



(12)发明专利申请

(10)申请公布号 CN 111413970 A

(43)申请公布日 2020.07.14

(21)申请号 202010193177.3

(22)申请日 2020.03.18

(71)申请人 天津大学

地址 300072 天津市南开区卫津路92号

(72)发明人 刘凯燕 金杰 何政蕊 郭如意

姜诗祺 陈志华 周梦伊

(74)专利代理机构 天津市北洋有限责任专利代

理事务所 12201

代理人 刘子文

(51)Int.Cl.

G05D 1/02(2020.01)

G01C 21/20(2006.01)

G01S 5/02(2010.01)

G01S 5/14(2006.01)

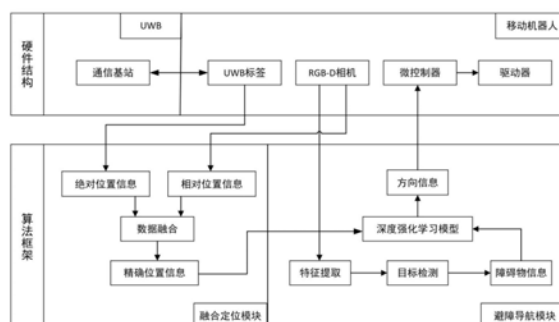
权利要求书1页 说明书5页 附图2页

(54)发明名称

超宽带与视觉融合的室内机器人定位与自主导航方法

(57)摘要

本发明公开一种超宽带与视觉融合的室内机器人定位与自主导航方法,包括以下步骤:(1)通过UWB定位系统获取并解算机器人绝对位置信息;(2)通过深度相机获取机器人的相对位置信息后;再通过融合定位方式得到机器人的精确位置信息;(3)通过避障导航模块获取方向信息控制驱动机器人完成导航;具体包括:(301)通过目标检测网络获取障碍物信息;(302)通过强化学习训练机器人完成避障导航。



1. 超宽带与视觉融合的室内机器人定位与自主导航方法,其特征在于,包括以下步骤:

(1) 通过UWB定位系统获取并解算机器人绝对位置信息;

(2) 通过深度相机获取机器人的相对位置信息后;再通过融合定位方式得到机器人的精确位置信息;

(3) 通过避障导航模块获取方向信息控制驱动机器人完成导航;具体包括:

(301) 通过目标检测网络获取障碍物信息;在避障导航模块中,将RGB-D相机获取的图像信息进行特征提取,放入到目标检测网络中检测出当前视图中的目标物体,然后根据当前图像的深度信息计算出目标物体与机器人的距离,由此得到当前的障碍物信息

(302) 通过强化学习训练机器人完成避障导航;在不依赖地图信息情况下,将得到的精确位置信息与障碍物信息作为先验知识,输入到深度强化学习模型中,将速度与转角的数值作为控制输出,然后送入到机器人微控制器中,驱动输出下一步的动作,完成机器人的自主导航。

2. 根据权利要求1所述超宽带与视觉融合的室内机器人定位与自主导航方法,其特征在于,步骤(1)中UWB定位系统由4个固定的基站点和1个可移动的标签点组成;将4个基站固定放置在房间的四个角,将1个UWB标签放置在机器人上;UWB定位系统通过组合标签点到四个基站的距离,最终计算出标签点在三维空间的坐标。

3. 根据权利要求1所述超宽带与视觉融合的室内机器人定位与自主导航方法,其特征在于,步骤(2)中机器人上装有RGB-D相机,用于直接获取空间点的深度信息,从而根据深度信息得到每张图片中每个像素点的3d位置。

4. 根据权利要求1所述超宽带与视觉融合的室内机器人定位与自主导航方法,其特征在于,步骤(2)中的融合定位方式为:

(201) 以UWB定位系统在定位过程中所使用的世界坐标系作为全局坐标系,将深度相机获取的相对位置信息经过空间变换,转换到世界坐标系下;

(202) 将两个位置坐标输入到一个3层($2 \times 3 \times 1$)神经网络中进行数据融合。

5. 根据权利要求1所述超宽带与视觉融合的室内机器人定位与自主导航方法,其特征在于,所述目标检测网络包括YOLO、Faster-RCNN。

超宽带与视觉融合的室内机器人定位与自主导航方法

技术领域

[0001] 本发明涉及室内机器人定位与导航技术领域,特别是涉及一种超宽带与视觉融合的室内机器人定位与自主导航方法。

背景技术

[0002] 室内定位导航解决的是导航技术最后一公里的问题,应用广泛,尤其在大型复杂的室内环境中,如博物馆、机场、超市、医院、地下矿井等区域,人们对位置服务有迫切需求。但是由于室内环境复杂,易受到遮挡和多径效应的影响,因此,获取精确的位置信息是当前的研究热点与难点。

[0003] 超宽带(Ultra Wide Band,UWB)定位技术是一种基于极窄脉冲无线技术,UWB射频信号与生俱来的物理特性使得UWB技术从一开始就被明确定义:能够实现实时、超精确、超可靠的定位和通信。它能够非常准确地测量无线电信号的飞行时间,从而实现厘米精度的距离/位置测量。UWB定位系统的精度可以达到15cm,测距范围达150m,数据更新频率最快可以达到10Hz。但是UWB定位技术的缺点是易受非视距误差的影响,室内环境下障碍物较为密集,移动标签与基站之间的信号无法直线传播,需要经过反射和衍射才能到达接收端,从而无法估算出准确的传播距离。

[0004] 自主定位导航是机器人实现智能化的前提之一,是赋予机器人感知和行动能力的关键因素。机器人实现自主定位导航,需要对周围环境进行分析、判断和选择,规划路径,进而完成导航任务。目前,自主定位导航技术有多种,视觉定位导航技术因其简单方便在各个领域中广泛应用。

[0005] 视觉定位导航主要借助视觉传感器完成,机器人借助单目、双目摄像头、深度摄像机、视频信号数字化设备或基于DSP(Digital Signal Processing)的快速信号处理器等其他外部设备获取图像,然后对周围的环境进行光学处理,将采集到的图像信息进行压缩,反馈到由神经网络和统计学方法构成的学习子系统,然后由子系统将采集到的图像信息与机器人的实际位置联系起来,完成定位。

[0006] UWB定位技术是一种绝对定位方式,可以获得三维空间的坐标信息,视觉定位是一种相对定位方式。因此,将两种不同信息源的数据融合在一起,可以实现优势互补,提供鲁棒性更好、精度更高的定位信息。

[0007] 深度强化学习适合于复杂环境下的连续决策任务,因此非常适合解决机器人问题,近几年,在机器人抓取、导航等任务中研究非常广泛。因此,结合深度强化学习可以完成移动机器人的导航控制问题,实现自主导航。

发明内容

[0008] 本发明的目的是为了克服现有技术中的不足,提供一种超宽带与视觉融合的室内机器人定位与自主导航方法,通过采用UWB技术与视觉信息融合的方式,达到了优势互补的效果,实现了高精度定位,同时将获取的位置信息与障碍物信息作为先验知识,结合深度强

化学习模型,实现无地图式的机器人自主导航,准确率高,鲁棒性好,且易于实现,可以应用到医疗,养老,商场等多个领域。

[0009] 本发明的目的是通过以下技术方案实现的:

[0010] 一种超宽带与视觉融合的室内机器人定位与自主导航方法,包括以下步骤:

[0011] (1) 通过UWB定位系统获取并解算机器人绝对位置信息;

[0012] (2) 通过深度相机获取机器人的相对位置信息后;再通过融合定位方式得到机器人的精确位置信息;

[0013] (3) 通过避障导航模块获取方向信息控制驱动机器人完成导航;具体包括:

[0014] (301) 通过目标检测网络获取障碍物信息;在避障导航模块中,将RGB-D相机获取的图像信息进行特征提取,放入到目标检测网络中检测出当前视图中的目标物体,然后根据当前图像的深度信息计算出目标物体与移动机器人的距离,由此得到当前的障碍物信息

[0015] (302) 通过强化学习训练机器人完成避障导航;在不依赖地图信息情况下,将得到的精确位置信息与障碍物信息作为先验知识,输入到深度强化学习模型—A3C (Asynchronous Advantage Actor-Critic) 算法框架中,将速度与转角的数值作为控制输出,然后送入到机器人微控制器中,驱动输出下一步的动作,完成机器人的自主导航。

[0016] 进一步的,步骤(1)中UWB定位系统由4个固定的基站点和1个可移动的标签点组成;将4个基站固定放置在房间的四个角,将1个UWB标签放置在移动机器人上;UWB定位系统通过组合标签点到四个基站的距离,最终计算出标签点在三维空间的坐标。

[0017] 进一步的,步骤(2)中移动机器人上装有RGB-D相机,用于直接获取空间点的深度信息,从而根据深度信息得到每张图片中每个像素点的3d位置。

[0018] 进一步的,步骤(2)中的融合定位方式为:

[0019] (201) 以UWB定位系统在定位过程中所使用的世界坐标系作为全局坐标系,将深度相机获取的相对位置信息经过空间变换,转换到世界坐标系下;

[0020] (202) 将两个位置坐标输入到一个3层($2 \times 3 \times 1$)神经网络中进行数据融合。

[0021] 进一步的,所述目标检测网络包括YOLO、Faster-RCNN

[0022] 与现有技术相比,本发明的技术方案所带来的有益效果是:

[0023] 1. 本发明融合UWB定位技术与视觉定位技术,弥补UWB定位技术易受非视距的影响,可以实现厘米级别的定位精度,得到更精确的定位。

[0024] 2. 本发明将位置信息和障碍物信息作为先验知识,结合深度强化学习模型进行机器人导航,实现无地图式的自主导航。

[0025] 3. 本方法适用于当前各类室内机器人定位与导航应用,尤其是在需要精确定位的复杂的室内场景中,比如商场、博物馆、医院、地下矿井等区域,鲁棒性强,稳定性好。同时本发明适用于无地图式的复杂场景中,可以实现自主导航。

附图说明

[0026] 图1是UWB室内定位示意图。

[0027] 图2是图片坐标系与相机坐标系的关系示意图。

[0028] 图3是强化学习框架图。

[0029] 图4是机器人融合定位与导航框架图。

具体实施方式

[0030] 以下结合附图和具体实施例对本发明作进一步详细说明。应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0031] 本发明提供一种超宽带与视觉融合的室内机器人定位与自主导航方法。该方案整体框架如图4所示。硬件结构主要分为两部分,UWB系统与移动机器人。方案主体围绕移动机器人展开。

[0032] 步骤1:通过UWB定位技术获取并解算机器人绝对位置信息。

[0033] 本实施例中UWB定位系统由4个固定的基站点和1个可移动的标签点组成。具体分布如图1所示。本实施例中将4个基站固定放置在房间的四个角,将1个UWB标签放置在移动机器人上。系统通过组合标签点到四个基站的距离,从而计算出标签点在三维空间的坐标。

[0034] 本实施例使用TOF的测距方式,飞行时间法(Timeofflight,TOF)是一种双向测距技术,它通过测量UWB信号在基站与标签之间往返的飞行时间来计算距离。基站与标签点之间的距离表示为:

$$[0035] \quad d_{at} = c(t_t - t_a) \quad (1)$$

[0036] 其中, c 为光速, t_t 为标签信号发射时间, t_a 为基站信号接收时间。

[0037] 标签与基站之间的距离 d_i ($i=1,2,3,4$)与标签在世界坐标系中的坐标位置有如下关系:

$$[0038] \quad \begin{cases} (x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2 = d_1^2 \\ \vdots \\ (x_4 - x_0)^2 + (y_4 - y_0)^2 + (z_4 - z_0)^2 = d_4^2 \end{cases} \quad (2)$$

[0039] 其中, (x_i, y_i, z_i) , ($i=1,2,3,4$)是基站的已知位置, (x_0, y_0, z_0) 为待求解的标签点位置。

[0040] 三个未知数和四个表达式(非线性相关)的方程无法直接求解,因此采用最小均方误差估计的方式求解:

$$[0041] \quad (\hat{x}_0, \hat{y}_0, \hat{z}_0) = \min_{(x_0, y_0, z_0)} \sum_{i=1}^4 [\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2} - d_i]^2 \quad (3)$$

[0042] 其中, $(\hat{x}_0, \hat{y}_0, \hat{z}_0)$ 为最小均方误差情况下标签点的估计位置。由此,获得了 t 时刻通过UWB定位技术解算获得的机器人在三维空间的估计位置。

[0043] 步骤2:通过融合定位模块得到机器人的精确位置信息。

[0044] 步骤(201):通过深度相机获取机器人的相对位置信息。

[0045] 移动机器人上装有RGB-D相机,可直接获取空间点的深度信息,从而根据深度信息得到每张图片中每个像素点的3d位置。

[0046] 具体流程为,首先标定相机,得到相机的内参与外参,然后设置坐标系如图2所示。其中, $o'-uv$ 是图片坐标系, $o-xyz$ 是相机坐标系,假设图片中的点为 (u, v) ,对应的三维点位置为 (x, y, z) ,则二者之间的转换关系为:

$$[0047] \quad l \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (4)$$

[0048] 其中, f_x, f_y 代表焦距, c_x, c_y 代表主点坐标 (相对于成像平面), l 为尺度因子, 表示从相机光心出去的射线都会落在成像平面的同一个点上。在 RGB-D 相机中, 从深度图中得到的距离 $dep(u, v)$ 与真实距离相差一个倍数 s , 因此得到尺度因子 l 为:

$$[0049] \quad l = \frac{dep(u, v)}{s} \quad (5)$$

[0050] 那么解算三维点位置为:

$$[0051] \quad \begin{cases} x = \frac{(u - c_x)z}{f_x} \\ y = \frac{(v - c_y)z}{f_y} \\ z = dep(u, v) / s \end{cases} \quad (6)$$

[0052] 完成相机标定后, 可以得到 c_x, c_y, f_x, f_y, l 的值。由此得到了通过视觉深度信息转换的机器人的相对位置信息。

[0053] 步骤 (202): 通过融合定位获取机器人的精确位置信息。

[0054] 由于 UWB 定位易受到非视距误差的影响, 本实施例采取两种定位技术融合的方式进行精确定位, 采用视觉相对位置信息进行辅助定位。具体的融合方式为:

[0055] 首先以 UWB 定位过程中所使用的世界坐标系作为全局坐标系, 将视觉解算出的相对位置信息经过空间变换, 转换到世界坐标系下。

[0056] 然后将两个位置坐标输入到一个 3 层 ($2 \times 3 \times 1$) 人工神经网络 (Artificial Neural Network) 中进行数据融合, 神经网络具有较强的容错性和自组织、自学习及自适应能力, 能够实现复杂的映射; 神经网络还具有强大的非线性处理能力, 能够很好的满足数据融合技术的要求。

[0057] 步骤 3: 通过避障导航模块获取方向信息控制驱动机器人完成导航。

[0058] 步骤 (301): 通过目标检测网络获取障碍物信息;

[0059] 在避障导航模块中, 将 RGB-D 相机获取的图像信息进行特征提取, 放入到目标检测网络如 YOLO、Faster-RCNN 等中检测出当前视图中的目标物体, 然后根据图像的深度信息计算出目标物体与移动机器人的距离, 由此得到当前的障碍物信息。

[0060] 步骤 (302): 通过强化学习训练机器人完成避障导航;

[0061] 在不依赖地图信息情况下, 将得到的精确位置信息与障碍物信息作为先验知识, 输入到深度强化学习模型中, 将速度与转角的数值作为控制输出, 然后送入到机器人微控制器中, 驱动输出下一步的动作, 完成机器人的自主导航。

[0062] 深度强化学习 DRL 将深度学习 (Deep Learning, DL) 与强化学习 (Reinforcement Learning, RL) 进行了融合, 其中 DL 主要负责利用神经网络的感知功能对输入的未知环境状态提取特征, 实现环境状态到状态动作值函数的拟合; 而 RL 则负责根据深度神经网络的输出和一定的探索策略完成决策, 从而实现状态到动作的映射, 其能较好地满足机器人的移

动需求。本实施例使用的深度强化学习模型为--A3C (Asynchronous Advantage Actor-Critic) 算法框架。

[0063] 机器人导航本质是机器人 (Agent) 与环境的交互任务, 在这个任务中机器人在时刻 t 从一组设定的运动动作中, 比如前进、后退、左转、右转等, 选择一个动作 $A_t = \{1, \dots, K\}$, 从一个状态 S_t 转移到另一个状态 S_{t+1} , 设定奖励规则, 从环境获得数值型奖励 (Reward) R , 动作 (Action) 指令被传递给机器人进行控制驱动, 完成与环境的交互, 即为一个强化学习过程, 强化学习框架如图3所示。

[0064] 具体地, 将每个时刻RGB-D深度图像作为马尔可夫状态, 故采集的深度图像集合作为马尔可夫状态空间。在构建状态空间过程中, 使用基于神经网络的A3C (Asynchronous Advantage Actor-Critic) 算法进行训练, 将深度图像作为机器人的控制输入, 将机器人的线速度与角速度作为输出来控制机器人进行运动规划, 使得机器人可以在一个连续的动作空间中选择接下来要做的动作。除此之外, 制定奖励值 R 的给予规则, 在机器人的运动过程中, 根据是否发生碰撞给予机器人奖励, 若是碰撞就给予较大的负奖励值; 结合障碍物信息, 离障碍物越近则给予越小的奖励值; 相反的, 若是机器人未碰到障碍物, 便给予正奖励值, 以鼓励无碰撞运动; 结合精确位置信息, 离终点越近则给予越大的奖励值, 以鼓励选择最短路径运动。

[0065] 在Actor-Critic算法的框架中, “Actor” 负责Policy Gradient的学习策略 $Q(s, a)$, 而“Critic” 作为策略评价估计的值函数 $V(s)$ 。策略和值函数互为依赖, 相互影响, 在训练过程中迭代优化。在A3C算法中, 定义优势函数 $A(s, a)$ 表示在状态 s 下选择的动作 a 的好坏程度。

$$[0066] \quad A(s, a) = Q(s, a) - V(s) \quad (7)$$

[0067] 由于在A3C中实际上并没有计算 Q 值, 在此使用折扣回报 R 作为 Q 值的估计值。

$$[0068] \quad R = r_t + \gamma \cdot r_{t-1} + \gamma^2 r_{t-2} \quad (8)$$

[0069] 其中, r_t 表示时刻 t 时的回报, γ 为权重值。则优势函数为:

$$[0070] \quad A(s, a) = R - V(s) \quad (9)$$

[0071] 这时, 即可得到值损失为折扣回报与状态值之间的均方差:

$$[0072] \quad \text{值损失} (L_v) = \sum (R - V(s))^2 \quad (10)$$

[0073] 则策略损失定义为:

$$[0074] \quad \text{策略损失} (L_p) = \log(\pi(s)) A(s, a) H(\pi) \quad (11)$$

[0075] 其中, $\pi(s)$ 为状态函数, 根据状态的输入来估计策略的选择。 $A(s, a)$ 为优势函数, $H(\pi)$ 为熵, 用于确保策略得到充分探索。熵值越小, 某一行为将会获得比其他行为更高的概率, 则机器人就会选择概率最大的行为。在损失函数中增加熵将会鼓励智能体进一步探索, 从而避免陷入局部最优。

[0076] 结合异步的优势演员评论家算法 (A3C) 进行训练, 克服了机器人输出有限的执行动作的局限性, 使得机器人可以在连续的速度与方向角度数值区间中输出执行动作。

[0077] 本发明并不限于上文描述的实施方式。以上对具体实施方式的描述旨在描述和说明本发明的技术方案, 上述的具体实施方式仅仅是示意性的, 并不是限制性的。在不脱离本发明宗旨和权利要求所保护的范围情况下, 本领域的普通技术人员在本发明的启示下还可做出很多形式的具体变换, 这些均属于本发明的保护范围之内。

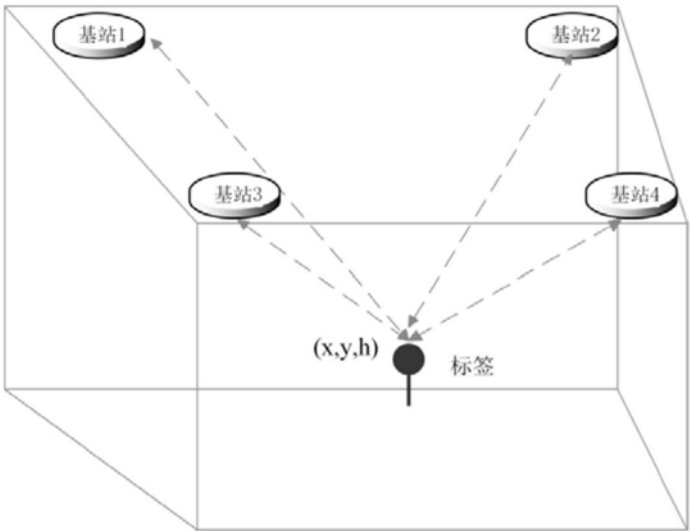


图1

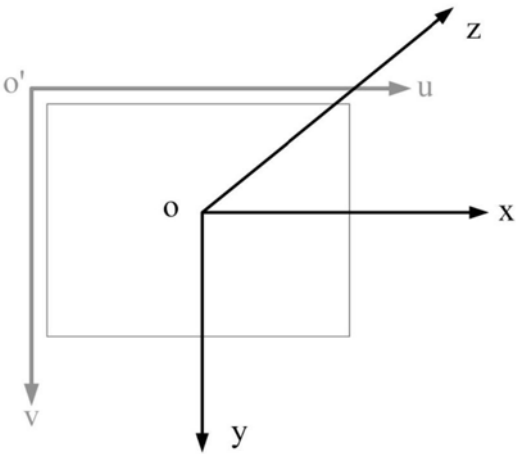


图2

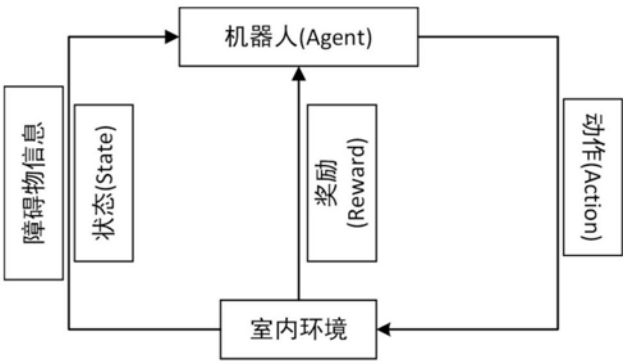


图3

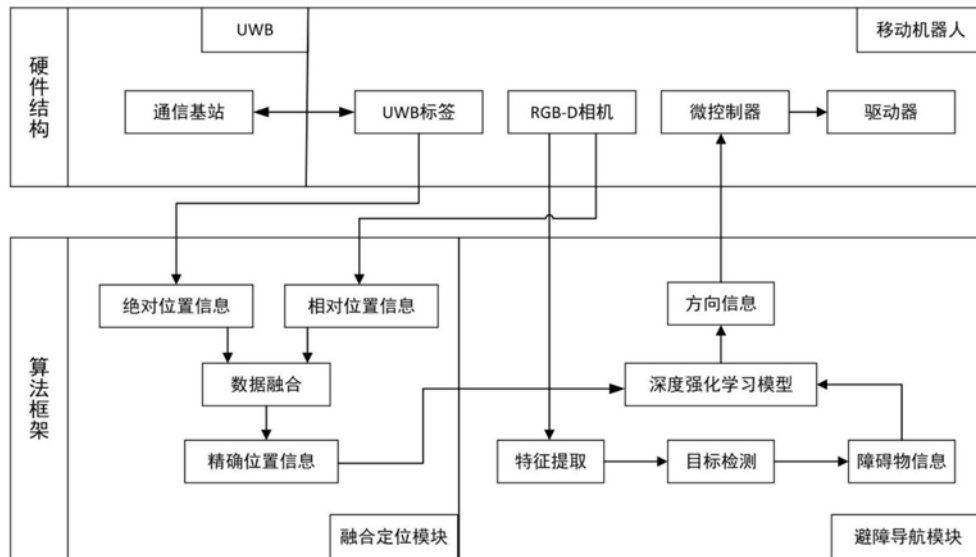


图4