步中,判断当前队列是否为空,若为空则算法退出,否则跳回继续执行第二步。利用区域生长算法感知可通行域得到的结果如图 7 所示,图中第一行为原始图像,第二行为利用区域生长算法求取可通行域得到的结果。

图 7 区域生长算法效果

2 与可通行区域结合的视觉 SLAM

有了环境中行人及可通行区域的信息之后,需要进一步将这些信息与机器人对自身的定位和环境地图的构建结合起来。本文采用了 ORB-SLAM 的整体算法框架,将视觉 SLAM 框架分为三个模块:跟踪模块,负责相机的位姿估计;制图模块,负责地图的绘制和位姿、地图的优化;闭环检测模块,负责闭环检测,消除机器人定位制图过程中的累积误差。在跟踪模块中主要进行特征提取,初始位姿估计,位姿优化,关键帧选择。在制图模块中主要进行冗余关键帧和地图点的剔除,语义信息的插入,新地图点的生成和局部优化工作。在闭环检测模块主要负责环路检测,环路融合,全局地图及位姿的优化。因为整体框架与 ORB-SLAM 一致,这里不再赘述与 ORB-SLAM 框架相同的算法策略,主要描述如何与前面提取到的行人信息与可通行区域信息结合,以提升算法效果及生成带有可通行区域信息的环境地图描述。

视觉 SLAM 部分的算法框架结构如图 8 所示。结合前述基于行人位置的道路可通行区域生成算法,主要对 ORB-SLAM 的原算法架构做了如下优化:在跟踪模块中,剔除了动态行人身上的特征点,并提出了更加鲁棒的恒速率模型+DLT 和贪心搜索

+DLT 特征匹配算法,加快了匹配的速度,同时提高匹配的精度;在基于局部地图进行位姿优化时,对不同的地图点赋予不同的权重,对质量较好的地图点赋予更高的权重,从而提高位姿优化的精度;在制图模块中,为每个关键帧保存一个点云结构,并且采用增量的方式将可通行区域信息添加进所绘制的环境地图,有效控制地图耗用内存的增长规模,有利用在大场景下实时运行。

图 8 同时定位与语义制图框架

2.1 基于混合模型的特征匹配算法

跟踪模块中采用特征法,使用 ORB 特征来跟踪相机在不同帧间的位姿变换。由于环境中存在的动态行人会降低 SLAM 算法精度,所以结合行人检测的结果,将动态行人身上的特征点剔除,如图 9 所示。

图 9 剔除动态特征点前后对比

在进行特征匹配搜索时,ORB-SLAM 采用的是恒速率模型,即第 $k-1$ 帧到第 k 帧之间的位姿变换和 $k-2$ 帧到 $k-1$ 帧之间的位姿变换相同。第 k 帧的初始位姿可以通过式(9)来估计

$$T_k = T_{k-2}^{-1} T_{k-1} T_{k-1} \quad (9)$$

但是在相机运动速率变化比较大且相机采样频率不高时,恒速率模型实际并不够准确。因此,为了满足实际中相机不同的运动状态,本文提出一种鲁棒的特征匹配算法,即恒速率+DLT 和贪心搜索+DLT 算法。

在进行特征匹配时,首先采用恒速率模型+DLT 的方法。与传统的恒速率模型不同,这里只使用恒速率模型搜索 50 对特征匹配,之后利用 DLT 算法求解出一个更加精确的位姿,以此位姿作为当前帧

的一个估计, 再利用重投影的方式搜索更多的匹配点。相比于传统的恒速率模型, 这里的算法可以在中间过程提供一个更加准确的位姿估计, 表面上看来这样会增加计算的时间, 但是, 根据此位姿估计进行重投影搜索, 我们可以在更小的范围内进行搜索, 这样会减小搜索的时间, 而且由于运动估计更加准确, 因此也会增加匹配的准确度。此外, 可以为后面优化求解相机的精确位姿提供更好的初值。

DLT 算法是一个用于解决包含尺度问题的最小二乘问题的算法。对于已知三维坐标的一点 X_i , 其世界坐标系下的齐次坐标为 $X_i = (X_w^i, Y_w^i, Z_w^i, 1)$, 其对应的像素点的齐次坐标 (x_i, y_i, w_i) , 它们之间的变换关系由式(10)决定, 其中 $P = K[R \ t]$ 为相机矩阵。由于同一点的两个齐次坐标表示的叉乘为零, 即 $(x_i, y_i, w_i)^T \times PX_i^T = 0$, 展开可以显式的得到式(11)

$$\begin{bmatrix} x_i \\ y_i \\ w_i \end{bmatrix} = K[R \ t] \begin{bmatrix} X_w^i \\ Y_w^i \\ Z_w^i \end{bmatrix} = P \begin{bmatrix} X_w^i \\ Y_w^i \\ Z_w^i \end{bmatrix} \quad (10)$$

$$\begin{bmatrix} \theta^T & -w_i X_i^T & y_i X_i^T \\ w_i X_i^T & \theta^T & -x_i X_i^T \\ -y_i X_i^T & x_i X_i^T & \theta^T \end{bmatrix} \begin{bmatrix} P^{1T} \\ P^{2T} \\ P^{3T} \end{bmatrix} = 0 \quad (11)$$

式中, P^{1T}, P^{2T}, P^{3T} 分别表示相机矩阵的各行矢量转置得到的列矢量, 式(11)表示的三个方程是线性相关的, 只能提供两个线性无关的方程, 而相机矩阵具有 11 个自由度, 所以需要 6 组这样的三维坐标与像素坐标的对应关系就可以完全确定相机矩阵 P 。由于这里有 50 组匹配, 所以采用随机采样一致性算法得到合适的相机矩阵 P , 并从中恢复出比恒速率模型估计更加精确的相机位姿 $[\hat{R} \ \hat{t}]$ 。然后我们以更加精确的相机位姿 $[\hat{R} \ \hat{t}]$ 作为指导, 利用重投影搜索的方式搜索更多的匹配点对。恒速率模型+DLT 算法的示意图如图 10 所示。

图 10 恒速率模型+DLT 算法示意图

如果恒速率模型+DLT 的方法匹配失败, 将采用贪心搜索+DLT 的方法来进行进一步匹配。首先

利用贪心搜索的方式搜索 100 对特征匹配, 然后利用 DLT 算法得到一个当前的位姿估计, 最后在这个位姿估计的基础上进行重投影搜索, 因为通过贪心搜索产生的匹配不如恒速率模型搜索产生的匹配对准确, 所以这里需要搜索更多的特征匹配对。

通过上面的两种方式进行特征匹配具有更高的鲁棒性, 而且会为后面的位姿优化提供一个更好的迭代初值, 通过更少的迭代次数, 可以产生更精确的结果。

2.2 加权 Bundle Adjustment 位姿优化

通过跟踪模块得到相机位姿初值以及帧间特征匹配对后, 需要利用这些特征匹配得到更加精确的相机位姿。这里, 采用 3D-2D 匹配的位姿优化过程。首先, 对于上一帧中观测到的三维特征点 (X_w^i, Y_w^i, Z_w^i) , 知道当前相机位姿 $[R \ t]$, 则该三维特征点在当前相机坐标系下的像素坐标 (u_e^i, v_e^i) 可由式(12)计算

$$Z_c^i \begin{bmatrix} u_e^i & v_e^i & 1 \end{bmatrix}^T = K[R \ t] \begin{bmatrix} X_w^i & Y_w^i & Z_w^i \end{bmatrix}^T \quad (12)$$

由于位姿存在误差所以预期的像素坐标和观测到的像素坐标存在误差 e_r^i , 如式(13)所示, 其中 (u^i, v^i) 表示该三维特征点实际对应的像素坐标。当前帧的位姿可以通过 BA 算法, 优化由式(14)表示的所有点产生的重投影误差实现

$$e_r^i = \begin{bmatrix} u^i & v^i \end{bmatrix}^T - \begin{bmatrix} u_e^i & v_e^i \end{bmatrix}^T \quad (13)$$

$$\min_{R, t} \sum_i (e_r^i)^T e_r^i \quad (14)$$

在 ORB-SLAM 的框架下, 它对每一个点对产生的重投影误差赋予了相同的权重, 其优化方式如式(14)所示。但实际上, 在环境中检测到的每个特征是应该具有不同的性质, 有的特征点处在特征点比较密集的地方, 有的特征点处于特征点比较稀疏的地方。有的特征点可以被多次观测到, 有的可能只有产生少数几次观测。所以应该考虑这样的因素, 为不同的特征匹配产生得到的重投影误差赋予不同的权重。考虑到优化的主要目的是消除定位的累积误差, 所以应当对重复出现的地图点在当前帧中的重投影误差项赋予更大的权值, 实际采用的优化目标如式(15)所示

$$\min_{R, t} \sum_i (e_r^i)^T e_r^i w_i \quad (15)$$

2.3 构建带有可通行区域信息的环境地图

地图的绘制是视觉 SLAM 算法的一个关键步骤, 传统基于特征的 SLAM 方法绘制的地图是一系列离散的特征点云, 这些特征点存在的唯一意义就

是用来定位。对于机器人来说,它可以通过将地图中的特征点和视野中的特征点进行匹配从而计算自身位姿,但是其并不能因此而知道环境中哪些地方是可以通行的,因此难以单靠视觉实现路径规划与自主导航。在对这些点云进行后续分析的时候,也很难从这种地图中得到更加有用的信息。因此,这里提出了一种绘制地图的方式,不但可以得到特征点点云地图,还包含了环境中的可通行区域信息。这样的地图对于机器人来说可以更方便地实现定位导航,对于人来说,可以使我们更好地了解环境中的状况。

地图点的生成、剔除以及局部地图点优化过程均采用 ORB-SLAM 原框架中的算法策略,只是在优化过程中依然对不同误差项分配不同的权重以提升算法精度,这里不再赘述。

前面对可通行区域检测到的是从 2D 图像中得到的 2D 信息,需要将其构建到 3D 点云地图当中。若图像坐标系中的某道路点像素坐标为 (u_i, v_i) , 其对应的深度是 d_i , 那么可以将其投影到三维空间中去, 其在相机坐标系下的三维坐标如式(16)~(18)所示

$$X_c = (u - u_0) \cdot Z_c / f_x \quad (16)$$

$$Y_c = (v - v_0) \cdot Z_c / f_y \quad (17)$$

$$Z_c = d / s \quad (18)$$

但是,如果只是在地图中将道路点通过 2D-3D 投影保存为 3D 离散点的集合,将会非常难以管理,并且难以对这样的地图进行操作。所以这里我们将每个关键帧所看到的道路区域保存为一个点云。另外,为了能有好的帧间跟踪效果,我们的关键帧选择策略使得相邻关键帧之间视角变换不会太大,这样会造成一个非常严重的问题,即相邻关键帧之间所看到的可通行区域必然有大量的重复。如果我们不对重复区域进行处理,那么这样的地图势必造成耗用内存的无限制增长,所以这里我们采用增量的方式绘制语义地图,每次为当前关键帧生成可通行区域的点云数据时,将当前帧的地面点点云和上一个关键帧中的点云比较,只保留新增的地面部分的点云数据,这样会大大减少语义地图占用的内存空间。图片中的可通行域到三维空间的映射如图 11 所示,图中的绿色部分表示可通行区域信息。

3 试验结果

3.1 基于行人位置的可通行域感知结果

为了验证所提出的基于行人位置的可通行域感

图 11 二维可通行域到三维空间的映射

知算法的效果,使用在不同的环境中得到的存在行人的图片上验证算法结果。图 12 中第一列是哈尔滨工业大学(深圳)校园园区中采集的图片,第二列是网络中获取的图片,第三列和第四列来源于数据集。图中第一行是行人检测结果,第二行是采样区域选择结果,第三行是可通行区域感知生成结果。这些图片中,有的没有明确的道路边缘信息,有的道路区域比较复杂,有的行人很少,有的行人很多,但所提出的检测算法在这些复杂的环境中都有比较好的效果,能够得到比较完整、准确的可通行路面区域。

图 12 可通行区域感知结果

3.2 SLAM 算法定位结果

机器人定位是机器人进行路径规划,自主导航的基础,所以机器人的定位精度是非常重要的指标。这里在 KITTI 公开数据集上对算法的定位精度进行衡量,并与 ORB-SLAM 得到的结果进行比较。图 13 为改进前后的算法在 KITTI-00 双目序列上的比较结果,图 13a 表示旋转误差,图 13b 表示平移误差,图中带方格的曲线为 ORB-SLAM 的误差曲线,带菱形的曲线为改进后的算法的误差曲线,从图中可以看出改进后的算法比 ORB-SLAM 具有更小的误差。在整个数据集共 11 个序列上进行了精度比较的综合结果显示改进后的 SLAM 算法比 ORB-SLAM 定位的精度提高 6%左右。

图 13 在 KITTI 数据集上的误差对比

3.3 机器人同时定位与可通行区域制图试验结果

最后在测试移动机器人在实际环境中运行的整体算法效果。我们分别在三个不同的场景下进行测试，得到了各个场景的语义地图，并且对算法的鲁棒性进行了验证。这里所使用的视觉传感器是 Carnegie Robotics 公司的传感器 Multisense-S7，该传感器集成双目视觉传感器和 IMU，双目之间的基线大约为 7 cm，传感器内部集成了 FPGA 进行加速运算，可以实时输出左右相机彩色图，深度图等信息，各输出图像之间在内部完成时间同步，输出帧率为 15 帧/s，图像大小为 1024×1024 。这里，直接利用左边相机的彩色图像和深度图像，将其作为深度相机来使用。试验中使用的机器人平台为 Robotnik 公司的轮式机器人 Summit-XL，整个试验平台如图 14 所示。

3.3.1 实际环境中的定位精度与地图构建

试验场景一是一条笔直的长 27.9 m 的人行小道，在真实的环境中我们得到的试验场景一所构建的点云地图如图 15 所示，左边是一个整体上的效果，右边是细节展示，图中的方框代表相机的位姿，

图 14 Summit-XL 机器人试验平台

星号点代表道路点，其余像素点代表特征点。机器人定位的结果显示路程为 27.67 m，累积误差为 23 cm。可以看到相比于传统的 SLAM 方法，我们提出的框架在保证定位精度和实时性的同时，可以得到更加丰富的环境信息，基于这些可通行域信息，机器人可以实现更加高级的功能。

图 15 试验场景一中的地图效果

试验环境二、三是一条林间小道，最后得到的地图实际效果如图 16 所示，图 16a、16b 分别对应

图 16 不同场景下的地图效果

试验场景二和三的点云地图。从图中可以看出最后能得到完整连续的可通行区域,而且得到的可通行区域可以反映出真实环境中的道路。

3.3.2 不同场景下的鲁棒性

在进行相机位姿跟踪时,由于机器人运动状态的复杂多变,ORB-SLAM 算法比较容易出现跟踪丢失的情况,这里将对改进算法的鲁棒性进行验证。在试验中,控制相机进行较剧烈的运动,并观察 ORB-SLAM 和改进后的算法的跟踪效果。试验的效果如图 17 所示,图中的标记点为成功匹配跟踪的特征点,图 17a、17c 对应 ORB-SLAM 的效果,图 17b、17d 对应改进后的算法效果。从试验中可以看出改进后的算法比 ORB-SLAM 具有更好的鲁棒性。

(a) ORB-SLAM 跟踪失败(一)

(b) 改进后的算法跟踪成功(一)

(c) ORB-SLAM 跟踪失败(二)

(d) 改进后的算法跟踪成功(二)

图 17 改进前后的跟踪效果对比

4 结论

本文提出了一种新的基于行人学习的环境可通行区域感知算法,并与视觉 SLAM 相结合进行机器人自主定位与可通行区域地图构建。针对人机共融环境下的非结构道路,采用行人学习的方式,使用

快速构建图像金字塔与自适应极大值抑制算法来定位行人并提取采样区域,构建道路的多变量高斯模型,之后采用区域生长算法得到环境中连通的全部道路区域。在视觉 SLAM 模块中,充分利用所提取的行人信息与可通行区域信息,在匹配过程中将动态行人身上的特征点剔除以提升算法精度,通过恒速率模型+DLT 算法和贪心搜索+DLT 算法结合的方式来进行快速特征匹配,在 Bundle Adjustment 优化过程中针对不同特征点给予不同的权重,并将可通行区域信息构建到环境点云地图当中去。试验结果表明,本算法可以快速准确地检测到环境中的行人并成功检测连通的可通行道路区域,也可以准确构建包含可通行区域的环境点云地图。与 ORB-SLAM 算法在 KITTI 数据集和真实场景下的对比试验表明,在搭载普通 CPU 计算资源的移动机器人平台上,本算法在保证运算实时性的基础上,有着更高的自主定位精度与对机器人剧烈运动状态下的鲁棒性。未来的研究将更加关注如何在机器人平台上快速检测其他多种环境语义信息,并与视觉 SLAM 技术相结合;也将关注包含这些语义信息的点云地图在机器人导航与路径规划方面的应用,使移动机器人在复杂的现实场景中可以更加自主地运动并执行实际任务。

参 考 文 献

- [1] RAMSTROM O, CHRISTENSEN H. A method for following unmarked roads[C]//Intelligent Vehicles Symposium, 2005. Proceedings IEEE. IEEE, 2005: 650-655.
- [2] TAN C, HONG T, CHANG T, et al. Color model-based real-time learning for road following[C]// Intelligent Transportation Systems Conference, 2006. ITSC'06. IEEE. IEEE, 2006: 939-944.
- [3] ALVAREZ J M Á, LOPEZ A M. Road detection based on illuminant invariance[J]. IEEE Transactions on Intelligent Transportation Systems, 2011, 12(1): 184-193.
- [4] LU X. Self-supervised road detection from a single image[C]// Image Processing (ICIP), 2015 IEEE International Conference on. IEEE, 2015: 2989-2993.
- [5] RASMUSSEN C. Texture-based vanishing point voting for road shape estimation[C]// British Machine Vision Conference. 2004: 405-430.
- [6] KONG H, AUDIBERT J Y, PONCE J. General road detection from a single image[J]. IEEE Transactions on Image Processing, 2010, 19(8): 2211-2220.
- [7] CHANG C K, SIAGIAN C, ITTI L. Mobile robot monocular vision navigation based on road region and

- boundary estimation[C]// Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on. IEEE, 2012: 1043-1050.
- [8] ALVAREZ J M, LOPEZ A M, GEVERS T, et al. Combining priors, appearance, and context for road detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 15(3): 1168-1178.
- [9] DENG F, ZHU X, HE C. Vision-based real-time traversable region detection for mobile robot in the outdoors[J]. Sensors, 2017, 17: 2101.
- [10] XIAO L, DAI B, LIU D, et al. Monocular road detection using structured random forest[J]. International Journal of Advanced Robotic Systems, 2016, 13(3): 101.
- [11] ALVAREZ J M, GEVERS T, LECUN Y, et al. Road scene segmentation from a single image[C]// European Conference on Computer Vision. Berlin: Springer, 2012: 376-389.
- [12] OLIVEIRA G L, BURGARD W, BROX T. Efficient deep models for monocular road segmentation[C]// Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on. IEEE, 2016: 4885-4891.
- [13] MUR-ARTAL R, MONTIEL J M M, TARDOS J D. ORB-SLAM: A versatile and accurate monocular SLAM system[J]. IEEE Transactions on Robotics, 2015, 31(5): 1147-1163.
- [14] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces[C]// Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on. IEEE, 2007: 225-234.
- [15] MOURIKIS A I, ROUMELIOTIS S I. A multi-state constraint Kalman filter for vision-aided inertial navigation[C]// Robotics and automation, 2007 IEEE international conference on. IEEE, 2007: 3565-3572.
- [16] LEUTENEGGER S, LYNEN S, BOSSE M, et al. Keyframe-based visual-inertial odometry using nonlinear optimization[J]. The International Journal of Robotics Research, 2015, 34(3): 314-334.
- [17] ENGEL J, SCHÖPS T, CREMERS D. LSD-SLAM: Large-scale direct monocular SLAM[C]// European Conference on Computer Vision. Springer, Cham, 2014: 834-849.
- [18] DOLLÁR P, APPEL R, BELONGIE S, et al. Fast feature pyramids for object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(8): 1532-1545.

作者简介: 邢志伟, 男, 1986 年出生, 博士研究生。主要研究方向为户外大尺度环境下移动机器人的自主定位与导航。

E-mail: williamxing@foxmail.com

朱晓蕊(通信作者), 女, 1977 年出生, 博士, 教授, 博士研究生导师。主要研究方向为移动机器人系统与技术, 2012 年获得国家科技进步二等奖。

E-mail: zhuxiaorui@hotmail.com

何超, 男, 1992 年出生, 硕士研究生。主要研究方向为机器人视觉定位, 图像处理。

E-mail: hithechao@gmail.com