

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:<https://orca.cardiff.ac.uk/id/eprint/169622/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Gao, Yu, Huo, Jing, Li, Wen-Nin, Wu, Jing , Lai, Yukun and Gao, Yang 2024. Object goal navigation based on path planning characteristics. CAAI Transactions on Intelligence Technology

Publishers page:

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



基于路径规划特点的语义目标导航方法

高宇¹, 霍静¹, 李文斌¹, 伍静², 来煜坤², 高阳¹

(1. 南京大学 计算机软件新技术全国重点实验室, 江苏 南京 210023; 2. 卡迪夫大学 计算机科学与信息学院, 英国 威尔士卡迪夫 CF10 3XQ)

摘要: 为了解决语义目标导航任务中存在的探索效率低、深度不精准等问题, 本文构建了一个解决语义目标导航任务的框架, 在语义地图构建模块中引入了深度图边缘处理以及地图纠错机制; 在探索模块中引入了覆盖范围最大化算法; 在路径规划模块中引入了替代点机制。本文在一个 3D 仿真环境下进行了实验。实验结果表明, 本文提出的解决方案明显提升了语义目标导航任务的性能。此外, 本文所提方法成功应用到了四足机器人上, 从而验证了在现实场景下的泛化性。

关键词: 人工智能; 视觉导航; 语义目标导航; 语义感知; 语义探索; 路径规划; 机器学习; 语义地图

Object Goal Navigation Based on Path Planning Characteristics

Gao Yu¹, Huo Jing¹, Li Wen-Bin¹, Wu Jing², Lai Yu-Kun², Gao Yang¹

(1. State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China; 2. School of Computer Science and Informatics, Cardiff University, Cardiff CF10 3XQ, England)

Abstract: To solve the problems of low exploration efficiency and imprecise depth in object goal navigation, this article constructs a framework to solve object goal navigation. Depth map edge processing and map error correction mechanisms were introduced in the semantic map construction module; a coverage maximization algorithm was introduced in the exploration module; alternative point mechanisms was introduced in the path planning module. This article conducted experiments in a 3D simulation environment. The experimental results show that the new solution proposed in this article significantly improves the performance of object goal navigation. In addition, the method proposed in this article was successfully applied to quadruped robots, thereby verifying its generalization in real-world scenarios.

Keywords: artificial intelligence; visual navigation; object goal navigation; semantic perception; semantic exploration; path planning; machine learning; semantic map

语义目标导航任务^[1]要求一个装有红绿蓝-深度 (RGB-Depth, RGBD) 相机的机器人 (智能体) 在未知地图环境中自主探索以找到特定类别的物体, 其对安防、智能巡检具有重要意义。智能体为完成语义目标导航任务所需要的能力可概括为场景建图、去哪里和如何去 3 类。场景建图是指针对传感器观测数据的记忆和整合能力, 除显式构建地图外也有使用循环神经网络^[2]等方法; 去哪里是指分析目标类别物体最有可能出现位置的能力, 以实现高效率探索; 如何去是指路径规划和轨迹跟踪能力, 以找到快速抵达目标位置的可行控制动作序列。去哪里是语义目标导航区别于其它导航任务的鲜明特点。

解决语义目标导航任务需要搭建涉及众多组

件的复杂系统, 每个组件的设定都会对最终性能产生很大影响, 导致对现有方法的对比分析和对改进之处的探索极为困难。为此, 本文创建了一个解决语义目标导航任务的标准框架, 并针对测试中发现的深度不精准、语义分割效果差、探索效率低、路径规划没有考虑任务特点等问题提出了新的解决方案从而进一步提升了性能。具体来说: 1) 针对探索效率低问题, 在探索模块中引入了基于路径规划特点的覆盖最大化算法; 2) 针对深度不精准和语义分割效果差问题, 在语义建图模块中引入了深度图边缘处理与地图纠错机制; 3) 针对路径规划没有考虑语义目标导航任务特点的问题, 引入了替代点机制和障碍物概率地图。

使用提出的框架在一个逼真的交互式 3D 室内

语义数据集 (Habitat-Matterport 3D Semantics Dataset, HM3DSem) [3-4] 下进行了实验。实验结果显示, 本文提出的 3 个改进明显提高了语义目标导航任务的性能指标。本文所提框架成功应用到了四足机器人上, 从而验证了该框架对现实世界具有泛化性。

1 语义目标导航相关工作

1.1 语义 SLAM

同时定位与地图构建 (Simultaneous Localization and Mapping, SLAM) 算法根据传感器数据来源的不同可以分为激光 SLAM^[5-7]和视觉 SLAM^[8-11]。激光 SLAM 根据激光测距结果直接构造障碍物地图, 视觉 SLAM 则寻找图片上的关键特征点并通过多个视角来确定相机位姿。SLAM 解决方案一般都集中在寻找像素级别的特征点, 即在单幅图像中提取特征点然后在多幅图像中对特征点进行匹配。这与人类通过判断物体在眼睛中的移动来进行定位极为不同, 因为人类定位针对于整个物体而 SLAM 解决方案针对于特征点。为此可以引申出使用语义信息来帮助寻找多幅图像中的关联或为 SLAM 的回环检测等引入更多信息, 语义信息为 SLAM 提供了更多判断相机移动的条件从而更好地定位。语义信息的引入也使得 SLAM 不仅能够构建点云地图、障碍物地图, 而且可以构建一个有语义标签的地图, 从而为下游任务提供了更广阔的空间。

语义目标导航任务主要用到语义 SLAM 提供的带有语义标签的地图。有了带有语义信息的二维地图或者三维地图, 智能体能够更加结构化地去分析不同类别物体的分布关系, 更好地去理解场景从而推理出目标类别物体在地图中各个位置的出现概率。对目标类别物体出现位置的推理是语义目标导航任务的核心。但是现有方法没有针对语义目标导航中存在的深度不精准和语义分割效果差问题进行优化, 这导致连续多帧的语义建图出现较大误差而无法完成语义导航任务。本文为此引入了深度图边缘处理及语义点云地图和二维地图纠错机制。

1.2 经典导航方法

导航能力是无人车、无人机、无人船等众多移动机器人所需要的基本能力之一。经典导航方法通常在 SLAM 构建的地图上使用包括全局和局部路径规划在内的路径规划模块产生路径, 然后使用控制算法对路径规划模块生成的路径进行跟踪和实

际行进。路径规划的经典算法有很多, 如迪克斯特拉算法 (Dijkstra algorithm)、A 星搜索算法 (A* search algorithm, A*)^[12]、概率路线图算法 (probabilistic road maps, PRM)^[13]、快速探索随机树算法 (rapidly exploring random trees, RRT)^[14]、人工势场法 (artificial potential fields, APF)^[15]等。Dijkstra 算法是基于图结构的能够保证最优解的路径规划算法, 但处理大地图时效率低下; A* 算法在 Dijkstra 算法的基础上引入衡量任何一个位置到目标位置距离的启发式函数从而使得探索方向具有一定的目的性; PRM 算法对地图进行稀疏采样从而将栅格地图转换为由少量采样点和可达边组成的图结构, 然后在图结构中再使用 A* 等算法寻找路径; RRT 和 PRM 一样同是基于随机采样的规划算法, 基本思想是从一个点出发向外探索扩展; 与前面几种算法不同的是, APF 更适用于动态环境和局部路径规划。但是现有算法都没有针对语义目标导航任务特点进行改进从而导致路径规划模块的失败率较高, 为此本文引入了替代点机制以提高路径规划算法对语义目标导航任务的鲁棒性。

1.3 基于学习的导航方法

经典导航算法所使用的基于网格的地图表示在精度和内存需求上具有天然的矛盾, 在动态环境中实时重新规划路径也需要大量计算。此外, 经典导航框架中从 SLAM 到路径规划, 计算误差会逐渐积累。为此, 有研究将深度强化学习等引入导航中, 从而得到了基于学习的导航方法^[16]。大多数研究均使用深度 Q 网络^[17]、异步优势演员评论家算法^[18]、近端策略优化算法^[19]、深度确定性策略梯度算法^[20]等通用强化学习算法, 但它们在状态设计、奖励函数设计、强化学习的使用方式等方面做了很多改进。状态设计涵盖了起点、目标点、障碍物位置等关键信息, 与测试环境息息相关。若只考虑导航的核心任务 (例如避障和到达目标点), 那么奖励会非常稀疏从而导致强化学习难以训练; 所以多数研究引入了手工设计的中间奖励函数, 包括碰撞、与最近障碍物的距离变化、与目标点的距离变化、时间步惩罚等。如何在导航中使用强化学习是基于学习的方法的核心问题, 本文将使用方式分为 3 种。第 1 种是在导航中直接使用深度强化学习^[21], 即将完整的导航过程描述为马尔科夫过程, 以传感器的观测数据作为状态, 直接学习到路径或控制动作序列的映射; 这种方法不再需要 SLAM、全局地图等, 但在复杂环境中容易落入局部陷阱。第 2 种是与经

典导航技术相结合^[22]，如先使用 PRM 对地图进行稀疏化，再在局部使用深度强化学习进行路径规划。第 3 种是分层深度强化学习，即将路径规划划分为静态避障、动态避障、趋向目标点等不同层次上的子任务，这主要是考虑到环境复杂度与状态空间巨大。但现有算法面对语义目标导航任务均存在探索效率低的问题，为此本文提出了基于路径规划特点的覆盖范围最大化算法。

2 基于路径规划特点的解决方案

2.1 语义目标导航任务定义

在语义目标导航任务中，智能体以随机的位置和方向被初始化在一个未知地图环境中，其目标是找到特定类别的物体，如床、厕所等。智能体需要依靠提供的 RGBD 相机、位置信息 (GPS) 和罗盘等传感器实现导航，也就是说视觉观察包括第 1 人称的 RGB 图和深度图。动作空间是离散的，由行进、左转、右转、停止共 4 个动作组成，行进意味着向前移动 0.25 m，左转和右转的幅度为 30°。当智能体认为它已经接近目标对象时，它需要采取“停止”操作；如果智能体采取“停止”操作且停止时与目标物体的距离小于阈值 1 m，则认为该回合是导航成功的。可见成功需要同时满足发出“停止”操作和与目标物体足够近 2 个条件。

2.2 基于语义建图的探索框架

本文构建了基于语义建图的探索框架，将探索模块、语义建图模块、路径规划模块等进行解耦合，为不同模块下不同算法的性能对比及模块之间的重要性对比提供了方便公平的对比框架。语义建图模块使用 GPS、罗盘、RGBD 图像等传感器数据构建语义、障碍物等地图；探索模块根据构建好的地图预测目标类别物体最有可能出现的位置从而得到长期目标；路径规划模块依据障碍物地图、智能体位置和长期目标规划出一条可行路径；轨迹跟踪模块依据路径等输出下一步智能体需要采取的动作。本文针对探索效率低问题，引入了基于路径规划特点的覆盖范围最大化算法；针对深度不精准和语义分割效果差问题，引入了深度图边缘处理及语义点云地图和二维地图纠错机制；针对没有考虑语义目标导航任务特点的问题，在路径规划模块中引入了替代点机制和障碍物概率地图。下面将详细介绍各个模块及对应改进。

2.3 基于路径规划特点的探索算法

探索模块负责选取路径规划模块的目标点，该目标点称为长期目标。长期目标的选择通常依据两点：1)到达长期目标所需经过区域多为未知以尽可能扩大探索面积；2)长期目标附近很可能存在目标类别物体以尽快完成任务。以往工作^[23-26]中长期目标点的选择通常使用强化学习来完成。以面向目标的语义探索算法 (Goal Oriented Semantic Exploration, SemExp)^[23]为例，其将障碍物地图、语义地图作为状态输入，将与目标类别距离变化和探索面积变化的加权和作为奖励函数，将近端策略优化 (Proximal Policy Optimization, PPO)^[19]作为策略网络进行强化学习训练。

本文通过实验发现该方法存在探索效率低问题，结果如表 1 所示，数据集描述以及成功率、带有路径长度加权的成功率 (success weighted by path length, SPL) 等指标说明详见第 3.2 节。使用 SemExp 论文公布的训练好的网络参数的成功率为 25.8%，SPL 为 12.8%；随机初始化的网络参数成功率为 25.3%，SPL 为 11.6%。二者指标接近，说明强化学习训练对探索效率的提高帮助有限。将强化学习状态输入的语义地图全部置零，成功率甚至达到了 26.4%，SPL 为 12.2%。这更加突出了探索效率低的问题。

为此，本文提出了基于路径规划特点的覆盖范围最大化算法。该算法将长期目标设置在地图的边界处：

$$p_{subgoal} = (x + \frac{\sqrt{2}K}{2} \cos \theta, y + \frac{\sqrt{2}K}{2} \sin \theta)$$

式中： $p_{subgoal}$ 为长期目标， (x, y, θ) 为当前智能体的位置和朝向， K 为地图尺寸。未探索区域被设置为无障碍物，路径规划会设计出一条涵盖无障碍物和未探索区域的路线并在行进中不断更新地图。当前路径在更新后地图中不再可通行时，会重新规划路线。

使用替代点的覆盖范围最大化算法

输入：仿真器 env，强制更新步数 LMAX，局部规划器 LPF，目标检测 OD

- 1) lstep=0, haveseen=false, subgoal=(0, 0)
- 2) obs=env.reset()
- 3) While not env.done:
- 4) lstep+=1
- 5) If lstep>LMAX or LPF(subgoal)失败:
- 6) 按 2.3 章节公式计算 subgoal

- 7) lstep=0
- 8) If not haveeseen:
- 9) goal=OD(obs)
- 10) If goal is not None:
- 11) 按 2.5 章节根据 goal 计算 subgoal
- 12) haveeseen=true
- 13) action=LPF(subgoal)
- 14) obs=env.step(action)

注意到，智能体不会再去访问已经探索过且没有其他支路的区域，因为这些区域不可能再涵盖到达目标点的路线。也就是说，智能体进入死胡同后会自动改变朝向从而发现其他未探索区域。这避免了在某个房间反复探索或是在 2 点间反复来回等低探索效率行为。本文设计的探索模块利用路径重新规划的特点使智能体覆盖范围最大化，从而提高了语义目标导航任务的探索效率。

表 1 SemExp 不同设置在 HM3DSem 测试数据集下的指标对比

模型设置	成功率(%)	SPL(%)
论文公布的网络参数	25.8	12.8
随机初始化网络参数	25.3	11.6
语义地图全部置零	26.4	12.2

2.4 边缘处理与地图纠错

语义建图模块负责地图构建，包括点云地图、障碍物地图、已探索区域地图和语义地图，具体描述如表 2 所示。点云地图通过对 RGBD 图像使用主动神经 SLAM (Active Neural SLAM, ANS)^[27]方法得到。

表 2 语义建图模块所建地图含义描述

名称	大小	说明
点云地图	4×M	M 表示点云数量，每个点均用 XYZ 坐标和一个语义标签表示
障碍物图	960×960	每个点代表 0.05 米宽的正方形区域；取值范围 0-1 的实数，表示有障碍物的概率；未观测区域视为无障碍物

已探索区域图	960×960	每个点代表 0.05 米宽的正方形区域；0 表示未探索，1 表示已探索
语义地图	960×960	每个点代表 0.05 米宽的正方形区域；取值范围 0-15 的整数，表示 15 个语义类别和其它类别

语义地图在语义目标导航任务中起着重要作用，是探索模块的主要输入。语义地图的质量直接决定了下游模块的性能。以往语义目标导航研究所使用的语义地图建立方法一般是先对 RGB 图像进行语义分割得到分割图，根据深度图像得到三维点云，然后根据分割图和深度图像的对应关系为每一个点标注一个语义标签，最后对得到的带有语义信息的三维点云在垂直维度上进行求和得到二维语义地图。使用每一帧的 RGB 图像和深度图像都能得到一幅二维语义地图。不同时间下得到的二维语义地图根据当时智能体的 GPS 和罗盘信息被拼接起来，这样就构建出了供下游模块使用的保存了全部所见帧信息的二维语义地图。

深度不精准和语义分割效果差为上述语义地图构建方法造成了严重问题。RGB 图像语义分割的些许差错可能会导致该帧生成的二维语义地图存在巨大错误，更严重的是二维语义地图错误会随着帧数增多而积累，最终导致拼接而成的二维语义地图出现严重偏差甚至失效。

造成 RGB 图像语义分割些许差错会导致二维语义地图产生巨大错误的根本来源是物体实例边缘处深度值的不精准。一方面，语义分割容易在物体实例边缘处的像素点上出错，将部分属于其他物体的像素点标注为该物体实例；另一方面，物体边缘的深度值会发生跳变，在二维图像中相邻的 2 个点在实际的三维空间中可能相距很远。

图 1 给出了一组深度图像、RGB 图像及对 RGB 图像进行语义分割得到的部分分割图，其中语义分割算法实际使用的是掩膜区域卷积神经网络 (Mask Region-based Convolutional Neural Network, Mask R-CNN)^[28]实例分割算法。深度图像中颜色深浅表示与相机的距离，最亮区域表示超过相机的最大测量距离；分割图给出了椅子和桌子的语义信息。上一段提到的物体实例边缘处的 2 个不利方面均在该组图中得到体现。以椅子腿为例，一方面，因为 RGB 图像的噪声和实例分割算法的性能限制，部分属于地面、桌子等的像素点被划分为椅子腿，这些像素

点将进入椅子所对应的语义地图层；另一方面，椅子与后方墙壁虽然在 RGB 图像和深度图像上像素点相邻，但其在三维空间中实际距离相差很远。这导致如果将部分墙壁像素点标注为椅子的语义信息，那么三维点云地图中与椅子相离很远的部分墙壁也将会被标注为椅子，而智能体到达该部分墙壁时会认为已找到椅子而采取停止动作。可见，在物体实例边缘处对深度图像和分割图进行处理是非常必要且重要的，但以往语义目标导航研究没有针对这一点进行优化。



图 1 一组深度图像、RGB 图像与对应的部分分割图

Fig.1 Depth image, RGB image, and corresponding partial segmentation image

为此，本文针对上述问题提出了 3 种基于形态学的方法以改进语义建图模块建图的质量。这 3 种方法分别应用于语义建图模块的 3 个步骤中，所以可以同时应用。

1)对语义分割得到的分割图进行腐蚀。具体地，使用 5×5 的全 1 卷积核对分割图进行卷积，分割图上像素值表示语义概率，如果一个像素点为概率 1 的特定语义点且位于语义物体内部，则经过卷积操作后的该点数值为 25。同时考虑语义点位置和概率，保留卷积后数值大于 23.5 的点作为最终分割图中的特定语义区域，用公式表示为

$$M_{SemNew} = (1_s * M_{Sem}) > 23.5$$

式中： M_{SemNew} 表示处理后的分割图， M_{Sem} 表示原始分割图， 1_s 表示 5×5 的全 1 矩阵， $*$ 表示卷积运算。可见这会缩小分割图中语义物体的面积从而减少语义分割在物体边缘处出错的概率，但不会影响语义目标导航任务的性能，因为解决语义目标导航任务只需要知道有一个特定语义物体在该区域而无需知道物体的摆放细节。这也有助于去除语义分割模型输出的异常孤立点。

2)对语义三维点云地图进行滤波，去除那些距同类语义物体距离远大于平均距离的语义点，从而确保去除掉由物体边缘深度值跳变导致的部分线段状点云，因为物体边缘形成的错误点云与其主体点云的距离很大。具体地，使用 K 维树

(K-Dimensional Tree, KD Tree) 来计算每个语义点周围 0.1 m 范围内相同语义点的数量，若数量大于 5 则保留该点，否则舍弃：

$$M_{PcdNew} = \{p \in M_{Pcd} :$$

$$\sum_{p' \in M_{Pcd}} I(dis(p, p') < 0.1) * I(label(p) = label(p')) > 5\}$$

式中： M_{PcdNew} 表示处理后的点云， M_{Pcd} 表示原始点云， dis 是 2 点之间的距离函数， $label$ 是点对应的语义标签函数， I 是指示函数。

3)对拼接而成的语义二维地图进行开运算。腐蚀操作类似于对分割图的处理，会去除掉地图中的孤立点，而膨胀操作恢复物体实例本来的尺寸并填充由于体素化而缺失的部分语义点。二维语义地图开运算保证了供下游模块使用的语义地图质量以及与原地图之间物体尺寸的匹配。

2.5 语义目标替代点

以往语义目标导航研究中使用的路径规划算法可分为 2 类：1)基于几何的经典路径规划算法，如 Sethian 等^[29]所使用的快速行进算法 (fast marching method, FMM)；2)基于学习的路径规划算法，如 Wijmans 等^[30]所使用的去中心化分布式近端策略优化算法 (Decentralized Distributed Proximal Policy Optimization, DDPPPO)。这些框架使用的均是标准路径规划算法，没有将语义目标导航的任务特点考虑进去。为此，本文以语义目标导航的任务特点为驱动对路径规划模块作了以下几点设计。

语义目标导航任务的一个显著特点是其长期目标 (也即路径规划的终点) 可能设置在障碍物上。这是合理的。如床在智能体所构建的障碍物地图中属于无法通过的障碍物，当语义目标类别设置为床时，标准的路径规划算法无法规划出一条无障碍路径，因为终点处于障碍物的包围之中。以往语义目标导航框架没有显式地考虑这一问题，而是简单地把所有目标语义点作为路径规划终点。这种方法能够解决同一语义物体被作为障碍物的问题，但对于被其他障碍物所包围的语义目标而言是无效的。如放置在桌子上的盆栽和嵌入墙壁的电视，二维障碍物地图中盆栽被桌子围绕，电视被墙壁围绕，导致从机器人当前位置无法规划出到任意一个语义目标点的路径。为解决该问题，本文提出了将语义目标附近可通行点作为替代点的方法，即将位于语义目标点周围且能够从智能体当前位置规划出一条路径的点作为长期目标。该方法解决了语义目标被障碍物包围而无法规划路径的问题。

求解替代点存在从最近替代点看不到目标物体的问题。以放置在桌子上的盆栽为例，按照寻找替代点的思路，替代点将被设置在桌子的4条边上，若桌子靠墙，则有一条边将被设置在墙的另一侧即另一间房间（因为无法从障碍物地图中判断从另一间房间是否能够看到盆栽）。这就导致长期目标有可能被设置在另一间房间而造成任务失败。本文使用下述方法来寻找替代点以规避该问题：将发现语义目标物体时的智能体位置与随机一个语义目标点进行连线，将连线上最靠近语义目标点的无障碍物点作为长期目标：

$$p_{subgoal} = \underset{p \in \text{Line}(p_{cur}, p_{goal}) \text{ AND } p \text{ is free}}{\text{argmin}} \text{dis}(p, p_{goal})$$

式中： $p_{subgoal}$ 表示语义目标替代点求解后的长期目标， p_{cur} 表示智能体发现语义目标时的位置， p_{goal} 表示发现的语义目标点位置。智能体在发现语义目标物体时的位置时肯定能够看到语义目标，所以二者连线是智能体能够完全观察的区域，能为其后续路径规划提供更多图像，并且从当前位置到连线上点的可通行概率更大。

除了引入语义目标驱动的替代点机制，本文还将更适用于实际的障碍物地图表示方法带入了语义目标导航框架中。以往的语义目标导航研究使用非0即1的障碍物地图表示方式，这种表示方式可能导致智能体在空旷处选择贴近障碍物的路径而这些路径实际无法通行，也可能导致在狭窄处穿越密集障碍物时路径规划失灵。以往的语义目标导航研究也没有考虑智能体的尺寸问题，只是把智能体作为一个质点。按照质点规划出来的路径往往难以通行。以往工作一般是构建所谓的碰撞传感器^[23,27,31-34]以记录实际不可通行的区域来解决智能体尺寸问题，但该方法在现实中是不可行的。为此，本文将障碍物概率地图引入到了语义目标导航任务中。本文使用 9×9 的全1矩阵作为卷积核二维障碍物地图进行卷积，卷积后地图中每个大于0的像素点均当作障碍物像素点，从而得到障碍物概率地图。这样不仅将智能体尺寸引入到了障碍物地图中，也显式地表示了障碍物的密集程度。数值较大的区域表示位于障碍物中心无法通行，数值较小的区域由于深度图像的误差则有可能可以通行。这就给了路径规划模块更大的选择空间，同时也可以去除了不符合实际的所谓“碰撞传感器”。图2显示了2种地图表示形式的可视化效果。右图表示概率地图，其中亮白区域数值较大不可能通行，灰白区域数值较小则有可能通行。智能体在无法找到全部

数值为0的路径时则会考虑数值较小的那些点。

最后，轨迹跟踪模块依据规划出的路径计算智能体下一步需要采取的动作。因为该任务中行进、左转、右转的幅度都是固定的，所以只需要将智能体转向下一个路径中间点并朝其前进即可。

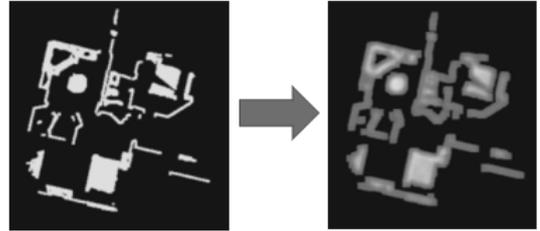


图2 障碍物地图处理示例

Fig.2 Example of obstacle map processing

3 实验设计与结果分析

3.1 数据集与仿真器

本文使用 HM3DSem 数据集。HM3DSem 是迄今为止最大的三维真实世界和室内空间数据集，具有密集的语义注释。它包含 142 646 个对象实例注释，216 个 3D 空间及这些空间中的 3 100 个房间。场景使用 142 646 个原始对象名称进行注释并被映射到 40 个 Matterport 类别中。每个场景平均包含 106 个类别和 661 件物品。图 3 展示了 HM3DSem 某个场景的预览图。

本文选用 Habitat^[35-36]作为仿真器。Habitat 仿真器会加载 HM3Dsem 数据集及语义目标导航任务数据集（包括对智能体初始化位姿、语义目标类别等的定义），提供了数据集和用户代码的交互接口（如对传感器数据的访问、对智能体运动的控制、返回任务的执行结果等）。在本文的实验环境中，智能体能够获得的信息有：RGBD 相机拍摄的第 1 人称 RGB 图像和深度图像、跟踪相机位置的里程计信息；智能体的动作空间为：前进 0.25 m、左转 30°、右转 30°、停止。



图3 HM3DSem 某个 3D 空间的预览图

Fig.3 Preview image of one 3D space in HM3DSem

3.2 评价指标及基准算法

本文使用成功率（success rate, SR）、带有路径长度加权的成功率（success weighted by path length, SPL）、发出停止动作时与目标的距离（distance to goal, DTG）3个评价指标。三个指标具体定义为

$$m_{SR} = \frac{1}{N} \sum_{i=1}^N S_i$$

$$m_{SPL} = \frac{1}{N} \sum_{i=1}^N S_i \frac{l_i}{\max(p_i, l_i)}$$

$$m_{DTG} = \frac{1}{N} \sum_{i=1}^N d_i$$

式中： N 表示回合数量， S_i 表示第 i 回合是否成功， l_i 表示第 i 回合能够完成任务的最短路径长度， p_i 表示第 i 回合实际采取路径的长度， d_i 表示第 i 回合发出停止动作时智能体与语义目标的距离。可见，SR和SPL越大越好，DTG越小越好。SPL综合考虑了成功率和效率。按照Anderson等^[1]的建议，本文将SPL作为主要指标，SR和DTG作为辅助指标。

为了评估模型效果，本文选取了前沿探索（Frontier-based Exploration, FrontierExp）^[37]、DDPPO^[30]、SemExp^[23]、对人类经验进行模仿学习的算法（HabitatWeb）^[38]共4种基准模型作为对比。下面是对基准模型的简要介绍。

FrontierExp选择离自身位置最近且可通行的frontier（即未访问区域和已访问区域的交界）作为长期目标。

DDPPO不进行建图，使用PPO直接学习传感器数据到控制动作的策略。DDPPO在PPO基础上实现了一种大规模强化学习训练框架。本文使用DDPPO提供的模型参数进行测试，其使用64块Tesla V100 GPU进行了3d训练，消耗了25亿帧样本。可见，DDPPO所需要的训练样本、训练时间和硬件资源都是巨大的。

SemExp由语义建图模块、强化学习长期策略选择和基于快速行进法（Fast Marching Method, FMM）的局部路径规划算法组成。本文所提系统框架正是对SemExp各个模块进行了解耦，依据测试中发现问题进行了相应改进。

HabitatWeb收集了大量人类完成语义目标导

航任务的示例并使用行为克隆对人类示例进行模仿学习。

3.3 与基准算法比较

图4以实验示意图的形式展示了本文所提的系统框架。本文在HM3DSem测试数据集下进行测试。HM3DSem测试集共有2000个测试回合，每回合任务必须在500步内完成。语义目标类别有椅子、床、植物、马桶、电视、沙发共6类。表3给出了本文所提框架与上述基准算法在HM3DSem测试集下6个类别及平均的实验结果。

从表3可以看出，在使用SPL指标的测评中，本文所提框架取得了优于其他所有模型的效果（0.236），比其他最好的HabitatWeb（0.193）高出0.043，提高了22.3%。在椅子、植物、马桶、电视、沙发这5个类上本文所提框架的SPL指标均最高，仅床类下HabitatWeb的SPL指标高于本文所提框架。这说明本文所提框架着实提高了语义目标导航任务的性能。在总体成功率方面，HabitatWeb（0.491）高于本文所提框架（0.441）；本文框架在植物、马桶、电视这3个类上成功率最高，HabitatWeb在椅子、床、沙发这3个类上成功率最高。高成功率低SPL说明HabitatWeb未能兼顾好成功率和效率，在提高成功率的同时极大地牺牲了效率。因为目标导航任务所涉及的场景复杂多变，人类示例很难穷尽所有情况，这导致HabitatWeb进行模仿学习时会遇到无法处理的状态而在同一片区域反复游荡，导致HabitatWeb完成任务的路径长度极度增加。而本文框架提出了基于路径规划特点的覆盖范围最大化算法，避免了在同一片区域反复探索或是在2点间反复来回等低效率行为，使得智能体覆盖范围最大化，从而提高了语义目标导航任务的探索效率，使得本文所提框架成功率在低于HabitatWeb的同时SPL指标却高于它。从以上分析可以看出，从兼顾成功率和效率的SPL指标来看，本文框架超过了以往各个方法；从成功率来看，本文框架超过了FrontierExp、DDPPO、SemExp，略低于HabitatWeb。不过HabitatWeb需要大量的人类示例，耗费大量的时间和资金。本文框架无需昂贵的数据收集过程和巨大的训练时间。

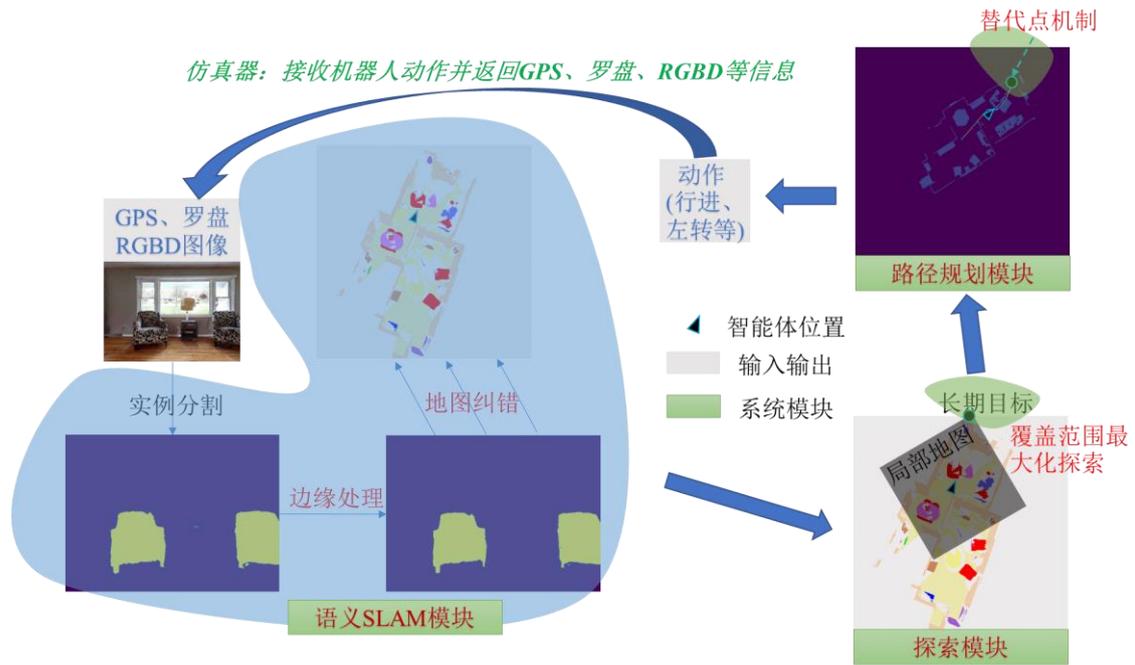


图 4 系统框架图

Fig.4 System framework diagram

表 3 HM3DSem 测试数据集下语义目标导航结果

Table 3 Object goal navigation results under HM3DSem test dataset

方法		FrontierExp	DDPPO	SemExp	HabitatWeb	本文方法
椅子	SPL	0.172	0.198	0.223	0.257	0.306
	SR	0.287	0.383	0.421	0.609	0.542
	DTG	3.485	2.431	2.659	1.631	2.222
床	SPL	0.159	0.198	0.097	0.255	0.222
	SR	0.282	0.355	0.180	0.648	0.422
	DTG	6.300	5.509	5.610	3.685	5.752
植物	SPL	0.107	0.006	0.076	0.066	0.121
	SR	0.179	0.023	0.179	0.226	0.240
	DTG	6.553	8.296	8.451	7.987	7.544
马桶	SPL	0.155	0.084	0.101	0.109	0.194
	SR	0.317	0.198	0.234	0.311	0.406
	DTG	5.390	6.241	5.486	5.864	4.942
电视	SPL	0.119	0.041	0.059	0.078	0.169
	SR	0.176	0.085	0.135	0.220	0.298
	DTG	7.811	7.864	7.609	7.609	7.167
沙发	SPL	0.219	0.171	0.148	0.251	0.290
	SR	0.355	0.361	0.298	0.627	0.534
	DTG	4.156	2.940	3.868	2.319	3.266
平均	SPL	0.164	0.140	0.128	0.193	0.236
	SR	0.285	0.279	0.258	0.491	0.441
	DTG	5.336	4.961	5.027	4.154	4.641

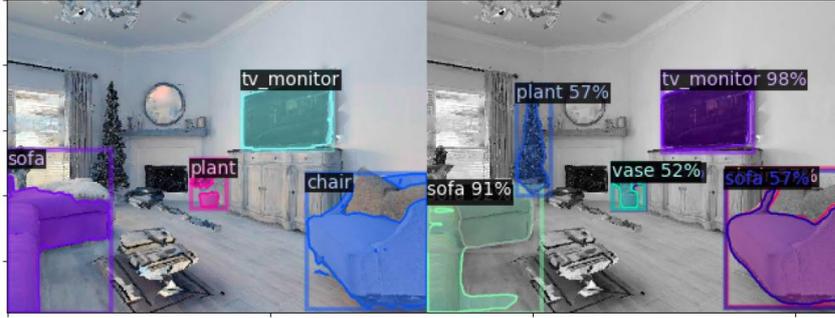


图 5 实例分割数据集可视化效果。左图为 Ground Truth，右图为实例分割预测结果。

Fig.5 Visualization of instance segmentation dataset

表 4 Mask R-CNN 在 HM3DSem 各个物体类别下的实例分割表现

Table 4 Instance segmentation performance of Mask R-CNN in various object categories of HM3DSem

类别	椅子	床	植物	马桶	电视	沙发
bbox AP 指标	28.77	61.27	12.10	83.88	7.94	36.10

从表 3 还可以看到，几乎所有方法在椅子、床、沙发、马桶这 4 个类上的表现优于在植物、电视这 2 个类上的表现。如 DDPPO 在椅子类上 SPL 有 0.198，而在植物类上 SPL 只有 0.006；HabitatWeb 在床类上 SPL 有 0.255，而在电视类上只有 0.078。本文分析认为造成这种巨大差异来源于不同类别的实例分割效果差异。为了验证该想法，本文在 HM3DSem 测试数据集中随机采样了 7 万张图片（其中包含 17.7 万个实例）并制作了实例分割数据集（如图 5 所示），使用 Mask R-CNN 算法对其进行了测试，测试结果如表 4 所示。从表 4 中可以看出，植物和电视的边界框平均精度指标（Bounding Box Average Precision, bbox AP）分别为 12.10 和 7.94，远低于其他物体类别。语义信息的不准确限制了植物和电视类的语义目标导航性能。另外，虽然本文框架在植物和电视类上的表现不如其他类别，但本文框架不同类别之间的差距低于其他算法。这说明本文框架对语义信息的容错率更高，更适合应用到复杂的现实任务中去。

3.4 消融实验

为了验证本文 3 个模块的有效性，本文进行了消融实验，结果如表 5 所示。其中去除深度图处理表示语义建图模块不再进行深度图边缘处理与语义点云地图和二维地图纠错；去除最大化探索表示探索模块不再使用覆盖范围最大化算法，更改成了基于 Frontier 的选择方法；去除替代点机制表示路径规划模块不再使用替代点机制和障碍物概率地图。从表 5 可以看出，去掉深度图处理、最大化探

索、替代点机制后，SPL 从 0.236 分别降为 0.219、0.198、0.213，SR 从 0.441 分别降为 0.407、0.420、0.397。这说明本文所提 3 个改进都不同程度地提高了语义目标导航任务的性能。另外可以看出最大化探索对 SPL 指标的提升贡献最大，再次证明了其较高的探索效率；深度图处理和替代点机制对成功率的贡献较大，说明其解决了部分噪音问题。

表 5 消融实验

Table 5 Ablation experiment

方法	去除深度图处理	去除最大化探索	去除替代点机制	本文方法	
椅子	SPL	0.275	0.281	0.247	0.306
	SR	0.488	0.518	0.439	0.542
	DTG	2.463	2.527	2.600	2.222
床	SPL	0.228	0.190	0.219	0.222
	SR	0.434	0.454	0.429	0.422
	DTG	5.556	5.619	5.829	5.752
植物	SPL	0.038	0.063	0.069	0.121
	SR	0.072	0.120	0.110	0.240
	DTG	7.866	7.342	8.138	7.544
马桶	SPL	0.186	0.143	0.194	0.194
	SR	0.395	0.365	0.400	0.406
	DTG	4.940	4.732	4.952	4.942
电视	SPL	0.175	0.139	0.174	0.169
	SR	0.295	0.274	0.284	0.298
	DTG	7.024	7.469	7.151	7.167
沙发	SPL	0.253	0.243	0.248	0.290
	SR	0.454	0.502	0.454	0.534

	DTG	3.337	3.344	3.479	3.266
	SPL	0.219	0.198	0.213	0.236
平均	SR	0.407	0.420	0.397	0.441
	DTG	4.656	4.684	4.801	4.641

3.5 现实世界实验

本文所提框架成功应用到了绝影 mini-lite 四足机器人(如图6所示)上去。该机器人使用了NVIDIA Xavier NX 主机, 装备了 16 线激光雷达、Intel RealSense D435i 相机等传感器。现实实验使用雷达进行定位, 建图要求与语义目标导航任务一致而未使用雷达。

图7展示了马桶为语义目标类别的一个回合所录制视频的部分截图。从左到右依次展示了: 机器人开始运动; 绕开障碍物; 探索卧室; 发现卧室没



开始运动 绕开障碍物 探索卧室 继续探索 发现并靠近目标

图7 马桶为语义目标类别的导航视频截图

Fig.7 Screenshot of navigation video with toilet as semantic goal

4 结束语

语义目标导航任务研究对于将机器学习成果应用到实际机器人中有着重要意义。本文构建了一个包含语义建图模块、探索模块、路径规划模块的系统框架。针对探索效率低问题, 在探索模块中引入了基于路径规划特点的覆盖范围最大化算法; 针对深度不精准问题, 在语义建图模块中引入了深度图边缘处理与地图纠错机制; 针对路径规划没有考虑语义目标导航任务特点的问题, 在路径规划模块中引入了替代点机制和障碍物概率地图。本文在HM3DSem数据集下进行了实验并应用到了实际四足机器人上。实验结果表明本文所提3个改进明显提高了语义目标导航任务的性能并对现实世界具有良好的泛化性。语义目标导航除了应用于室内场景, 还能应用于隧道、洞穴、城市等室外场景下的巡检、勘探、安防等任务。与室内场景中无需考虑机器人运动特性不同, 在拥有复杂地形的室外场景中进行语义目标导航需要结合机器人特点。为此, 下一步将研究机器人操作和语义目标导航的融合应用。

有目标物体后继续向前探索; 发现并靠近马桶后停止。四足机器人实验证明了本文框架对现实世界具有泛化性。

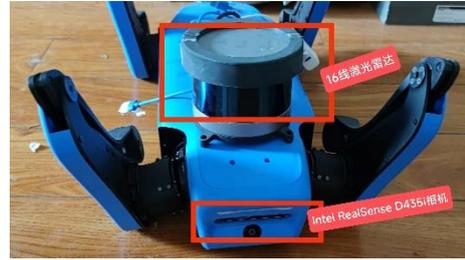


图6 现实实验所用的四足机器人

Fig.6 Quadruped robot used in real-world experiments

参考文献:

- [1] ANDERSON P, CHANG A, CHAPLOT D S, et al. On evaluation of embodied navigation agents[EB/OL]. 2018: arXiv: 1807.06757. <https://arxiv.org/abs/1807.06757.pdf>.
- [2] 王作为, 徐征, 张汝波, 等. 记忆神经网络在机器人导航领域的应用与研究进展[J]. 智能系统学报, 2020, 15(5): 835-846.
- [3] WANG Zuowei, XU Zheng, ZHANG Rubo, et al. Research progress and application of memory neural network in robot navigation[J]. CAAI transactions on intelligent systems, 2020, 15(5): 835-846.
- [4] YADAV K, RAMRAKHYA R, RAMAKRISHNAN S K, et al. Habitat-matterport 3D semantics dataset[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2023: 4927-4936.
- [4] RAMAKRISHNAN S K, GOKASLAN A, WIJMANS E, et al. Habitat-matterport 3D dataset (HM3D): 1000 large-scale 3D environments for

-
- embodied AI[EB/OL]. 2021: arXiv: 2109.08238. <https://arxiv.org/abs/2109.08238.pdf>.
- [5] ZHANG Ji, SINGH S. LOAM: lidar odometry and mapping in real-time[C]//Robotics: Science and Systems X. Robotics: Science and Systems Foundation, 2014: 1-9.
- [6] VIZZO I, GUADAGNINO T, MERSCH B, et al. KISS-ICP: in defense of point-to-point ICP—simple, accurate, and robust registration if done the right way[J]. IEEE robotics and automation letters, 2023, 8(2): 1029-1036.
- [7] 沈斯杰, 田昕, 魏国亮, 等. 基于 2D 激光雷达的 SLAM 算法研究综述[J]. 计算机技术与发展, 2022, 32(1): 13-18, 46.
- SHEN Sijie, TIAN Xin, WEI Guoliang, et al. Review of SLAM algorithm based on 2D lidar[J]. Computer technology and development, 2022, 32(1): 13-18, 46.
- [8] MUR-ARTAL R, TARDÓS 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.
- [9] CVIŠIĆ I, ČESIĆ J, MARKOVIĆ I, et al. SOFT-SLAM: Computationally efficient stereo visual simultaneous localization and mapping for autonomous unmanned aerial vehicles[J]. Journal of field robotics, 2018, 35(4): 578-595.
- [10] 张毅, 陈起, 罗元. 室内环境下移动机器人三维视觉 SLAM[J]. 智能系统学报, 2015, 10(4): 615-619.
- ZHANG Yi, CHEN Qi, LUO Yuan. Three dimensional visual SLAM for mobile robots in indoor environments[J]. CAAI transactions on intelligent systems, 2015, 10(4): 615-619.
- [11] 王霞, 左一凡. 视觉 slam 研究进展[J]. 智能系统学报, 2020, 15(5): 825-834.
- WANG Xia, ZUO Yifan. Advances in visual SLAM research [J]. CAAI transactions on intelligent systems, 2020, 15(5): 825-834.
- [12] HART P E, NILSSON N J, RAPHAEL B. A formal basis for the heuristic determination of minimum cost paths[J]. IEEE transactions on systems science and cybernetics, 1968, 4(2): 100-107.
- [13] KAVRAKI L E, SVESTKA P, LATOMBE J C, et al. Probabilistic roadmaps for path planning in high-dimensional configuration spaces[J]. IEEE transactions on robotics and automation, 1996, 12(4): 566-580.
- [14] 李娟, 张韵, 陈涛. 改进 RRT 算法在未知三维环境下 AUV 目标搜索中的应用[J]. 智能系统学报, 2022, 17(2): 368-375.
- LI Juan, ZHANG Yun, CHEN Tao. Application of the improved RRT algorithm to AUV target search in an unknown 3D environment[J]. CAAI transactions on intelligent systems, 2022, 17(2): 368-375.
- [15] WARREN C W. Global path planning using artificial potential fields[C]//Proceedings, 1989 International Conference on Robotics and Automation. Piscataway: IEEE, 2002: 316-321.
- [16] 朱少凯, 孟庆浩, 金晟, 等. 基于深度强化学习的室内视觉局部路径规划[J]. 智能系统学报, 2022, 17(5): 908-918.
- ZHU Shaokai, MENG Qinghao, JIN Sheng, et al. Indoor visual local path planning based on deep reinforcement learning[J]. CAAI transactions on intelligent systems, 2022, 17(5): 908-918.
- [17] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Playing atari with deep reinforcement learning[EB/OL]. 2013: arXiv: 1312.5602. <https://arxiv.org/abs/1312.5602.pdf>.
- [18] MNIH V, BADIA A P, MIRZA M, et al. Asynchronous methods for deep reinforcement learning[EB/OL]. 2016: arXiv: 1602.01783. <https://arxiv.org/abs/1602.01783.pdf>.
- [19] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal policy optimization algorithms[EB/OL]. 2017: arXiv: 1707.06347. <https://arxiv.org/abs/1707.06347.pdf>.
- [20] LILLICRAP T P, HUNT J J, PRITZEL A, et al. Continuous control with deep reinforcement learning[EB/OL]. 2015: arXiv: 1509.02971. <https://arxiv.org/abs/1509.02971.pdf>.
- [21] 黄东晋, 蒋晨凤, 韩凯丽. 基于深度强化学习的三维路径规划算法[J]. 计算机工程与应用, 2020, 56(15): 30-36.
- HUANG Dongjin, JIANG Chenfeng, HAN Kaili. 3D path planning algorithm based on deep reinforcement learning[J]. Computer engineering and applications, 2020, 56(15): 30-36.
- [22] FAUST A, OSLUND K, RAMIREZ O, et al. PRM-RL: long-range robotic navigation tasks by

-
- combining reinforcement learning and sampling-based planning[C]//2018 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE, 2018: 5113-5120.
- [23] CHAPLOT D S, GANDHI D P, GUPTA A, et al. Object goal navigation using goal-oriented semantic exploration[J]. *Advances in neural information processing systems*, 2020, 33: 4247-4258.
- [24] YANG Wei, WANG Xiaolong, FARHADI A, et al. Visual semantic navigation using scene priors[EB/OL]. 2018: arXiv: 1810.06543. <https://arxiv.org/abs/1810.06543.pdf>.
- [25] MAKSYMETS O, CARTILLIER V, GOKASLAN A, et al. THDA: treasure hunt data augmentation for semantic navigation[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2022: 15354-15363.
- [26] MAYO B, HAZAN T, TAL A. Visual navigation with spatial attention[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2021: 16893-16902.
- [27] CHAPLOT D S, GANDHI D, GUPTA S, et al. Learning to explore using active neural SLAM[EB/OL]. 2020: arXiv: 2004.05155. <https://arxiv.org/abs/2004.05155.pdf>.
- [28] HE Kaiming, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//2017 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2017: 2980-2988.
- [29] SETHIAN J A. A fast marching level set method for monotonically advancing fronts[J]. *Proceedings of the national academy of sciences of the United States of America*, 1996, 93(4): 1591-1595.
- [30] WIJMANS E, KADIAN A, MORCOS A, et al. DD-PPO: learning near-perfect PointGoal navigators from 2.5 billion frames[EB/OL]. 2019: arXiv: 1911.00357. <https://arxiv.org/abs/1911.00357.pdf>.
- [31] RAMAKRISHNAN S K, JAYARAMAN D, GRAUMAN K. An exploration of embodied visual exploration[J]. *International journal of computer vision*, 2021, 129(5): 1-34.
- [32] GEORGAKIS G, BUCHER B, SCHMECKPEPER K, et al. Learning to map for active semantic goal navigation[EB/OL]. 2021: arXiv: 2106.15648. <https://arxiv.org/abs/2106.15648.pdf>.
- [33] ZHU Minzhao, ZHAO Binglei, KONG Tao. Navigating to objects in unseen environments by distance prediction[C]//2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE, 2022: 10571-10578.
- [34] LIANG Yiqing, CHEN Boyuan, SONG Shuran. SSCNav: confidence-aware semantic scene completion for visual semantic navigation[C]//2021 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE, 2021: 13194-13200.
- [35] SAVVA M, KADIAN A, MAKSYMETS O, et al. Habitat: a platform for embodied AI research[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2020: 9338-9346.
- [36] SZOT A, CLEGG A, UNDERSANDER E, et al. Habitat 2.0: training home assistants to rearrange their habitat[EB/OL]. 2021: arXiv: 2106.14405. <https://arxiv.org/abs/2106.14405.pdf>.
- [37] YAMAUCHI B. A frontier-based approach for autonomous exploration[C]//Proceedings 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA'97. 'Towards New Computational Principles for Robotics and Automation'. Piscataway: IEEE, 2002: 146-151.
- [38] RAMRAKHYA R, UNDERSANDER E, BATRA D, et al. Habitat-web: learning embodied object-search strategies from human demonstrations at scale[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 5163-5173.