



(12)发明专利申请

(10)申请公布号 CN 111578940 A

(43)申请公布日 2020.08.25

(21)申请号 202010330835.9

(22)申请日 2020.04.24

(71)申请人 哈尔滨工业大学

地址 150001 黑龙江省哈尔滨市南岗区西
大直街92号

(72)发明人 高会军 李湛 薛喜地 孙维超
杨学博 于兴虎

(74)专利代理机构 北京高沃律师事务所 11569
代理人 刘凤玲

(51)Int.Cl.

G01C 21/20(2006.01)

G01C 21/00(2006.01)

G01S 17/86(2020.01)

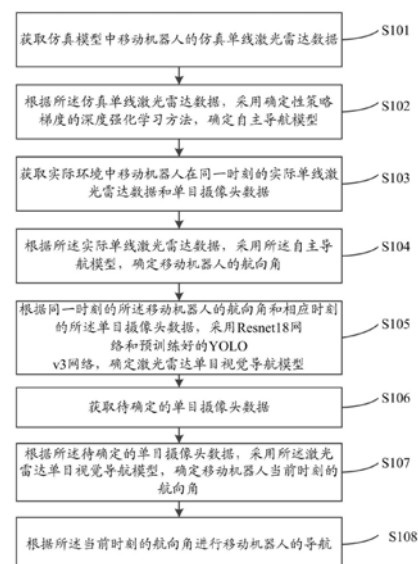
权利要求书2页 说明书8页 附图3页

(54)发明名称

一种基于跨传感器迁移学习的室内单目导航方法及系统

(57)摘要

本发明涉及一种基于跨传感器迁移学习的室内单目导航方法及系统。所述方法包括：根据仿真单线激光雷达数据，采用确定性策略梯度的深度强化学习方法，确定自主导航模型；获取实际环境中移动机器人在同一时刻的实际单线激光雷达数据和单目摄像头数据；根据实际单线激光雷达数据，采用所述自主导航模型，确定移动机器人的航向角；根据同一时刻的移动机器人的航向角和相应时刻的单目摄像头数据，采用Resnet18网络和预训练好的YOLO v3网络，确定激光雷达单目视觉导航模型；根据待确定的单目摄像头数据，采用所述激光雷达单目视觉导航模型，确定移动机器人当前时刻的航向角；根据当前时刻的航向角进行移动机器人的导航。本发明提高了搭载单目摄像头的移动机器人导航的准确性。



1. 一种基于跨传感器迁移学习的室内单目导航方法,其特征在于,包括:

获取仿真模型中移动机器人的仿真单线激光雷达数据;所述仿真模型在Webots开源仿真环境里搭建;所述移动机器人的型号为大疆Mavic Pro,所述移动机器人上方安装LMS 291单线激光雷达;所述仿真单线激光雷达数据为深度值;

根据所述仿真单线激光雷达数据,采用确定性策略梯度的深度强化学习方法,确定自主导航模型;所述自主导航模型以所述激光雷达数据为输入,以所述移动机器人的航向角为输出;

获取实际环境中移动机器人在同一时刻的实际单线激光雷达数据和单目摄像头数据;

根据所述实际单线激光雷达数据,采用所述自主导航模型,确定移动机器人的航向角;

根据同一时刻的所述移动机器人的航向角和相应时刻的所述单目摄像头数据,采用Resnet18网络和预训练好的YOLO v3网络,确定激光雷达单目视觉导航模型;所述激光雷达单目视觉导航模型以所述单目摄像头数据为输入,以对应时刻的航向角为输出;

获取待确定的单目摄像头数据;

根据所述待确定的单目摄像头数据,采用所述激光雷达单目视觉导航模型,确定移动机器人当前时刻的航向角;

根据所述当前时刻的航向角进行移动机器人的导航。

2. 根据权利要求1所述的一种基于跨传感器迁移学习的室内单目导航方法,其特征在于,所述根据所述仿真单线激光雷达数据,采用确定性策略梯度的深度强化学习方法,确定自主导航模型,具体包括:

利用 $R = -k \cdot d_{\min}^{-1} + R_p$ 确定回报函数;其中, R 为回报函数, k 为一个常数取值为0.5, d_{\min} 为当前移动机器人的安全裕度, R_p 为当前移动机器人的动态性能激励项,

$$R_p = \begin{cases} -0.3 & \text{if } v \approx 0 \\ +0.4 & \text{if } v \neq 0 \text{ and } w \neq 0 \\ +0.1 & \text{else} \end{cases}; \text{其中, } v \text{ 为线速度, } w \text{ 为角速度。}$$

3. 根据权利要求1所述的一种基于跨传感器迁移学习的室内单目导航方法,其特征在于,所述根据同一时刻的所述移动机器人的航向角和相应时刻的所述单目摄像头数据,采用Resnet18网络和预训练好的YOLO v3网络,确定激光雷达单目视觉导航模型,具体包括:

根据同一时刻的所述移动机器人的航向角对所述单目摄像头数据打标签,得到样本图像。

4. 根据权利要求3所述的一种基于跨传感器迁移学习的室内单目导航方法,其特征在于,所述根据同一时刻的所述移动机器人的航向角和相应时刻的所述单目摄像头数据,采用Resnet18网络和预训练好的YOLO v3网络,确定激光雷达单目视觉导航模型,具体还包括:

获取所述Resnet18网络和预训练好的YOLO v3网络;

将所述样本图像分别输入所述Resnet18网络和预训练好的YOLO v3网络;

将所述预训练好的YOLO v3网络输出的行人信息向量合并到所述Resnet18网络的全连接层,得到激光雷达单目视觉导航模型。

5. 一种基于跨传感器迁移学习的室内单目导航系统,其特征在于,包括:

第一获取模块,用于获取仿真模型中移动机器人的仿真单线激光雷达数据;所述仿真模型在Webots开源仿真环境里搭建;所述移动机器人的型号为大疆Mavic Pro,所述移动机器人上方安装LMS 291单线激光雷达;所述仿真单线激光雷达数据为深度值;

自主导航模型确定模块,用于根据所述仿真单线激光雷达数据,采用确定性策略梯度的深度强化学习方法,确定自主导航模型;所述自主导航模型以所述激光雷达数据为输入,以所述移动机器人的航向角为输出;

第二获取模块,用于获取实际环境中移动机器人在同一时刻的实际单线激光雷达数据和单目摄像头数据;

移动机器人的航向角确定模块,用于根据所述实际单线激光雷达数据,采用所述自主导航模型,确定移动机器人的航向角;

激光雷达单目视觉导航模型确定模块,用于根据同一时刻的所述移动机器人的航向角和相应时刻的所述单目摄像头数据,采用Resnet18网络和预训练好的YOLO v3网络,确定激光雷达单目视觉导航模型;所述激光雷达单目视觉导航模型以所述单目摄像头数据为输入,以对应时刻的航向角为输出;

第三获取模块,用于获取待确定的单目摄像头数据;

移动机器人当前时刻的航向角确定模块,用于根据所述待确定的单目摄像头数据,采用所述激光雷达单目视觉导航模型,确定移动机器人当前时刻的航向角;

导航模块,用于根据所述当前时刻的航向角进行移动机器人的导航。

6. 根据权利要求5所述的一种基于跨传感器迁移学习的室内单目导航系统,其特征在于,自主导航模型确定模块具体包括:

回报函数确定单元,用于利用 $R = -k \cdot d_{\min}^{-1} + R_p$ 确定回报函数;其中,R为回报函数,k为一个常数取值为0.5, d_{\min} 为当前移动机器人的安全裕度, R_p 为当前移动机器人的动态性能激励项, $R_p = \begin{cases} -0.3 & \text{if } v \approx 0 \\ +0.4 & \text{if } v \neq 0 \text{ and } w \neq 0 \\ +0.1 & \text{else} \end{cases}$;其中,v为线速度,w为角速度。

7. 根据权利要求5所述的一种基于跨传感器迁移学习的室内单目导航系统,其特征在于,所述激光雷达单目视觉导航模型确定模块具体包括:

样本图像确定单元,用于根据同一时刻的所述移动机器人的航向角对所述单目摄像头数据打标签,得到样本图像。

8. 根据权利要求7所述的一种基于跨传感器迁移学习的室内单目导航方法,其特征在于,所述激光雷达单目视觉导航模型确定模块具体还包括:

第一获取单元,用于获取所述Resnet18网络和预训练好的YOLO v3网络;

激光雷达单目视觉导航模型确定单元,用于将所述样本图像分别输入所述Resnet18网络和预训练好的YOLO v3网络,并将所述预训练好的YOLO v3网络输出的行人信息向量合并到所述Resnet18网络的全连接层,得到激光雷达单目视觉导航模型。

一种基于跨传感器迁移学习的室内单目导航方法及系统

技术领域

[0001] 本发明涉及室内视觉导航领域,特别是涉及一种基于跨传感器迁移学习的室内单目导航方法及系统。

背景技术

[0002] 无人机、无人车等可移动机器人能够为人类社会带来极大的便利,以无人机为例,从最初的实验性试飞到民用航拍,再到近些年来的自主导航甚至分布式定位及建图,都极大提高了人们的工作效率。以小型无人机、无人车为代表的移动机器人体积小、机动性高等特点使其在室内及相似复杂环境中能够更加方便地执行任务。因此,如果移动机器人能够在复杂陌生室内环境中自主智能导航,将极大的扩展其应用范围。然而,移动机器人的室内自主导航算法的研发具有很大的挑战性,自从自主导航算法诞生以来,其在环境结构不确定、光线不稳定,以及室内的行人具有形态差异大、行走方向随意性高等场景下自适应稳定导航一直是无人机、无人车等移动机器人自主导航领域中的难点之一。因此研究室内环境下的自主导航方法具有重大意义。

[0003] 自然界的生物经过长期进化,环境的多变性反复强化了生物对环境的适应能力。生物科学研究表明,生物在适应一个环境时,是通过不断与环境进行交互,并且获得环境的反馈来不断强化生物的某项能力。例如动物的行走、捕猎、群体运动等等,生物都是在不断地试错的过程中来不断改善和提升自身的决策策略。该类型学习过程亦属于演化进化的范畴,其特点是没有对环境进行精确建模,仅仅是通过不断交互,来获取环境对自身的奖励或者惩罚来不断改进自身的策略。以无人机为例,对于其自主导航,亦可以用这种类型的学习方式来得在未知环境中的自主导航能力。在此过程中,将无人机类比作生物,无人机不断试错学习的过程类比作生物的学习过程,生物能够获取到自然界反馈的回报,因此对于无人机在自主学习导航的过程中,可以人为的设定对于无人机学习自主导航有利的奖励和惩罚来激励无人机更快适应这个环境。这也是强化学习算法的基本思想,可以看出,强化学习属于端到端的学习类型,输入一个环境状态,直接输出一个决策动作。

[0004] 目前,根据路径规划是否基于轨迹法,可将移动机器人导航算法分为两大类,分别是基于轨迹规划的方法和基于动作响应的(端到端决策)方法。对于基于轨迹规划的方法,其优点在于通过动态规划,可以求解出一条全局最优路径,其结果更加可控,更加稳定,但是该类方法带来的弊端是算法耗时长,比较依赖计算机的算力,而且往往基于轨迹规划的算法的泛化能力会相对比较弱一些。对于基于动作响应的方法,此类方法的优点在于其为端到端决策类型的算法,通常算法决策耗时比较短,同时目前基于端到端的算法的泛化能力普遍要优于基于轨迹的方法,但是基于端到端的方法的短板也非常明显,那就是其决策并不一定是全局最优解。

[0005] 基于动作响应的导航算法又可分为自主学习类和非自主学习两类,非自主学习类的算法诸如人工势场法等算法,但由于当移动机器人距离障碍物过近或者离目标位置太远将会导致局部震荡等一系列问题,虽然后续有改进版本的人工势场法,但也使得其参数变

得很难整定,实际应用起来很难。对于自主学习类的移动机器人导航算法,目前主要是基于深度强化学习和深度学习的方法。目前应用深度强化学习进行室内导航较为成功的案例有诸如基于状态值估计的Generalization through Simulation Integrating Simulated and Real Data into Deep Reinforcement Learning for VisionBasedAutonomous Flight (GTS),该方法为移动机器人室内导航的端到端决策方法。

[0006] 然而对于该深度强化学习(GTS)的导航算法,首先,其利用仿真环境训练出来的状态值估计网络,其对动作空间里所有的动作采取的是利用状态值估计网络逐个评估的方法,这意味着在决策时会加大响应延迟;其次,其采用离散动作空间,这意味着决策动作也将变得离散,动作不是那么平滑;最后,该方法(GTS)将仿真环境训练好的网络移植到实物移动机器人上,由于网络模型过于简单,在实物效果上表现为导航裕度不够、路径不稳定。

[0007] 目前基于单目视觉的端到端移动机器人导航算法几乎都是在无人的环境下进行飞行试验,因为单目摄像头难以求取深度信息,而且,行人运动的特点是随机性比较强,同时人们喜欢聚焦、围观,这使得该类算法的实用性大大下降,甚至无法使用。

发明内容

[0008] 本发明的目的是提供一种基于跨传感器迁移学习的室内单目导航方法及系统,提高搭载单目摄像头的移动机器人导航的准确性。

[0009] 为实现上述目的,本发明提供了如下方案:

[0010] 一种基于跨传感器迁移学习的室内单目导航方法,包括:

[0011] 获取仿真模型中移动机器人的仿真单线激光雷达数据;所述仿真模型在Webots开源仿真环境里搭建;所述移动机器人的型号为大疆Mavic Pro,所述移动机器人上方安装LMS 291单线激光雷达;所述仿真单线激光雷达数据为深度值;

[0012] 根据所述仿真单线激光雷达数据,采用确定性策略梯度的深度强化学习方法,确定自主导航模型;所述自主导航模型以所述激光雷达数据为输入,以所述移动机器人的航向角为输出;

[0013] 获取实际环境中移动机器人在同一时刻的实际单线激光雷达数据和单目摄像头数据;

[0014] 根据所述实际单线激光雷达数据,采用所述自主导航模型,确定移动机器人的航向角;

[0015] 根据同一时刻的所述移动机器人的航向角和相应时刻的所述单目摄像头数据,采用Resnet18网络和预训练好的YOLO v3网络,确定激光雷达单目视觉导航模型;所述激光雷达单目视觉导航模型以所述单目摄像头数据为输入,以对应时刻的航向角为输出;

[0016] 获取待确定的单目摄像头数据;

[0017] 根据所述待确定的单目摄像头数据,采用所述激光雷达单目视觉导航模型,确定移动机器人当前时刻的航向角;

[0018] 根据所述当前时刻的航向角进行移动机器人的导航。

[0019] 可选的,所述根据所述仿真单线激光雷达数据,采用深度确定性策略梯度的学习算法,确定自主导航模型,具体包括:

[0020] 利用 $R = -k \cdot d_{\min}^{-1} + R_p$ 确定回报函数;其中,R为回报函数,k为一个常数取值为0.5,

d_{\min} 为当前移动机器人的安全裕度, R_p 为当前移动机器人的动态性能激励项,

$$R_p = \begin{cases} -0.3 & \text{if } v \approx 0 \\ +0.4 & \text{if } v \neq 0 \text{ and } w \neq 0; \text{其中, } v \text{ 为线速度, } w \text{ 为角速度。} \\ +0.1 & \text{else} \end{cases}$$

[0021] 可选的,所述根据同一时刻的所述移动机器人的航向角和相应时刻的所述单目摄像头数据,采用Resnet18网络和预训练好的YOLO v3网络,确定激光雷达单目视觉导航模型,具体包括:

[0022] 根据同一时刻的所述移动机器人的航向角对所述单目摄像头数据打标签,得到样本图像。

[0023] 可选的,所述根据同一时刻的所述移动机器人的航向角和相应时刻的所述单目摄像头数据,采用Resnet18网络和预训练好的YOLO v3网络,确定激光雷达单目视觉导航模型,具体还包括:

[0024] 获取所述Resnet18网络和预训练好的YOLO v3网络;

[0025] 将所述样本图像分别输入所述Resnet18网络和预训练好的YOLO v3网络;

[0026] 将所述预训练好的YOLO v3网络输出的行人信息向量合并到所述Resnet18网络的全连接层,得到激光雷达单目视觉导航模型。

[0027] 一种基于跨传感器迁移学习的室内单目导航系统,包括:

[0028] 第一获取模块,用于获取仿真模型中移动机器人的仿真单线激光雷达数据;所述仿真模型在Webots开源仿真环境里搭建;所述移动机器人的型号为大疆Mavic Pro,所述移动机器人上方安装LMS 291单线激光雷达;所述仿真单线激光雷达数据为深度值;

[0029] 自主导航模型确定模块,用于根据所述仿真单线激光雷达数据,采用确定性策略梯度的深度强化学习方法,确定自主导航模型;所述自主导航模型以所述激光雷达数据为输入,以所述移动机器人的航向角为输出;

[0030] 第二获取模块,用于获取实际环境中移动机器人在同一时刻的实际单线激光雷达数据和单目摄像头数据;

[0031] 移动机器人的航向角确定模块,用于根据所述实际单线激光雷达数据,采用所述自主导航模型,确定移动机器人的航向角;

[0032] 激光雷达单目视觉导航模型确定模块,用于根据同一时刻的所述移动机器人的航向角和相应时刻的所述单目摄像头数据,采用Resnet18网络和预训练好的YOLO v3网络,确定激光雷达单目视觉导航模型;所述激光雷达单目视觉导航模型以所述单目摄像头数据为输入,以对应时刻的航向角为输出;

[0033] 第三获取模块,用于获取待确定的单目摄像头数据;

[0034] 移动机器人当前时刻的航向角确定模块,用于根据所述待确定的单目摄像头数据,采用所述激光雷达单目视觉导航模型,确定移动机器人当前时刻的航向角;

[0035] 导航模块,用于根据所述当前时刻的航向角进行移动机器人的导航。

[0036] 可选的,自主导航模型确定模块具体包括:

[0037] 回报函数确定单元,用于利用 $R = -k \cdot d_{\min}^{-1} + R_p$ 确定回报函数;其中, R 为回报函数, k 为一个常数取值为0.5, d_{\min} 为当前移动机器人的安全裕度, R_p 为当前移动机器人的动态性

能激励项, $R_p = \begin{cases} -0.3 & \text{if } v \approx 0 \\ +0.4 & \text{if } v \neq 0 \text{ and } w \neq 0 \\ +0.1 & \text{else} \end{cases}$; 其中, v 为线速度, w 为角速度。

[0038] 可选的, 所述激光雷达单目视觉导航模型确定模块具体包括:

[0039] 样本图像确定单元, 用于根据同一时刻的所述移动机器人的航向角对所述单目摄像头数据打标签, 得到样本图像。

[0040] 可选的, 所述激光雷达单目视觉导航模型确定模块具体还包括:

[0041] 第一获取单元, 用于获取所述Resnet18网络和预训练好的YOLO v3网络;

[0042] 激光雷达单目视觉导航模型确定单元, 用于将所述样本图像分别输入所述Resnet18网络和预训练好的YOLO v3网络, 并将所述预训练好的YOLO v3网络输出的行人信息向量合并到所述Resnet18网络的全连接层, 得到激光雷达单目视觉导航模型。

[0043] 根据本发明提供的具体实施例, 本发明公开了以下技术效果:

[0044] 本发明所提供的一种基于跨传感器迁移学习的室内单目导航方法及系统, 通过在仿真环境中使用虚拟单线激光雷达作为传感器, 通过基于确定性策略梯度 (DDPG) 的深度强化学习方法, 得到一个稳定的自主导航模型, 然后将单线激光雷达与单目摄像头采集的现实环境数据逐帧绑定, 利用自主导航模型, 对现实环境数据进行实时标注, 并采用标注后的数据集进一步训练得到激光雷达单目视觉导航模型, 从而实现从只使用激光雷达的仿真环境向只使用单目摄像头的现实环境的跨传感器迁移学习。进而通过只单目摄像头数据获取搭载单目摄像头的移动机器人的导航角, 进一步提高搭载单目摄像头的移动机器人导航的准确性。

附图说明

[0045] 为了更清楚地说明本发明实施例或现有技术中的技术方案, 下面将对实施例中所需要使用的附图作简单地介绍, 显而易见地, 下面描述中的附图仅仅是本发明的一些实施例, 对于本领域普通技术人员来讲, 在不付出创造性劳动性的前提下, 还可以根据这些附图获得其他的附图。

[0046] 图1为本发明所提供的一种基于跨传感器迁移学习的室内单目导航方法流程示意图;

[0047] 图2为单线激光雷达数据示意图;

[0048] 图3为自主导航模型结构示意图;

[0049] 图4为激光雷达单目视觉导航模型结构示意图;

[0050] 图5为本发明所提供的一种基于跨传感器迁移学习的室内单目导航系统结构示意图。

具体实施方式

[0051] 下面将结合本发明实施例中的附图, 对本发明实施例中的技术方案进行清楚、完整地描述, 显然, 所描述的实施例仅仅是本发明一部分实施例, 而不是全部的实施例。基于本发明中的实施例, 本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例, 都属于本发明保护的范围。

[0052] 本发明的目的是提供一种基于跨传感器迁移学习的室内单目导航方法及系统,提高搭载单目摄像头的移动机器人导航的准确性。

[0053] 为使本发明的上述目的、特征和优点能够更加明显易懂,下面结合附图和具体实施方式对本发明作进一步详细的说明。

[0054] 图1为本发明所提供的一种基于跨传感器迁移学习的室内单目导航方法流程示意图,如图1所示,本发明所提供的一种基于跨传感器迁移学习的室内单目导航方法,包括:

[0055] S101,获取仿真模型中移动机器人的仿真单线激光雷达数据;所述仿真模型在Webots开源仿真环境里搭建;所述移动机器人的型号为大疆Mavic Pro,所述移动机器人上方安装LMS 291单线激光雷达;所述仿真单线激光雷达数据为深度值;单线激光雷达数据如图2所示。

[0056] S102,根据所述仿真单线激光雷达数据,采用确定性策略梯度的深度强化学习方法,确定自主导航模型;所述自主导航模型以所述激光雷达数据为输入,以所述移动机器人的航向角为输出。所述强化学习方法包含状态空间设计、动作空间设计、网络设计,回报函数设计四大因素。

[0057] 1、状态空间设计:状态空间即强化学习的输入状态(state)。激光雷达数据如图2中绿色射线所示,激光雷达在 $0\sim 180^\circ$ 范围内发射出180条射线,并返回射线的长度值(深度值)。在这里对激光雷达的 $0\sim 180^\circ$ 范围内的数据进行采样,每隔 9° 采样一次。由此组成一个新的激光雷达的数据数组,长度为20。此时激光雷达数据如式子(1)所示。其中newLidarData即为新的激光雷达数据, $l_1, l_2, \dots, l_{18}, l_{19}, l_{20}$ 分别表示这20个采样后的激光雷达数据。将新的激光雷达数据当作强化学习状态state输入,如式子(1)所示。

[0058] $state = newLidarData = [l_1, l_2, \dots, l_{18}, l_{19}, l_{20}]$ (1)

[0059] 2、动作空间设计:在这里强化学习策略的输出动作空间定义为 $0\sim 180^\circ$ 范围内的航向角度数值,为连续值。

[0060] 3、网络结构:结构示意图如图3所示,“|”后面的数字表示当前层神经元个数。DDPG算法类似对抗生成网络,有两个网络,一个是表演者Actor,一个是评论者Critic,Actor网络输入状态state然后输出动作,Critic输入Actor发出的动作a以及状态state,输出一个评分Q,用来评价Actor输出的动作的好坏。Actor网络的代价函数如式子(2)所示,Critic网络的代价函数如式子(3)所示。式子(2)表示Actor的目标是最大化Critic的评分。式子(3)为经典的时间差分(TD)代价函数,其中r为环境回报, γ 为折扣率,常数,一般为0.99。同时, $Q(s')$ 、 $Q(s)$ 分别表示上一个时刻和下一个时刻的评分Q值。

[0061] $actor_loss = -Q$ (2)

[0062] $critic_loss = r + \gamma \cdot Q(s') - Q(s)$

[0063] 4、回报函数:利用 $R = -k \cdot d_{min}^{-1} + R_p$ 确定回报函数;其中,R为回报函数,k为一个常数取值为0.5, d_{min} 为当前移动机器人的安全裕度, R_p 为当前移动机器人的动态性能激励项,

$$R_p = \begin{cases} -0.3 & \text{if } v \approx 0 \\ +0.4 & \text{if } v \neq 0 \text{ and } w \neq 0 \\ +0.1 & \text{else} \end{cases}; \text{其中, } v \text{ 为线速度, } w \text{ 为角速度。}$$

[0064] d_{min} 为式(1)中所述的20个激光雷达数据中的最小值,表征移动机器人当前的安

全裕度。

[0065] 回报函数包括两大部分:其一,碰撞惩罚部分,当移动机器人发生碰撞时,其直接给予移动机器人一个-100的惩罚,这种形式虽然能够起作用,但是其回报函数数值突变太过明显,非常不利于神经网络的拟合以及策略决策的平滑性。因此将回报函数设计成 $-k \cdot d_{\min}^{-1}$ 形式。 $-k \cdot d_{\min}^{-1}$ 表示移动机器人距离所有障碍物越远,所受的惩罚越小,此项用以惩罚移动机器人的碰撞,同时还能激励移动机器人尽量往道路的正中间飞行,安全裕度更大,同时 $-k \cdot d_{\min}^{-1}$ 的形式使得移动机器人在避障方面的惩罚函数是连续函数,使得神经网络更加易于拟合以及策略决策的平滑性更高;其二,移动机器人动态性能的激励部分,本部分将移动机器人的角速度和线速度激励融合为一个分段函数,同时分段函数值的差异不是很大,因此不会对神经网络的拟合造成太大的难度。

[0066] R_p 的表达式中第一项表示若移动机器人的决策线速度 v 越低,则说明移动机器人极有可能处于静止不动或者原地打转的情况,此时给予惩罚。第二项表示若移动机器人的线速度 v 和角速度 w 均不为0,则说明移动机器人正在按照预期在建筑物内漫游并且存活,则给予一定的奖励。第三项表示当移动机器人线速度 v 不为0时,且偏航角速度 w 为0,则给予0.1的正激励,目的是为了激励移动机器人积极漫游。值得注意的是,一般情况下数值优先级是 $|k \cdot d_{\min}^{-1}| > 0.4 > 0.3 > 0.1$,这也就意味着子任务的优先级为:避障>积极漫游。这个优先级设置保证移动机器人在稳定避障的同时,能够保证以最大的安全裕度在建筑物内积极漫游,这样就极大提升了移动机器人的“存活”时间。

[0067] 在具体的实施例中,计算机配置为GPU 1080Ti;CPU i78700K的情况下训练大约4h后策略趋于稳定。

[0068] S103,获取实际环境中移动机器人在同一时刻的实际单线激光雷达数据和单目摄像头数据。将单线激光雷达与单目摄像头固定在一起,并用固定后的设备进行同一时刻的实际单线激光雷达数据和单目摄像头数据的采集。采集回来的数据集格式为(image, lidarData),其中image和lidarData分别是同一时刻摄像头采集的图片和激光雷达采集的数据,图片大小为640*480,RGB格式。

[0069] S104,根据所述实际单线激光雷达数据,采用所述自主导航模型,确定移动机器人的航向角。

[0070] 具体的确定过程为:

[0071] 利用公式 $\text{direction}_n = \text{mid}(\max(\text{lidarData_blank})_n)$ 确定移动机器人的航向角。

[0072] 式中, $\max(\text{lidarData_blank})_n$ 代表求取 t_n 时刻单线激光雷达覆盖范围内连续空白区域面积最大的区域, $n=1,2,\dots$ 。

[0073] $\text{mid}(\max(\text{lidarData_blank})_n)$ 代表求取区域 $\max(\text{lidarData_blank})_n$ 的中心航向角,将求取出的 $\text{mid}(\max(\text{lidarData_blank})_n)$ 作为 t_n 时刻移动机器人的航向角 direction_n 。

[0074] S105,根据同一时刻的所述移动机器人的航向角和相应时刻的所述单目摄像头数据,采用Resnet18网络和预训练好的YOLO v3网络,确定激光雷达单目视觉导航模型;所述激光雷达单目视觉导航模型以所述单目摄像头数据为输入,以对应时刻的航向角为输出。

[0075] 根据同一时刻的所述移动机器人的航向角对所述单目摄像头数据打标签,得到样本图像,其格式为(image,direction),direction为航向角。

[0076] 获取所述Resnet18网络和预训练好的YOLO v3网络。

[0077] 将所述样本图像分别输入所述Resnet18网络和预训练好的YOLO v3网络。将所述预训练好的YOLO v3网络输出的行人信息向量合并到所述Resnet18网络的全连接层,得到激光雷达单目视觉导航模型,并如图4所示。

[0078] 利用预训练好的YOLO v3网络输出行人信息向量合,其具体为:

[0079] $person_n = (x_{t1}, y_{t1}, x_{b1}, y_{b1}, x_{t2}, y_{t2}, x_{b2}, y_{b2} \dots x_{ti}, y_{ti}, x_{bi}, y_{bi} \dots x_{tI}, y_{tI}, x_{bI}, y_{bI})$ 。

[0080] 式中, $(x_{ti}, y_{ti}, x_{bi}, y_{bi})$ 代表 t_n 时刻图像中第 i 个行人预测框在图像中的左上角顶点坐标和右下角顶点坐标, $i=1, 2, \dots, I$, I 代表 t_n 时刻图像中的行人总个数, $person_n$ 代表 t_n 时刻图像中全部行人预测框在图像中的左上角顶点坐标和右下角顶点坐标组成的集合。

[0081] 激光雷达单目视觉导航模型包括Resnet18网络、训练好的YOLO v3网络、隐藏层1~4和输出层,其中Resnet18网络的神经元个数为7168,训练好的YOLO v3网络输出的行人信息个数为40,隐藏层1、隐藏层2、隐藏层3、隐藏层4和输出层的神经元个数分别为7208、256、128、16和1。

[0082] 激光雷达单目视觉导航模型的隐藏层1~4均采用ReLU激活函数,输出层采用Linear激活函数。

[0083] 激光雷达单目视觉导航模型的代价函数采用均方误差,代价函数的表达式如公式(3)所示:

$$[0084] \quad Loss = \frac{1}{m} \sum_{j=1}^m (y_j - \hat{y}_j)^2。$$

[0085] 式中, $Loss$ 代表代价函数, y_j 代表第 j 张样本图像对应的激光雷达单目视觉导航模型的输出结果, \hat{y}_j 代表第 j 张样本图像对应的航向角, m 代表一个批次的样本图像数量。

[0086] S106,获取待确定的单目摄像头数据。

[0087] S107,根据所述待确定的单目摄像头数据,采用所述激光雷达单目视觉导航模型,确定移动机器人当前时刻的航向角。

[0088] S108,根据所述当前时刻的航向角进行移动机器人的导航。

[0089] 图5为本发明所提供的一种基于跨传感器迁移学习的室内单目导航系统结构示意图,如图5所示,本发明所提供的一种基于跨传感器迁移学习的室内单目导航系统,包括:第一获取模块501、自主导航模型确定模块502、第二获取模块503、移动机器人的航向角确定模块504、激光雷达单目视觉导航模型确定模块505、第三获取模块506、移动机器人当前时刻的航向角确定模块507和导航模块508。

[0090] 第一获取模块501用于获取仿真模型中移动机器人的仿真单线激光雷达数据;所述仿真模型在Webots开源仿真环境里搭建;所述移动机器人的型号为大疆Mavic Pro,所述移动机器人上方安装LMS 291单线激光雷达;所述仿真单线激光雷达数据为深度值;

[0091] 自主导航模型确定模块502用于根据所述仿真单线激光雷达数据,采用确定性策略梯度的深度强化学习方法,确定自主导航模型;所述自主导航模型以所述激光雷达数据为输入,以所述移动机器人的航向角为输出;

[0092] 第二获取模块503用于获取实际环境中移动机器人在同一时刻的实际单线激光雷达数据和单目摄像头数据;

[0093] 移动机器人的航向角确定模块504用于根据所述实际单线激光雷达数据,采用所述自主导航模型,确定移动机器人的航向角;

[0094] 激光雷达单目视觉导航模型确定模块505用于根据同一时刻的所述移动机器人的航向角和相应时刻的所述单目摄像头数据,采用Resnet18网络和预训练好的YOLO v3网络,确定激光雷达单目视觉导航模型;所述激光雷达单目视觉导航模型以所述单目摄像头数据为输入,以对应时刻的航向角为输出;

[0095] 第三获取模块506用于获取待确定的单目摄像头数据;

[0096] 移动机器人当前时刻的航向角确定模块507用于根据所述待确定的单目摄像头数据,采用所述激光雷达单目视觉导航模型,确定移动机器人当前时刻的航向角;

[0097] 导航模块508用于根据所述当前时刻的航向角进行移动机器人的导航。

[0098] 自主导航模型确定模块502具体包括:回报函数确定单元。

[0099] 回报函数确定单元用于利用 $R = -k \cdot d_{\min}^{-1} + R_p$ 确定回报函数;其中,R为回报函数,k为一个常数取值为0.5, d_{\min} 为当前移动机器人的安全裕度, R_p 为当前移动机器人的动态性能激励项, $R_p = \begin{cases} -0.3 & \text{if } v \approx 0 \\ +0.4 & \text{if } v \neq 0 \text{ and } w \neq 0 \\ +0.1 & \text{else} \end{cases}$;其中,v为线速度,w为角速度。

[0100] 所述激光雷达单目视觉导航模型确定模块505具体包括:样本图像确定单元、第一获取单元和激光雷达单目视觉导航模型确定单元。

[0101] 样本图像确定单元用于根据同一时刻的所述移动机器人的航向角对所述单目摄像头数据打标签,得到样本图像。

[0102] 第一获取单元用于获取所述Resnet18网络和预训练好的YOLO v3网络;

[0103] 激光雷达单目视觉导航模型确定单元用于将所述样本图像分别输入所述Resnet18网络和预训练好的YOLO v3网络,并将所述预训练好的YOLO v3网络输出的行人信息向量合并到所述Resnet18网络的全连接层,得到激光雷达单目视觉导航模型。

[0104] 本说明书中各个实施例采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似部分互相参见即可。对于实施例公开的系统而言,由于其与实施例公开的方法相对应,所以描述的比较简单,相关之处参见方法部分说明即可。

[0105] 本文中应用了具体个例对本发明的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本发明的方法及其核心思想;同时,对于本领域的一般技术人员,依据本发明的思想,在具体实施方式及应用范围上均会有改变之处。综上所述,本说明书内容不应理解为对本发明的限制。

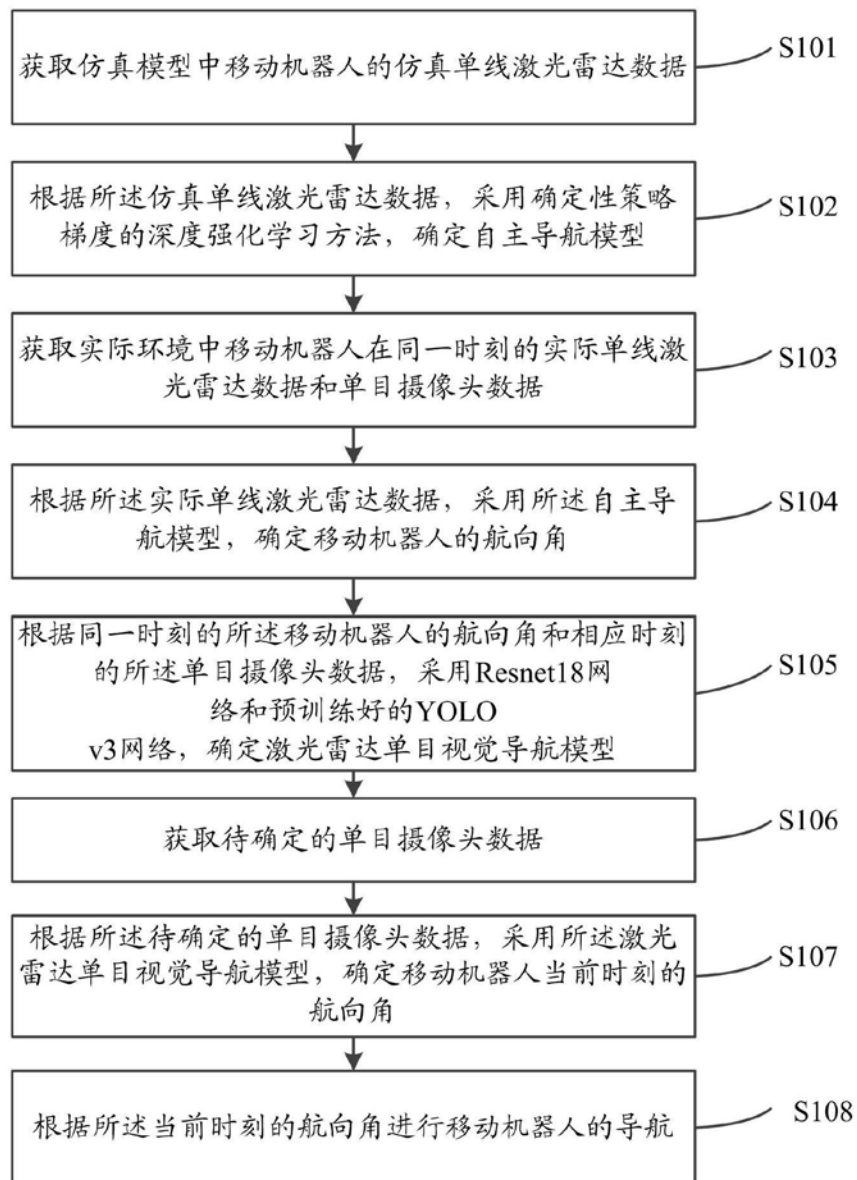


图1

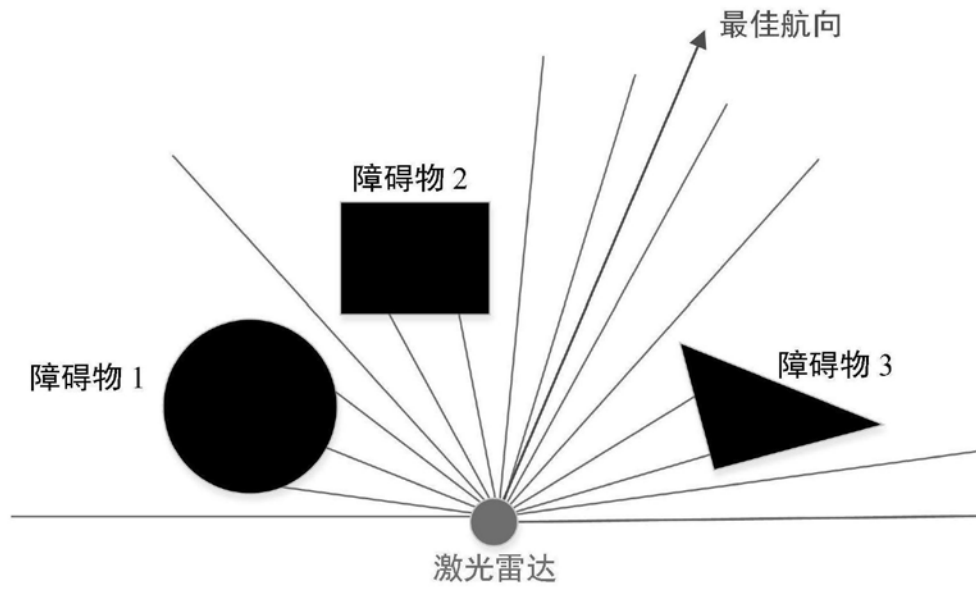


图2

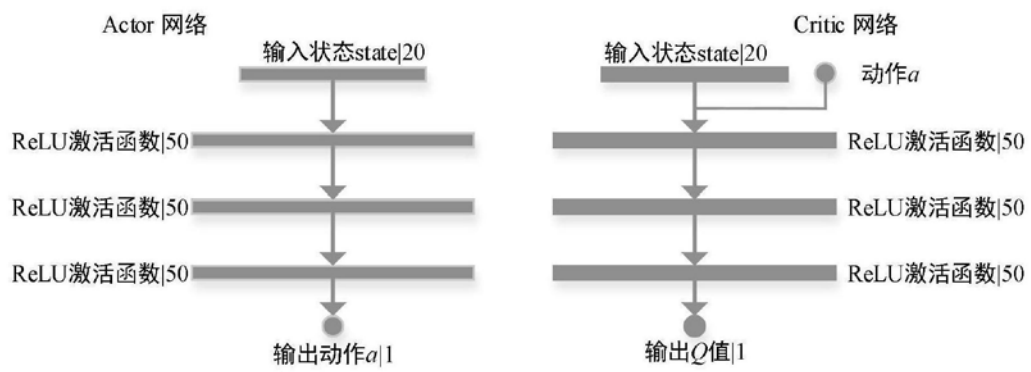


图3

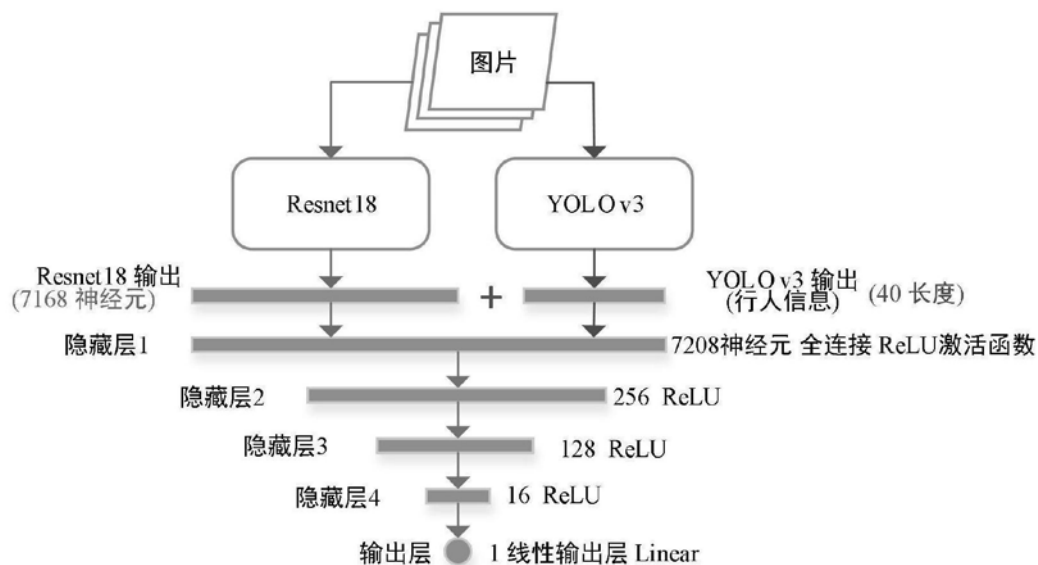


图4



图5