

动态场景下基于语义分割的视觉SLAM方法

杜晓英¹, 袁庆霓^{1,2,3}, 齐建友¹, 王晨¹, 杜飞龙¹, 任澳¹

(1. 贵州大学现代制造技术教育部重点实验室, 贵州 贵阳 550025; 2. 贵州大学机械工程学院, 贵州 贵阳 550025;
3. 贵州大学省部共建公共大数据国家重点实验室, 贵州 贵阳 550025)

摘要: 针对在动态场景下视觉同步定位与建图(SLAM)鲁棒性差、定位与建图精度易受动态物体干扰的问题, 设计一种基于改进DeepLabv3plus与多视图几何的语义视觉SLAM算法。以语义分割网络DeepLabv3plus为基础, 采用轻量级卷积网络MobileNetV2进行特征提取, 并使用深度可分离卷积代替空洞空间金字塔池化模块中的标准卷积, 同时引入注意力机制, 提出改进的语义分割网络DeepLabv3plus。将改进后的语义分割网络DeepLabv3plus与多视图几何结合, 提出动态点检测方法, 以提高视觉SLAM在动态场景下的鲁棒性。在此基础上, 构建包含语义信息和几何信息的三维语义静态地图。在TUM数据集上的实验结果表明, 与ORB-SLAM2相比, 该算法在高动态序列下的绝对轨迹误差的均方根误差值和标准差(SD)值最高分别提升98%和97%。

关键词: DeepLabv3plus网络; 视觉同步定位与建图; 多视图几何; 动态场景; 语义地图

源代码链接: <https://github.com/yinyinyiy/improved-DeepLabplus-slam.git>

中图分类号: TP391.41

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0067370

Visual SLAM Method Based on Semantic Segmentation in Dynamic Scenes

DU Xiaoying¹, YUAN Qingni^{1,2,3}, QI Jianyou¹, WANG Chen¹, DU Feilong¹, REN Ao¹

(1. Key Laboratory of Advanced Manufacturing Technology, Ministry of Education, Guizhou University, Guiyang 550025, Guizhou, China;
2. School of Mechanical Engineering, Guizhou University, Guiyang 550025, Guizhou, China;
3. State Key Laboratory of Public Big Data Jointly Built by Provincial and Ministerial Governments,
Guizhou University, Guiyang 550025, Guizhou, China)

[Abstract] A semantic visual SLAM algorithm based on an improved semantic segmentation network DeepLabv3plus and multiview geometry is designed to address the issues of poor robustness and susceptibility to interference from dynamic objects in visual Synchronous Localization And Map (SLAM) construction in dynamic scenes. Based on the semantic segmentation network DeepLabv3plus, a lightweight convolutional network MobileNetV2 is used for feature extraction, and depthwise separable convolutions are used instead of standard convolutions in the Atrous Spatial Pyramid Pooling (ASPP) module. Simultaneously, an attention mechanism is introduced to propose an improved semantic segmentation network DeepLabv3plus. Combining the improved semantic segmentation network DeepLabv3plus with multiview geometry, a dynamic point detection method is proposed to enhance the robustness of visual SLAM in dynamic scenes. On this basis, a three-dimensional semantic static map containing both semantic and geometric information is constructed. The experimental results on the TUM dataset demonstrate that compared with ORB-SLAM2, the highest Root Mean Square Error (RMSE) and Standard Deviation (SD) values increased by more than 98% and 97%, respectively.

[Key words] DeepLabv3plus network; visual Synchronous Localization And Map (SLAM); multiview geometry; dynamic scenes; semantic map

0 引言

同步定位与建图(SLAM)^[1]能够在未知环境下完成地图的自主构建与定位, 是工业机器人、自主导航、无人驾驶等应用中的基础技术之一。目前

SLAM采用的传感器主要是激光雷达、相机以及惯性测量单元(IMU)^[2], 视觉SLAM是一种以相机作为传感器的SLAM, 如ORB-SLAM2^[3]、LSD-SLAM^[4]、DSO^[5]等视觉SLAM方法。

研究视觉SLAM在真实动态场景下的性能表现

收稿日期: 2023-04-06 修回日期: 2023-05-30

基金项目: 国家自然科学基金(52165063, 52065010); 贵州省科技厅资助项目([2022]重点024, [2022]一般140, [2023]一般094, [2023]一般025); 贵州大学实验室开放资助项目(SYSKF2023-089)。

通信作者 E-mail: qnyuan@gzu.edu.cn。

已成为 SLAM 领域中的趋势,如张慧娟等^[6]先计算获得变换矩阵,之后提取线特征,并对其进行静态权重的评估,最后通过余下的静态特征进行相机位姿估计完成跟踪任务。杨世强等^[7]对传统 SLAM 系统进行改进,提出一种基于几何约束的动态目标检测算法。DAI 等^[8]采用 Delaunay 三角剖分方法为地图点建立与图类似的结构,以判断其邻接关系,之后,将多个关键帧之间观测不一样的边缘移除,最终,完成动态物和静态背景的分隔。以上算法虽然可以提高动态环境下视觉 SLAM 的鲁棒性,但是其构建的地图缺少丰富的语义信息。随着深度学习的发展和计算机性能的不提高,目标检测与语义分割为解决动态环境下 VSLAM 算法鲁棒性差的问题提供另 1 种技术路线。YU 等^[9]提出 DS-SLAM 算法,利用语义分割网络 SegNet 和运动一致性检测方法剔除动态特征点。BESCOS 等^[10]采用实例分割网络 Mask R-CNN 分割先验动态目标,并利用多视图结合方法检测潜在运动物。ZHONG 等^[11]采用目标检测方法 SSD^[12]检测帧中的动态物体,并对动态特征点进行运动概率传播,在跟踪线程中剔除动态点。XIAO 等^[13]同样使用 SSD 检验先验运动物体,并通过选择性跟踪算法处理动态特征点。然而,无论是 SegNet 还是 Mask R-CNN^[14]网络,都存在计算量大、运行时间长的问题。虽然语义分割可以检测动态物体,改善动态物体对 SLAM 的影响,但是如何选择 1 个精度高且实时性较好的语义分割网络完成动态特征点的检测,是动态 SLAM 需要解决的首要问题。

针对以上问题,本文结合轻量化语义分割网络和视觉 ORB-SLAM2,提出一种面向动态场景的视觉 SLAM 方法,在分割出具体动态对象和静态对象的基础上,去除动态目标并将剩余的静态特征用于位姿估计,提高定位精度。本文主要贡献:1)提出改进的 DeepLabv3plus 语义分割网络,采用轻量级卷积网络 MobileNetV2 提取特征,将深度可分离卷积应用到空洞空间金字塔池化(ASPP)模块中,在网络中插入 SE 注意力机制,减少模型参数量;2)设计一种语义分割结合多视图几何的动态点检测方法,利用改进后的 DeepLabv3plus 识别检测先验动态物体,通过多视图几何检测剩余的运动物体,并剔除动态目标,有效减少动态物体的影响,提高视觉 SLAM 的定位精度;3)在点云地图处理的基础上,结合语义信息构建语义点云地图,并将点云地图转换为八叉树地图,便于后续机器人完成路径规划任务。

1 算法框架

ORB-SLAM2 是 1 个开源视觉 SLAM 算法,因其良好的代码规范性与拓展性,被众多研究人员使用。

ORB-SLAM2 主要组成部分为跟踪线程、建图线程、回环检测线程。本文以 ORB-SLAM2 为基础框架,提出一种动态场景下基于语义分割的视觉 SLAM 方法,具体内容包括:在视觉 ORB-SLAM2 的 RGB-D 相机模式中,在原有的前端里程计、局部建图、回环检测 3 个线程基础上,添加语义分割模块、动态特征点检测模块和构建语义八叉树地图的线程。视觉 SLAM 总体框架如图 1 所示。首先,RGB-D 相机获取的 RGB 图像传入跟踪线程,在这个线程中采用 ORB 方式提取图像特征,包含当前帧的关键点和描述子;然后,通过语义分割网络对 RGB 图像进行像素级的语义分割,分割出具体的对象并筛选出运动对象,对动态点进行初步的去除,如行走的人等;最后,利用多视图几何检测并去除剩余动态物体,将剩余的静态特征用于位姿估计;最终,在语义地图构建线程中利用语义分割提取的语义信息生成点云地图并转换为八叉树地图。

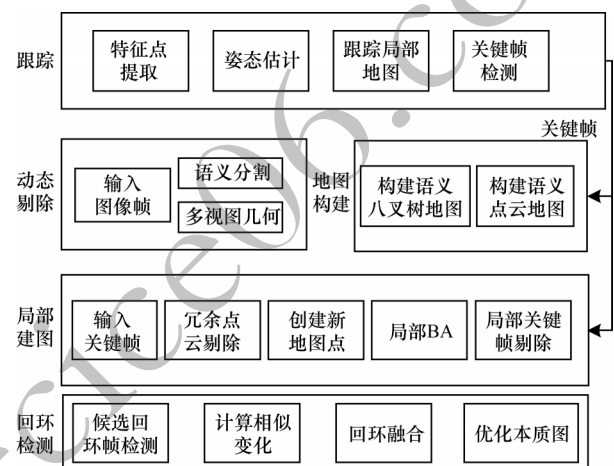


图 1 视觉 SLAM 总体框架

Fig.1 Overall framework of visual SLAM

1.1 改进的 DeepLabv3plus 语义分割网络

本文以 DeepLabv3plus^[15]为基础框架,设计一种轻量化网络。首先,采用 MobileNetV2^[16]完成特征提取;然后,将深度可分离卷积应用到 ASPP 模块中;最后,在网络中插入 SE 注意力机制^[17]。图 2 所示为本文改进的 DeepLabv3plus 的网络结构,与原网络结构相比,主干网络选择 MobileNetV2 代替原来的 Xception^[18],并采用迁移学习加载预训练模型,在 ASPP 模块中使用深度可分离卷积,在保证分割精度的同时缩短运行时间。考虑到注意力机制能够自动调整网络中表现较好和表现较差的特征通道权重,本文还在网络中插入 SE 注意力机制,使网络性能尽可能达到最佳,提升训练效果。

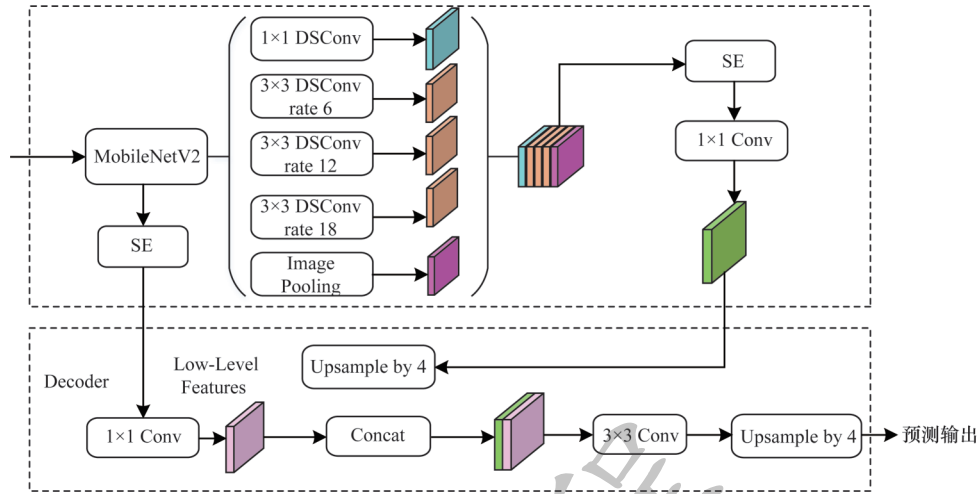


图 2 改进的 DeepLabv3plus 网络结构

Fig.2 Structure of improved DeepLabv3plus network

1.2 多视图几何结合的动态点检测方法

研究人员利用语义分割网络只能检验先验运动目标对象(如人),忽略了椅子、书本等静态物体发生被动运动时对 SLAM 产生的影响,如人们手中的书本或者人为推动的椅子等,应将其视为动态目标对象,却被视为静态对象,这将会对 SLAM 产生较大影响。因此,本文进一步采用多视图几何方法进行动态特征点检测:将地图点云投影到当前帧,并利用视点差异和深度值变化大小将目标对象区分为动态目标和静态目标。多视图几何原理示意图如图 3 所示,对于每个输入帧,选择多个之前与输入图像帧重合度较高的帧,把关键帧 K_f 中的关键点 p 投影到当前帧 C_f ,得到投影点 p' 和投影深度 D_{proj} ,每个关键点对应的 3D 点是 P ,然后计算视差角 α (p 的反投影与 p' 之间形成的夹角 α)。从 TUM 数据集上的测试结果可以看出,当 $\alpha > 30^\circ$,可判断其为动态点。同时,本文还须计算深度值差 $\Delta D = D_{proj} - D'$, D' 表示当前帧中关键点深度,若 $\Delta D = 0$,该点被认为是静态的,若 $\Delta D > 0$,则该点被认为是动态的。

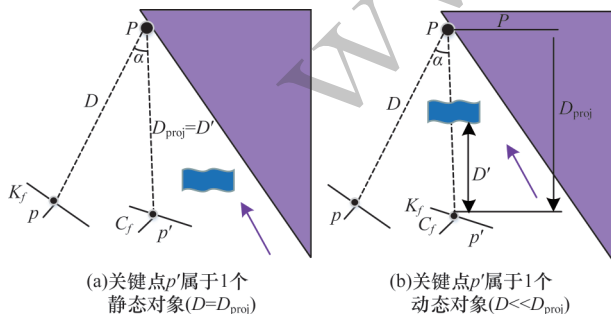


图 3 多视图几何原理示意图

Fig.3 Schematic diagram of multiview geometry principles

1.3 语义地图的构建

在语义地图构建线程中,借助 PCL 库^[19],结合关键帧和深度图生成点云,利用当前帧的位姿及其点云进行点云后续处理,完成点云地图的构建,并在点云地图中进行语义信息的标注。然而点云地图虽然给人很直观的感觉,但是点云地图存在存储空间大、位置信息冗余以及不能直接用于导航等问题。与点云地图相比,八叉树地图^[20]同样具有点云地图的直观性,但存储空间却远远小于点云地图,同时八叉树地图可以用于物流仓储机器人后续的路径规划,适用于各种导航算法,如 A*、D* 导航^[21]算法。因此,本文对点云地图做进一步处理,将点云地图转换为八叉树地图,完成语义八叉树地图的构建。但是在建图过程中,因存在传感器的噪声和运动物体产生的误差,导致某个节点在每个时间点上状态都不一样。因此,本文选用概率方式说明某一节点被占据还是未被占据。假设使用 $x \in [0, 1]$ 进行说明,设 x 初始值为 0.5,若该节点状态一直是被占据的,那么 x 值会不断增大,反之,该节点是未被占据的,那么 x 值会不断减小。若 x 出现不断增大或减小的情况,则会出现 x 值跳出 $[0, 1]$ 区间的情况,给数据处理带来不便。因此,采用概率对数值描述节点是否被占用,设 $y \in \mathbb{R}$ (实数集)表示概率对数值,占用概率 p 的数值范围为 $[0, 1]$,则 logit 变换公式如下:

$$y = \log_a(p) = \log_a\left(\frac{p}{1-p}\right) \quad (1)$$

式(1)可逆变换为:

$$p = \text{logit}^{-1}(y) = \frac{1}{1 + \exp(-y)} \quad (2)$$

假设某节点 n 在 T 时刻的观测概率是 $P(n|Z_{1:T})$, Z 表示观测数据,则其被占据的概率 $P(n|Z_{1:T})$ 表示:

$$P(n|Z_{1:T}) = \left[1 + \frac{1 - P(n|Z_T)}{P(n|Z_T)} \frac{1 - P(n|Z_{1:T-1})}{P(n|Z_{1:T-1})} \right] - \frac{P(n)}{1 - P(n)} \quad (3)$$

其中： $P(n)$ 表示节点 n 被占据的先验概率； $P(n|Z_{1:T-1})$ 表示 n 节点从起始到 $T-1$ 时刻的估计概率。本文将先验概率 $P(n)$ 设为 0.5, 那么式(3)转换为概率对形式为 $L(n|Z_{1:T})$, 表示节点 n 从起始到时间 T 的概率对数值, 则 $T+1$ 时刻观测概率如下:

$$L(n|Z_{1:T+1}) = L(n|Z_{1:T-1}) + L(n|Z_T) \quad (4)$$

其中： $L(n|Z_{1:T-1})$ 与 $L(n|Z_T)$ 分别表示节点 n 在 T 时刻前和 T 时刻被占据的概率对数值。由式(4)可知, 当某一节点被重复观测到被占据时, 其概率对数值随之增加, 否则减少。根据获得的信息, 能动态调整该节点的占据概率, 对八叉树地图不断进行更新。

2 实验结果与分析

实验搭配的平台配置如下: CPU为AMD R7-4800H, GPU型号为NVIDIA RTX3060, OS为64位Ubuntu 16.04系统的笔记本。语义分割实验部分使用的软件配置: PyTorch 1.7.1, CUDA 11.0, CUDNN 8.0.5.39, Python 3.6, 显存大小为16 GB。

2.1 语义分割实验测试

改进后的DeepLabv3plus在PASCAL Voc数据集^[22]上进行训练并验证。表1所示为DeepLabv3plus与本文改进的语义分割算法的实验测试结果。从表1可以看出, 本文设计的语义分割模型的平均交并比为73.4%, 模型大小为13.1 MB, 时间约21 ms。

表1 不同模型的语义分割性能对比
Table 1 Semantic segmentation performance comparison among different models

网络模型	平均交并比/%	模型大小/MB	运行时间/ms
DeepLabv3plus	79.6	439.0	189.4
改进的 DeepLabv3plus	73.4	13.1	21.0

图4所示为DeepLabv3plus与改进后的语义分割模型的分割效果对比。从图4可以看出, 改进后的DeepLabv3plus大致轮廓的分割结果与原模型相同, 在分割对象细节方面存在微小差别, 如图4中的第2行图中2个人之间的间隙部分。相比DeepLabv3plus, 改进后的DeepLabv3plus模型大小减少约97%, 单张图片运行时间缩短约89%。

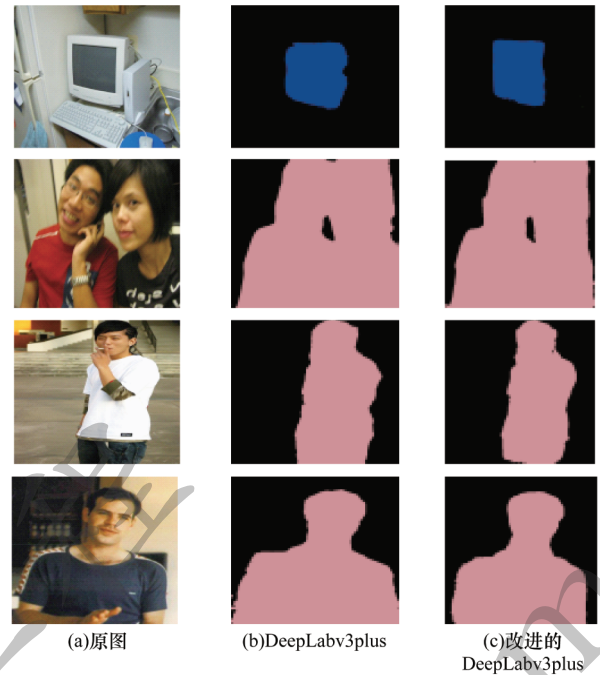


图4 改进前后DeepLabv3plus的分割结果对比

Fig.4 Comparison of segmentation results of DeepLabv3plus before and after improvement

2.2 定位误差实验

本文将从TUM^[23]数据集中选取高动态场景walking序列共4组序列进行定位误差实验。为了便于区分, 本文用fr3、w、half分别代表freiburg3、walking、halfsphere, 并将其作为序列名称。将本文算法分别与DS-SLAM、DLP-SLAM以及ORB-SLAM2进行比较, 其中DLP-SLAM是指以ORB-SLAM2为基础框架, 将DeepLabv3plus算法与多视图几何相结合进行动态物体剔除的视觉SLAM。

本文以绝对轨迹误差(ATE)和相对位姿误差(RPE)^[24]作为性能指标, 分别计算其对应的均方根误差(RMSE)和标准差(SD)。RMSE记录了估计值和真实值之间误差, 其值越小, 代表所估计的轨迹越接近真实轨迹。SD代表轨迹估计的离散度。

表2和表3所示为本文算法与ORB-SLAM2、DLP-SLAM以及DS-SLAM算法在绝对轨迹误差和相对位姿误差方面的实验结果。从表2和表3可以看出: 与传统ORB-SLAM2相比, 在高动态walking序列下, 本文算法的绝对轨迹误差的RMSE值和SD值最高分别提升98%和97%以上; 本文算法相对位姿误差的RMSE和SD指标值提升幅度在52%~74%之间。与DLP-SLAM、DS-SLAM相比, 本文算法的绝对轨迹误差和相对位姿误差的RMSE值也有所提升。

表 2 不同算法的绝对轨迹误差对比结果

Table 2 Comparison results of absolute trajectory errors among different algorithms

算法	fr3/w/xyz		fr3/w/rpy		fr3/w/static		fr3/w/half	
	RMSE	SD	RMSE	SD	RMSE	SD	RMSE	SD
ORB-SLAM2	0.578 2	0.243 6	0.869 2	0.467 5	0.391 3	0.164 4	0.813 1	0.313 8
DLP-SLAM	0.014 9	0.007 6	0.309 9	0.143 9	0.007 0	0.003 2	0.029 3	0.016 1
DS-SLAM	0.043 7	0.030 2	0.441 4	0.219 9	0.007 9	0.003 8	0.032 5	0.017 3
本文算法	0.013 9	0.006 8	0.051 3	0.033 7	0.007 8	0.003 8	0.027 8	0.014 0

表 3 不同算法的相对位姿误差平移部分的对比结果

Table 3 Comparison results of relative pose error translation parts among different algorithms

算法	fr3/w/xyz		fr3/w/rpy		fr3/w/static		fr3/w/half	
	RMSE	SD	RMSE	SD	RMSE	SD	RMSE	SD
ORB-SLAM2	0.044 1	0.025 5	0.053 2	0.048 1	0.014 9	0.010 9	0.039 2	0.034 2
DLP-SLAM	0.011 6	0.006 9	0.022 9	0.016 0	0.005 4	0.002 9	0.013 9	0.008 9
DS-SLAM	0.019 3	0.015 6	0.025 1	0.018 5	0.005 2	0.002 9	0.014 5	0.009 5
本文算法	0.011 3	0.006 5	0.025 0	0.018 3	0.005 4	0.002 9	0.013 9	0.008 9

图 5 所示为在高动态 fr3/w/rpy 序列中本文算法、ORB-SLAM2 以及 DLP-SLAM 在 x 、 y 、 z 轴上的位移估计值和真实值比较(彩色效果见《计算机工程》官网 HTML 版),其中蓝色线条代表 ORB-SLAM2,虚线代表实际值(ground truth)^[25],橘色实线代表 DLP-SLAM,绿色代表本文算法。

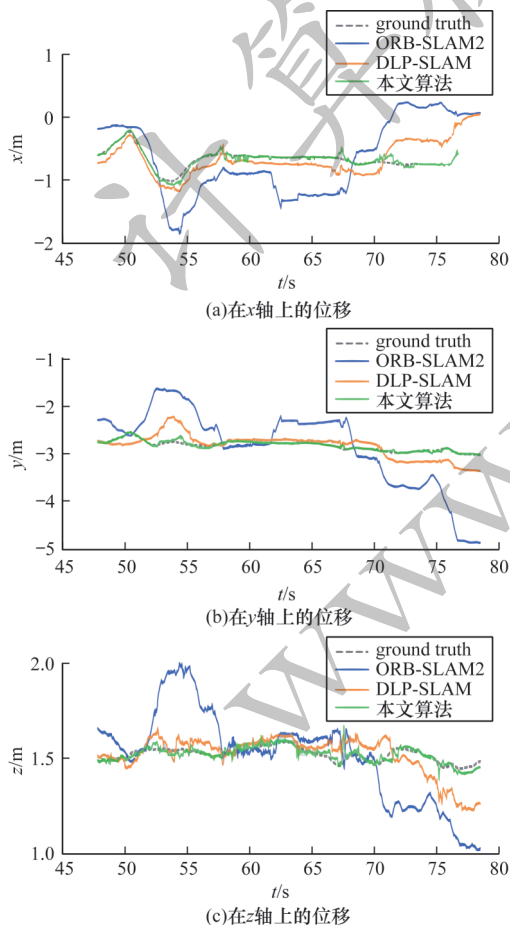
图 5 不同算法在 fr3/w/rpy 序列下的 x 、 y 、 z 轴位移对比Fig.5 Comparison of x , y , and z -axis displacements among different algorithms in the fr3/w/rpy sequence

图 6 所示为各算法姿态角估计值和真实值的比较(彩色效果见《计算机工程》官网 HTML 版)。

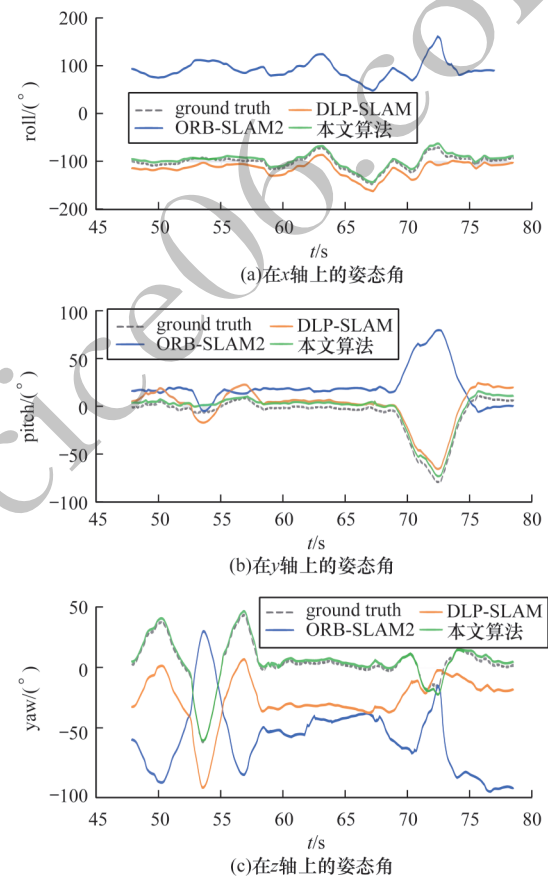


图 6 不同算法在 fr3/w/rpy 序列下的姿态角对比

Fig.6 Comparison of attitude angles among different algorithms in fr3/w/rpy sequences

从图 5 和图 6 可以看出,绿色实线和虚线几乎重合在一起,蓝色线条与虚线的距离最远,其误差最大,橘色实线与虚线之间也存在较小偏差。这表明无论是在 x 、 y 、 z 轴上的轨迹还是姿态角(yaw、pitch、roll)上的轨迹,与真实轨迹相比,ORB-SLAM2 生成的轨迹发生了较大幅度的漂移,DLP-SLAM 发生较小漂移,

而本文算法生成的轨迹与真实值几乎重合。本文算法在动态场景下定位与建图的精度更加准确,鲁棒性更好。其原因为原有算法假设环境是静态的,而运动物体的存在会影响位姿估计的结果,甚至引起系统故障的发生。本文算法利用改进的 DeepLabv3plus 以及多视图几何检测动态点,并剔除动态点,之后利用剔除掉动态点之后的静态点进行位姿估计,使获得的位姿估计更加准确,避免引起系统故障。

图 7 分别所示为在高动态序列 fr3/w/xyz 下 ORB-SLAM2、DLP-SLAM 及本文算法得到的相机运动轨迹与 ground truth 之间的偏差(彩色效果见《计算机工程》官网 HTML 版)。黑色线条代表 ground truth,蓝色线条代表预测轨迹,红色线条代表相机预测轨迹与 ground truth 之间的偏差,红色线条越短代表偏差越小。从图 7 可以看出,本文算法运动轨迹图中红色线条最短,位姿估计更加准确,说明在动态环境下定位误差较小,能有效降低动态物体对 VSLAM 的影响。

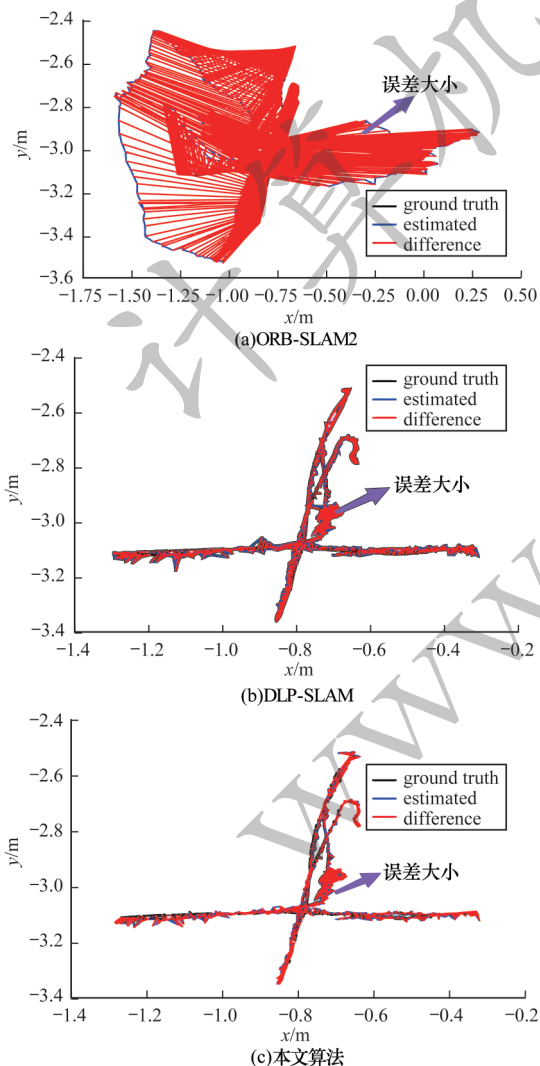


图 7 不同算法在高动态序列 fr3/w/xyz 下的轨迹
Fig.7 Trajectories among different algorithms under high dynamic sequence fr3/w/xyz

为进一步说明本文算法的有效性,本文在实验室真实环境下进行实验验证。图 8 所示为移动机器人的真实运动轨迹,边框表示移动机器人在真实环境中的实际运动轨迹。移动机器人平台利用搭载的 KinectV2 相机^[26]在搭建的实验环境中根据先前设计好的路线运动,并进行环境感知,记录环境数据。随后,在 Ubuntu16.04 中,将获取的数据通过 ROS 工具制作成 TUM 数据集的格式,进而验证本文算法的有效性和可行性。

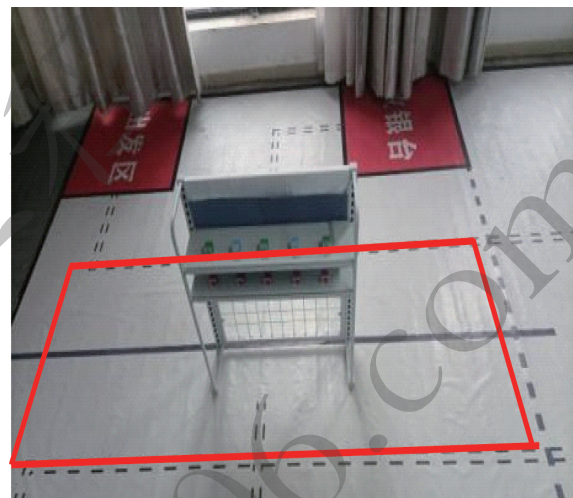


图 8 移动机器人的真实运动轨迹

Fig.8 Real motion trajectory of the mobile robot

图 9 所示为本文算法在真实环境中的特征提取。从图 9 可以看出,本文算法在进行特征提取时剔除了人和人手中物体共 2 个动态物体。



图 9 在真实场景下的特征提取

Fig.9 Feature extraction in the real scene

图 10 所示为本文算法和 ORB-SLAM2 算法在真实环境下的三维轨迹对比。从图 10 可以看出,由于实验场景中有运动物体,因此 ORB-SLAM2 算法生成的轨迹与实际运动轨迹相比出现了较大幅度的波动,而本文算法生成的轨迹与实际运动轨迹基本相符,波动幅度较小。

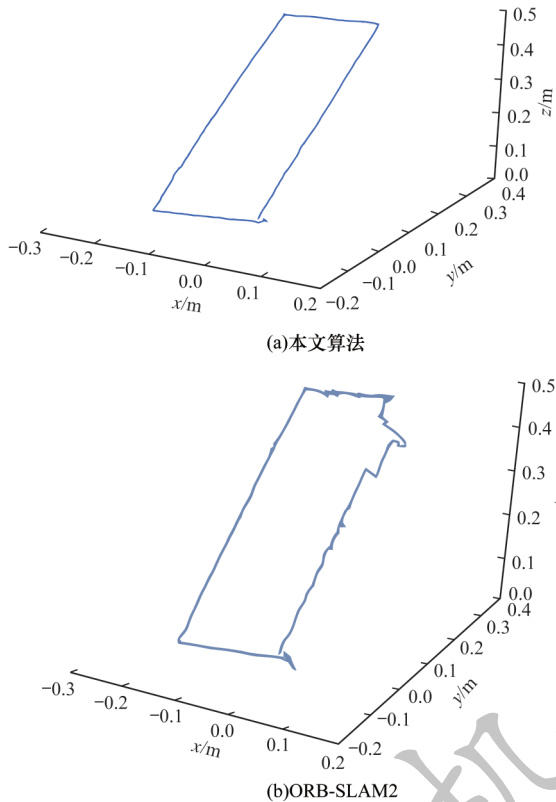


图10 本文算法和ORB-SLAM2的三维轨迹对比

Fig.10 Three-dimensional trajectory comparison between the proposed algorithm and ORB-SLAM2

2.3 建图测试

本文选取TUM数据集中walking序列的1个子序列作为测试对象,如图11所示,此序列中存在2个不断运动的行人。

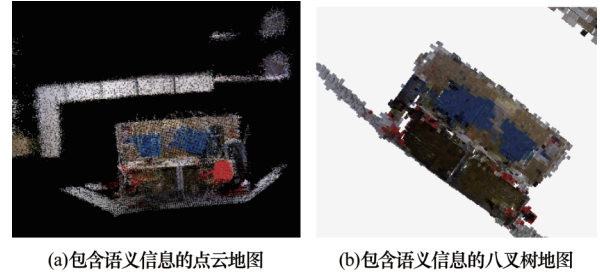


图11 fr3/w/static子序列

Fig.11 fr3/w/static subsequence

图12所示为本文方法建立的地图(彩色效果见《计算机工程》官网HTML版)。为了可视化,本文使用蓝色代表显示屏的像素,红色代表椅子的像素,其余像素使用物体本身颜色。从图12可以看出,本文算法建立的语义地图中已剔除动态物体,不存在显著的动态物体如行走的人,说明较好地完成地图构建。从图12中像素分布区域可以看出,显示屏中大部分像素得到了很好的语义赋予,但是椅子相对语义赋予较差。这是因为在PASCAL VOC数据集上训练时,该数据集中的椅子和TUM数据集中的椅子存在较大差异,因此本文算法没有得到较优的椅子分割结果。另外,为了说明八叉树地图所占内存空间小于点云地

图,本文记录了fr3/w/xyz、fr3/w/static序列构建的点云和八叉树地图所占内存大小,如表4所示。



(a)包含语义信息的点云地图 (b)包含语义信息的八叉树地图

图12 本文算法生成的点云和八叉树地图

Fig.12 Point cloud and octree maps generated by the proposed algorithm

表4 不同序列的存储空间对比

Table 4 Storage space comparison among different sequences 单位:MB

序列	八叉树地图	点云地图
fr3/w/xyz	1.20	2.8
fr3/w/static	0.79	2.1

3 结束语

本文提出动态场景下基于语义分割的视觉SLAM算法。通过对DeepLabv3plus进行改进,采用MobileNetV2进行特征提取,对ASPP模块进行修改,并引入SE注意力机制,提高分割速度,通过改进后的DeepLabv3plus和多视图几何方法检测动态特征点并剔除。实验结果表明,本文算法与ORB-SLAM2相比,运行时间延长,与DLP-SLAM相比运行时间有所降低,并且其精度依然较高,利用语义信息生成静态的语义八叉树地图,节省了大量的存储空间,同时生成的地图可直接用于机器人的路径规划中。在去除动态特征点时因分割的不完整导致部分动态点未被识别,因此后续将选择分割精度更高的网络。

参考文献

- [1] 陈首彬. 激光LiDAR/视觉融合的SLAM(LV-SLAM)关键技术研究[J]. 测绘学报, 2023, 52(1): 169.
CHEN S B. Research on key technologies of laser LiDAR/visual fusion SLAM(LV-SLAM) [J]. Journal of Surveying and Mapping, 2023, 52(1): 169. (in Chinese)
- [2] LIU Z, SHI D, LI R H, et al. PLC-VIO: visual-inertial odometry based on point-line constraints [J]. IEEE Transactions on Automation Science and Engineering, 2022, 19(3): 1880-1897.
- [3] MUR-ARTAL R, TARDOS J D. ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras [J]. IEEE Transactions on Robotics, 2017, 33(5): 1255-1262.
- [4] ENGEL J, SCHÖPS T, CREMERS D. LSD-SLAM: large-scale direct monocular SLAM [C]//Proceedings of European

- Conference on Computer Vision. Berlin, Germany; Springer, 2014; 834-849.
- [5] ENGEL J, KOLTUN V, CREMERS D. Direct sparse odometry[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(3): 611-625.
- [6] 张慧娟, 方灶军, 杨桂林. 动态环境下基于线特征的RGB-D视觉里程计[J]. *机器人*, 2019, 41(1): 75-82.
- ZHANG H J, FANG Z J, YANG G L. RGB-D visual odometer in dynamic environments using line features[J]. *Robot*, 2019, 41(1): 75-82. (in Chinese)
- [7] 杨世强, 范国豪, 白乐乐, 等. 基于几何约束的室内动态环境视觉SLAM[J]. *计算机工程与应用*, 2021, 57(16): 203-212.
- YANG S Q, FAN G H, BAI L L, et al. Geometric constraints-based visual SLAM under dynamic indoor environment [J]. *Computer Engineering and Applications*, 2021, 57(16): 203-212. (in Chinese)
- [8] DAI W C, ZHANG Y, LI P, et al. RGB-D SLAM in dynamic environments using point correlations[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(1): 373-389.
- [9] YU C, LIU Z X, LIU X J. DS-SLAM: a semantic visual SLAM towards dynamic environments[C]//*Proceedings of International Conference on Intelligent Robots and Systems*. Washington D. C., USA; IEEE Press, 2018: 1168-1174.
- [10] BESCOS B, FÁCIL J M, CIVERA J, et al. DynaSLAM: tracking, mapping and inpainting in dynamic scenes[J]. *IEEE Robotics and Automation Letters*, 2018, 3(4): 4076-4083.
- [11] ZHONG F W, WANG S, ZHANG Z Q. Detect-SLAM: making object detection and SLAM mutually beneficial[C]//*Proceedings of IEEE Winter Conference on Applications of Computer Vision*. Washington D. C., USA; IEEE Press, 2018: 1001-1010.
- [12] LIU W, ANGUELOV D, ERHAN D. SSD: single shot multibox detector[EB/OL]. [2023-03-01]. https://link.springer.com/content/pdf/10.1007/978-3-319-46448-0_2.pdf.
- [13] XIAO L H, WANG J G, QIU X S, et al. Dynamic-SLAM: semantic monocular visual localization and mapping based on deep learning in dynamic environment[J]. *Robotics and Autonomous Systems*, 2019, 117: 1-16.
- [14] HE K M, GKIOXARI G, DOLLÁR O, et al. Mask R-CNN[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 42(2): 386-397.
- [15] HU Z, ZHAO J, LUO Y, et al. Semantic SLAM based on improved DeepLabv3⁺ in dynamic scenarios [J]. *IEEE Access*, 2022, 10: 21160-21168.
- [16] BOTTLENECKS M I R A. MobileNetV2: inverted residuals and linear bottlenecks [C]//*Proceedings of Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA; IEEE Press, 2018: 18-23.
- [17] ZHANG X, LI J, HUA Z. MRSE-Net: multiscale residuals and SE-attention network for water body segmentation from satellite images [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022, 15: 5049-5064.
- [18] SHARMA S, KUMAR S. The Xception model: a potential feature extractor in breast cancer histology images classification [J]. *ICT Express*, 2022, 8(1): 101-108.
- [19] PAN Z, HOU J, YU L. Optimization RGB-D 3D reconstruction algorithm based on dynamic SLAM [J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 1-13.
- [20] ZHOU P, LIU Y, MENG Z. PointSLOT: real-time simultaneous localization and object tracking for dynamic environment [J]. *IEEE Robotics and Automation Letters*, 2023, 8(5): 2645-2652.
- [21] CAI Y Q, ZHOU W J, ZHANG L T, et al. DHFNet: dual-decoding hierarchical fusion network for RGB-thermal semantic segmentation [J]. *The Visual Computer*, 2023, 6: 1-11.
- [22] CHEN L, WANG Y, MIAO Z, et al. Transformer-based imitative reinforcement learning for multi-robot path planning [J]. *IEEE Transactions on Industrial Informatics*, 2023, 46: 1-10.
- [23] XIE H, ZHANG D, WANG J, et al. Semi-direct multimap SLAM system for real-time sparse 3D map reconstruction [J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 1-13.
- [24] ROSINOL A, LEONARD J J, CARLONE L. Probabilistic volumetric fusion for dense monocular SLAM [C]//*Proceedings of Winter Conference on Applications of Computer Vision*. Washington D. C., USA; IEEE Press, 2023: 3097-3105.
- [25] CHENG S, SUN C, ZHANG S, et al. SG-SLAM: a real-time RGB-D visual SLAM toward dynamic scenes with semantic and geometric information [J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 1-12.
- [26] WANG J, XU M, ZHAO G, et al. Feature- and distribution-based LiDAR SLAM with generalized feature representation and heuristic nonlinear optimization [J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 1-15.