



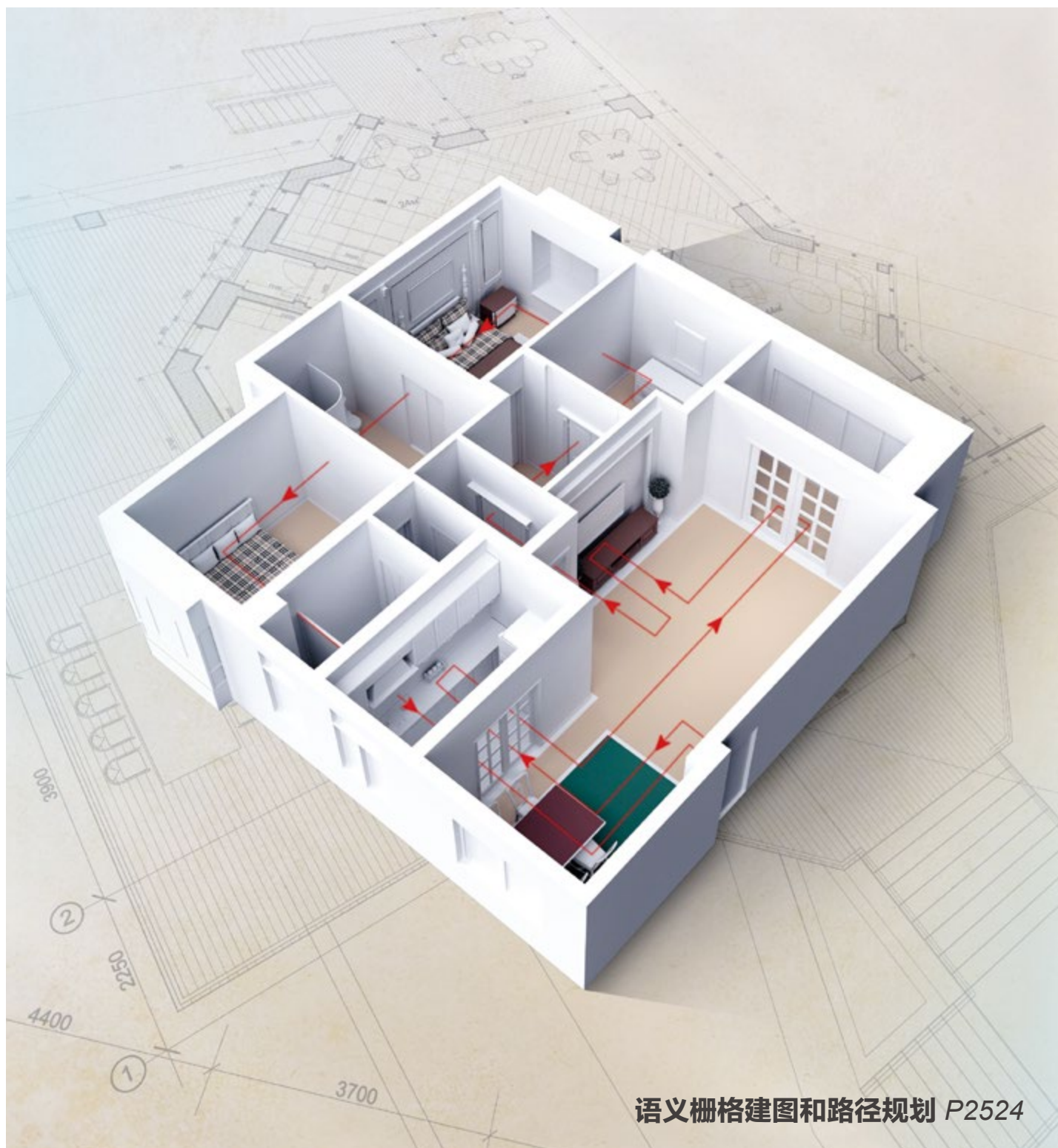
JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2021
10
VOL.26

ISSN1006-8961
CN11-3758/TB



语义栅格建图和路径规划 P2524



第26卷第10期（总第306期）
2021年10月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 吴一戎
编辑出版 《中国图象图形学报》编辑出版委员会
通信地址 北京市海淀区北四环西路19号
邮 编 100190
电子信箱 jig@aircas.ac.cn
电 话 010-58887035
网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian | Monthly, Started in 1996

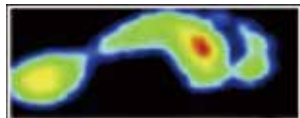
Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China
Zip code 100190
E-mail jig@aircas.ac.cn
Telephone 010-58887035
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



穿鞋足迹序列的足迹能量图组表达与识别(第2357页)



时空图卷积网络与注意机制的视频目标分割(第2376页)



双光流网络指导的视频目标检测(第2473页)

综述

图像分类的深度卷积神经网络模型综述

张珂, 冯晓晗, 郭玉荣, 苏昱坤, 赵凯, 赵振兵, 马占宇, 丁巧林 2305

中国图像工程25年

章毓晋 2326

数据集论文

FGSC-23: 面向深度学习精细识别的高分辨率光学遥感图像舰船目标数据集

姚力波, 张筱晗, 吕亚飞, 孙炜玮, 李孟洋 2337

图像处理和编码

并行生成网络的红外—可见光图像转换

余佩伦, 施佳, 王晗 2346

图像分析和识别

穿鞋足迹序列的足迹能量图组表达与识别

王新年, 于丹, 张涛 2357

时空图卷积网络与注意机制的视频目标分割

姚睿, 夏士雄, 周勇, 赵佳琦, 胡伏原 2376

多支路协同的RGB-T图像显著性目标检测

蒋亭亭, 刘昱, 马欣, 孙景林 2388

改进R-FCN模型的小尺度行人检测

刘万军, 董利兵, 曲海成 2400

边缘与区域不一致性引导下的图像拼接篡改检测网络

蒋小玉, 刘春晓 2411

图像理解和计算机视觉

各向异性导向滤波的红外与可见光图像融合

刘明威, 王任华, 李静, 焦映臻 2421

卷积稀疏与细节显著图解析的图像融合

杨培, 高雷阜, 訾玲玲 2433

计算机图形学

利用距离与内能极小平滑链接Bézier曲线

李军成, 刘成志, 刘珊君 2450

NCIG 2020

结合置信度加权融合与视觉注意机制的前景检测

成科扬, 孙爽, 王文杉, 师文喜, 李鹏, 詹永照 2462

双光流网络指导的视频目标检测

尉婉青, 禹晶, 史薪琪, 肖创柏 2473

生成对抗网络驱动的图像隐写与水印模型

郑钢, 胡东辉, 戈辉, 郑淑丽 2485

融合语义—表观特征的无监督前景分割

李熹, 马惠敏, 马洪兵, 王弈冬 2503

场景视点偏移的激光雷达点云分割

郑阳, 林春雨, 廖康, 赵耀, 薛松 2514

激光—相机系统语义栅格建图和路径规划

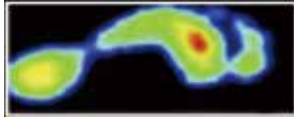
丁梦远, 郭迟, 黄凯 2524

融合词袋连通图的图像检索特征选择

李国祥, 王继军, 马文斌 2533

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Shoeprint sequence representation and recognition using shoeprint energy map set (P2357)



Spatial-temporal video object segmentation with graph convolutional network and attention mechanism(P2376)



Dual optical flow network-guided video object detection(P2473)

Review

Overview of deep convolutional neural networks for image classification

Zhang Ke, Feng Xiaohan, Guo Yurong, Su Yukun, Zhao Kai, Zhao Zhenbing, Ma Zhanyu, Ding Qiaolin 2305

Twenty-five years of image engineering in China

Zhang Yujin 2326

Dataset

FGSC-23: a large-scale dataset of high-resolution optical remote sensing image for deep learning-based fine-grained ship recognition

Yao Libo, Zhang Xiaohan, Lyu Yafei, Sun Weiwei, Li Mengyang 2337

Image Processing and Coding

Infrared-to-visible image translation based on parallel generator network

Yu Peilun, Shi Quan, Wang Han 2346

Image Analysis and Recognition

Shoeprint sequence representation and recognition using shoeprint energy map set

Wang Xinnian, Yu Dan, Zhang Tao 2357

Spatial-temporal video object segmentation with graph convolutional network and attention mechanism

Yao Rui, Xia Shixiong, Zhou Yong, Zhao Jiaqi, Hu Fuyuan 2376

Multi-path collaborative salient object detection based on RGB-T images

Jiang Tingting, Liu Yu, Ma Xin, Sun Jinglin 2388

Small-scale pedestrian detection based on improved R-FCN model

Liu Wanjun, Dong Libing, Qu Haicheng 2400

Edge and region inconsistency-guided image splicing tamper detection network

Jiang Xiaoyu, Liu Chunxiao 2411

Image Understanding and Computer Vision

Infrared and visible image fusion with multi-scale anisotropic guided filtering

Liu Mingwei, Wang Renhua, Li Jing, Jiao Yingzhen 2421

Image fusion method of convolution sparsity and detail saliency map analysis

Yang Pei, Gao Leifu, Zi Lingling 2433

Computer Graphics

Smoothing linked Bézier curves by distance and internal energy minimization

Li Juncheng, Liu Chengzhi, Liu Shanjun 2450

NCIG 2020

Foreground detection via fusing confidence by weight and visual attention mechanism

Cheng Keyang, Sun Shuang, Wang Wenshan, Shi Wenxi, Li Peng, Zhan Yongzhao 2462

Dual optical flow network-guided video object detection

Yu Wanjing, Yu Jing, Shi Xinqi, Xiao Chuangbai 2473

End-to-end image steganography and watermarking driven by generative adversarial networks

Zheng Gang, Hu Donghui, Ge Hui, Zheng Shuli 2485

Semantic-apparent feature-fusion-based unsupervised foreground segmentation method

Li Xi, Ma Huimin, Ma Hongbing, Wang Yidong 2503

LiDAR point cloud segmentation through scene viewpoint offset

Zheng Yang, Lin Chunyu, Liao Kang, Zhao Yao, Xue Song 2514

Semantic grid mapping and path planning combined with laser-camera system

Ding Mengyuan, Guo Chi, Huang Kai 2524

Feature selection method for image retrieval based on connected graphs and bag of words

Li Guoxiang, Wang Jijun, Ma Wenbin 2533

中图法分类号: TP242 文献标识码: A 文章编号: 1006-8961(2021)10-2524-09

论文引用格式: Ding M Y, Guo C and Huang K. 2021. Semantic grid mapping and path planning combined with laser-camera system. Journal of Image and Graphics, 26(10): 2524-2532 (丁梦远, 郭迟, 黄凯. 2021. 激光—相机系统语义栅格建图和路径规划. 中国图象图形学报, 26(10): 2524-2532) [DOI:10.11834/jig.200372]

激光—相机系统语义栅格建图和路径规划

丁梦远¹, 郭迟^{1,2*}, 黄凯¹

1. 武汉大学人工智能研究院, 武汉 430072; 2. 武汉大学卫星导航定位技术研究中心, 武汉 430072

摘要: **目的** SLAM(simultaneous localization and mapping)是移动机器人在未知环境进行探索、感知和导航的关键技术。激光 SLAM 测量精确,便于机器人导航和路径规划,但缺乏语义信息。而视觉 SLAM 的图像能提供丰富的语义信息,特征区分度更高,但其构建的地图不能直接用于路径规划和导航。为了实现移动机器人构建语义地图并在地图上进行路径规划,本文提出一种语义栅格建图方法。**方法** 建立可同步获取激光和语义数据的激光—相机系统,将采集的激光分割数据与目标检测算法获得的物体包围盒进行匹配,得到各物体对应的语义激光分割数据。将连续多帧语义激光分割数据同步融入占据栅格地图。对具有不同语义类别的栅格进行聚类,得到标注物体类别和轮廓的语义栅格地图。此外,针对语义栅格地图发布导航任务,利用路径搜索算法进行路径规划,并对其进行改进。**结果** 在实验室走廊和办公室分别进行了语义栅格建图的实验,并与原始栅格地图进行了比较。在语义栅格地图的基础上进行了路径规划,并采用了语义赋权算法对易移动物体的路径进行对比。**结论** 多种环境下的实验表明本文方法能获得与真实环境一致性较高、标注环境中物体类别和轮廓的语义栅格地图,且实验硬件结构简单、成本低、性能良好,适用于智能化机器人的导航和路径规划。

关键词: 智能机器人;语义栅格地图;激光 SLAM;目标检测;路径规划

Semantic grid mapping and path planning combined with laser-camera system

Ding Mengyuan¹, Guo Chi^{1,2*}, Huang Kai¹

1. Artificial Intelligence Research Institute, Wuhan University, Wuhan 430072, China;

2. GNSS Center, Wuhan University, Wuhan 430072, China

Abstract: **Objective** Intelligent mobile robots are widely used in industry, logistics, home service, and other fields. From complex industrial robots to simple sweeping robots, the ability of simultaneous localization and mapping is essential. Taking the common low-cost sweeping robot as an example, the scheme adopted is to obtain the distance information between the robot and an object in a plane through 2D lidar, and establish occupation grid map by using laser SLAM (simultaneous localization and mapping) to support robot navigation, path planning, and other functions. With the increasing demand of intelligent service, the geometric map with simple information cannot meet the needs of people. Tasks such as “cleaning the chair” and “going to the refrigerator” require the robot to perceive the environment from the geometric level to the content level. In addition to describing the geometric contour of the environment through grid map, the intelligent robot should have

收稿日期: 2020-07-16; 修回日期: 2020-11-13; 预印本日期: 2020-11-20

* 通信作者: 郭迟 guochi@whu.edu.cn

基金项目: 国家十三五重点研发计划项目(2018YFB1305001); 武汉市科技计划应用基础前沿项目(2019010701011413)

Supported by: China National Key Research and Development Program during the 13th Five-year Plan Period(2018YFB1305001)

the functions of target recognition, semantic segmentation, or scene classification to obtain semantic information. SLAM is a key technology for a mobile robot to explore, perceive, and navigate in an unknown environment, which can be divided into laser SLAM and visual SLAM. Laser SLAM is accurate and convenient for robot navigation and path planning, but it lacks semantic information. The image of visual SLAM can provide rich semantic information with a higher feature discrimination, but the map constructed by visual SLAM cannot be directly used for path planning and navigation. A semantic grid mapping method based on laser camera system is proposed in this paper to realize the construction of semantic map for mobile robot and path planning. **Method** To construct a semantic map that can be used for path planning, this paper uses a monocular camera assisted lidar to extract the object level features and the bounding box segmentation matching algorithm to obtain the semantic laser segmentation data as well as participate in the construction of the map. When the robot constructs the occupation grid map, the grid that stores the semantic information is called the semantic grid. The semantic map updates the occupation probability of the grid corresponding to each object and the semantic information in the grid. Then, through the steps of global optimization, semantic grid clustering, and semantic grid annotation, the semantic grid map with object category and contour is obtained. In addition, semantic tasks are published on the semantic grid map, the semantic weighting algorithm is used to identify the easily moving objects in the environment, and the path planning is improved. This system is mainly divided into three parts: semantic laser data extraction, semantic grid mapping and path planning. The input of the system includes the scanning data of 2D lidar and the pictures of monocular camera, and the output is the semantic grid map that can be used for path planning. Semantic laser segmentation data extraction is based on the laser camera system. The laser radar provides the scanning data in a certain height plane in the space. The module projects the laser segmentation in the camera field of view to the image and matches with the detection frame to obtain the semantic laser segmentation data. When the laser segmentation data are acquired, the semantic grid mapping should be carried out simultaneously. When the laser data are used to construct the grid map, the corresponding objects in the grid are marked with semantics and contour according to the semantic information. Moreover, density-based spatial clustering of applications with noise (DBSCAN) clustering algorithm is used to cluster the grids of the same object category and obtain the grid set representing each independent object. Finally, according to the semantic information of objects in the grid set, the corresponding object positions occupying the grid map are marked with different colors and words, and the semantic grid map is obtained. The semantic grid map contains geometric information and content information of environment, which can provide specific semantic objects for robot path planning and assist robot navigation. Adaptive Monte Carlo localization (AMCL) is used to locate a mobile robot in a 2D environment, and the pose information of a robot is determined. The global path planning uses A* algorithm to plan the global path from the starting point to the target given any target in the map and optimizes it according to the semantic weighting algorithm. **Result** The hardware of the system consists of a mobile robot and a control machine. The robot platform is composed of a kobuki chassis, an Robo Sense (RS) lidar, and a monocular camera. The laptop of the control computer is configured with 2.5 GHz main frequency, Intel Core i5-7300 HQ processor, GTX 1050ti GPU, and 8 G memory. Semantic mapping and path planning system are deployed as software facilities and run in the robot operating system (ROS) environment. The test system is Ubuntu 16.04. First, the performance of semantic laser data extractions is evaluated, and the detection accuracy of different object detection, laser segmentation, and bounding box segmentation matching is measured. Second, the semantic grid mapping experiment realizes the semantic grid mapping of mobile robot in the corridor and office. Results show that the semantic map in this paper integrates the object-level semantics well on the basis of occupying the grid map, enriches the map content, and improves the readability of the map. Finally, in the path planning experiment, the semantic grid map can perceive the environment content, provide semantic information for robot navigation and path planning, and support the intelligent service of robot. Compared with the route in the original map, the path after semantic discrimination is more flexible and more suitable for the situation of mobile objects in map building. **Conclusion** Experiments in various environments prove that the method proposed can obtain a semantic grid map with a high consistency with the real environment and labeled target information, and the experimental hardware structure is simple, low cost with good performance, and suitable for the navigation and intelligent robotic path plan.

Key words: intelligent robot; semantic grid map; laser SLAM (simultaneous localization and mapping); object detection; path planning

0 引言

智能化移动机器人广泛应用于工业、军事、物流和家用服务等领域。从复杂的工业型机器人,到简单的扫地机器人,完成同步定位与建图(simultaneous localization and mapping, SLAM)、路径规划和自主导航的能力必不可少。

以市面上常见的低成本扫地机器人为例,采用的方案多是通过2D激光雷达获取其与某一平面内物体间的距离信息,利用激光SLAM建立占据栅格地图以支持机器人自主导航、路径规划等功能(Blanco等,2007)。随着智能化服务需求的日益增加,信息量简单的几何地图已不能满足人们的需求。譬如“清扫椅子那边”、“到冰箱那里”,这些任务需要机器人从几何层面上升到内容层面来感知环境。除通过栅格地图完成对环境的几何轮廓描述外,智能机器人还应具备目标识别(Mousavian等,2019)、语义分割(Chen等,2019)或场景分类(Kumari等,2020)等获取语义信息的功能,语义SLAM不仅了解环境的几何信息,还能感知到环境中的内容信息(Doherty等,2019)。

SLAM是移动机器人在未知环境进行探索、感知和导航的关键技术(Hess等,2016),可分为激光SLAM和视觉SLAM两类。激光SLAM测量精确,便于机器人导航和路径规划,但缺乏语义信息(Li等,2020)。而视觉SLAM的图像能提供丰富的语义信息,特征区分度更高,但其构建的地图不能直接用于路径规划和导航(Sualeh和Kim,2019)。

为了实现移动机器人构建语义地图并在地图上进行路径规划,本文提出一种基于激光—相机系统的语义栅格建图方法,通过包围盒—分割匹配算法将激光雷达数据与相机图像进行物体级别的信息融合,采用激光SLAM建立语义栅格地图,并在此地图上发布语义任务,实现全局路径规划。

本文的主要贡献如下:

1)提出了一种语义栅格建图方法,通过包围盒—分割匹配算法将激光扫描数据精确对应到环境中检测到的各个语义物体,提取语义激光分割数据,并实时同步建立标注物体类别和轮廓的语义栅格地图。

2)在语义栅格地图上进行路径规划,针对建图

中易移动物体对机器人导航和路径规划产生的干扰。提出了一种语义赋权方法进行改善。

3)在移动机器人平台上部署激光SLAM算法Cartographer(Hess等,2016)和目标检测算法YOLOv3(you only look once v3)(Redmon和Farhadi,2018)进行语义建图,并采用A*算法(Šišlak等,2009)进行路径规划。通过多种室内环境下的实验验证本文系统的真实效果。

1 语义栅格建图和路径规划系统设计

1.1 系统框架

为构建可用于路径规划的语义地图,本文以单目相机辅助激光雷达提取物体级别的特征,通过包围盒—分割匹配算法将激光扫描数据精确对应到各个语义物体,得到语义激光分割数据,并同步参与建图。在机器人构建占据栅格地图时,将保存语义信息的栅格称为语义栅格,语义建图更新各物体对应栅格的占据概率和栅格中的语义信息,最后通过全局优化、语义栅格聚类 and 语义栅格标注等步骤,得到标注物体类别和轮廓的语义栅格地图。此外,在语义栅格地图上发布语义任务,定位机器人在地图中的位置,使用A*算法(Šišlak等,2009)进行全局路径规划。对环境中的易移动物体,采用语义赋权算法识别,对路径规划进行改进。

系统结构如图1所示,主要分为语义激光数据提取、语义栅格建图和路径规划3部分。系统输入包括2D激光雷达的扫描数据和单目相机的图像,输出为可用于路径规划的语义栅格地图。

1.2 语义激光数据提取

传统SLAM算法采用的单一激光或视觉传感器不能同时满足获取语义和构建栅格地图的任务。为了构建语义栅格地图,对于环境中的每一个可识别物体,都需要得到其激光数据和匹配的语义信息。

语义激光分割数据提取基于激光—相机系统,结构如图2所示,激光雷达给出了空间中某一高度平面内的扫描数据,模块将处于相机视野内的激光分割投影到图像上并与检测框匹配,得到语义激光分割数据。

为了实现匹配,首先要标定激光和相机的相对位置,得到目标检测的包围盒和激光分割数据,根据位置关系,将它们转换在同一坐标系下进行匹配

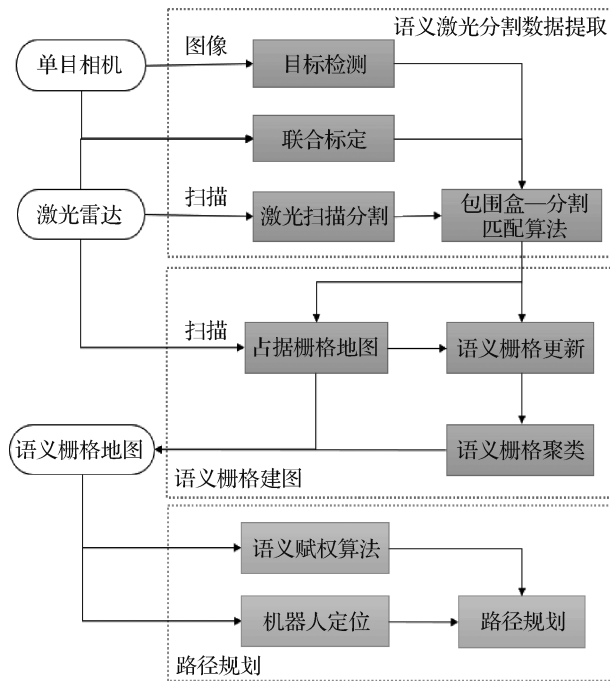


图1 语义栅格建图和路径规划系统

Fig. 1 Semantic grid mapping and path planning system

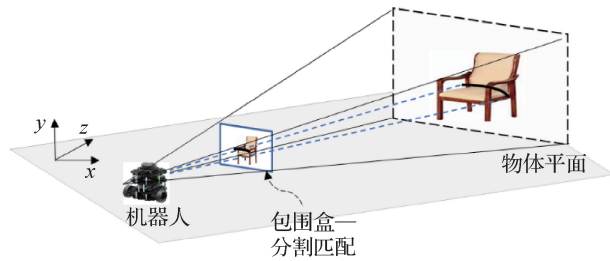


图2 激光雷达—相机系统

Fig. 2 Lidar-camera system

(de Alvis 等, 2019)。

首先,联合标定激光雷达和单目相机获取其相对位置,获得 (K, D) 和 (R, T) 。 (K, D) 的获取采用传统的棋盘格平面标定法(Park 等, 2020), D 是广角镜头的径向畸变参数。像素坐标系下坐标 P_c 、相机坐标系下坐标 P_c 以及激光雷达坐标系下的坐标 P_l 可通过以下变换关系转换到同一坐标系

$$P_c = R \cdot P_l + T \quad (1)$$

$$P = K \cdot P_c \quad (2)$$

式中, K 为相机内参矩阵, R 表示激光雷达坐标系到相机坐标系的相对旋转关系, T 表示两坐标系的相对平移关系。

其次,利用目标检测算法获取可识别物体的成像位置,同时得到物体的语义,即类别和置信度。系

统采用 YOLOv3 (Redmon 和 Farhadi, 2018) 算法,可得到图像中的物体类别标签 l , 置信度 s , 所处图像位置信息 (x, y) 和物体大小信息 (h, w) 。由物体所处的位置信息和大小信息,可得到物体在图像中的边界信息,即物体的包围盒 $Box = \{(x, y), (h, w)\}$, 其类别标签 l 和置信度 s 即为物体的语义信息,检测结果表示为

$$Det = \{(l_1, s_1, Box_1), \dots, (l_n, s_n, Box_n)\} \quad (3)$$

然后,获取激光扫描分割数据。激光分割是将激光测距数据分成若干组相邻的扫描点,这些扫描点将被视为与目标相对应的激光分割的候选点。分割过程参考 Leutenegger 等人(2015)的工作,首先根据欧氏距离将激光扫描点分为若干分组,然后采用迭代端点拟合算法将分组分裂,保留其中具有足够扫描点的分组作为初始分割,利用最小二乘回归计算初始分割的拟合直线,最后将属于同一物体的初始分割合并,合并后可得到某物体精确的激光分割数据,用 Seg_i 表示, i 表示激光扫描时环境中物体的数量。因此,分割结果可表示为

$$Seg = \{Seg_1, Seg_2, \dots, Seg_n\} \quad (4)$$

最后是包围盒—分割匹配,将目标检测的结果 Det 与激光分割 Seg 相匹配,得到目标物体对应的精确激光分割,也就是将激光扫描分割数据精确对应到各个语义物体,得到语义激光分割数据。由于激光雷达和相机的频率不同,需要将获得的激光分割数据和语义信息进行时间同步后,根据算法 1 对它们进行匹配,对于每一个可识别物体,找到其唯一对应的激光分割数据,完成语义信息与激光扫描分割数据的融合。算法分以下几步进行:

1) 对于任一激光扫描分割 Seg_i , 利用已知 K 、 R 和 T 计算其在视觉图像上的投影 Pro_i ;

2) 循环匹配每一个包围盒 Box_i , 当投影包含扫描点数大于最小阈值、投影与原始激光扫描分割包含扫描点数的百分比大于最小阈值时,可找到投影唯一对应的包围盒 Box_k , 计算投影 Pro_i 在物体包围盒 Box_k 上的投影 Pro_i^k 作为物体对应的精确分割;

3) 将激光扫描分割的位置信息与语义信息封装得到语义激光分割数据 Sl_s 。

算法 1 包围盒—分割匹配算法

输入: 分割 Seg , 目标检测 Det , 联合标定 K , R, T ;

输出: 语义激光扫描分割数据 Sl_s ;

- 1) 初始化 Sls , 给定参数 $min_points, thread$;
- 2) 删除超过相机 FOV (field angle of view) 的 Seg ;
- 3) for 每个 Seg_i do
- 4) 初始化在图像上的投影 Pro_i ;
- 5) for 在 Seg_i 的每个 P_{lj} do
- 6) $P_k = K(R \cdot P_{lj} + T)$;
- 7) if P_k 在图像中 then
- 8) $Pro_i += P_k$;
- 9) end if
- 10) end for
- 11) for 在 Det 里的每个 (l_k, s_k, Box_k)
- 12) 得到在 Box_k 中的 Pro_i 部分: Pro_i^k
- 13) if $|Pro_i^k| < min_points$ && $\frac{|Pro_i^k|}{|Pro_i|} > thread$ then
- 14) 计算标定标引 $RayID$;
- 15) $Sls += (l_k, s_k, RayID)$;
- 16) end if
- 17) end for
- 18) end for
- 19) return Sls

1.3 语义栅格建图

获取语义激光分割数据时要同步进行语义栅格建图,系统在激光数据构建栅格地图时,根据语义信息对栅格中对应物体做语义和轮廓的标记。可分栅格地图构建、栅格语义更新以及语义栅格聚类3步。

栅格地图用栅格集合表示环境某一平面的几何结构,传统的栅格仅保存占据概率。本文在赋予栅格初始占据概率的同时,赋予初始物体类别(暂无语义)构成语义栅格。

语义栅格会更新栅格获得的语义信息,累计获取栅格的可信语义。若地图原始语义栅格中无待融合栅格的物体语义信息,则将该物体信息直接赋予该栅格;若栅格中为相同物体的语义信息,则对其标签概率求平均值并更新其置信度。

本文采用基于密度的 DBSCAN (density-based spatial clustering of applications with noise) 聚类算法 (Tran 等, 2013), 判定零散语义栅格或处于某一簇中的少量不同类别栅格为噪声, 将相同物体类别的栅格进行聚类, 得到代表各独立物体的栅格集合。最后, 按照栅格集中的物体语义信息, 对占据栅格

地图的对应物体位置做不同颜色和文字的标记, 得到语义栅格地图。

1.4 路径规划

语义栅格地图含有环境的几何信息和内容信息, 能为机器人路径规划提供具体的语义物体作为目标, 辅助机器人进行导航。在地图已知的情况下, 机器人需要在地图中定位自己的位置, 同时发布目标位置来做全局路径规划, 绕过地图中障碍物到达目的地。

利用自适应蒙特卡罗定位 (adaptive Monte Carlo localization, AMCL) 对移动机器人在 2 维环境下进行概率定位 (Hanten 等, 2017), 确定一个机器人的位姿信息。全局路径规划采用 A* 算法, 并根据语义赋权算法对其进行优化。A* 算法的估算函数为

$$F(n) = G(n) + H(n) \quad (5)$$

式中, $G(n)$ 是从起点出发到达任一顶点 n 的实际距离, $H(n)$ 是任意顶点 n 到目标位置点的估算距离, 算法计算从起点到目标位置的路径, 使代价 $F(n)$ 最小。

系统结构如图 3 所示。

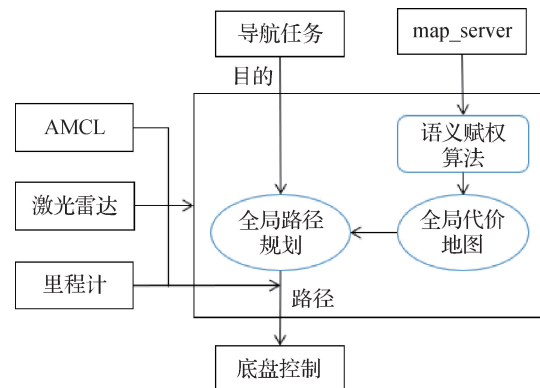


图3 路径规划系统结构

Fig. 3 Path planning system structure

SLAM 建立的地图刻画的是实际环境中某时刻的状态 (Xiao 等, 2019), 当其他时刻环境中的物体位置发生变化, 如人走动或雨伞被拿走, SLAM 地图会产生与实际环境不一致的问题, 其定位往往会失效, 对机器人导航产生干扰。尤其是建图在含有较多易移动物体的环境中, 机器人导航很容易不准确。

为降低环境中易移动物体对路径规划的干扰, 本文提出了语义赋权算法, 主要对语义栅格地图中的语义物体进行判别, 将背包、雨伞和人等视为环境中的易移动物体, 即建图时存在但实际环境中可能不在原地的物体, 通过改变其在全局代价地图中的

权重,在全局路径规划时可以得到通过易移动物体的路线。实际运行过程中,若易移动物体仍在原地,则机器人会重新规划路线绕行。

首先,语义赋权算法会对栅格地图中的语义物体进行判别,确认其是否为易移动物体。 α 表示物体还在原地的概率, D_{cost} 表示绕开该物体的路径代价, P_{cost} 表示物体不在原地可直接通过的代价,则路径期望代价为

$$E_{\text{cost}} = D_{\text{cost}} \cdot \alpha + P_{\text{cost}} \cdot (1 - \alpha) \quad (6)$$

若期望代价与绕开该物体代价的差值大于实验设定的阈值 β ,则将该物体判断为易移动物体,修改其在全局代价地图中的权重值,将其看做空地,否则仍视为障碍物。

全局代价地图 *costmap* 通过 *map_server* 获取语义栅格地图中的静态地图。静态地图中每个栅格 *grid* 存放地图的障碍物信息,有1或0两种取值,分别表示该栅格状态为障碍物或空闲。判别后的易移动物体需将其在静态地图中的状态由障碍改为空闲。具体过程如算法2所示。

算法2 语义赋权算法

输入:语义栅格数据 *grid*,物体代价表 O_{cost} ;

输出:全局代价地图 *costmap*;

- 1) 初始化 *grid*、 O_{cost} ,给定参数 β ;
- 2) for 每个 *grid_i*
- 3) if $|D_{\text{cost}} - E_{\text{cost}}| > \beta$ then
- 4) 从 *costmap* 中得到 *static_map*
- 5) 令 *grid_i.cost* = *Free_Space.cost*;
- 6) end if
- 7) end for
- 8) return *costmap*

经语义赋权算法得到的全局代价地图会传给路径规划算法使用,其效果如图4所示。

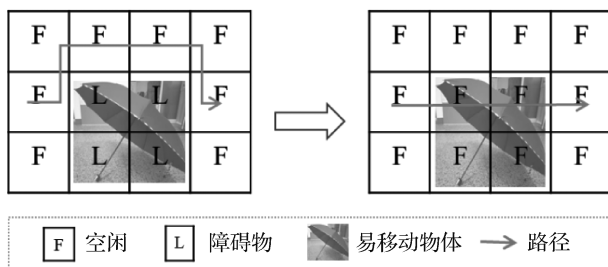


图4 语义赋权算法示意图

Fig. 4 Semantic cost method schematic diagram

2 实验

系统的硬件设施由移动机器人和控制机组成。机器人平台由 Kobuki 底盘、激光雷达 RS-LiDAR、单目相机组成,控制机笔记本配置为主频 2.5 GHz、Intel Core i5-7300HQ 处理器、GTX 1050Ti GPU 和 8 GB 内存。同时部署语义建图和路径规划系统作为软件设施,运行在 ROS (robot operating system) 环境下,测试系统为 Ubuntu 16.04。

2.1 语义激光分割数据提取实验

如图5所示,分别是目标检测、激光扫描分割以及包围盒—匹配的示意图,对应环境中识别的相关物体,将其激光分割数据匹配到物体包围盒上,实现了激光和语义的融合。



图5 包围盒—分割匹配示意图

Fig. 5 Box-seg match schematic diagram

实验对语义激光数据提取的过程进行了性能评估,表1是本文系统在相同条件下对不同物体进行目标检测、激光分割和包围盒—分割匹配的检测精度,其中包围盒—分割匹配是本文算法1的检测结

表1 不同物体的检测精度结果

Table 1 Precision results of various objects

目标物体	目标检测	激光分割	包围盒—分割匹配
椅子	99.2	88.5	95.4
微波炉	98.3	90.5	93.7
冰箱	98.5	84.7	92.2
雨伞	97.6	81.2	89.5
平均精度	98.4	86.2	92.7

果,经过包围盒—分割匹配处理,采用语义激光数据可以提高单一激光雷达分割的精度。

此外,由于雷达和物体距离的增加会导致激光分割数据的失败,从而使包围盒—分割匹配精度降低。实验计算物体上分配到的激光点数量与实际计算中的激光点数量的百分比,由于分割时会删除一些孤立的点,因此该值应当小于 1,物体在进行校准、投影和测量中不可避免地会产生误差。表 2 展示了不同距离下物体激光分割的精度。

表 2 不同距离下物体的分割结果
Table 2 Segmentation results of objects at different distances

距离/m	椅子	冰箱	微波炉	雨伞	/%
1	88.2	96.4	92.5	89.4	
2	85.3	98.3	98.2	87.7	
3	84.5	95.1	96.1	87.9	
4	85.6	95.6	96.5	81.8	
5	82.6	94.2	fail	60.8	
平均精度	85.2	95.9	—	81.5	

注:fail 表示检测失败,“—”表示平均精度无法计算。

2.2 语义栅格建图实验

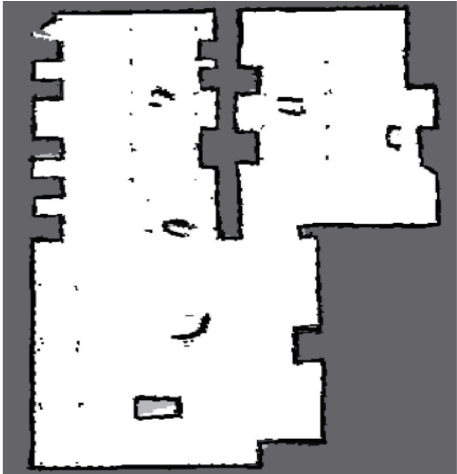
语义栅格建图实现了移动机器人在走廊和办公室内的语义栅格建图。

图 6(a)是在实验室利用单一激光传感器根据传统的激光 SLAM 算法 Cartographer(Hess 等,2016)建立的占据栅格地图,该地图只能得到环境中的几何信息,不能识别环境中的物体。图 6(b)是采用本文系统建图后得到的语义栅格地图,并对其中标注的语义物体和环境中的对应真实物体进行了对比,结果证明本文的语义地图在占据栅格地图的基础上很好地融入了物体级别语义,丰富了地图内容,提高了地图的可读性。

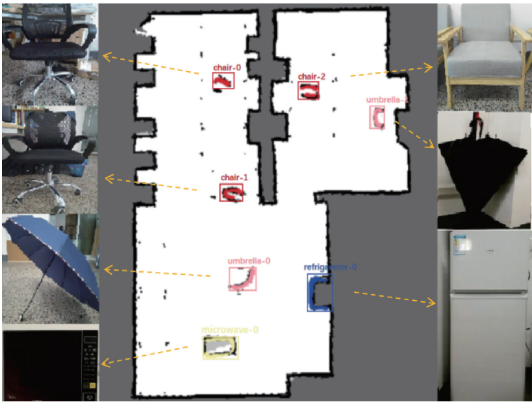
图 7 展示了移动机器人在走廊上进行语义栅格建图的具体过程。移动机器人首先根据激光雷达对环境的扫描数据建立占据栅格地图,同时根据相机获取环境中的物体的语义信息,如图 7 中的微波炉和椅子,并将其标示在地图上,最后可得到含有物体的类别和轮廓的语义栅格地图。

2.3 路径规划实验

经过语义栅格建图实验得到了语义栅格地图,



(a) 原始栅格地图



(b) 语义栅格地图与真实环境物体对比

图 6 实验室语义栅格建图

Fig. 6 Semantic grid map of laboratory

((a) original grid map; (b) comparison of semantic grid map and real environment objects)

根据地图中的语义信息可以实现基于语义的路径规划。如图 8 所示,系统直接发布了“去冰箱处”和“去椅子(chair-2)处”的语义任务,机器人能够避开地图中的障碍物,分别规划到达冰箱和椅子处的一条路径。结果证明本文的语义栅格地图能感知环境内容,为机器人导航和路径规划提供语义信息,支持机器人智能化服务。

针对易移动物体,本文在采用 Cartographer 建立的原始栅格地图和本系统建立的语义栅格地图上进行了对比实验。

如图 9 所示,由于建图时门口有把雨伞,在原始栅格地图中,机器人不能识别雨伞,必须绕行到达目的地。而根据语义赋权算法判别雨伞为易移动物体,实际环境中很可能不在原地,因此为机器人规划了一条穿过原地的路线。但当实际环境中雨伞仍停

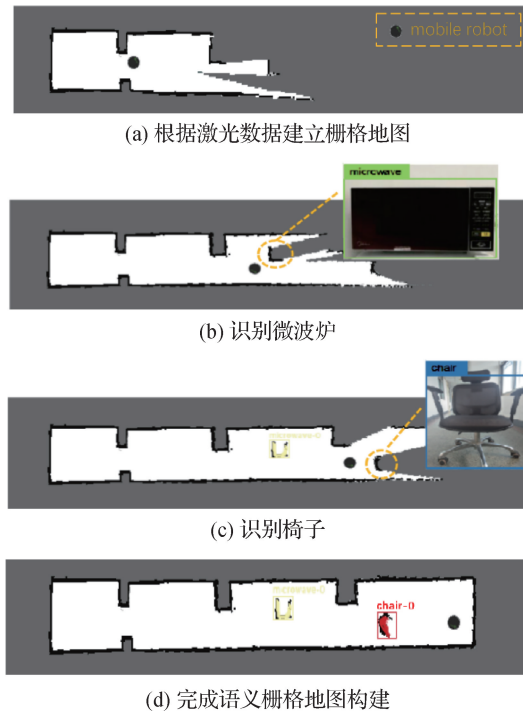


图7 走廊语义建图过程

Fig. 7 Semantic mapping in a corridor

- ((a) build a grid map based on laser data;
 (b) identify the microwave oven; (c) identify the chair;
 (d) complete semantic grid map construction)



图8 语义栅格地图下的路径规划结果

Fig. 8 Path planning results under semantic grid map

留在原地时,机器人可以重新规划一条绕行雨伞从当前位置出发到达目的地的路线。与原始地图中路线相比,经过语义判别后的路径更加灵活,更适应建图存在易移动物体的情况。

此外,实验在仿真环境 gazebo 下搭建了一段 20 m 长的走廊,在其中随机布置了 20 个易移动物体,通过语义栅格建图得到包含了物体语义信息的语义栅格地图。给定机器人运动的起点和终点,在实际环境中每次随机移开一定数目的物体,分别采用原始路径规划 A* 算法与经过语义赋权的路径规划算法

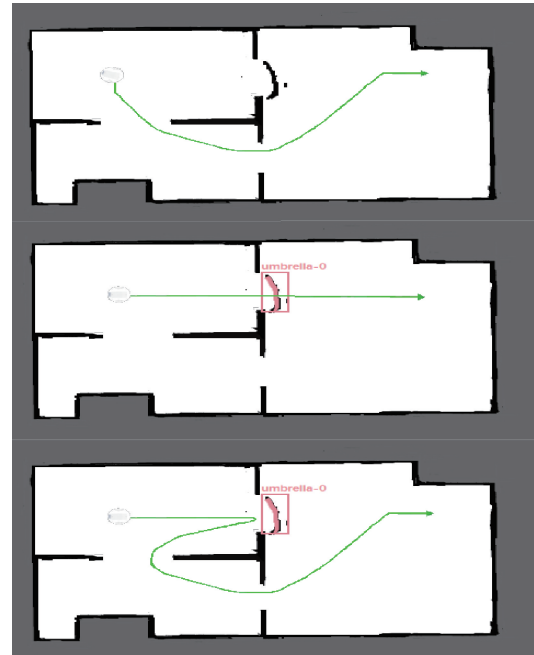


图9 易移动物体的路径规划对比图

Fig. 9 Comparison of path planning with movable objects

进行机器人导航,并计算其路径长度。

表3是多次实验后的路径长度对比,可以看出,未经语义判别的路径规划只会按照已建的栅格地图中的物体进行避障运动,无论实际环境中障碍物是否还存在,经过语义判别后的路径相对灵活,路径长度也较短,更适合在较多易移动物体的环境下使用。

表3 易移动物体的路程对比

Table 3 Comparison of distances for easily moving objects

	移开物体数目					
	3	6	9	12	15	18
A* 算法路程/m	27.2	27.2	27.2	27.2	27.2	27.2
语义赋权算法后路程/m	26.5	25.3	24.6	23.8	22.5	21.2

3 结 论

提出了一种基于激光—相机系统的语义建图方法。通过包围盒—分割匹配和同步建图,将语义激光分割数据融入占据栅格地图中,实现了各物体对应语义激光分割的提取和语义栅格地图的标注,得到与真实环境一致的语义栅格地图。此外,系统在语义栅格地图的基础上,利用全局路径规划和语义赋权算法,给定语义任务,对机器人的运动轨迹进行

规划,并减少语义地图中易移动物体对路径规划的影响。实验证明了机器人在语义栅格地图下的自主定位和路径规划功能。

本文不足之处在于语义栅格建图时场景较小、物体的种类较少,主要原因是实验条件有限,所用室内小型机器人更适合在小场景下运行。因此,下一步的研究方向是尝试把工作转移到中型机器人上,扩展算法的应用条件,在较大场景下建立语义栅格地图,改善大场景下机器人定位易漂移、建图回环效果差的问题。

参考文献 (References)

- Blanco J L, Gonzalez J and Fernandezmadrigal J A. 2007. A new method for robust and efficient occupancy grid-map matching//Proceedings of Iberian Conference on Pattern Recognition & Image Analysis. Heidelberg, Germany: Springer; 194-201 [DOI: 10.1007/978-3-540-72849-8_25]
- Chen X Y L, Milioto A, Palazzolo E, Giguère P, Behley J and Stachniss C. 2019. SuMa ++: efficient LiDAR-based semantic SLAM//Proceedings of 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems. Macau, China: IEEE; 4530-4537 [DOI: 10.1109/IROS40897.2019.8967704]
- de Alvis C, Shan M, Worrall S and Nebot E. 2019. Uncertainty estimation for projecting lidar points onto camera images for moving platforms//Proceedings of 2019 International Conference on Robotics and Automation. Montreal, Canada: IEEE; 6637-6643 [DOI: 10.1109/ICRA.2019.8794424]
- Doherty K, Fourie D and Leonard J. 2019. Multimodal semantic SLAM with probabilistic data association//Proceedings of 2019 International Conference on Robotics and Automation. Montreal, Canada: IEEE; 2419-2425 [DOI: 10.1109/ICRA.2019.8794244]
- Hanten R, Buck S, Otte S and Zell A. 2017. Vector-AMCL: vector based adaptive Monte Carlo localization for indoor maps//Proceedings of Intelligent Autonomous Systems 14. Shanghai, China: Springer; 403-416 [DOI: 10.1007/978-3-319-48036-7_29]
- Hess W, Kohler D, Rapp H and Andor D. 2016. Real-time loop closure in 2D LIDAR SLAM//Proceedings of 2016 IEEE International Conference on Robotics and Automation. Stockholm, Sweden: IEEE; 1271-1278 [DOI: 10.1109/ICRA.2016.7487258]
- Kumari S, Jha R R, Bhavsar A and Nigam A. 2020. Indoor-outdoor scene classification with residual convolutional neural network//Proceedings of the 3rd International Conference on Computer Vision and Image Processing. Singapore, Singapore: Springer; 325-337 [DOI: 10.1007/978-981-32-9291-8_26]
- Leutenegger S, Lynen S, Bosse M, Siegwart R and Furgale P. 2015. Keyframe-based visual-inertial odometry using nonlinear optimization. International Journal of Robotics Research, 34(3): 314-334 [DOI: 10.1177/0278364914554813]
- Li B, Wang Y Q, Zhang Y, Zhao W J, Ruan J Y and Li P. 2020. GP-SLAM: laser-based SLAM approach based on regionalized Gaussian process map reconstruction. Autonomous Robots, 44: 947-967 [DOI: 10.1007/s10514-020-09906-z]
- Mousavian A, Toshev A, Fiser M, Kosecka J, Wahid A and Davidson J. 2019. Visual representations for semantic target driven navigation [EB/OL]. [2020-07-16]. <https://arxiv.org/pdf/1805.06066.pdf>
- Park C, Moghadam P, Kim S, Sridharan S and Fookes C. 2020. Spatio-temporal camera-LiDAR calibration: a targetless and structureless approach. IEEE Robotics and Automation Letters, 5(2): 1556-1563 [DOI: 10.1109/LRA.2020.2969164]
- Redmon J and Farhadi A. 2018. YOLOv3: an incremental improvement [EB/OL]. [2020-07-16]. <https://arxiv.org/pdf/1804.02767.pdf>
- Šišlák D, Volf P and Pěchouček M. 2009. Accelerated A* path planning//Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems. Budapest, Hungary: [s. n.]; 1133-1134
- Sualeh M and Kim G W. 2019. Simultaneous localization and mapping in the epoch of semantics: a survey. International Journal of Control, Automation and Systems, 17(3): 729-742 [DOI: 10.1007/s12555-018-0130-x]
- Tran T N, Drab K and Daszykowski M. 2013. Revised DBSCAN algorithm to cluster data with dense adjacent clusters. Chemometrics and Intelligent Laboratory Systems, 120: 92-96 [DOI: 10.1016/j.chemolab.2012.11.006]
- Xiao L H, Wang J G, Qiu X S, Rong Z and Zou X D. 2019. Dynamic-SLAM: semantic monocular visual localization and mapping based on deep learning in dynamic environment. Robotics and Autonomous Systems, 117: 1-16 [DOI: 10.1016/j.robot.2019.03.012]

作者简介



丁梦远,1995年生,女,硕士研究生,主要研究方向为智能机器人和语义SLAM。

E-mail:2712202808@qq.com



郭迟,通信作者,男,教授,主要研究方向为北斗应用、无人系统导航及位置服务。

E-mail:guochi@whu.edu.cn

黄凯,男,硕士研究生,主要研究方向为语义地图和激光SLAM。E-mail:1278943938@qq.com