

# 体现人工智能的调查：从模拟器到研究任务

段家飞、余参孙、回李谭、朱红、栗谭

摘要——从“互联网人工智能”时代到“具体化人工智能”时代的范式转变，在这个时代，人工智能算法和代理不再主要由互联网的网策划的图像、视频或文本数据集中学习。相反，他们通过与环境的互动，从一种类似于人类的自我中心感知中学习。因此，对具身人工智能模拟器的需求显著增长，以支持各种具身人工智能研究任务。这种对具身人工智能日益增长的兴趣有利于对人工智能一般智能（AGI）的更大追求，但目前还没有对这一领域进行当代和全面的调查。本文旨在为具身人工智能的模拟器和研究提供两个百科全书式的七个调查。通过评估9个当前的9个具身人工智能模拟器和我们提出的7个调查。最后，本文介绍了具身人工智能的三个主要研究任务——视觉探索、视觉导航和具身问答（QA），涵盖了最先进的方法、评估指标和数据集。最后，通过调查发现的新见解，为任务选择提供建议，并为未来的发展方向提供建议。

索引术语—体现人工智能，计算机视觉，3D模拟器。

## I. 介绍

RECENT在深度学习、强化学习、计算机图形学和机器人技术方面的进步，引起了人们对开发通用人工智能系统越来越大的兴趣。因此，从互联网上学习的“互联网人工智能”转向“具体化人工智能”，使人工智能能够通过与其周围环境的互动来学习。体现人工智能是一种信念，即真正的智能可以从代理与环境[1]的互动中产生。但就目前而言，具体化的人工智能是关于整合

这项工作得到了新加坡科学、技术和研究机构（A\*STAR）根据其AME项目资助计划（奖励#A18A2b0046）和新加坡国家研究基金会根据其NRF ISF联合呼叫（奖励NRF2015-NRF-ISF001-2541）的支持。

J. 段曾就职于新加坡南阳科技大学电气与电子工程学院，新加坡639798（电子邮件：duan0038@e.ntu.edu.sg）。

S. 余曾就读于新加坡科技设计大学，新加坡487372（电子邮件：samsonyu@sutd.edu.sg）。

H. L. 谭，H. 朱和C. 谭在新加坡信息通信研究所工作，新加坡138632（电子邮件：{段加飞，香港，zhuh, cheston-tan}@i2r.a-star.edu.sg）。

手稿接受于2021年12月4日，IEEE-TETCI。©2021 IEEE。允许个人使用本材料。IEEE许可必须获得所有其他用途，在任何当前或未来的媒体，包括转载/转载这些材料广告或促销目的，创建新的集体作品，转售或再分配服务器或列表，或重用的任何版权组件在其他作品。

传统的智能概念，从视觉、语言和推理到一个人体的体现，以帮助解决虚拟环境中的人工智能问题。

人们对具身人工智能的兴趣日益增长，这使得旨在忠实复制物理世界的具身人工智能模拟器取得了重大进展。这些模拟世界可以作为虚拟测试平台来训练和测试具体化的人工智能框架

在将它们部署到现实世界之前。这些具体化的人工智能模拟器还促进了基于任务的数据集[2]，[3]的收集，这在现实世界中收集是很乏味的，因为它需要大量的人工劳动来复制与虚拟世界中相同的设置。虽然在具体化AI[4]–[6]领域有几篇调查论文，但它们大多是过时的，因为它们是在现代深度学习时代之前发表的，开始于2009年[7]–[10]。据我们所知，只有一篇关于具体化导航[11]的调查论文。

为了解决当代关于具身人工智能新兴的综合调查论文的稀缺性，我们提出了这篇关于具身人工智能的调查论文，从其模拟器到研究任务。本文涵盖了过去四年开发的9个具体的人工智能模拟器：深度思维实验室[12]，AI2-THOR [13]，[14]小屋，虚拟家庭[15]，VR厨房[16]，生境-Sim[17]，iGibson[18]，SAPIEN[19]，和ThreeDWorld [20]。所选择的模拟器是为通用智能任务设计的，不像游戏模拟器[21]只用于训练强化学习代理。这些具体化的人工智能模拟器在计算机模拟中提供了现实世界的真实表示，主要是为房间或公寓提供了对环境的一些形式的约束。这些模拟器中的大多数最小限度地包括一个物理引擎、Python API和可以在环境中进行控制或操作的人工代理。具身人工智能模拟器已经产生了一系列潜在的具身人工智能研究任务，如视觉探索、视觉导航和具身QA等。我们将专注于这三个任务，因为大多数现有的论文[11]、[22]、[23]要么关注这些任务，要么利用

为这些任务引入的模块，为更复杂的任务建立模型，如视听导航。这三个任务也在日益复杂的问题上联系在一起。视觉探索在视觉导航[22]，[24]中是一个非常有用的组件，用于现实情况下的[25]，[26]，而体现的QA进一步涉及到建立在视觉和语言导航之上的复杂QA功能。由于语言是一种常见的形态，而视觉QA是人工智能中的流行任务，因此具身QA是具身人工智能的自然方向。这三个任务

本文所讨论的已经在9个提出的具体化人工智能模拟器中至少一个实现。然而, 本文将不涉及Sim2Real [27]–[29]和物理世界中的机器人技术。

这些模拟器是基于计算机视觉和模式识别会议(CVPR)年度体现AI研讨会[30]的体现AI挑战中的具体化人工智能模拟器选择的。研究任务然后来自于这些模拟器的直接引用。

为此, 我们将通过回顾人工智能模拟器的当代和全面的调查, 并将模拟器的发展从模拟器到研究。在第一节中, 本文概述了本调查论文的概述结构。在第二节中, 本文对9个具体化的人工智能模拟器进行了基准测试, 以理解它们提供现实主义、可伸缩性、交互性, 从而在具体化的人工智能研究中使用。最后, 在第三部分仿真器的基础上, 本文探讨了具身人工智能的三个主要研究任务——视觉探索、视觉导航和具身问题回答(QA), 涵盖了最先进的方法、评估和数据集。最后, 本文将在第四节中建立模拟器、数据集和研究任务以及具体化人工智能模拟器和研究中的现有挑战之间的相互联系。本调查论文全面探讨了具体化人工智能的新兴领域, 并进一步揭示了具体化人工智能的新见解和挑战。此外, 通过本文, 我们试图利用人工智能研究人员选择理想的具身人工智能模拟器为他们感兴趣的研究任务。

## II. 用于实体人工智能的模拟器

在本节中, 我们将在补充材料中介绍具体化AI模拟器的背景, 并比较具体化AI模拟器的特征并在第二节-A节中进行了讨论。

### A. 体现人工智能模拟器

本节介绍了九个具体的人工智能模拟器的背景: 深度思维实验室, AI2-THOR, SAPIEN, 虚拟家庭, VR厨房, 三, dWorld, 小屋, igi布森, 和栖息地-Sim。读者可以参考补充材料, 以了解更多关于各自的模拟器的细节。在本节中, 本文将基于7个技术特征, 综合比较9个具身人工智能模拟器。

参考[13]、[20]、[31], 这七个技术特征被选择作为评估具体化人工智能模拟器的主要特征, 因为它们涵盖了准确复制环境、交互和物理世界状态所需的基本方面, 从而为化身测试智能提供了合适的测试平台。参考表一, 这七个特性分别是: 环境、物理、对象类型、对象属性、控制器、动作和多代理。

环境: 构建实体人工智能模拟器环境主要有两种方法: 基于游戏的场景构建(G)和基于世界的场景构建(W)。参考图。1、构建了基于游戏的场景

而基于世界的场景则是由对对象和环境的真实世界扫描来构建的。一个完全由3D资产构建的3D环境通常具有内置的物理特性和对象类, 与来自真实世界扫描的环境的3D网格相比, 这些特性被分割得很好。3D资产的清晰对象分割使其很容易建模为具有可移动关节的铰接对象, 如PartNet [32]中提供的3D模型。相比之下, 真实世界的环境和对象扫描提供了更高的保真度和更准确的真实世界表示, 有助于将代理性能更好地从模拟转移到真实世界。如表1所示, 除了栖息地-sim和iGibson之外, 大多数模拟器都有基于游戏的场景, 因为基于世界的场景构建显然需要更多的资源。

物理: 模拟器不仅要构建现实的环境, 还要构建代理和物体之间的现实交互, 或者真实的物体来模拟真实世界的物理属性。我们研究模拟器的物理特征, 我们将其大致划分为基本物理特征(B)和高级物理特征(A)。参考图。2、基本的物理特征包括碰撞、刚体动力学和重力建模, 而先进的物理特征包括布, fluid和软体物理学。由于大多数具体化的人工智能模拟器使用内置的物理引擎构建基于游戏的场景, 它们配备了基本的物理特性。另一方面, 对于像ThreeDWorld这样的模拟器, 他们的目标是了解复杂的物理环境如何塑造环境中的人工代理的决策, 他们配备了更先进的物理能力。对于专注于基于交互式导航的任务的模拟器来说, 基本的物理特性通常是非常有效的。

对象类型: 如图所示。用于创建模拟器的对象有两个主要源。第一种类型是数据集驱动的环境, 其中对象主要来自现有的对象数据集, 如SUNCG [33]数据集、矩阵端口3D数据集[34]和Gibson数据集[35]。第二种类型是资产驱动的环境, 其中的对象来自网络, 如Unity 3D游戏资产存储。这两个源之间的区别是对象数据集的可持续性。数据集驱动的对象比资产驱动的对象进行收集的成本更高, 因为任何人都可以在线为3D对象模型做出贡献。然而, 要确保资产驱动对象中的三维对象模型的质量, 要比确保数据集驱动对象中的质量更困难。根据我们的回顾, 基于游戏的具体化人工智能模拟器更有可能从资产存储中获得对象数据集, 而基于世界的模拟器则倾向于从现有的三维对象数据集导入对象数据集。

对象属性: 一些模拟器只启用具有基本交互性的对象, 如碰撞。高级模拟器使具有更多正粒度交互性的对象, 如多状态更改。例如, 当一个苹果被切片时, 它会经历一个状态变化为苹果片。因此, 我们将这些不同层次的对象交互分类为具有可交互对象(I)和多状态对象(M)的模拟器。参考表一, 一些模拟器, 如AI2-THOR和VRKitchen, 可以实现多种状态变化,

表i  
具体化的人工智能模拟器的总结。环境：基于游戏的场景构建(G)和基于世界的场景构建(W)。物理：基本物理特性(B)和高级物理特性(A)。对象类型：数据集驱动的环境(D)和对象资产驱动的环境(O)。对象属性：可交互的对象(I)和多状态对象(M)。控制器：直接PYTHON API控制器(P)、虚拟机器人控制器(R)和虚拟现实控制器(V)。操作：导航(N)，原子动作(A)和人机交互(H)。多智能体：基于虚拟角色的(AT)和基于用户的(U)。这七个特性可以是进一步分为三个次要评价特征：真实性、可伸缩性和交互性。

年	浮雕模拟器	人工智能	环境 (现实主义)	物理学 (现实主义)	对象类型 (可扩展性)	对象财产 (相互作用)	控制器 (相互作用)	行动 (相互作用)	多代理 (相互作用)
2016	深度思维实验室		G	-	-	-	P, R	N	-
2017	ai2托尔		G	B	O	I, M	P, R	A, N	U
2018	小木屋		G	B	O	I, M	P	A, N	-
2018	虚拟主页		G	-	O	I, M	R	A, N	-
2019	VR厨房		G	B	O	I, M	P, V	A, N, H	-
2019	生境Sim		W	-	D	-	-	N	-
2019	我吉布森		W	B	D	I	P, R	A, N	U
2020	萨皮恩		G	B	D	I, M	P, R	A, N	-
2020	威胁DWorld		G	B, A	O	I	P, R, V	A, N, H	在

表二  
在环境配置、仿真引擎、技术规范和呈现性能。

浮雕人工智能模拟器	环境结构	模拟引擎	技术规范	渲染性能
深度思维实验室	定制的环境	地震II竞技场发动机	6核英特尔至强CPU和一个NVIDIA Quadro K600 GPU	158 fps/thread
ai2托尔	120间，4个类别	Unity三维引擎	Intel (R) 至强(R) CPU E5-2620 v4和NVIDIA Titan X	240 fps/thread
小木屋	58间，10间房子	Unity三维引擎	-	-
虚拟主页	6间公寓，带多个联合房间	Unity三维引擎	-	定制的帧率
VR厨房	16个厨房	虚幻引擎4	Intel (R) 酷睿 (TM) i7-7700K处理器和NVIDIA Titan X	15 fps/thread
生境Sim	多个数据集	-	Xeon E5-2690 v4 CPU和英伟达泰坦Xp GPU	10,000 fps/thread
我吉布森	吉布森V1	-	现代GPU	1000 fps/thread
萨皮恩	定制的环境	PhysX物理引擎和ROS	英特尔i7-8750 CPU和英伟达GeForce RTX 2070 GPU	700 fps/thread
威胁DWorld	定制的环境	Unity三维引擎	Intel i7-7700K GPU; NVIDIA GeForce GTX 1080	168 fps/thread

提供一个平台来理解物体在现实世界中采取行动时将如何反应和改变它们的状态。

控制器：参考图。4、用户与模拟器之间有不同类型的控制器接口，从直接的Python API控制器(P)和虚拟机器人控制器(R)到虚拟现实控制器(V)。机器人实施例允许现有的现实世界机器人的虚拟交互，如通用机器人5 (UR5) 和涡轮机机器人V2，并可以直接使用ROS接口进行控制。虚拟现实控制器接口提供了更多身临其境的人机交互，并便于使用现实世界的对手进行部署。例如，iGibson和AI2-THOR等模拟器主要为视觉导航设计，也配备了虚拟机器人控制器，以便在现实世界中部署，如iGibson的Castro [36]和RoboTHOR [37]。

动作：在人工智能模拟器中，人工代理的动作能力的复杂性不同，从只能执行主要导航操作到通过虚拟现实界面进行更高级的人机动作。本文将它们分为机器人操作：导航(N)、原子动作(A)以及人机交互作用(H)。导航是最低的一层，也是所有具体化的人工智能模拟器[38]的共同特征。它是由代理在其虚拟环境中导航的能力所决定的。原子作用为人工代理提供了一种对感兴趣的对象执行基本的离散操作的方法，这在大多数具体化的人工智能模拟器中都可以找到。人机交互是一种方法

虚拟现实控制器的结果，因为它使人类能够控制虚拟代理，以学习和交互与模拟世界的实时[16]。大多数基于大规模导航的模拟器，如AI2-THOR、iGibson和习惯sim，倾向于有导航、原子动作和ROS [13]、[17]、[35]，使它们能够在执行点导航或对象导航等任务时更好地控制和操作环境中的对象。另一方面，模拟器如ThreeDWorld和VR厨房[16]，[20]属于人机交互类别，因为它们的构建提供了一个高度真实的基于物理的模拟和多种状态变化。这只有在人机交互中才有可能实现，因为当与这些虚拟对象交互时，需要人类水平的灵活性。

多智能体：参考表I，只有少数模拟器，如AI2-THOR，iGibson和ThreeDWorld，配备了多智能体设置，因为目前涉及多智能体强化学习的研究很少。一般来说，模拟器需要丰富的对象内容，才能构建这样的多代理特性，用于对抗和协作训练[39]，[40]。由于缺乏多智能体支持的模拟器，在这些具体化的人工智能模拟器中使用多智能体特性的研究任务较少。

对于基于多智能体强化学习的训练，它们目前仍在OpenAI健身房环境[41]中完成。有两种不同的多智能体设置。首先是ThreeDWorld [20]中基于角色的(AT)多代理，它允许人工代理和模拟之间的交互

化身。第二种是AI2THOR [13]中基于用户的(U)多代理,它可以扮演双学习网络的角色,在模拟中与其他人工代理交互学习,实现共同的任务[42]。



图1. 基于游戏的场景(G)和基于世界的场景(W)之间的比较。基于游戏的场景(G)关注于由三维对象资产构建的环境,而基于世界的场景(W)则是基于对环境的真实扫描而构建的。

### B. 嵌入式人工智能模拟器的比较

基于这七个特征和艾伦人工智能[31]研究所对具身人工智能的研究,我们提出了一套针对模拟器的二级评估特征。它包括三个关键特性:真实性、可伸缩性和交互性,如表I所示。三维环境的现实性可以归因于模拟器的环境和物理特性。环境模拟真实世界的物理外观,而物理模拟真实世界中复杂的物理属性。三维环境的可伸缩性可以归因于对象类型。该扩展可以通过为数据集驱动的对象收集更多的真实世界的3D扫描图,或为资产驱动的对象购买更多的3D资产来完成。交互性归因于对象属性、控制器、动作和多代理。

基于具体化AI模拟器的二次评价特征,得出了表1和图中的7个主要特征。6、具有上述三个次要特征的模拟器(e.g. AI2-THOR, iGibson和生境-sim)更受欢迎,并广泛应用于不同范围的具身人工智能研究任务。此外,还对所有具体化的人工智能模拟器进行了全面的定量比较,从而比较了各模拟器的环境状况和技术性能。环境协调特性在很大程度上依赖于模拟器创建者建议的应用程序,而其他特性如技术规格和渲染性能很大程度上是由于用于创建所使用的模拟引擎。与其他模拟器相比,AI2-THOR具有最大的环境条件,而生境-sim和iGibson是在图形渲染性能方面表现最好的两名。表二所示的定量性能基准进一步证明了这三种具体化人工智能模拟器的优越性和复杂性。这些对具体的人工智能模拟器的比较进一步加强了本文建立的七个主要评价指标和三个次要评价的重要性,以帮助选择理想的模拟器。

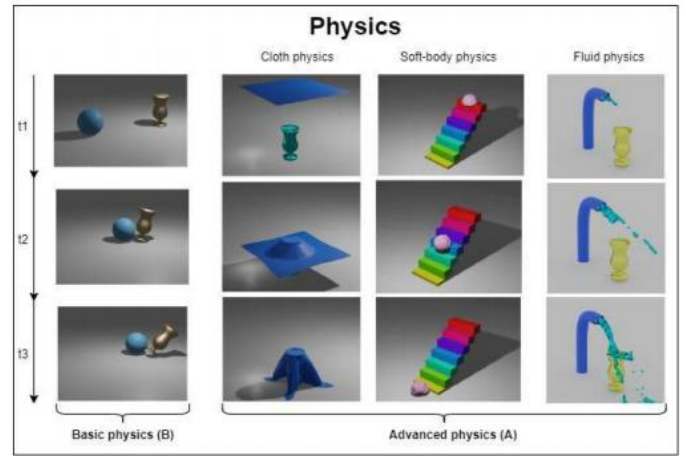


图2. 比较基本的物理特征,如刚体和碰撞(B)和高级物理特征(A),其中包括布,软身体和fluid物理学。

### III. 实体人工智能研究

在本节中,我们将讨论依赖于前一节中调查的9个具体化人工智能模拟器的各种人工智能研究任务。最近具身人工智能研究的增加有多种动机。从认知科学和心理学的角度来看,体现假说[1]认为,智力来自于与环境的相互作用,并且是感觉运动活动[66]的结果。直觉上,人类并不仅仅通过“互联网人工智能”范式来学习,其中大多数体验都是随机的和被动的。e. 外部管理)。人类也通过积极的感知、运动、互动和交流来学习。从人工智能的角度来看,与传统的学习方法相比,当前的具体化人工智能的研究任务允许对[44]的映射和导航,以及对传感器噪声的鲁棒性更强。嵌入式人工智能还支持存在性和可能的更高的性能,因为像深度、语言[59]和音频[67]等各种模式可以通过基于学习的方法轻松集成。

具身人工智能研究任务的三种主要类型分别是视觉探索、视觉导航和具身QA。我们将专注于这三个任务,因为大多数现有的具体化人工智能论文要么关注这些任务,要么利用为这些任务引入的模块来为更复杂的任务构建模型,如视听导航。随着从探索到QA的发展,任务的复杂性也会增加。在转向视觉导航和非体现QA之前,我们将从视觉探索开始。这些任务都构成了下一个任务(s)的基础,形成了一个具体化的人工智能研究任务的金字塔结构,如图所示。5、进一步暗示了具体化人工智能的自然方向。我们将突出每个任务的重要方面,从摘要、方法、评估指标开始。这些任务细节见表三。



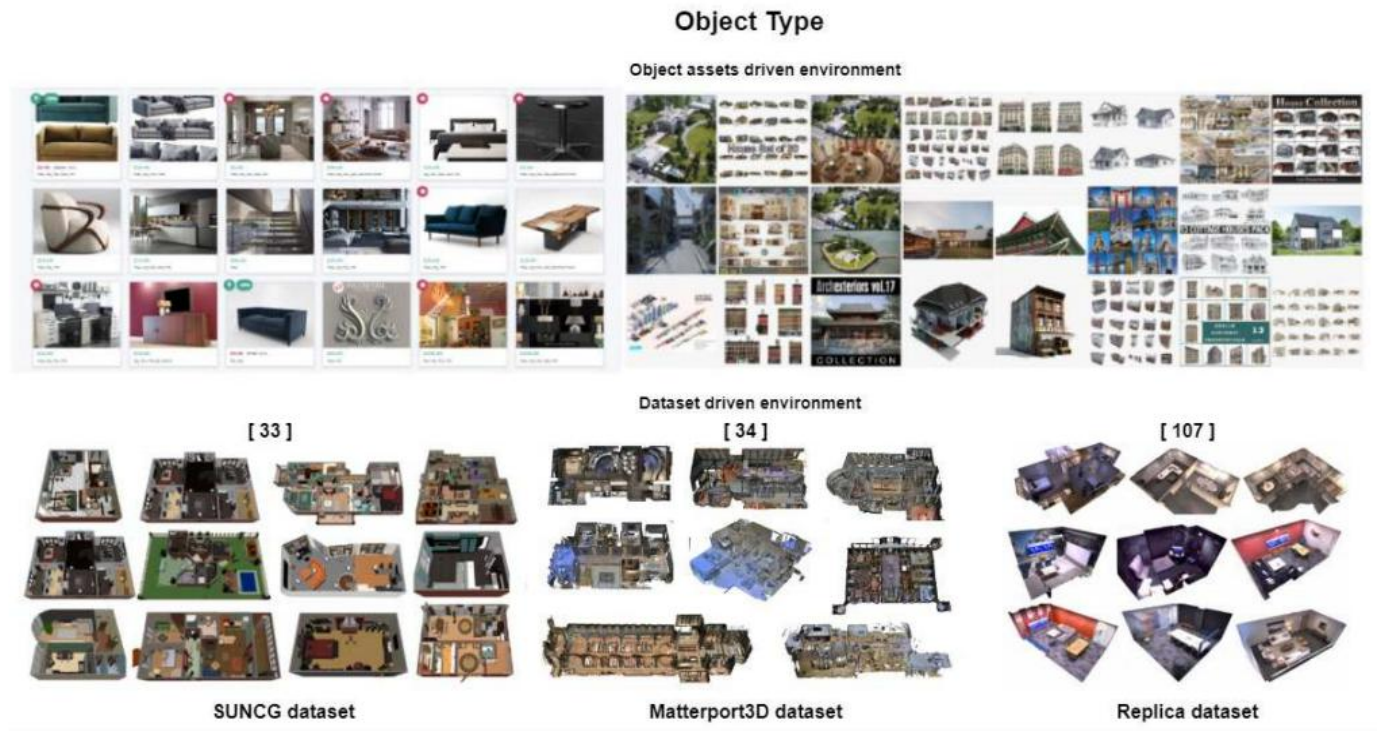


图3. 基于从资产市场获得的三维对象，构建了由三维对象数据集构建的数据集驱动环境(D)与对象资产驱动环境(O)之间的比较。

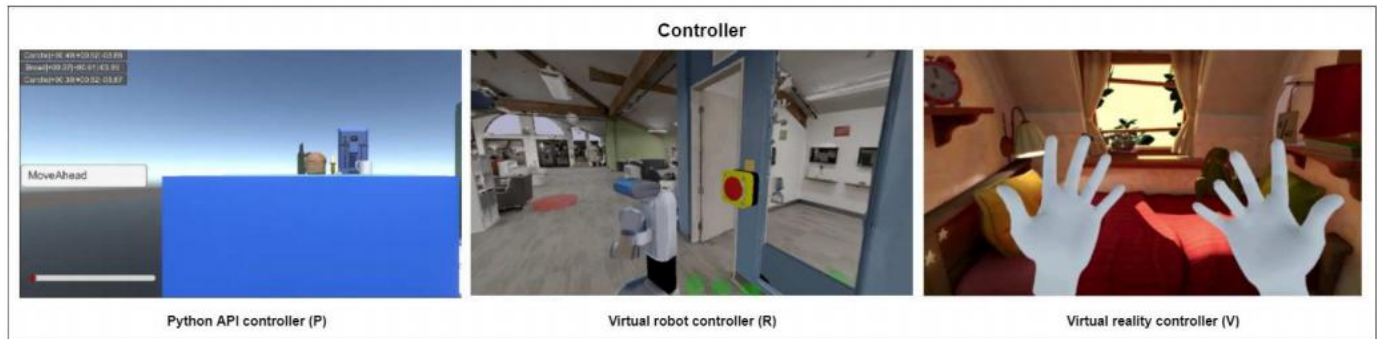


图4. 比较直接Python API控制器(P)，机器人实施例(R)，它指的是具有虚拟副本的真实世界的机器人，最后虚拟现实控制器(V)。

#### A. 视觉探索

在视觉探索[24], [68], 代理收集关于3D环境的信息, 通常通过运动和感知, 更新其内部环境模型[11], [22], 这可能对下游任务有用, 如视觉导航[24], [25], [69]。其目的是尽可能有效地做到这一点。用尽可能少的步骤)。内部模型可以是拓扑图[26]、语义图[46]、占用图[45]或空间记忆[70]、[71]等形式。这些基于地图的体系结构可以捕获几何学和语义, 与反应性和递归神经网络策略[72]相比, 允许更高效的策略学习和规划[45]。视觉探索通常是在导航任务之前或同时进行的。在第一个情况下, 视觉探索将内部记忆作为先验

这对于下游导航任务中的路径规划很有用。代理可以在一定的预算范围内自由地探索环境(e.g. 有限的步骤数)之前, 开始导航[11]。在后一种情况下, 代理在导航一个看不见的测试环境[48]、[73]、[74]时构建映射, 这使其与下游任务更紧密地集成。在本节中, 我们基于现有的视觉探索调查论文[22], [24], 包括更近期的工作和方向。在经典的机器人技术中, 探索是通过被动或主动的同步定位和映射(SLAM) [24], [45]来构建环境的地图。然后, 将此地图与导航任务的定位和路径规划一起使用。SLAM是研究非常充分的[75], 但纯几何方法有改进的空间。由于它们依赖于传感器, 它们很容易受到测量噪声[24]的影响, 并且需要广泛的正弦调谐。另一方面, 学习-

表三

对具体化的人工智能研究任务的总结。评价指标: 目标访问量 (ATV), 下游任务 (D), 成功与否 (TL / EL), 成功率 (SR), 路径长度比率 (PLR), 甲骨文成功率 (OSR), 轨迹/事件长度进行加权 (TL / EL), 到成功的距离/导航错误 (DTS / NE / dt), 目标进展 (GP / d), ORACLE PATH成功率 (OPSR), 在一个事件中的任何一点到目标的最小距离 (d分), 事件代理结束导航的百分比最大事件长度 (%stop), 代理在包含目标对象的房间中终止的问题的百分比 (%rt), 代理至少一次进入包含目标OBJECT的房间的问题的百分比 (%re), 交叉点目标对象 (IOU), 基于IOU的HIT精度 (ht), qa预测中地面真实答案的平均排名 (MR) 和Qa精度 (ACC)。

任务	方法和类别	出版	年	模拟器	数据集	评价指标
视觉探索	好奇号	夏普特等人。[43]	2020	生境Sim	Matterport3D, 吉布森V1	装甲运输车
	覆盖	夏普特等人。[44]	2020	生境Sim	Matterport3D, 吉布森V1	ATV, d
	重构	罗摩克里希南等人。[45]	2020	生境Sim	Matterport3D, 吉布森V1	ATV, d
		罗摩克里希南等人。[22]	2020	生境Sim	匹配端口3D	ATV, d
视觉导航	点导航	纳拉西姆汉等。[46]	2020	生境Sim	匹配端口3D	ATV, d
		维曼斯等人。[47]	2019	生境Sim	Matterport3D, 吉布森V1	SPL, SR
		乔治卡基斯等。[48]	2019	生境Sim	匹配端口3D	SR, PLR
		Ye等人。[49]	2020	生境Sim	吉布森V1	SPL, SR
		夏普特等人。[44]	2020	生境Sim	Matterport3D, 吉布森V1	SPL, SR
		罗摩克里希南等人。[45]	2020	生境Sim	Matterport3D, 吉布森V1	SPL, SR
		罗摩克里希南等人。[22]	2020	生境Sim	匹配端口3D	SPL
		纳拉西姆汉等。[46]	2020	生境Sim	匹配端口3D	SPL, SR
		克劳迪娅等。[50]	2020	我吉布森	吉布森V1	高级的
		沃斯基曼等人。[51]	2019	ai2托尔	-	SPL, SR
	对象导航	坎帕里等人。[52]	2020	生境Sim	匹配端口3D	SPL, SR, DTS
		杜等人。[53]	2020	ai2托尔	-	SPL, SR
		夏普特等人。[54]	2020	生境Sim	Matterport3D, 吉布森V1	SPL, SR, DTS
		沈等人。[55]	2020	我吉布森	吉布森V1	高级的
		Wahid等人。[56]	2020	-	吉布森V1	SPL, SR
		杨等人。[57]	2020	ai2托尔	-	SPL, SR
	视觉与语言导航	安德森等人。[58]	2018	-	房间到房间	SR, OSR, TL, NE
		朱等人。[59]	2020	-	房间到房间	SPL, SR, OSR, TL, NE
		朱等人。[60]	2020	-	协同视觉和对话框导航	SR, OSR, GP, OPSR
		Das等人。[61]	2018	-	设备质量分析	dt, d, d <sub>△</sub> 分, %rt, %re, %停止, 先生
体现问题回答	问题回答	Das等人。[62]	2018	-	设备质量分析	dt <sub>△</sub> , d, Acc
	多目标问题回答	余等人。[63]	2019	-	MT-EQA	dt, d, %rt <sub>△</sub> T, %stop, IoU, ht, Acc
	交互式问答	戈登等人。[64]	2018	ai2托尔	IQUAD V1	EL, Acc
		Tan等人。[65]	2020	ai2托尔	IQUAD V1	EL, Acc

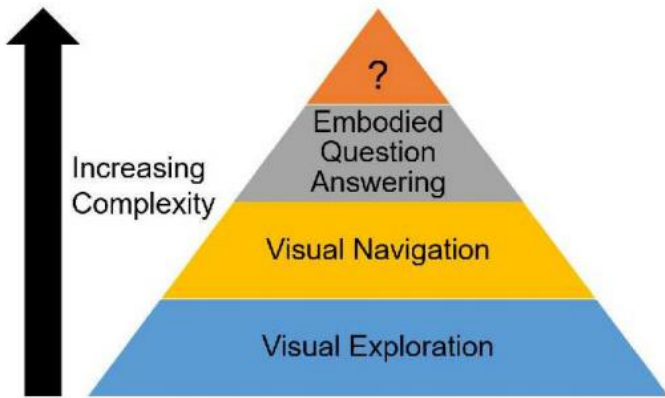


图5. 金字塔层次结构的各种具体化人工智能研究任务的复杂性。

通常使用RGB和/或深度传感器的基于方法对[24], [44]的噪声更稳健。此外, 视觉探索中基于学习的方法允许人工代理合并语义理解(例如。环境中的对象类型)[45], 并推广其对以前看到的环境的知识, 以帮助以无监督的方式理解新环境。这减少了对人类的依赖, 从而提高了熟练程度。学习以映射的形式创建有用的环境内部模型可以提高代理的性能[45], 无论它是在之前完成的(即。未指定的下游任务)或与下游任务并发执行。在代理必须探索随着时间[76]动态展开的新环境的情况下, 智能探索也会特别有用, 比如救援机器人和深海探索机器人。

1) 方法: 在本节中, 视觉探索中的非基线方法通常被形式化为部分观察到的马尔可夫决策过程(POMDPs)[77]。POMDP可以由一个7元组(S、A、T、R、O、V)表示, 状态空间S、动作空间a、转移分布T、奖励函数R、观察空间、观察分布O、折扣因子 $V \in [0, 1]$ 。一般来说, 这些方法被视为POMDP[22]中的一种特殊的奖励功能。

基线视觉探索有一些常见的基线[22]。对于随机动作[17], 代理从所有动作的均匀分布中采样。对于正向行动, 它是选择正向行动。对于正向动作+, 代理选择正向动作, 但如果发生碰撞则左转。对于边界探索, 它使用地图[24], [78]迭代地访问自由空间和未探索空间之间的边缘。

好奇。在好奇的方法中, 代理寻找难以预测的状态。将预测误差作为强化学习[79], [80]的奖励信号。这关注的是内在奖励和动机, 而不是来自环境的外部奖励, 这在外在奖励是稀疏[81]的情况下是有益的。通常有一个前向动力学模型, 使损失最小化:  $L(s^t + 1, s_t + 1)$ 。在这种情况下,  $s^t + 1$ 如果代理采取操作 $a_t$ 当它处于 $s$ 状态时 $t$ , 而 $s_t + 1$ 是代理最终进入的下一个状态。在最近的工作[79]中已经列出了对好奇心的实际考虑, 例如使用近端策略优化(PPO)进行策略优化。在最近的工作[43]中, 好奇心已经被用来生成更高级的地图, 比如语义地图。随机性对好奇心方法提出了一个严重的挑战, 因为正向动力学模型可以利用随机性[79]来处理高预测误差(即。高的

报酬这可能是由于“噪声电视”问题或代理执行动作[81]时的噪声等因素引起的。一个提出的解决方案是使用一个逆动力学模型[68]来估计作用 $a_{t-1}$ 被代理从它以前的状态转移到 $s_{t-1}$ 到其当前状态 $s_t$ ，这可以帮助代理了解其操作在环境中可以控制什么。虽然这种方法试图解决由环境引起的随机性，但它在解决由代理的行为导致的随机性方面可能是无效的。一个例子是代理使用遥控器随机改变电视频道，允许它在没有进展的情况下积累奖励。为了特别地解决这个更具挑战性的问题，最近提出了一些方法。随机蒸馏网络[82]是一种预测随机初始化神经网络输出的方法，因为答案是其输入的确定性函数。另一种方法是通过不一致[81]进行探索，其中激励主体去探索在前向动力学模型集合的预测之间存在不一致或方差最大的行动空间。该模型收敛于均值，这减少了集成的方差，并防止了它陷入随机性陷阱。

覆盖在覆盖方法中，代理试图使它直接观察到的目标数量最大化。通常，这是在[22]，[24]，[44]环境中看到的区域。由于代理使用以自我为中心的观察，它必须基于可能阻塞的3D结构进行导航。最近的一种方法结合了经典的和基于学习的方法[44]。它使用分析路径规划器和一个学习过的SLAM模块来维护一个空间地图，以避免高样本培训端到端政策所涉及的复杂性。这个方法还包括噪声模型，以提高物理现实主义，以推广到现实世界的机器人。另一个最近的工作是一个场景记忆转换器，它在策略网络[72]中的场景记忆上使用了来自变形器模型[83]的自注意力机制。场景记忆嵌入和存储了所有遇到的观察结果，与需要归纳偏差的类似地图的记忆相比，允许更大的存在性和可伸缩性。■

重构在重建方法中，代理试图从观察到的视图重新创建其他视图。过去的工作集中在360度全景和CAD模型的像素级重建，这通常是人类拍摄的照片[45]的策划数据集。最近的工作已经将这种方法应用于具身人工智能，这更为复杂，因为该模型必须从主体的自我中心的观察和控制自己的传感器（即。主动感知）。在最近的一项工作中，agent使用其以自我为中心的RGB-D观测结果来重建可见区域以外的占用状态，并随着时间的推移汇总其预测，形成一个准确的占用图[45]。占用预期是一种像素级分类任务，在摄像机前的 $V \times V$ 单元局部区域的每个单元被分配其被探索和占据的概率。与覆盖方法相比，预测占用状态允许代理处理不能直接观察到的区域。另一项最近的工作的重点是语义重建，而不是像素级重建[22]。代理

旨在预测像“门”这样的语义概念是否存在于采样的查询位置。使用K-means方法，查询位置的真正重建概念是与其特征表示的J个最近的聚类质心。如果代理获得了有助于其预测的视图，那么它就会得到奖励。采样查询视图的真实重构概念。

2) 评估指标：访问的目标数量。考虑了不同类型的目标，如区域[44]、[86]和有趣的对象[72]、[87]。访问的区域度量有一些变体，例如m的绝对覆盖面积 $m^2$ 以及在场景中被探索的区域的百分比。

对下游任务的影响。视觉探索性能也可以通过其对视觉导航等下游任务的影响来衡量。这种评价度量类别在最近的研究中更为常见。利用视觉探索输出的下游任务的示例(i.e. 地图)包括图像导航[26]，[73]，点导航[11]，[44]和对象导航[53]，[54]，[56]。更多细节关于这些导航任务可以在第三-B节中找到。

3) 数据集：对于视觉探索，一些流行的数据集包括马特波特3D和Gibson V1。Matterport3D和Gibson V1都是逼真的RGB数据集，具有有用的具体化AI信息，如深度和语义分割。栖息地-sim模拟器允许使用这些具有额外功能的数据集，如可协调的代理和多个传感器。Gibson V1也增强了交互和现实机器人控制等功能，以形成iGibson。然而，最近的3D模拟器，如第二节中提到的，都可以用于视觉探索，因为它们至少都提供了RGB观测。

## B. 视觉导航

在视觉导航中，代理在有无外部先验或自然语言指令的情况下，将三维环境导航到目标。许多类型的目标已经被用于这个任务，如点、对象、图像[88]、[89]和区域[11]。在本文中，我们将把点和对象作为视觉导航的目标，因为它们是最常见和最基本的目标。它们可以进一步与感知输入和语言等专业技术结合，以实现更复杂的视觉导航任务，如先验导航、视觉和语言导航，甚至是嵌入式QA。在点导航[49]下，代理任务是导航到特定点，而在对象导航[38] [52]中，代理任务是导航到特定类的对象。

而经典的导航方法，[90]通常是由手工工程的子组件组成，如定位、映射[91]、路径规划[92]、[93]和运动。具体化人工智能的视觉导航旨在从数据中学习这些导航系统，从而减少案例专业的手工工程，从而简化与下游任务的集成，与数据驱动的学习方法具有优越的性能，如问题回答[23]。[44]也有一些混合方法，旨在结合两个世界的优点。如前在第二节中所述，基于学习的方法对传感器测量噪声更具鲁棒性，因为它们使用RGB和/或深度传感器，并且能够合并

对环境的语义理解。此外，它们使代理能够概括其对以前看到的环境的知识，以帮助以无监督的方式理解新环境，从而减少人类的努力。

随着近年来研究的增加，基本点导航和对象导航任务中的视觉导航也面临挑战，以基准和加速具体化AI [38]的进展。最引人注目的挑战是iGibsonSim2真正的挑战，栖息地挑战[36]和机器人挑战。对于每个挑战，我们将描述2020年版本的挑战，这是本文的最新版本。在所有这三个挑战中，该药物仅限于以自我为中心的RGB-D观察。对于iGibsonSim2真正的挑战2020，特殊的任务是点导航。73个高质量的吉布森3D场景用于训练，而卡斯特罗场景，重建一个真实世界的公寓，将用于训练、开发和测试。有三种情况：当环境没有障碍物时，包含代理可以与之交互的障碍物，和/或由其他移动代理填充。对于2020年的栖息地挑战，有点导航和目标导航任务。Gibson数据集分割的3D场景用于点导航任务，而原始数据集[11]，[34]指定的61/11/18训练/验证/测试室分割的马特港3D场景用于对象导航任务。对于2020年的机器人挑战，只有对象导航任务。培训和评估分为三个阶段。在第一阶段，代理在60个模拟公寓上进行训练，并在其他15个模拟公寓上验证其性能。在第二阶段，该代理将在四个模拟公寓和现实世界的对应公寓上进行评估，以测试其对现实世界的推广。在最后一阶段，该代理将对10套真实世界的公寓进行评估。

在本节中，我们将以现有的视觉导航为基础调查论文[11]，[23]，[74]，包括更多最近的作品。

1) 类别：点导航是最近视觉导航文献中最基本的和最流行的任务之一。在点导航中，代理的任务是导航到距离特定点[11]一定混合距离内的任何位置。通常，代理在环境中的原点(0, 0, 0)处初始化，混合的目标点由相对于原点/初始位置[11]的3D坐标(x, y, z)指定。要想成功地完成任务，人工代理将需要拥有多种多样的技能集，如视觉感知、情景记忆构建、推理/计划和导航。代理通常配备了GPS和指南针，允许它访问它们的位置坐标，并隐式地访问它们相对于目标位置[17]，[49]的方向。目标的相对目标坐标也可以是静态的。只给出一次，在这一集的开始)或动态的(i. e. 在每个时间步长中给出)[17]。最近，由于在室内环境中不完美的定位，2020栖息地挑战转向了更具挑战性的任务[47]，基于RGBD的在线定位，没有GPS和罗盘。

在最近的文献中，有许多基于学习的点导航方法。早期的作品之一是[74]

使用一个端到端的方法，在一个现实的自主导航设置中处理点导航(即。没有地面真相地图，没有地面真相，没有代理的姿态)与不同的感官输入。基本导航算法是直接未来预测(DFP) [94]，其中相关的输入，如彩色图像，深度图和四次最近观测的动作，由适当的神经网络处理(e. g. 用于感官输入的卷积网络)，并连接到一个双流完全连接的动作期望网络。输出是对所有行动和未来时间步长的未来测量预测。

作者还介绍了Belief DFP (BDFP)，它旨在通过在未来的测量预测中引入一种中间映射表示，使DFP的黑盒策略更易于解释。这是受到神经网络中的注意机制，以及强化学习中的后继表示[95]、[96]和特征[97]的启发。实验表明，BDFP在大多数情况下优于DFP，经典的导航方法通常优于具有RGB-D输入的基于学习的导航方法。[98]提供了一种更模块化的方法。对于点导航，SplitNet的架构包括一个视觉编码器和多个用于不同辅助任务的解码器。自我运动预测)和政策。这些解码器旨在学习有意义的表征。使用相同的PPO算法[99]和行为克隆训练，SplitNet可以在以前未见过的环境中优于类似的端到端方法。

另一项工作提出了一种模块化架构，用于室内环境中同步映射和目标驱动导航。在这项工作中，作者建立在MapNet [71]的基础上，包括2。具有语义信息功能的5D内存，并为导航策略训练一个LSTM。他们表明，在以前未见过的环境中，这种方法优于没有映射[100]的学习LSTM策略。

随着2019年栖息地挑战的引入及其标准化评估、数据集和传感器设置，最近的方法已经通过2019年栖息地挑战进行了评估。最初的工作来自于人居署背后的团队，并使用了PPO算法、演员-评论家模型结构和一个CNN来为视觉输入生成嵌入。后续工作提供了一个“存在证明”，在模拟[47]中看不见的环境中，对于代理的GPS、指南针和巨大的学习步骤(2.5亿步相比，栖息地的首次PPO工作可以达到接近完美的结果)。具体来说，最佳代理的性能在最短路径预测值的3-5%以内。本工作使用了一种改进的广义优势估计PPO[101]算法，该算法适用于资源密集型模拟环境中的分布式强化学习，即分散分布式近端策略优化(DD-PPO)。在每个时间步长，代理接收一个以自我为中心的观察(深度或RGB)，嵌入CNN，利用其GPS和指南针更新目标位置相对于其当前位置，然后初始输出下一个动作和值函数的估计。实验表明，该药剂长时间持续改进，结果与最短路径的预测基本吻合。



下一步的工作旨在通过增加辅助任务[49]的样本率和时间效率来改进这一资源密集型的工作。使用与之前工作相同的DD-PP0基线架构,这项工作增加了三个辅助任务:操作条件对比预测编码(CPC-A)[102],逆动态[68]和时间距离估计。作者尝试了不同的表示组合的方法。在4000万帧时,性能最好的代理可以达到与之前的工作相同的性能,速度要快5.5倍,甚至提高了性能。RGB和RGBD追踪[44]的2019年栖息地挑战赛的获胜者提供了一种混合解决方案,它结合了经典和基于学习的方法,因为基于端到端学习的方法在计算上都很昂贵。这项工作以一种模块化方式将学习整合到一个“经典的导航管道”中,从而隐含地整合了低级导航中的障碍和控制知识。该体系结构由一个学习过的神经SLAM模块、一个全局策略、一个局部策略和一个分析路径规划器组成。神经SLAM模块使用观测和传感器预测地图和代理姿态估计。全局策略总是将目标坐标作为长期目标输出,并使用解析路径规划器将其转换为短期目标。最后,一个地方政策被训练来实现这个短期目标。模块化设计和分析规划的使用有助于显著减少训练过程中的搜索空间。

对象导航是最直接的任务之一,但在具身人工智能中也是最具挑战性的任务之一。对象导航关注于在未探索的环境[38]中导航到其标签指定的对象的基本思想。代理将在随机位置初始化,并执行任务,在该环境中绑定对象类别的实例。物体导航通常比点导航更复杂,因为它不仅需要许多相同的技能集,如视觉感知和情景记忆构建,而且还需要语义理解。这些都使得对象导航任务更具挑战性,但解决这个问题也很值得。

物体导航的任务可以通过适应来演示或学习,这有助于在没有任何直接监督的环境下概括导航。[51]通过元强化学习方法实现了这项工作,因为代理学习了一个自我监督的交互损失,这有助于鼓励有效的导航。与传统的导航方法不同,代理在推理过程中冻结学习模型,这项工作允许代理学习以自我监督的方式适应自己,并在之后调整或纠正错误。这种方法可以防止代理在实现之前犯太多错误,并做出必要的纠正。另一种方法是在执行导航规划之前学习对象之间的对象关系。本工作[53]实现了一个对象关系图(ORG),它不是来自外部先验知识,而是在视觉探索阶段构建的知识图。该图由对象关系,如类别相关性和空间相关性。使用先验的导航侧重于以多模态输入的形式注入语义知识或先验的思想

如知识图或音频输入,或帮助训练在可见和看不见的环境中为具体化的人工智能代理的导航任务。过去的工作[57]将人类知识先验集成到深度强化学习框架中,表明人工代理可以利用类人的语义/功能先验,帮助代理学习在看不见的环境中导航和识别看不见的物体。这样的例子基于这样一种理解,即感兴趣的物品,比如在厨房里缠绕一个苹果,人类会倾向于查看逻辑位置来开始我们的搜索。这些知识被编码在一个图网络中,并在一个深度强化学习框架中进行训练。

还有其他一些使用人类先验的例子,比如人类感知和捕获音频信号模态和物体的物理位置之间的对应关系的能力,从而执行到信号源的导航。在这项工作[103]中,人工代理选择多种感官观察结果,如目标物体的视觉和声音信号,并找出从其起始位置到声音源的最短导航轨迹。这项工作通过拥有一个视觉感知映射器、声音感知模块和动态路径规划器来实现它。

视觉和语言导航(VLN)是一项通过自然语言指令来导航环境的任务。这项任务的挑战性方面是同时按顺序感知视觉场景和语言。VLN仍然是一项具有挑战性的任务,因为它需要代理根据过去的操作和指令[11]来预测未来的行动。此外,代理可能无法将它们的运行轨迹与自然语言指令无缝地对齐。虽然视觉和语言导航和视觉问题回答(VQA)看起来相似,但这两个任务有主要的差异。这两个任务都可以表述为视觉基础的,序列到序列的转编码问题。然而,VLN序列要长得更多,需要不断输入和操作摄像机视点的能力,而VQA输入一个输入问题并生成答案。我们现在可以给机器人一个自然语言指令,并期望他们执行任务[2],[3],[58]。这些是通过递归神经网络方法[58]的进步实现的,用于联合解释视觉和自然语言输入和数据集,这些数据集旨在简化三维环境中导航任务和执行任务的指令过程。

VLN的一种方法是辅助推理导航框架[59]。它处理了轨迹复述、轨迹估计、角度预测和跨模态匹配四个辅助推理任务。代理学习对之前的动作进行推理,并预测任务的未来信息。

视觉对话导航是VLN的最新扩展,因为它旨在培训一个代理来开发与人类持续进行自然语言对话的能力,以帮助其导航。目前在这一领域的工作[60]使用了一个跨模态记忆网络(CMN),它通过单独的语言记忆和视觉记忆模块来记忆和理解与过去的导航动作相关的有用信息,并进一步使用它来做出导航决策。

2) 评价指标: 视觉导航使用(1)路径长度(SPL)和(2)成功率作为主要评价指标[11]。根据路径长度加权的成功可以定义为:

$$1 - \frac{\sum_{i=1}^N S_i}{N} \quad \text{最大值} (n_i, l_i) \cdot S_i \text{ 是一个事件的成功指标}$$

$i$ ,  $p_i$  是代理的路径长度,  $l_i$  为最短路径长度,  $N$  为事件数。值得注意的是, 有一些已知的问题与路径长度[38]的成功加权。成功率是指代理在时间预算[74]内达到目标的事件的比例。除上述两个指标外, 还有其他不太常见的评价指标[11]、[48]、[52]、[54]、[74], 即: (3) 路径长度比, 即预测路径与最短路径长度的比值, 仅对成功事件计算; (4) 成功距离/导航误差, 分别测量代理的初始位置与最近对象或目标位置周围的成功阈值边界之间的距离。

除了上述四个指标外, 还有另外两个指标用于评估VLN代理。它们是: (1) oracle成功率, 即代理在最接近目标的点上停止的速率; (2) 轨迹长度。一般来说, 对于VLN任务, 最好的度量仍然是SPL, 因为它考虑了所考虑的路径, 而不仅仅是目标。

对于视觉对话导航, 除了成功率和oracle成功率外, 还有两个指标: (1) 目标进度, 代理向目标位置的平均进度; (2) oracle路径成功率, 代理的成功率沿着最短的路径在离目标最近的点停止。

3) 数据集: 在视觉探索中, 马特波特3D和Gibson V1是最流行的数据集。值得注意的是, Gibson V1中的场景较小, 通常情节较短(从开始位置到目标位置的GDSP较低)。同时还使用了AI2-THOR模拟器/数据集。

与其他的视觉导航任务不同, VLN需要一种不同类型的数据集。大多数VLN工作使用房间到房间(R2R)数据集与马特港3D模拟器[104]。它由21,567条导航指令组成, 平均长度为29个单词。在视觉-对话导航[59]中, 使用了合作视觉和对话导航(CVDN)[105]数据集。它包括2050个人对人的对话和7000多个马特波特3D模拟器的7000轨迹。

### C. 体现问题回答

最近的具身人工智能模拟器的具身问题回答(QA)任务是通用智能系统的一个重大进展。为了在物理体现的状态下执行QA, 人工智能代理需要拥有广泛的人工智能能力, 如视觉识别、语言理解、问题回答、常识推理、任务规划和目标驱动导航。因此, 所体现的QA可以被认为是最繁重的和目前具身人工智能研究的复杂任务。

1) 类别: 对于具体化的QA(EQA), 这是一种通用的框架, 它将任务划分为两个子任务: 导航任务和QA任务。导航模块是必要的, 因为代理需要探索环境

在回答有关它们的问题之前, 请先查看它们。例如, [61]提出了规划器-控制器导航模块(PACMAN), 它包括一个导航模块的层次结构, 包括一个选择动作(方向)的规划器和一个决定跟随每个动作移动多远的控制器。一旦代理决定停止, QA模块将通过使用其路径上的帧序列来执行。导航模块和视觉问答模块首先单独训练, 然后由强化[106]联合训练。[62]和[63]通过神经模块化控制(NMC)进一步改进了PACMAN模型, 其中高级主策略建议由子策略执行语义子目标。

多目标具身QA(MT-EQA)[63]是一种更为复杂的具身QA任务, 它研究其中具有多个目标的问题。“卧室里的苹果比客厅里的橘子大吗?”, 这样代理就必须导航到“卧室”和“客厅”来定位“苹果”和“橙色”, 然后进行比较来回答问题。

交互式问答(IQA)[64]是解决AI2-THOR环境中体现QA任务的另一项工作。IQA是EQA的一个扩展, 因为代理与对象的交互对于成功地回答某些问题是至关重要的。代理需要打开冰箱来回答存在的问题“冰箱里有鸡蛋吗?”。[64]提出使用层次交互式记忆网络(HIMN), 这是一个控制器的层次结构, 帮助系统在多个时间尺度上操作、学习和推理, 同时降低每个子任务的复杂性。一个以空间为中心的空间门控循环单元(GRU)作为一个记忆单元, 用来保留环境的空间和语义信息。计划模块将控制其他模块如导航运行一个\*搜索的最短路径目标, 扫描仪执行旋转检测新图像, 机械器调用执行操作改变环境的状态, 最后一个回答, 回答这个问题发布到AI代理。[65]从多主体的角度研究了IQA, 其中几个主体共同探索一个互动场景来回答一个问题。[65]提出了多层结构记忆和语义记忆作为场景记忆, 由多个代理共享重建三维场景, 然后进行QA。

2) 评估指标: 体现的QA和IQA涉及两个子任务: 1) 导航和2) 问答, 这两个子任务基于不同的指标进行评估。

导航性能评估方法为: (1) 导航终端到目标的距离。e. 导航错误(dT); (2) 从初始位置到目标到初始位置的距离的变化。目标进展(d); (3) 在事件的任何点与目标的最小距离(d<sub>Δ</sub>); (4) 事件代理的百分比在达到最大事件长度(%停止)之前终止了导航以进行应答; (5) 代理在包含目标对象的房间中终止的问题的百分比(%rT); (6) 代理至少进入一次包含目标对象的房间的问题的百分比(%re); (7) 目标目标(Iou);

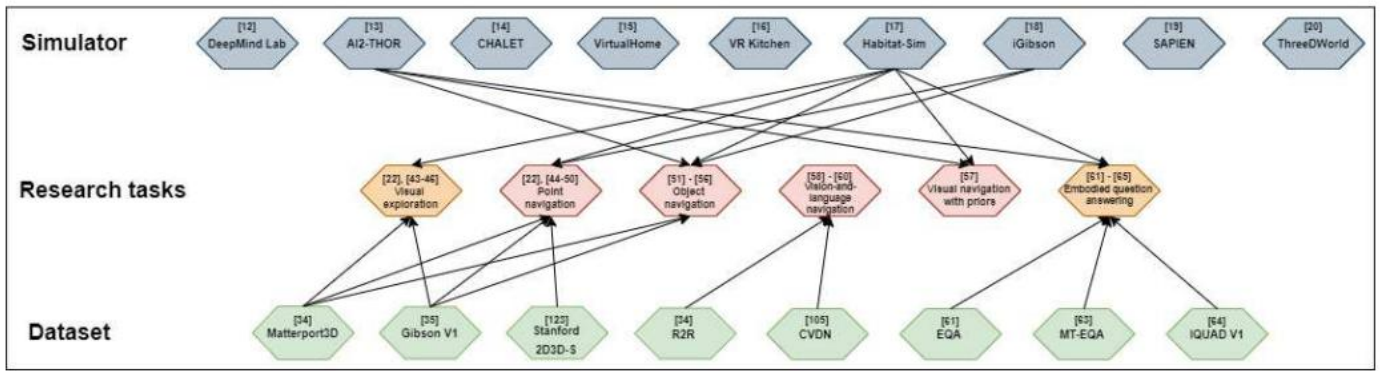


图6. 嵌入式人工智能模拟器之间的联系与研究。（上）9个最新的具体化人工智能模拟器。（中）由9个具身人工智能模拟器完成的各种具身人工智能研究任务。红色的研究任务属于视觉导航类别，其余黄色的研究任务属于其他研究类别。（下图）用于评估9个具体化人工智能模拟器之一的研究任务的评估数据集

(8) 命中精度基于IoU (hT); (9) 每集的长度, 即。轨迹长度。度量(1)、(2)和(9)也被用作视觉导航任务的评估度量。

QA绩效的评估方法为: (1) 平均等级 (MR) 预测中的真实答案; (2) 准确性。

3) 数据集: EQA[61]数据集基于House3D, 这是流行的SUNCG [33]数据集的一个子集, 具有合成的房间和布局, 类似于复制数据集[107]。House3D将SUNCG的静态环境转换为一个虚拟环境, 在那里代理可以使用物理约束进行导航(e.g. 它不能通过墙壁或物体)。为了测试代理在语言基础、常识推理和导航方面的能力, [61]在CLEVR [108]中使用了一系列函数程序来综合关于对象及其属性的问题和答案。颜色、存在性、位置和相对介词)。在750个环境中, 共有5000个问题, 涉及到7个独特的房间类型中的45个独特的对象。

对于MT-EQA [63], 作者介绍了MT-EQA数据集, 该数据集包含6种类型的组合问题, 它们比较了多个目标(对象/房间)之间的对象属性属性(颜色、大小、距离)。

对于IQA [64], 作者注释了一个大规模的数据集, IQAD V1, 该数据集由75,000个多项选择题组成。与EQA数据集类似, IQAD V1也有一些问题关于对象的存在性、计数性和空间关系。

#### 增值见解和挑战

##### A. 洞察具体化的人工智能

图中的互连。#6影响了模拟器对研究任务的适用性。基于无花果。6、栖息地-sim和iGibson都支持在视觉探索和一些列视觉导航任务中的研究任务, 这表明了高保真度的重要性, 这来自于基于世界的场景模拟器。然而, 由于它们独特的独特特性, 使它们更适合于非具体化的人工智能独立任务, 如深度强化学习, 一些模拟器目前还没有连接到任何具体化的研究任务。尽管如此, 它们仍然符合被归类为具体化人工智能模拟器的标准。

相反, 由于这些任务的交互性, 诸如具身问题回答和先验视觉导航等研究任务需要具身人工智能模拟器具有多状态对象属性。因此, AI2-THOR无疑是模拟器的选择。最后, VLN是目前唯一没有使用9个具体化人工智能模拟器中的任何一个, 而是使用矩阵端口3D模拟器[104]。这是因为以前在VLN中的工作不需要在其模拟器中的交互性特性; 因此, 马特波特3D模拟器技巧。然而, 随着VLN任务的进一步推进, 我们可以预期在VLN任务中需要交互, 因此需要使用具身的AI模拟器。此外, 与传统的强化学习模拟环境[41]不同, [109]专注于任务专业训练, 而具体化的人工智能模拟器提供了一个训练环境来训练类似于在物理世界中进行的各种不同的任务。

此外, 基于第三节中对具身人工智能研究任务的调查, 我们提出了一个金字塔结构, 其中每个具身人工智能研究任务对下一个都有贡献。

例如, 视觉探索有助于视觉导航的发展, 而视觉导航有助于具体化QA的创作。这种累积方法也与任务日益增加的复杂性相关联。基于具身人工智能研究的可预见趋势, 我们假设具身人工智能研究金字塔的下一个进展是基于任务的交互式问题回答(TIQA), 它旨在将任务与回答特殊问题结合起来。例如, 这样的问题可以是“一个鸡蛋煮沸需要多长时间?”“橱柜里有苹果吗?”。这些问题无法通过传统的[61], [64]的方法来回答。它们要求具体化的代理执行与这些问题相关的特殊任务, 以解锁在回答这些QA问题时具有重要意义的见解。我们假设的TIQA代理可以执行一系列一般的家务任务, 这使他们能够推断出有用的环境信息, 这些信息帮助他们获得QA问题的答案至关重要。TIQA可能是在模拟中推广任务规划和开发通用人工智能的关键, 之后可以部署到

现实世界。

### B. 嵌入式人工智能模拟器所面临的挑战

目前体现的人工智能模拟器在功能和保真度方面都达到了一个水平, 这使它们有别于那些用于强化学习的传统模拟器。即使具身人工智能模拟器的差异如此飙升, 具身人工智能模拟器在从现实性、可扩展性到交互性等领域仍存在着一些现有的挑战。

现实主义。它着重于模拟器的可靠性和物理特性。具有高视觉保真度和逼真物理特性的模拟器被机器人社区高度追捧, 因为它们为各种机器人任务提供了理想的测试台, 如导航和交互任务[110], [111]。然而, 目前还缺乏既具有世界场景又具有先进物理特性的实体人工智能模拟器。

对于完整性来说, 基于世界的场景的模拟器在模拟真实任务[27], [112]方面无疑将超过基于游戏的场景模拟器。尽管有这样的观察, 只有栖息地-Sim[17]和iGibson [18]是基于世界的场景模拟器。基于世界的场景模拟器的缺乏是具体化人工智能代理模拟到真实任务的瓶颈, 这进一步阻碍了具体化人工智能研究在现实世界部署中的可转移性。对于物理学, 基于物理的预测模型[113]-[115]强调了具有先进物理特性的具体化人工智能模拟器的重要性, 因为它们为训练提供了一个理想的测试平台, 训练具体化人工智能代理执行复杂的物理交互任务[2], [3], [116]。尽管需要一个先进的基于物理的具身人工智能模拟器, 但目前只有一个模拟器, ThreeDWorld [20]是这个标准。因此, 目前严重缺乏具有布、uid和软体物理等先进物理特性的实体人工智能模拟器。我们相信, 三维重建技术和物理引擎[117]-[119]的进步将提高具体化人工智能的真实性。

可扩展性。与基于图像的数据集[7]不同, [120]可以很容易地从众包或互联网中获得。收集大规模的基于世界的三维场景数据集和三维对象资产[121]-[123]的方法和工具很少。这些三维场景数据集对于构建各种具体化的人工智能模拟器是至关重要的。目前收集真实的3D场景数据集的方法需要通过摄影测量[124]扫描物理房间, 如马特港3D扫描仪, 网格式[125], 甚至移动3D扫描应用程序。然而, 它们在商业上并不适用于收集大规模的3D物体和场景扫描。这在很大程度上是由于用于摄影测量的3D扫描仪成本昂贵且无法获得。因此, 可伸缩性的瓶颈在于开发用于高保真3D对象或场景扫描的工具。希望随着基于3D学习的方法[126], [127]的进一步发展, 旨在从单个或几个图像渲染3D对象网格, 甚至通过场景生成方法[128], 我们将能够扩大大规模3D数据集的收集过程。

互动。与具体化的人工智能中的功能对象进行正粒度操作交互的能力

模拟器在复制与真实世界对象[129]的人类级交互时至关重要。大多数基于游戏的场景模拟器[13], [16], [19], [20]都提供了正粒度的对象操作功能和符号交互功能(e.g. <下拉对象X在Y>动作)或简单的“点和选择”。然而, 由于基于游戏的场景模拟器的性质, 在这个环境中执行的许多研究任务将选择与其正粒度对象操作[3]相比的符号交互能力, 除了一些同时使用了[2], [130]。

另一方面, 基于世界的场景模拟器[17], [18]的代理具有大肌肉运动控制能力, 而不是符号交互能力。然而, 这些模拟器中的对象属性在表面上很大程度上是可交互的, 这允许大运动控制, 但缺乏多状态对象类, 即对象所具有的状态变化的数量。因此, 有必要在其对象属性中的对象功能和所体现的人工智能代理在环境中可以执行的操作的复杂性方面取得平衡。

毫无疑问, 主流模拟器如AI2-THOR [13]、iGibson [18]和Sim[17]确实为推进各自体现的人工智能研究提供了一个极好的环境。然而, 它们确实有自己的优势和局限性需要克服。随着计算机图形学和计算机视觉的发展, 以及创新的真实世界数据集的引入, 真实到模拟领域的自适应是改进具身人工智能模拟器的明确途径之一。真实到模拟的概念围绕着捕捉真实世界的信息, 如触觉感知[131], 人类水平的运动控制[132]和音频输入[133], 以及视觉感官输入, 并将它们集成起来, 以开发更现实的具身人工智能模拟器, 可以有效地建立桥梁物理世界和虚拟世界。

### C. 体现式人工智能研究中的挑战

嵌入式人工智能研究任务标志着从“网络人工智能”到自主嵌入式学习代理的复杂性的增加, 具有多种传感器模式和潜在的长轨迹[22], [34]。这导致了代理的记忆和内部表征成为极其重要的[11], [22], [56]。长轨迹和多种输入类型也表明了健壮的内存体系结构的重要性, 它允许代理专注于其环境的重要部分。近年来, 有许多不同类型的记忆被使用, 如循环神经网络[47], [49], [51], [56], [58], [61]-[63], 基于注意力的记忆结构, [52], [60], [72], 预期占用图[45], 占用地图[22]和语义地图[43], [46], [48], [64], [65], , 一些论文非常强调他们的记忆结构的新颖性[22], [45], [60], [72]。然而, 虽然已知循环神经网络在捕获具体化AI [56]和[72]的长期依赖方面受到限制, 但由于缺乏对记忆结构的研究, 目前仍然很难同意哪种记忆类型(s)是更好的[11]。在具体化的人工智能研究任务中, 复杂性也有所增加, 这从视觉的发展中可以看出



探索VLN和体现QA, 其中分别添加了语言理解和QA等新组件。每一个新组件都会导致人工智能代理的难度和更长的训练, 特别是目前的方法通常是完全基于学习的。这一现象导致了两个有希望的进展, 以减少搜索空间和样本复杂性, 同时提高鲁棒性: 混合方法结合了经典和基于学习的算法[44], [74]和先验知识合并[23], [57]。此外, 消融研究更难管理更复杂的任务的[31], 因为具体化人工智能中的每个新组件使其更难测试它的贡献。

Agent的性能, 因为它被添加到现有的组件集上, 并且体现的AI模拟器在特性和问题上差异显著。研究任务的数量也迅速增加。因此, 虽然一些基本的任务, 如视觉探索, 得到了更多的关注, 因此有更多的方法来解决它们, 但更新和更小众的任务, 如MT-EQA, 却更少得到解决。新的任务通常会在重要方面引入新的考虑因素, 如方法、评估度量[22]、输入类型和模型组件, 如表三所示, 因此比视觉探索等更简单的任务需要更多的评估。

最后, 缺乏对多代理设置的关注, 这有助于提供有用的新任务[65]。这种焦点的缺乏可以归因于直到最近还缺乏具有多智能体特性的模拟器。用于协作和通信的多代理系统在现实世界的[134]、[135]中很普遍存在, 但目前[31]受到的关注相对较少。随着具有多智能体功能的[13]、[20]、[55]的模拟器的增加, 多智能体是否支持还有待观察。对多智能体算法的支持)是非常有效的。

## 结论

具身人工智能模拟器的最新进展是具身人工智能研究进展的关键驱动力。为了了解具身人工智能模拟器和研究的趋势和差距, 本文提供了一个关于具身人工智能模拟器和研究的当代和全面的概述。本文调查了9个具身人工智能模拟器及其在服务和推动具身人工智能研究任务的近期创新方面的联系。通过根据7个功能对9个具体化人工智能模拟器进行基准测试, 我们试图理解它们对现实性、可伸缩性和交互性的提供, 从而在具体化人工智能研究中使用。支持具身人工智能研究金字塔的三个主要任务——视觉探索、视觉导航和具身QA, 都根据他们的方法、评估指标和数据集进行了检验。这是回顾和基准现有的方法, 在各种人工智能模拟器中处理这些类别的人工智能研究任务的具体化人工智能。此外, 本文使我们能够揭示模拟器、数据集和研究任务之间的深刻关系。在本文的帮助下, 新领域的人工智能研究者将能够选择最适合的具身人工智能模拟器, 并为推动具身人工智能的发展做出贡献。

## 参考文献

- [1] L. 史密斯和M. 《具身认知的发展: 婴儿的六门课》, 《艺术生活》, 第1卷. 11日, 没有. 1-2, pp.13 - 29, 2005.
- [2] J. 段, S. 余, H. L. Tan和C. Tan, “动作集: 基于3d环境中的任务的数据收集和增强的交互式端到端平台”, 2020年IEEE图像处理国际会议 (ICIP). IEEE, 2020, pp.1566 - 1570.
- [3] M. 希里达尔, J. Thomason, D. 戈登, Y. Bisk, W. 韩, R. Mottaghi, L. Zettlemoyer和D. 福克斯, “阿尔弗雷德: 解释日常任务中扎根指令的基准”, 发表在IEEE/CVF计算机视觉和模式识别会议论文集, 2020年, 页. 10 740 - 10 749.
- [4] R. Pfeifer和F. 饭田, “体现人工智力: 趋势与挑战”, 体现人工智力. 施普林格, 2004年, 页. 1-26.
- [5] J. 豪格, 《艺术智慧: 正是这个想法, 剑桥, 马, 布拉德福德》, 1985年.
- [6] R. 法伊弗和J. C. 邦加德, “身体如何塑造我们的思维方式——一种关于智能的新观点”, 2006年.
- [7] J. 邓, W. 董, R. Socher, L.-J. 李, K. 李和L. 费飞, “图像集: 大型层次图像数据库”, 2009年IEEE计算机视觉和模式识别会议. IEEE, 2009, pp. 248 - 255.
- [8] Y. Lecun. 孟加拉和G. 《深度学习》, 《自然》, 第1卷. 521, 没有. 7553, pp. 436 - 444, 2015.
- [9] K. 他, X. 张, S. 任和J. 孙先生, “图像识别的深度残差学习”, IEEE计算机视觉与模式识别会议论文集, 2016年, 页. 770 - 778.
- [10] D. 银, J. Schrittwieser, K. 西蒙尼安, 我. 安东诺格鲁, A. 黄, A. Guez, T. 休伯特, L. 贝克, M. 赖, A. 博尔顿等人., “掌握没有人类知识的围棋游戏”, 《自然》, 卷. 550, 没有. 7676, pp. 354 - 359, 2017.
- [11] P. 安德森. 张, D. S. 夏普拉特. Dosovitskiy, S. 古普塔, 五. Koltun]. Kosecka, J. 马利克, R. Mottaghi, M. 萨夫瓦等人., “关于具体化导航代理的评估”, arXiv预印本arXiv: 1807.06757, 2018.
- [12] C. 比蒂, J. Z. Leibo, D. Teplyashin, T. 病房, M. 温赖特, H. K. 突特利, Lefrancq, S. 绿色, V. Valde's. Sadik等., “心灵实验室”, arXiv预印本arXiv: 1612.03801, 2016.
- [13] E. Kolve, R. Mottaghi, W. 韩, E. 范德比尔特, L. 韦斯, A. Herrasti, D. 戈登, Y. 朱, A. Gupta和A. 法哈迪, “爱2-托尔: 视觉ai的交互式3d环境”, arXiv预印本arXiv: 1712.05474, 2017.
- [14] C. 燕, D. 米斯拉. 班尼特. Walsman, Y. Bisk和Y. Artzi, “小屋: 康奈尔学院代理学习环境”, ArXiv预印本arXiv: 1801.07357, 2018.
- [15] X. Puig, K. Ra, M. Boben, J. 李, T. 王, S. 软管和A. “虚拟之家: 通过节目模拟家庭活动”, 在IEEE计算机视觉与模式识别会议论文集, 2018年, 第3页. 8494 - 8502.
- [16] X. 高, R. 龚, T. 舒, X. 谢, S. 王和南卡罗来纳. 朱, “Vr厨房: 面向任务学习的交互式3d虚拟环境”, arXiv预印本arXiv: 1903.05757, 2019.
- [17] M. Savva. 科迪亚, 哦. Maksymets, Y. 赵, E. Wijmans, B. Jain J. 斯特劳布, J. 刘, 五. Koltun]. 马利克等人., “栖息地: 体现ai研究的平台”, 发表在IEEE国际计算机视觉会议论文集, 2019年, 页. 9339 - 9347.
- [18] F. 夏, W. B. 沈, C. 李, P. Kasimbeg, M. E. 查普曼, A. Toshev, R. Mart 'In-Mart' In和S. “交互式吉布森基准: 混乱环境下交互式导航的基准”, 《IEEE机器人与自动化信件》, 第1卷. 5、没有. 2, pp. 713 - 720, 2020.
- [19] F. 向, Y. 秦, K. Mo, Y. 夏, H. 朱, F. 刘, M. 刘, H. 江, Y. 元, H. 王等人., “Sapien: 一个模拟的基于部分的交互环境”, 发表在IEEE/CVF计算机视觉和模式识别会议论文集, 2020年, 页. 11097 - 11 107.
- [20] C. 干, J. 施瓦茨. 改变, M. 施里夫, J. 特拉尔, J. 德弗雷塔斯, J. 库比利乌斯. Bhandwadar, N. 哈伯, M. Sano等., “威胁世界: 交互式多模态物理模拟平台”, arXiv预印本arXiv: 2007.04954, 2020年.
- [21] M. G. Bellemare, Y. 纳达夫, J. Veness和M. 保龄球, “街机学习环境: 为一般代理的评估平台”, 艺术智能研究杂志, 第1卷. 47, pp. 253 - 279, 2013.
- [22] S. K. 拉马克里希南, D. Jayaraman和K. 格劳曼, “一种对具体化的视觉探索的探索”, 国际版 计算机视觉杂志, 2021年. 在线可用: <https://doi.org/10.1007/s11263-021-01437-z>

- [23] X. Ye和Y. 杨, “从视觉到移动: 一个关于室内视觉导航学习的调查(vin)”, arXiv预印本arXiv: 2002.11310, 2020年。
- [24] T. 陈, S. Gupta和A. Gupta, “学习导航的探索政策”, arXiv预印本arXiv: 1903.01959, 2019年。
- [25] N. 萨维诺夫, 多索维茨基和V. Koltun, “导航的半参数拓扑记忆”, 国际学习表示会议 (ICLR), 2018年。
- [26] E. 比奇宁, J. Dibangoye, O. 西蒙宁和C. 沃尔夫, “学习规划不确定的拓扑地图”, arXiv预印本arXiv: 2007.05270, 2020年。
- [27] A. 卡迪亚, J. 特朗, A. 戈卡斯兰, 克莱格, Wijmans, S. 李, M. Savva, 切尔诺瓦和D. Batra, “Sim2真实预测: 模拟评估能预测真实世界的性能吗?”, *“IEEE机器人和自动化信件”*, 卷. 5, 没有. 4, pp. 6670–6677, 2020。
- [28] X. B. 彭, M. Andrychowicz, W. Zaremba和P. Abbeel, “机器人控制的动态随机化模拟真实转移”, 2018年IEEE机器人与自动化国际会议 (ICRA), IEEE, 2018, 页. 1–8。
- [29] J. 托宾, 方, A. 雷, J. 施耐德, W. Zaremba和P. “用于转移深度神经网络域随机化”, 在2017年IEEE/RSS/国际Con-  
尊重智能机器人和系统 (IROS)。IEEE, 2017, 页. 23–30。  
[30] CVPR, “体现ai工作室”, <https://embodied-ai.org/>, 1月。
- [2020. [31] L. Weihs, J. 萨尔瓦多, K. 科塔尔, 美国. Jain, KH. 曾, R. 莫塔吉和A. “艾伦阿克特: 一个具体化的ai研究的框架”, arXiv, 2020。
- [32] K. Mo, S. 朱, A. X. 张, L. 易, S. Tripathi, L. J. Guibas和H. Su, “Partnet: 正粒度和层次分级的三维对象理解的大规模基准”, 发表在IEEE计算机视觉和模式识别会议论文集, 2019年, 页. 909–918。
- [33] S. 宋, F. 余, A. 曾, A. X. 常, M. Savva和I. “来自单个深度图像的语义场景完成”, 第30届IEEE计算机视觉与模式识别会议论文集, 2017年。
- [34] A. 张, A. 戴, Funkhouser, M. 哈尔伯, M. Niebner, M. Savva, 宋, A. 曾和Y. 张, “马特波特3d: 从室内环境中的rgb-d数据中学习”, 2017年3D视觉 (3DV) 国际会议. IEEE计算机学会, 2017年, 第3页. 667–676。
- [35] F. 夏, A. R. Zamir, Z. 他, A. Sax, J. 马利克和S. 萨瓦雷斯, “吉布森环境: 现实世界的感知”, *IEEE计算机视觉与模式识别会议论文集*, 2018年, 页. 9068–9079。
- [36] 阿比舍克Kadian\*, 乔安妮·张强\*, A. 戈卡斯兰, 克莱格, Wijmans, S. 李, M. Savva, 切尔诺瓦和D. 巴特拉, “我们在模拟环境中取得了真正的进展吗?” 测量体现视觉导航中Sim2的真实差距, 在arXiv: 1912.06321, 2019年。
- [37] M. Deitke, W. 韩, A. 赫拉斯蒂, 肯巴维, E. Kolve, R. Mot-塔吉, J. 萨尔瓦多, D. 施文克, E. 范德比尔特, M. 沃林福德等人, “机器人: 一个开放的模拟到真实的ai平台”, *IEEE/CVF计算机视觉和模式识别会议论文集*, 2020年, 页. 3164–3174。
- [38] D. 巴特拉, 戈卡斯兰, 科姆巴维, O. Maksymets, R. Mottaghi, M. Savva, Toshev和E., “重新回顾: 关于主体导航到物体的评估”, arXiv预印本arXiv: 2006.13171, 2020。
- [39] U. Jain, L. 韦斯, E. 科尔夫, A. 法哈迪, 拉泽布尼克, 肯巴维和A. “亲切同步: 超越多主体任务补充材料。”
- [40] U. Jain, L. 韦斯, E. Kolve, M. Rastegari, S. 拉泽布尼克, 法哈迪, G. 施温和A. 肯巴维, “两个身体问题: 协作视觉任务完成”, 在CVPR, 2019年, 前两位作者贡献相等。
- [41] G. 布罗克曼, 五. 张, L. 佩特森, J. 施耐德, J. 舒尔曼, J. 唐和W. 弗雷姆巴, “Openai健身房”, arXiv预印本arXiv: 1606.01540, 2016年。
- [42] U. Jain, L. 韦斯, E. Kolve, M. Rastegari, S. 拉泽布尼克, 法哈迪, G. 施温和A. “两个身体问题: 协同视觉任务完成”, *《IEEE计算机视觉与模式识别会议论文集》*, 2019年, 页. 6689–6699。
- [43] D. S. Chaplot, H. 江, S. Gupta和A. Gupta, “对主动视觉学习的语义好奇心”, 发表在ECCV, 2020年。
- [44] D. S. Chaplot, D. 甘地, S. 古普塔, Gupta和R. 萨拉库蒂诺夫, “学习探索使用主动神经网络满贯”, *国际学习表征会议 (ICLR)*, 2020年。
- [45] S. K. 罗摩克里希南, Z. Al-Halah和K. 格劳曼, “对高效探索和导航的占用率预期”, arXiv预印本arXiv: 2008.09285, 2020年。
- [46] M. 纳拉西姆汉, E. Wijmans, X. 陈, T. 达雷尔, D. 巴特拉, D. 帕里克和A. 辛格, “看到非场景: 学习房间导航的模态语义地图”, 在欧洲计算机视觉会议上, 施普林格, 2020年, 页. 513–529。
- [47] E. 维曼斯, 卡迪亚, A. 莫科斯, 李, 我. 埃萨, D. 帕里克, M. Savva和D. d-ppo: 从2学习近乎完美的目标导航器. 50亿帧, arXiv, pp. arXiv: 1911.2019。
- [48] G. 乔治卡基斯, Y. 李和J. 柯塞卡卡, “同步映射和目标驱动导航”, arXiv预印本arXiv: 1911.07980, 2019年。
- [49] J. 是的, D. 巴特拉, E. Wijmans和. Das, “辅助任务加速学习点目标导航”, arXiv预印本arXiv: 2007.04561, 2020。
- [50] C. P. erez-D’ Arpino, C. 刘, P. 戈贝尔, R. Mart ‘In-Mart’ In和S. “机器人在受限的行人环境中使用强化学习进行导航”, arXiv预印本arXiv: 2010.08600, 2020年。
- [51] M. 沃斯曼, K. 埃萨尼, M. 拉斯特加里, A. 法哈迪和R. “学习如何学习: 使用元学习的自适应视觉导航”, 发表在IEEE计算机视觉和模式识别会议论文集, 2019年, 页. 6750–6759。
- [52] T. Campari, P. Eccher, L. Seraini和L. 《利用物体目标导航的特定场景特性》, ECCV研讨会, 2020年。
- [53] H. 杜, X. Yu和L. 郑, “学习对象关系图与视觉导航的初步政策”, 欧洲计算机视觉会议, 施普林格, 2020年, 页. 19–34。
- [54] D. S. Chaplot, D. 甘地, Gupta和R. “使用面向目标的语义探索的对象目标导航”, arXiv预印本arXiv: 2007.00643, 2020年。
- [55] B. 沈, F. 夏, C. 李, R. MartIn-Martin, L. 风扇, G. 王, S. Buch, C. D’ Arpino, S. 斯里瓦斯塔瓦, L. P. Tchammi, K. 维尼奥, L. Fei-Fei和S. “大型现实场景中交互式任务的模拟环境”, 2021年IEEE/RSS/智能机器人与系统国际会议 (IROS), 2020年。
- [56] A. 瓦希德, 石头, K. 陈, B. 希特和A. Toshev, “使用分布式软演员评论家学习对象条件探索”, arXiv预印本arXiv: 2007.14545, 2020。
- [57] W. 杨, X. 王, A. 法哈迪, Gupta和R. “使用场景先验的视觉语义导航”, arXiv预印本arXiv: 1810.06543, 2018年。
- [58] P. 安德森, 问. 吴, D. Teney, J. 布鲁斯, M. 约翰逊, N. S. undhauf, 我. 里德, S. 古尔德和A. 范登·亨格勒, “视觉与语言导航: 在真实环境中解释基于视觉的导航指令”, *《IEEE计算机视觉与模式识别会议论文集》*, 2018年, 页. 3674–3683。
- [59] F. 朱, Y. 朱, X. 张和X. 梁, “具有自我监督辅助推理任务的视觉语言导航”, *IEEE/CVF计算机视觉与模式识别会议论文集*, 2020年, 页. 10 012–10 022。
- [60] Y. 朱, F. 朱, Z. 詹, B. 林, J. 焦, X. 张和X. 梁, “通过探索跨模态记忆的视觉对话导航”, 在教授中 *IEEE/CVF计算机视觉与模式识别会议论文集*, 2020年, 第3页. 10 730–10 739。
- [61] A. 达斯, S. 达塔, G. Gkioxari, S. 李, D. 帕里克和D. Batra, “体现问题回答”, 发表在IEEE计算机视觉和模式识别研讨会论文集, 2018年, 页. 2054–2063。
- [62] A. 达斯, G. Gkioxari, S. 李, D. 帕里克和D. Batra, “具体化问题回答的神经模块化控制”, *《机器学习会议 (CoRL) 论文集》*, 2018年。
- [63] L. Yu, X. 陈, G. Gkioxari, M. 坦萨尔, T. L. Berg和D. Batra, “多目标体现问题回答”, *IEEE计算机视觉和模式识别会议, CVPR 2019, 美国长滩, 2019年6月16日, 20日*, 页. 6309–6318。
- [64] D. 戈登, 肯巴维, M. Rastegari, J. 雷德蒙, D. 福克斯和A. 法哈迪, “Iqa: 交互环境中的视觉问题回答”, 发表在IEEE计算机视觉与模式识别会议论文集, 2018年, 页. 4089–4098。
- [65] S. 棕褐色, W. 香, H. 刘, D. 郭和F. Sun, “多代理在互动环境中体现问题回答”, 在ECCV 2020–第16届欧洲会议, 格拉斯哥, 英国, 2020年8月23–28日, 论文集, A. Vedaldi, H. 比肖夫, T. 布罗克斯和J. 弗拉姆, Eds., 2020年, 页. 663–678。
- [66] P. D. 阮, Y. K. 乔治, E. Kayhan, Eppe, V. V. 哈夫纳和S. “机器人中“主动自我”的感觉运动表征学习: 模型调查”, “KI-K 自由智能, 卷. 35岁, 没有. 1, pp. 9–35, 2021。

- [67] C. 陈, 美国耆那教, C. 五味子, S. V.A. 加里, Z. Al-HalahV. K. Ithapu, P. 罗宾逊和K. 格劳曼, “声音空间: 三维环境中的视听导航”, 计算机视觉-ECCV2020: 第16届欧洲会议, 格拉斯哥, 英国, 2020年8月23-28日, 论文集, 第六部分, 第16页。施普林格, 2020年, 页。17 - 36.
- [68] D. 帕萨克, P. 阿格拉瓦尔, A. Efros和T. 达雷尔, “自我监督预测的好奇心驱动探索”, IEEE计算机视觉和模式识别研讨会论文集, 2017年, 页。16 - 17.
- [69] S. 古普塔, D. Fouhey, S. 莱文和J. Malik, “统一地图和基于地标的视觉导航表示”, arXiv预印本arXiv: 1712.08125, 2017.
- [70] S. 古普塔, J. 戴维森, 莱文, R. Sukhankar和J. Malik, “视觉导航的认知映射与规划”, IEEE计算机视觉与模式识别会议论文集, 2017年, 页。2616 - 2625.
- [71] J.F. 亨利克斯和A. 维达尔迪, “Mapnet: 映射环境的异中心空间记忆”, IEEE计算机视觉和模式识别会议论文集, 2018年, 页。8476 - 8484.
- [72] K. 方, A. Toshev, L. Fei-Fei和S. 萨瓦雷斯, “长期任务中具体化代理的场景记忆转换器”, 发表在IEEE计算机视觉与模式识别会议论文集, 2019年, 页。538 - 547.
- [73] L. Mezhghani, S. 苏赫巴塔尔, S. Szlam, J. Joulin和P. 博亚诺夫斯基, “学习在没有任何监督的情况下在逼真的环境中进行视觉导航”, arXiv预印本arXiv: 2004.04954, 2020年.
- [74] D. 米什金, A. 多索维茨基和V. Koltun, “对复杂3d环境中的经典和学习导航的基准测试”, arXiv预印本arXiv: 1901.10915, 2019.
- [75] C. 卡迪纳, L. Carbone, H. 卡里洛, Y. Latif, D. 斯卡拉穆扎, J. 内拉, 我, 里德和J. 伦纳德, “同步定位和映射的过去、现在和未来: 走向稳健感知时代”, IEEE机器人学报, 第1卷。32岁, 没有。6, pp. 1309 - 1332, 2016.
- [76] S.K. 拉马克里希南, D. Jayaraman和K. 格劳曼, “通过主动观察完成, 探索性观察行为的出现”, 《科学机器人》, 第1卷。4, 没有。30, 2019.
- [77] W. S. “对部分观察到的马尔科夫决策过程的算法方法的调查”, 《运筹学年鉴》, 第1卷。28日, 没有。1, pp. 47 - 65, 1991.
- [78] B. 山内, “一种基于前沿的自主探索方法”, 发表在1997年IEEE机器人和自动化计算智能国际研讨会CIRA '97. “迈向机器人技术和自动化的新的计算原则” IEEE, 1997, 页。146 - 151.
- [79] Y. 布尔达, H. 爱德华兹, D. 帕萨克, Storkey, T. 达雷尔和A. A. Efros, “好奇心驱动学习的大规模研究”, arXiv预印本arXiv: 1808.04355, 2018年.
- [80] R. Houthoofd, X. 陈, Y. 段, J. 舒尔曼, F. De Turck和P. 文摘, “价值: 差分信息最大化探索”, 《神经信息处理系统的进展》, 2016年, 页。1109 - 1117.
- [81] D. Pathak, D. 甘地和A. Gupta, “通过分歧的自我监督探索”, 在ICML, 2019年.
- [82] Y. 布尔达, H. 爱德华兹, Storkey和O. Klimov, “随机网络蒸馏的探索”, arXiv预印本arXiv: 1810.12894, 2018.
- [83] A. Vaswani, N. Shazeer, N. 帕尔马, J. Uszkoreit, L. 琼斯, N. 戈麦斯L凯泽和我, “注意就是你所需要的”, 《神经信息处理系统的进展》, 2017年, 页。5998 - 6008.
- [84] S.K. 罗摩克里希南和K. 格劳曼, “积极视觉探索的伙伴政策学习”, 《欧洲计算机视觉会议 (ECCV) 论文集, 2018年, 页。413 - 430.
- [85] S. 宋, A. 曾, A. X. 常, M. Savva, 萨瓦雷斯和T. “Im2pano3d: 推断360结构和语义”, IEEE计算机视觉和模式识别会议论文集, 2018年, 页。3847 - 3856.
- [86] N. 萨维诺夫, Raichuk, R. 马尼尼, D. 文森特, M. Pollefeys, T. Lillicrap和S. 盖利, “通过可达性的情景好奇心”, arXiv预印本arXiv: 1810.02274, 2018.
- [87] N. 哈伯, D. Mrowca, S. 王, L. F. Fei-Fei和D. L. Yamins, “学习与内在动机, 自我意识的代理玩”, 在神经信息处理系统的进展, 2018年, 页。8388 - 8399.
- [88] Y. 朱, R. Mottaghi, E. Kolve, J. Lim, 古普塔, L. 费菲和A. Farhadi, “使用深度强化学习的室内场景中的目标驱动视觉导航”, 2017年IEEE机器人与自动化国际会议 (ICRA). IEEE, 2017, 页。3357 - 3364.
- [89] D. S. Chaplot, R. 萨拉库蒂诺夫, Gupta和S. Gupta, “视觉导航的神经拓扑大满贯”, 发表在IEEE/CVF计算机视觉和模式识别会议论文集, 2020年, 页。12 875 - 12 884.
- [90] F. Bonin, A. Ortiz和G. 奥利弗, “移动机器人的视觉导航: 调查”, 《智能与机器人系统杂志, 第1卷。53岁, 没有。3, p. 263, 2008.
- [91] J. 富恩特斯-帕切科, J. Ruiz-Ascencio和J. M. “视觉同步定位和地图: 一个调查”, 艺术艺术 情报审查, 卷。43岁, 没有。1, pp. 55 - 81, 2015.
- [92] L. E. 卡拉基, P. Svestka, J. -C. Latombe和M. H. “高维协调空间中路径规划的概率路线图”, 《IEEE关于机器人与自动化的事务》, 第1卷。12日, 没有。4, pp. 566 - 580, 1996.
- [93] S. M. 拉瓦莱和J. 库夫纳, “快速探索随机树: 进展与前景”, 算法与计算机人: 新方向, 不。5, pp. 293 - 308, 2001.
- [94] A. 多索维茨基和V. 科尔顿, “通过预测未来学习行动”, arXiv预印本arXiv: 1611.01779, 2016.
- [95] P. “改进时间差异学习的泛化: 后继表示”, 神经计算, 第1卷。5, 没有。4, pp. 613 - 624, 1993.
- [96] Y. 朱, D. 戈登, E. Kolve, D. 福克斯, L. 费菲, A. 古普塔, R. 莫塔吉和A. 法哈迪, “使用深度后继表示的视觉语义规划”, IEEE计算机视觉国际会议论文集, 2017年, 页。483 - 492.
- [97] A. Barreto, W. 达布尼, R. 穆诺斯, J. J. 狩猎, T. Schaul, H. P. 范·哈塞尔和D. “强化学习中迁移的继任者特征”, 神经信息处理系统的进展, 卷。30, pp. 4055 - 4065, 2017.
- [98] D. 戈登, 卡迪亚, D. 帕里克, J. 霍夫曼和D. Batra, “分裂网: 具体化视觉导航的Sim2sim和任务2任务转移”, IEEE计算机视觉国际会议论文集, 2019年, 页。1022 - 1031.
- [99] J. 舒尔曼, F. 沃尔斯基, 达里瓦尔, Radford和O. “近端策略优化算法”, arXiv预印本arXiv: 1707.06347.
- [100] A. 帕萨维安, Toshev, M. “Fiser J. “科塞克阿. 瓦希德和J. 戴维森, “语义目标驱动导航的视觉表示”, 2019年机器人与自动化国际会议 (ICRA). IEEE, 2019, 页。8846 - 8852.
- [101] J. 舒尔曼, P. 莫里茨, 莱文, M. 乔丹和P. 摘要, “使用广义优势估计的高维连续控制”, arXiv预印本arXiv: 1506.02438, 2015.
- [102] Z. D. 郭, M. G. 阿扎尔, B. Piot, B. A. Pires和R. Munos, “神经预测信念表征”, arXiv预印本arXiv: 1811.06407, 2018.
- [103] C. 甘, Y. 张, J. 吴, B. 龚和J. B. 泰南鲍姆, “看、听和行动: 走向视听具体化导航”, 2020年IEEE机器人和自动化国际会议 (ICRA). IEEE, 2020, pp. 9701 - 9707.
- [104] P. 安德森, 问, 吴, D. Teney, J. 布鲁斯, M. 约翰逊, N. 德斯豪夫, 我, 里德, S. 古尔德和A. 范登·亨格尔, “视觉和语言导航: 在真实环境中解释基于视觉基础的导航指令”, 发表在IEEE计算机视觉和模式识别 (CVPR) 会议论文集, 2018年.
- [105] J. Thomason, M. 穆雷, M. Cakmak和L. “视觉与对话导航”, 机器人学习会议 (CoRL), 2019年.
- [106] R. J. “连接主义强化学习的简单统计梯度跟踪算法”, 机器学习, 卷。8, 没有。3-4, pp. 229 - 256, 1992.
- [107] J. Straub, T. 惠兰, L. 妈妈, Y. 陈, E. Wijmans, S. 绿色, J. J. 恩格尔, R. Mur-Artal, C. 任, S. Verma等., “复制品数据集: 室内空间的数字复制品”, arXiv预印本arXiv: 1906.05797, 2019年.
- [108] J. 约翰逊, B. Hariharan, L. 范德马顿, L. 费菲, C. 劳伦斯·齐特尼克和R. 摘要, “毕业论文: 组合语言和基本视觉推理的诊断数据集”, 发表在IEEE计算机视觉与模式识别会议论文集, 2017年, 页。2901 - 2910.
- [109] E. 托多罗夫, T. Erez和Y. Tassa, “MuJoCo: 基于模型的物理引擎”, 2012年IEEE/RSJ智能机器人和系统国际会议. IEEE, 2012, 页。5026 - 5033.
- [110] C. 高, J. 陈, S. 刘, L. 王, 问, 张和Q. 吴, “漫游对象感知知识推理的远程体现参考

- “IEEE/CVF计算机视觉和模式识别会议论文集”, 2021页, 页。3064 - 3073.
- [111] J. 太阳, DA. .-黄, B. 卢, 是的。.-刘, B. 周和A. “平板: 在程序任务中的变压器的视觉接地计划”, 2021年。
- [112] F. 萨德吉和S. 莱文, “机器人: 真实图像”, 机器人: 科学与系统十三, 麻省理工学院, 2017年7月12-16日。#
- [113] D. M. 熊, E. 王, D. Mrowca, F. J. 粘合剂, H. -Y. F. 东, R. Pramod, C. Holdaway, S. 陶, K. 史密斯, L. Fei-Fei等。Physion: 评估人类和机器视觉的物理预测, ” arXiv预印本arXiv: 2106.08261, 2021。
- [114] J. 段, S. Y. B. 吉安和C. Tan: “空间: 在三维环境中的物理交互和因果学习的模拟器”, arXiv预印本arXiv: 2108.06180, 2021年。
- [115] J. 段, S. Yu, S. 茯苓。温和C. Pip: 物理互动通过带有跨度选择的心理意象进行预测, ” arXiv预印本arXiv: 2109.04683, 2021。
- [116] T. 纳加拉扬和K. 格劳曼, “学习为三维环境中的交互探索提供了景观”, 神经信息处理系统的进展, 33, 2020年。
- [117] M. 王, Y. 邓, X. 香港, A. H. 普拉萨德, S. 熊和B. 朱, “薄粒光滑粒子流体动力学uid”, arXiv预印本arXiv: 2105.07656, 2021。#
- [118] A. 库兹涅佐夫, K. Mullia, Z. 徐, M. “哈桑和R. “新大学: 多分辨率神经网络”, 《ACM图形学学报》(TOG), 第1卷。40岁, 没有。4, pp. 1 - 13, 2021。
- [119] S. R. 里希特, H. A. 阿尔海贾和V. 科尔顿, “增强摄影现实主义增强”, arXiv预印本arXiv: 2105.04619, 2021年。
- [120] T. -Y. 林, M. 梅尔, S. 贝隆吉, J. 海斯, P. Perona, D. Ramanan, P. 娃娃和C. L. 齐特尼克, “微软可可: 背景中的共同物体”, 在欧洲计算机视觉会议上。施普林格, 2014, 页。740 - 755。
- [121] J. Straub, T. 惠兰, L. 妈妈, Y. 陈, E. Wijmans, S. 绿色, J. J. 恩格尔, R. Mur-Artal, C. 任, S. Verma等。 “复制品数据集: 室内空间的数字复制品”, arXiv预印本arXiv: 1906.05797, 2019年。
- [122] S. K. 罗摩克里希南, Gokaslan, E. Wijmans, O. Maksymets. 克拉克, J. 特纳, E. Undersander, W. Galuba. 韦斯特伯里。X. 常, M. 萨夫瓦, Y. 赵和D. Batra, “栖息地-矩阵港3d数据集 (hm3d): 1000个大规模3d环境”, 2021年。
- [123] I. Armeni, S. 萨克斯, A. R. Zamir和S. “用于室内场景理解的联合2d-3d语义数据”, arXiv预印本arXiv: 1702.01105, 2017。
- [124] E. M. 米哈伊尔, J. S. 伯特利和J. C. “现代摄影测量概论”, 《纽约, 第1卷。19, 2001。
- [125] [125], 搅拌机——一个3D建模和渲染包, [125], 2018。在线可用: <https://github.com/alicevision/meshroom>
- [126] A. Yu, V. 是的, 我。Tancik和A. 金泽, “像素元: 来自一张或几张图像的神经辐射元”, 发表在IEEE/CVF计算机视觉和模式识别会议论文集, 2021年, 页。4578 - 4587。
- [127] R. 马丁-布鲁拉, N. 拉德万, M. S. Sajjadi, J. T. 巴伦。多索维茨基和D. 达克沃斯, “野外的神经: 不受约束的照片收藏的神经辐射”, IEEE/CVF计算机视觉和模式识别会议论文集, 2021年, 页。7210 - 7219。
- [128] Z. 郝, A. 马利亚, S. 贝隆吉和M. -Y. 刘, 《高夫: 我的世界的无监督3d神经渲染》, arXiv预印本arXiv: 2104.07659, 2021年。
- [129] Y. 朱, T. 高, L. 风扇, S. 黄, M. 埃德蒙兹。刘, F. 高, C. 张, S. 齐, Y. N. 吴等人。 , 《黑暗, 超越深度: 一种向具有人类常识的认知ai的范式转变》。6, 没有。3, pp. 310 - 345, 2020。
- [130] M. 罗曼, J. 萨尔瓦多。肯巴维和R. “通过学习与物体互动来学习物体”, 神经信息处理系统的进展, 33, 2020年。
- [131] R. Bhirangi, T. Hellegren, C. 马吉迪和A. 古普塔, “树脂金: 多用途, 可更换, 持久的触觉皮肤”, 在第五届机器人学习年会, 2021年。
- [132] B. 史密斯, C. 吴, H. 文, P. Peluse, Y. 谢赫, J. K. 霍奇金斯和T. “用弹性约束密集的手表面跟踪”, ACM图形交易 (TOG), 第1卷。39日, 没有。6, pp. 1 - 14, 2020。

- [133] C. 陈, 美国耆那教, C. 五味子, S. V. A. 加里, Z. Al-HalahV. K. Ithapu, P. 罗宾逊和K. 格劳曼, “视听具体化导航”, 环境, 卷。97, p. 103, 2019。
- [134] L. 帕纳特和S. “合作, 多智能体学习: 现状”, 自主代理和多智能体系统, 卷。11日, 没有。3, pp. 387 - 434, 2005。
- [135] YC. .-刘, J. 田, 纽约。.-马, N. Glaser, CW. .-郭和Z. Kira, “Who2com: 通过可学习的握手交流的协作感知”, 在2020年IEEE机器人与自动化国际会议 (ICRA)。IEEE, 2020, pp. 6876 - 6883。

## V. 传记部分

段家飞获得了学士。(最高公司2021年毕业于新加坡南洋科技大学电气与电子工程学院。他目前在新加坡科学、技术和研究机构 (A\*STAR) 的信息通信研究所担任研究工程师。他目前的研究兴趣是体现了人工智能和计算认知科学。



于参孙获得了英国英语英语学士学位。信息学位2020年, 新加坡技术与设计大学的系统技术与设计。他目前在新加坡科学、技术和研究机构 (A\*STAR) 的高性能计算研究所担任研究工程师, 研究基础人工智能研究和实体人工智能。



许丽诚收到了B元奖。塞舌尔应用程度2007年毕业于新加坡国立大学 (NUS)。她获得了博士学位。2017年毕业于新加坡国立大学, 获得电气和计算机工程学位。自2007年以来, 她一直在新加坡的信息通信研究所工作。她目前的研究兴趣包括计算机视觉、多模态深度学习、增量学习和联邦学习。



朱红获得博士学位。学位在com-2014年, 新加坡南洋技术大学的计算机工程专业。他目前是新加坡新加坡信息通信研究所的研究科学家。他的研究兴趣包括多媒体内容分析和分割。







切斯顿·谭接受了B. Sc。（最高荣誉）  
电气工程的毕业生  
以及加州大学计算机科学分校，  
以及博士学位。学位从  
大脑与认知科学系  
麻省理工学院。他现在是  
美国信息通信研究所的资深科学家  
搜索（I2R），科学、技术和  
研究（A\*STAR），新加坡。