

元宇宙中的人工智能技术与应用

武强¹, 季雪庭², 吕琳媛¹

(1. 电子科技大学基础与前沿研究院复杂系统与社会计算研究中心, 四川 成都 610054;

2. 杭州师范大学阿里巴巴商学院, 浙江 杭州 311121)

摘要: 元宇宙整合并应用多种数字技术, 从而产生了一种虚拟与现实、数字与应用交叉融合的互联网社会形态。人工智能是基于海量数据来模仿人类的智力, 从而执行任务并对其自身进行迭代改进的系统(机器)。在构建元宇宙的过程中, 人工智能技术不仅推动了元宇宙中的其他关键技术(人机交互、通信、机器人等)的发展, 还能在元宇宙中直接进行内容的创作, 从而链接起现实世界和虚拟世界。通过对元宇宙和人工智能的概念和技术进行梳理, 系统介绍了人工智能技术对构建元宇宙关键技术的促进作用, 并详细说明了人工智能技术在元宇宙中的应用过程。最后, 对人工智能技术在元宇宙中的发展和未来趋势进行了展望。

关键词: 元宇宙; 人工智能; 内容创作

中图分类号: TP39

文献标志码: A

doi: 10.11959/j.issn.2096-6652.202241

Artificial intelligence technologies and applications in the metaverse

WU Qiang¹, JI Xueting², LYU Linyuan¹

1. Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu 610054, China

2. Alibaba Business School, Hangzhou Normal University, Hangzhou 311121, China

Abstract: Metaverse integrates and applies a variety of digital technologies, resulting in an Internet social form that integrates virtual and reality, digital and application. Artificial intelligence (AI) are systems and machine that imitate human intelligence to perform tasks and iteratively improve themselves based on the information gathered. In the process of constructing the metaverse, AI technologies not only vigorously promote the development of crucial metaverse technologies (human-computer interaction, communication, robotics, etc.) but also enables direct content creation in the metaverse, organically connecting the real and virtual worlds. By sorting out the concepts and representative technologies of the metaverse and AI, the optimization progress of AI technologies for constructing key technologies in the metaverse was introduced and the application process of AI in the metaverse was detailed. Finally, the AI technology's development and application trends in the metaverse prospected.

Key words: metaverse, artificial intelligence, content creation

0 引言

美国互联网巨头“元(Meta)”(原脸书(Facebook))首席执行官扎克伯格(Zuckerberg)指出“你可以把元宇宙(metaverse)看作一个具身性的互联网。在这里, 你不再仅仅浏览内容, 而是你本身就在内容之中”。现阶段, 元宇宙的概念受

到极大关注。2020年, 罗布乐思(Roblox)公司上市, 第一次把元宇宙这一概念写入招股书。同年, 以太坊2.0正式启动, 其能够提供更好的可扩展性、安全性和可编程性, 标志着区块链基础设施里程碑式升级。2020—2021年, 国内外巨头元、英佩游戏、英伟达、微软、腾讯、网易、字节跳动等纷纷布局元宇宙概念相关产业。

收稿日期: 2022-07-18; 修回日期: 2022-08-13

通信作者: 吕琳媛, linyuan.lv@uestc.edu.cn

随着人工智能 (artificial intelligence, AI)、人机交互 (human-computer interaction, HCI)、区块链、云计算等技术的快速发展和终端硬件的迭代升级,目前元宇宙的发展和演化速度已经超出了人们的预期。元宇宙极有可能成为“多维度、全感官、沉浸式”的新互联网形态^[1]。元宇宙整合了多种数字技术,从而产生了一种虚拟与现实、数字与应用交叉融合的互联网社会形态^[2]。在使用这些新技术构建元宇宙的过程中,人机交互技术(以扩展现实技术为代表)提供了沉浸式体验,数字建模技术(以数字孪生技术为代表)生成了现实世界的镜像,去中心化的可信数字经济体系(以区块链技术^[3]为代表)在经济系统、社交系统、身份系统上密切融合并关联虚拟世界与现实世界。与此同时,元宇宙允许每个参与的用户进行内容的生产和编辑。因此,元宇宙本身并非一种单一的技术,而是建立在多种技术基础上的、关于未来世界构想的数字化应用。事实上,从技术角度来看,今天的人类才刚刚迈入元宇宙的大门。

AI 是基于收集的数据和信息不断对自身的能力进行改进的系统或机器,从而可以在一定程度上模拟人类智能来执行任务^[4]。AI 技术已被广泛应用于人们的生活中。例如,聊天机器人使用 AI 更快速、高效地理解用户问题并提供更有效的解答;智能助手使用 AI 来解析主人的语音或者文本中的关键信息,从而更好地为主人进行个性化服务;推荐系统可以根据用户的观看习惯自动推荐视频短片等^[5]。从无人驾驶到人脸识别,人们无时无刻不在使用 AI 技术来高效解决现有问题。现阶段, AI 技术的发展趋势是“通用性”和“可解释性”。

元宇宙技术是由多个技术发展、演进并交叉应用的技术组合^[6],在构建元宇宙的过程中, AI 技术不仅推动了人机交互、通信、机器人等元宇宙关键技术的发展,而且能够在元宇宙中直接进行内容创作等,其本身已成为构建元宇宙的关键技术之一,链接起现实世界和虚拟世界,如图 1 所示。

现阶段,国内外对 AI 的研究日益深入。因此,对 AI 技术构建元宇宙和促进元宇宙其他关键技术的发展进行总结和梳理十分必要。本文首先介绍元宇宙的概念、关键技术及架构以及 AI 概念和应用,总结 AI 对交互、通信、机器人等元宇宙关键技术的支撑和提升作用。其次,阐述 AI 技术在元宇宙中数字孪生、数字虚拟人、个性化数字化身、无障碍沟通等典型场景中的应用。最后,对 AI 技

术在元宇宙中的未来发展趋势进行展望。

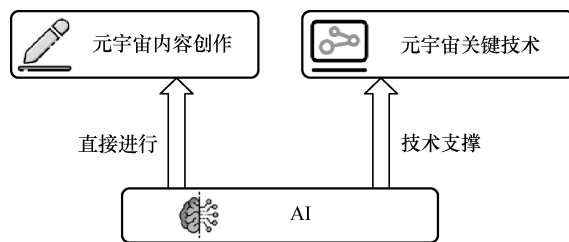


图1 AI 技术对构建元宇宙的作用

1 相关概念

1.1 元宇宙的概念

元宇宙这个词源于尼尔·斯蒂芬森 (Neal Stephenson) 的文学作品《雪崩》^[7],该书描述了一个名为“矩阵”的虚拟数据空间。书中的元宇宙是一个通过个人终端和虚拟现实 (virtual reality, VR) 护目镜访问的 3D 虚拟空间,与现阶段 VR 有很多共同之处。虽然当前元宇宙重新受到多方关注且发展迅速,但是关于元宇宙的定义还未形成统一的概念。例如,一些观点认为元宇宙是一个虚拟世界,是平行于现实世界的存在。在元宇宙中,所有现实生活中的人都有一个网络分身 (avatar)。另外一些观点认为,元宇宙通过具有链接感知的互联网和共享特征的虚拟空间,增强虚拟和现实的联系。尽管对元宇宙的定义和认识不同,但如果以现实世界为参照,可以大致提炼出元宇宙的基本元素,即“协作关系、生产资料、技术体系、经济体系”。因此,需要在元宇宙中对这些基本元素进行改造或者重构。

1.2 元宇宙关键技术

构建元宇宙的技术纷繁复杂。本文通过总结和整理相关技术,将组成元宇宙的技术分为 8 个类别^[2]: ①区块链技术; ②交互技术; ③通信技术; ④云和边缘计算; ⑤高性能计算; ⑥物联网和机器人技术; ⑦网络技术; ⑧人工智能技术。

同时,本文将元宇宙技术架构分为基础设施层和实际应用层,如图 2 所示。

基础设施层包括硬件基础设施、计算机网络、云计算和高性能计算。

实际应用层在基础设施层之上,以 AI 技术为支撑,以区块链技术为去中心化的可信任计算能力,将人机交互、通信技术、机器人技术等元宇宙核心技术进行整合和提升。

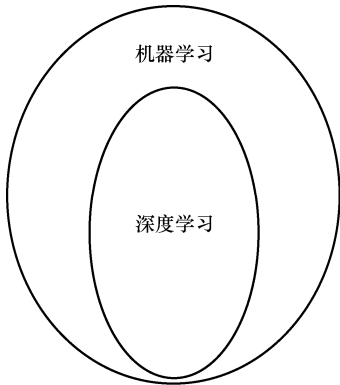


图5 机器学习与深度学习的关系

1.4.1 机器学习

机器学习，顾名思义，就是让机器具有学习能力。机器学习通过设计算法来挖掘数据中蕴含的潜在规律，得出相应的结果。机器学习系统通过不断降低自身得到的结果和人为结果之间的误差，对算法模型的参数进行调优，从而形成更强的数据分析能力。但是，这种学习不会让机器产生推理、联想，甚至意识。机器学习主要应用高等数学和计算机科学挖掘数据中蕴含的潜在规律和经验，代替人做预测、分类、数据生成等工作，而这些工作以往只能依靠人的智力才能完成。例如在语音转文本任务中，机器学习基于输入的一段语音信号数据，识别并输出对应的文字内容；在图片识别任务中，机器学习基于输入的图片数据，分析并输出该图片的类别；在对话系统中，机器学习基于现阶段对话文本的内容，联系上下文并回复对话内容；在电视游戏中，机器学习基于不断变化的屏幕图片，计算并得

到下一步游戏任务的动作。可以将机器学习的训练过程看作找到一个最适合具体场景任务的“函数”的过程。但是，传统的机器学习需要针对数据进行专业的预处理工作，这需要较强的专业能力和较长的处理时间。

1.4.2 深度学习

深度学习是机器学习中的一类方法，在学术界和产业界都获得了巨大成功。深度学习模仿人类大脑的生理结构和思维过程，利用神经级元组成的多层神经级网络进行学习。深度学习通过多层神经级网络来获取数据的特征信息，最后得到预测、分类等结果。实际上，与机器学习寻找最适合具体场景任务的“函数”的过程类似，深度学习通过设计多层神经网络结构(3层或以上)来训练参数，进而获得“函数”。相比于机器学习繁杂的数据预处理过程，深度学习可以直接使用原始数据，获得像人一样的基础感知能力——听、读、看，即语音识别^[13]、文本识别^[14]、图片识别^[15]。深度神经网络的基础结构(如图6所示)由输入层(input layer)、隐藏层(hidden layer)、输出层(output layer)组成，其中，输入层进行数据的输入，隐藏层(中间层)的神经网络表示数据信息并挖掘数据所蕴含的规律，输出层输出最后结果。

1.4.3 AI 技术研究进展

现阶段，基于深度学习的 AI 技术应用十分广泛。笔者将这些技术按照应用分为计算机视觉、自然语言处理、生成模型(图片、语言)、强化学习

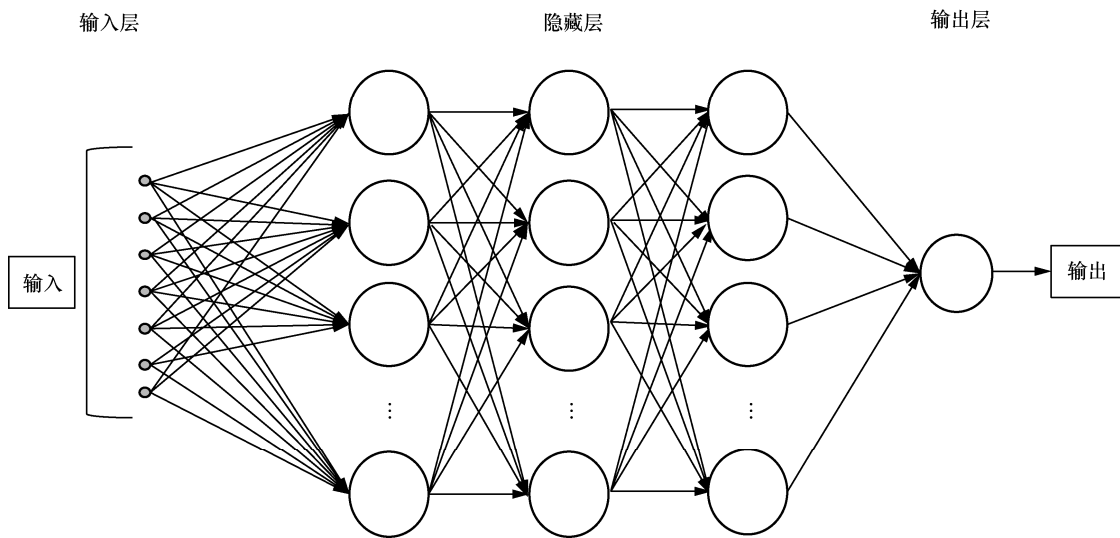


图6 深度神经网络的基础结构

(reinforcement learning, RL) 4 个领域。

(1) 计算机视觉

计算机视觉是 AI 应用的一个重要领域, 其让计算机和系统能够从图片、视频和其他视觉输入中获取有意义的信息, 并根据该信息采取行动或提供建议。如果说 AI 赋予计算机思考的能力, 那么计算机视觉赋予计算机发现、观察和理解的能力。在计算机视觉中, 以卷积神经网络 (convolutional neural network, CNN) 为代表的技术极大地推进了计算机视觉的发展。CNN 是一种前馈神经网络, 它的人工神经元可以响应一部分覆盖范围内的周围单元, 对图片处理有出色表现。人们目前已经习以为常的人脸识别、车牌识别等业务都是以 CNN 模型为基础的。因此, 在元宇宙中, 计算机视觉技术的应用也极其广泛, 可以成为元宇宙的“眼睛”, 识别这个数字世界。

(2) 自然语言处理

自然语言处理使用机器学习来剖析文本的结构和含义。借助自然语言处理应用, 可以分析文本并提取关于人物、地点和事件的信息, 以更好地理解社交媒体内容的情感和客户对话。现阶段, 前沿自然语言处理的 AI 方法是基于迁移学习的大模型, 即在海量的语料中进行训练, 再平滑迁移到机器翻译、机器对话等任务中。迁移学习根据已有经验解决相似任务, 将训练好的内容应用到新的任务上, 即将源域 (被迁移对象) 应用到目标域 (被赋予经验的领域)。目前迁移学习的应用场景非常广泛。在自然语言处理中, Brown T B 等人^[16]提出的超大规模预训练的自然语言模型 GPT-3, 就是使用海量文本数据预训练语言模型, 然后在不同的语言任务 (翻译、对话等) 中微调神经网络, 模型性能获得了大幅提升。因此, 自然语言处理技术在元宇宙中的应用潜力巨大, 可以成为元宇宙中的“嘴巴”, 在这个数字世界中交流。

(3) 生成模型

生成模型的输入通常是数据 (图片、文本) 的性质, 而输出是性质对应的数据 (图片、文本)。这种生成模型相当于构建了数据的分布, 因此利用这类模型, 人们可以完成数据自动生成 (采样)、信息补全等工作。在训练过程中, 生成模型已知观察变量 X 和隐含变量 z , 对 $p(z|X)$ 概率密度函数进行建模, 即它根据输入的观察变量 X 得到隐含变量 z 出现的可能性。在具体生成数据的过程中, 生成模

型对 $p(X|z)$ 进行建模, 输入是隐含变量 z , 输出是观察变量 X 。生成式对抗网络 (generative adversarial networks, GAN)^[17]就是一种典型的生成模型, GAN 包含两个部分: 生成网络和鉴别网络。二者互相对抗, 共同演进, 在此过程中两个网络生成和鉴别的水平都越来越高。最后, 生成网络生成的内容 (图片或者文字) 足以达到以假乱真的水平, 鉴别网络无法识别。因此, 生成模型技术在元宇宙中可以担当“设计师”的角色, 设计这个数字世界。

(4) 强化学习

强化学习是机器学习中的一个领域, 强调如何基于环境行动, 以取得最大化的预期利益。深度强化学习是使用神经网络构建强化学习主体的方法。在现实问题中, 场景多种多样, 环境、行为、回报很难穷尽。但是, 即使面对没有见过的情况, 具有神经网络的强化学习算法也能做出决策。著名的 AlphaGo 击败人类围棋世界冠军, AI 击败多人德州扑克选手^[18]等, 就是深度强化学习的典型应用。在元宇宙中, 强化学习可以承担相应的资源决策、环境优化等工作。

除了以上 4 个领域, 在现实世界中, 很多数据是属于用户的, 虽然现有的 AI 技术已经非常适合处理这些数据, 但是, 需要考虑保护数据的私密性。联邦学习^[19]可以安全高效地实现数据合作。联邦学习系统是面向多客户端的模型训练系统, 各个客户端在参与训练时, 数据保留在本地, 不会被发送到其他客户端或中心服务器。中心服务器通过对客户端发送的本地模型进行整合, 最终完成共享模型的训练。在元宇宙中, 考虑到保护用户隐私和扩展计算资源, 联邦学习技术会大有用处。

2 人工智能推动元宇宙关键技术发展

2.1 人工智能与人机交互技术

(1) 元宇宙中的人机交互技术

人机交互^[20]技术被用来研究人和计算机系统之间的交互关系。人机交互的发展历史就是人和计算机互相适应的历史, 其中交互信息的输入从精准的命令到模糊的语音或者手势, 发生了巨大的变化。人机交互的主要过程是通过人机交互界面完成的, 主要有两个阶段: 命令行、图形界面。近些年, 随着交互技术的不断发展, 人机界面逐渐向自然交互的方式发展, 例如交互过程与人的身体、行为习惯相匹配。因此, 出现了触摸交互和 3D 交互界面。

人类对外界情况的感知来自各种感触信号,包括视觉、听觉、触觉、味觉信号等。通过时间序列、空间位置、信号强度等因素组合这些信号,并作用于人的感觉器官,就是人类感知的过程。类似地,在元宇宙中,研究人员利用传感装置模拟并产生这些感触信号,将各种信号的次序、位置、强弱、彼此间的组合关系写入软件,并整合信号、软件、设备,形成一个数字交互系统。实际上,这就是元宇宙中人机交互技术的目标。人机交互技术在元宇宙中的应用体现在以下3个方面。

①三维图形实时生成。利用计算机模型生成三维图形是实现元宇宙视觉交互的基础,而实时生成这些三维图形却并不容易。例如在飞行虚拟系统中,想要达到实时生成的目的,需要很高的图片刷新频率、图片质量和逼真的虚拟环境。因此实现三维图形实时生成具有一定的难度。

②立体显示。用户带上VR眼镜后立即进入虚拟现实系统。实际上,用户两只眼睛分别看到不同的图片。奇数帧和偶数帧的图片分别对应不同的眼睛。因此,奇数帧、偶数帧之间的不同会造成视觉差,从而产生立体的可视效果。广角立体显示技术让人们能够感受到逼真、立体的画面,在视觉感知方面,虚拟现实已经十分成熟,当用户穿戴相应的设备后,就能在虚拟的环境中感受到真实的视觉体验。

③传感反馈。在数字显示系统中,用户除了视觉外无法感受物体的形状。但是,在虚拟现实系统中,一系列传感设备可以让用户对虚拟世界中的物体进行触觉感知。例如,用户通过虚拟现实系统看到一个虚拟的花瓶,并且还能感受到花瓶的触感和温度。这就是传感反馈技术在人机交互中的应用效果。

(2) 人工智能推动人机交互技术发展

交互式人工智能(conversational artificial intelligence, CAI)是指通过对话、交互体现出来的智能,通常智能系统与用户或环境进行交互,并在交互中实现学习与建模。在数字化时代,以语音、手势等为媒介的交互方式正逐步成为人机交互的主流。AI技术的迅猛发展也为人机交互提供了理论和技术支持。作为AI与人机交互的结合体,交互式人工智能进一步降低了人机交互的技术门槛,增加了用户人数,在多领域展现出广阔的应用前景与巨大的商业价值。

作为AI和人机交互技术融合的产物,交互式

人工智能让机器以一种人性化的方式理解人类语言与情感,并提供智能响应^[21-22]。当前,交互式人工智能为传统行业带来显著的价值提升,其优势体现在高效率、标准化。未来,交互式人工智能的发展将会侧重多样化、定制化。

近年来,以手势交互为代表的人机交互方式在AI技术的推动下取得了长足进步,实现了交互模式更加自然、交互效果更加智能^[23]。随着可穿戴设备的快速发展,如何使人机交互在普适计算环境中协同发展,成了一个关键问题。针对该问题,我国科研人员研究出有效解决智能眼镜、智能手表的文本输入问题的方法^[24-25]。采用问答系统、语音识别和语音合成等AI技术的以语音为媒介的交互方式,已经被广泛应用,例如银行的客服机器人。

现阶段,在人机交互过程中,AI技术不断提供更自然的人机互动方式,甚至已经开始尝试理解人类的情绪。因此,交互式技术在AI技术的支撑下,将在元宇宙中发挥重要的人机沟通、人机协调的作用。

2.2 人工智能推动通信技术发展

(1) 元宇宙中的通信技术

人们通常将通信技术视为元宇宙的数字“高速公路”。实际上,通信就是指信息的传输和交换。人们的日常生活离不开现代通信技术的支撑。社会生产力的不断发展对通信技术提出了更高的标准,也推动了通信技术的蓬勃发展。元宇宙的沉浸感、低延迟、随时随地接入等特点对无线通信技术提出了极高的要求,如何实现随时随地的低延迟的连接,成为当前的技术难点。

元宇宙打破了隔在通信双方之间的时空束缚,让交流双方置身于同一个虚拟空间,进行实时的沟通和互动,使双方切身体会到真实的氛围感和沉浸感,极大地提高了沟通效率和交互体验。通信技术架起了现实世界与元宇宙间的桥梁,只有在高性能通信技术的支撑下,才能给予用户在元宇宙中游戏、社交、生产、办公等场景的舒适体验。

(2) 人工智能推动通信技术发展

在过去的几十年中,从最早的寻呼机到现在的智能手机,移动通信网络发生了巨大的变革,目前已经发展到第六代(6G)网络。移动通信与AI的融合越来越紧密。其原因在于,移动通信网络本身正日益复杂化,通信业务生态也日益多样化。如今,语音甚至视频通信已经不能满足人们的需求,例如元宇宙具有随时随地的接入需求。因此,通信业务

的场景日趋复杂。并且,传统的人工规则制定和执行已经无法适应网络个性化服务等场景。因此,通信系统需要一套非人工干预、全智能处理的技术体系来支撑通信网络的快速迭代和发展。

以5G和6G通信为例。5G系统复杂度高,存在多个位置模式,现有的解决方案难以准确建立模型。针对该问题,AI技术可以通过建立深度神经网络挖掘数据特征,自动适应和识别未知场景和模式,降低系统复杂度^[26]。在5G网络正式商用后,6G网络的研发已经全面展开。在6G网络中,AI技术不仅仅是一种优化工具,而是全新的网络架构中的基础设施,AI将体现在接口协议栈的设计中^[27]。具体来讲,在6G网络中,AI技术可以支撑多类型资源跨域管理、无线资源的智能调度、实时的业务需求匹配,还可以实现接入网和核心网网元的智能化管理。

2.3 人工智能推动机器人技术

(1) 元宇宙中的机器人技术

机器人技术综合了计算机、控制论、机构学、信息和传感技术、AI、仿生学等多学科,是当前产业和学术界研究活跃、应用广泛的领域。机器人技术的发展分为3个阶段:第一阶段是人遥控操作机器人的行为;第二阶段是机器人按照提前设定的程序内容进行特定任务的操作;第三阶段是在传感设备获取的环境信息基础上,基于AI技术进行环境识别和任务理解,最后智能做出相应决策行为。

“机器人+元宇宙”的想法是进一步扩大人类影响力,为移动性和机器人技术在元宇宙中找到一个角色。通过将机器人链接到元宇宙,人们将实现在现实世界和虚拟现实之间的自由移动。相比元宇宙提供的身临其境的现场感,“机器人+元宇宙”的体验将更进一步,机器人将成为身体感官的延伸,让人们能够通过“机器人+元宇宙”重塑和丰富日常生活。“元宇宙”机器人系统将在顶层实现虚拟世界与现实世界的联动,在这个系统中,虚拟数字人和现实世界中的机器人实现了互动。在现实世界中,物理机器人可以同步接收和执行虚拟人的虚拟命令,机器人在现实世界中的状态也将在虚拟世界中实时显示。

(2) 机器人技术与人工智能

AI技术的机器人应用案例如下。在货运仓库中,机器人需要根据客户需求,将数百万种产品放入箱子。传统机器人无法处理如此广泛的物品类

别,而基于计算机视觉AI技术的机器人,很容易训练、识别不同种类的箱子,从而在不需要人工干预的情况下完成仓储拣货。因此,基于AI技术的智能机器人是未来的主要发展方向。以深度学习为代表的AI技术可以增强智能机器人在各方面的能力。现阶段,机器人应用场景已经从室内拓展到室外,以深度学习为代表的AI技术被广泛应用于工业服务、家庭服务以及多机器人协作^[28]等。AI在机器人技术中的应用体现在以下3个方面。

①机器人场景识别。传统的机器人只需要在设定程序下工作,因此不需要了解其周围的环境。但是,现阶段出现了更多的智能化需求,需要机器人根据特定的环境做出与之相适应的行为。因此,机器人需要感知并理解它所处的环境,并提升智能化决策能力。基于计算机视觉的环境识别和理解是机器人对当前工作环境进行构建的关键步骤之一。该技术不仅需要实现三维地图与语义信息的关联,还要分类识别场景中的景物^[29]。

②机器人抓取。机器人学习抓取动作是智能机器人研究的重要部分^[30]。研究内容包括抓取位置识别、运动规划和抓取姿势控制等。具体来讲,在给定抓取目标任务后,智能机器人基于深度神经网络获取场景特征,并推断出针对不同形状、不同摆放方向物体的抓取定位和操作行为。

③工业机器人。可以使用工业机器人代替人完成一些重复性和危险性高的任务。例如,消防机器人可以代替消防员进入火灾现场;工业分拣机器人代替工人完成繁杂的分拣工作;自主作业的挖掘机器人^[31]基于AI技术对自身所处环境进行地图建模,然后自主进行挖掘施工。

3 人工智能技术在元宇宙中的应用

现阶段,AI在元宇宙的应用主要有4个方面:数字孪生、数字虚拟人、个性化数字化身和无障碍沟通。

3.1 数字孪生

在元宇宙中,用户的位置是随时变换的。与之对应的虚拟场景也需要随之变换,以保证其与用户的互动。在这个过程中,大量的图形、阴影变化都需要AI来实时做出判断和响应。除此之外,在工业领域,人们还希望通过数字孪生技术^[32]让元宇宙和现实世界能够相互影响,使得改变其中任何一者都会导致另一个的改变。数字孪生是对物理实体或

系统进行高度完整性的数字复制，并可以与现实世界保持交互作用的数字克隆体，它们可以在元宇宙里对物理实体进行实时处理。那么，如何才能能在元宇宙里面产生一个与现实世界中的物体高度一致，并且可以随现实状况即时改变的数字孪生体呢？这在很大程度上需要依靠 AI 技术来实现。

数字孪生是客观世界中的事物及其发展规律被软件数字定义后的一种结果。相比于设计图纸，数字孪生体最大的优势在于对实体对象的数字化动态仿真。数字孪生技术基于被数字建模的物体（本体）上传感器反馈的数据，使数字孪生体会根据数据实时发生变化，与本体同步。实际上，本体的传感器数据和历史数据都可以传输到数字孪生体上。基于本体的实时状态和传感器返回的数据，AI 技术可以比较容易地用深度学习方法在元宇宙中制造出真实物体的数字孪生体，并且在实物发生变化时，数字孪生体也同时发生变化。通过这种方法，可以成功地实现元宇宙与现实世界之间的虚实互动。

3.2 数字虚拟人

元宇宙中有很多数字虚拟人。这些数字虚拟人需要有一定的与人进行交互的能力，以满足人们与之进行互动的需要。而要实现这些数字虚拟人的智能化，仍然需要应用 AI 技术。为了大幅提升数字虚拟人的训练效果，经常会用到强化学习方法^[33]。这种学习方法让智能体在复杂的环境中不断试错，并根据返回结果的正误来给予“奖励”或“处罚”，借助这种思路，就可以在很短时间内让 AI 的水平有很大的提升，人们所熟知的 AlphaGo 实际上就应用了强化学习的方法。现在，强化学习已经被广泛应用于多个领域中。数字虚拟人由 3D 图片软件或其他模拟仿真工具制作而成。以数据形式存在的人与类人角色都可以算作数字人或者数字角色。

具体来说，虚拟环境需要适合人的感知需求，针对数字虚拟人的探索性研究主要包括以下 3 个方面。

(1) 运动感知：通过方向、位移、行程进行感知，实时计算数字虚拟人的动作、肢体位置。

(2) 力的感知：数字虚拟人在执行任务过程中，虽然不能看到施加力或者被施加力的具体数值，却能感觉到该力的舒适程度。

(3) 视域感知：观察范围感知。例如，不同身高、体重的数字虚拟人在虚拟三维环境中以不

同位置静止观察（在控制室内，遥望台上）或动态观察（随飞机、汽车、舰船运动时）到的视域范围不同。

3.3 个性化数字化身

在元宇宙中，每一个用户都需要通过一个数字化身体验元宇宙生活。因此，为了创建更逼真的虚拟环境，丰富的化身是必要的。但是，在很多元宇宙项目中，创作者只提供少数几个特定的模型或只允许玩家创建只有几个可选子模型（如鼻子、眼睛、嘴巴等）的完整化身。因此，不同玩家的化身形象高度雷同，这在很大程度上损失了一部分元宇宙本身的乐趣。针对以上问题，一个直观的方法是，建立一个人体特征库，然后随机从中抽取一些特征来让 AI 组合出一个化身。但这种方法存在问题，即如果不加限制，AI 组合出的图片可能会不尽如人意。为了解决以上问题，GAN 就派上了用场。以人脸生成为例，GAN 中的生成器会提取人脸库中的部分信息组合出一张新的人脸，然后把这张图片交给鉴别器，判别这是不是一张人脸。这样，生成器网络将被不断训练，直到这些“假图片”不被鉴别器网络识别，然后对鉴别器网络进行训练，以提高其识别精度。最终得到了一个性能良好的、可以用于人脸生成的神经网络。

3.4 无障碍沟通

元宇宙中的化身来自不同的国家，在现实世界中使用不同的语言和文字进行交流。但是，元宇宙需要全球“通用”的语言进行人与人之间的实时交流和沟通。AI 技术能够编辑和生成图片以响应文本输入，让不同文化、地域背景的人能够更好地理解文本内容；更重要的是，AI 技术可以实时翻译全球主要语言的声音和文字信息^[34]。例如，两个化身（分别来自中国和美国）在元宇宙中进行交流，AI 技术可以在极短的时间内将中英文互转，使两个化身之间的对话没有任何语言障碍。

4 应用趋势展望

未来，在元宇宙中，虚拟世界和现实世界的界限将越来越模糊。在这个过程中，将一系列技术点“连点成线”并进行产业聚合，是打破虚拟和现实的界限，促进虚拟和现实融合的重要力量。现在，随着算力持续提升，AI 技术、人机交互技术、通信技术、机器人技术等创新逐渐融合，元宇宙也一步一步走向成熟。在元宇宙的世界里，

人类将不再受到现实世界的限制。人与人的交互也将不再停留在文字、音/视频的社交，实时互动、交错时空的互动都可以实现，这必然会诞生新的生活方式。但在奔向“元宇宙”前，必须打造出实现虚实结合的基础设施。显然，AI就是那个制造虚拟世界的“脚手架”^[35]。

4.1 人工智能将成为内容创作的主要途径

AI可以大幅提高内容创作效率，元宇宙中需要文字、图片、视频等内容，这些内容如果全部靠人力去创作，工作量极大。因此，AI将成为内容创作的主要途径。一方面，利用AI自动化生成内容或进行内容增强。例如，从简单的随机物体摆放到全自动生成场景、建筑、物品、外形等，或通过AI算法增强内容呈现质量，扩展人们的视觉边界。另一方面，AI辅助内容创作。例如，基于GAN模型的设计能力，在图片创作中可以将只有一些轮廓和颜色的草图，转化为分辨率高的图片。用户只需要画出一些线条、颜色，然后指定特定部分为内容元素（例如“草地”“河流”“天空”等），AI就可以自动“脑补”填充细节。此外，还有AI配音、2D快转3D或由AI代替人类测试员对元宇宙世界进行安全漏洞检测并生成测试报告等，其中部分内容生成能力已经实现。

4.2 人工智能将进一步链接现实世界和虚拟世界

随着数字化转型的深入，现实世界的各种要素都可以被“搬到”虚拟世界中，同时虚拟世界创设的内容又可以通过不同的载体投射到现实世界中，进而对现实世界产生影响。在这个过程中，AI就是现实世界和虚拟世界的“纽带”。例如，在电力网络中构建基于平行系统与元宇宙思想的平行电网系统，形成虚实融合的智能大电网管理与控制方法^[36]。

通常来说，建立虚实融合的元宇宙世界需要3个步骤。

第一步：场景的数字化、像素化和3D化。

第二步：要素的结构化。结合智能感知和分析，将第一步数字化得到的大量非结构化数据理解抽取为对人类有意义的元要素，即转化为结构化数据。在现实世界的各种场景中，有80%以上的结构化应用是低频、长尾的，因此需要AI技术推动要素的结构化。

第三步：流程的可交互化。只有基于可交互的流程，才能去做业务流程的重塑和自动化。基于上

一步得到的结构化数据，进一步实现应用端的决策智能。

4.3 人工智能将提升元宇宙的“生活”体验

AI技术为元宇宙中的“数字原住民”植入“AI大脑”，使他们拥有独立的沟通和决策能力，可以通过判断人类意图和需求，适时、准确地给予回应。“数字原住民”可以通过不断学习和自我强化，升级成为数字世界的AI代理人，为人们提供各类咨询和服务。未来，这些AI代理人或将成为人们在元宇宙中获取服务和信息的AI助手。除此之外，元宇宙还是一个理想的AI模型训练环境，可以为AI这种“数字原住民”提供天然的训练场。原因有二：一是元宇宙中有大量的标注过的数据，数据质量高且带指向性，有利于互联网巨头开展AI深度学习与机器学习；二是元宇宙拥有无边无际的数字环境和海量用户，可随时为AI提供反馈，指导AI助手的进一步升级。

4.4 人工智能将全方位、多领域地在元宇宙中发挥重要作用

(1) 商业领域

想象一下，未来的电子商务将不再是一个静态网页，而是一个虚拟世界购物中心，你的化身沿着街道穿过各种店铺，在那里你会受到一位销售助理的欢迎，她会引导你到体育用品商店。然后，你使用加密货币购买了一双运动鞋。你的化身立即通过AI技术试穿了这双鞋，很快，你将在现实世界中收到一双一模一样的鞋。元宇宙为销量的增加和用户体验的提升提供了巨大的机会。

(2) 工作领域

AI技术将在绝大部分脑力劳动中全面代替人类工作。司机、文员、导购、导游、个人助理等重复性高的职业都会被AI代替。在元宇宙中，以数字创作工作为代表的元宇宙构建工作也是AI擅长的。但是，AI代替人类进行逻辑推理、艺术设计，甚至创造发明仍需时日。

(3) 社会治理

作为与现实世界对应的虚拟世界，元宇宙将是一个人与人交互的全新场所。在这里，来自五湖四海的人们将汇聚在一起，彼此交流、彼此协作，形成各种新的组织。和现实中一样，一旦有了组织，各种与组织相关的治理问题就会随之出现。AI技术将成为社会治理的主要技术。以游戏《我的世界》为例，开发者将需要所有用户都共同遵守的一些规

范直接写入程序；与此同时，开放权限，允许各服务器根据自己的实际需要，通过插件（plugin）的方式对规则进行补充。但是为了防止插件作弊，需要 AI 技术进行实时侦测和检查。

5 结束语

近年来，AI 技术取得了前所未有的进步。随着元宇宙的兴起和快速发展，其对人机交互、网络通信、机器人等关键技术的性能要求逐渐提升。在构建元宇宙的过程中，AI 技术不仅极大地推动了元宇宙关键技术的发展，而且能够在元宇宙中直接进行内容创作，链接起现实世界和虚拟世界。本文通过对元宇宙和 AI 的概念和代表技术的梳理，总结了 AI 技术对构建元宇宙关键技术（人机交互、网络通信、机器人技术）的推动和提升，并且详细介绍了 AI 在元宇宙中的数字孪生、数字虚拟人、个性化数字化身、无障碍沟通的应用。最后，从内容创作、链接“虚实”和提升元宇宙体验 3 个方面，对 AI 技术在元宇宙中的发展和应用趋势进行了展望。

参考文献：

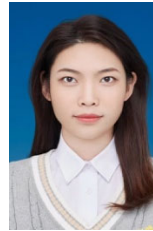
- [1] 胡泳, 刘纯懿. “元宇宙社会”: 话语之外的内在潜能与变革影响[J]. 南京社会科学, 2022(1): 106-116.
HU Y, LIU C Y. “Metaverse society”: transformative potential beyond discourse[J]. Nanjing Journal of Social Sciences, 2022(1): 106-116.
- [2] 陈永伟, 吕琳媛. 元宇宙漫游指南[M]. 上海: 上海人民出版社, 2022.
CHEN Y W, LYU L Y. The hitchhiker’s guide to the metaverse[M]. Shanghai: Shanghai People’s Publishing House, 2022.
- [3] 袁勇, 王飞跃. 区块链技术发展现状与展望[J]. 自动化学报, 2016, 42(4): 481-494.
YUAN Y, WANG F Y. Blockchain: the state of the art and future trends[J]. Acta Automatica Sinica, 2016, 42(4): 481-494.
- [4] GOODFELLOW I J, BENGIO Y, COURVILLE A. Deep learning[M]. [S.l.]: MIT Press, 2016.
- [5] ZHANG S, YAO L N, SUN A X, et al. Deep learning based recommender system[J]. ACM Computing Surveys, 2020, 52(1): 1-38.
- [6] 袁园, 杨永忠. 走向元宇宙: 一种新型数字经济的机理与逻辑[J]. 深圳大学学报(人文社会科学版), 2022, 39(1): 84-94.
YUAN Y, YANG Y Z. Embracing the metaverse: mechanism and logic of a new digital economy[J]. Journal of Shenzhen University (Humanities & Social Sciences), 2022, 39(1): 84-94.
- [7] 尼尔斯蒂芬森. 雪崩[M]. 郭泽, 译. 成都: 四川科学技术出版社, 2009.
STEPHENSON N. Snow crash[M]. Translated by Guo Z. Chengdu: Sichuan Science and Technology Press, 2009.
- [8] SILVER D, HUANG A, MADDISON C J, et al. Mastering the game of Go with deep neural networks and tree search[J]. Nature, 2016, 529(7587): 484-489.
- [9] SILVER D, SCHRITTWIESER J, SIMONYAN K, et al. Mastering the game of Go without human knowledge[J]. Nature, 2017, 550(7676): 354-359.
- [10] LIANG H Y, TSUI B Y, NI H, et al. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence[J]. Nature Medicine, 2019, 25(3): 433-438.
- [11] ZHAO S, WEI W, ZOU D, et al. Multi-view intent disentangle graph networks for bundle recommendation[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(4): 4379-4387.
- [12] YURTSEVER E, LAMBERT J, CARBALLO A, et al. A survey of autonomous driving: common practices and emerging technologies[J]. IEEE Access, 2020, 8: 58443-58469.
- [13] LI B, CHANG S Y, SAINATH T N, et al. Towards fast and accurate streaming end-to-end ASR[C]//Proceedings of 2020 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2020: 6069-6073.
- [14] DEVLIN J, CHANG M, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[J]. arXiv preprint, 2018, arXiv:1810.04805.
- [15] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 770-778.
- [16] BROWN T B, MANN B, RYDER N, et al. Language models are few-shot learners[J]. arXiv preprint, 2020, arXiv:2005.14165.
- [17] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of the Conference and Workshop on Neural Information Processing Systems. New York: Curran Associates, 2014: 2672-2680.
- [18] BROWN N, SANDHOLM T. Superhuman AI for multiplayer poker[J]. Science, 2019, 365(6456): 885-890.
- [19] 王飞跃, 王艳芬, 陈慧竹, 等. 联邦生态: 从联邦数据到联邦智能[J]. 智能科学与技术学报, 2020, 2(4): 305-311.
WANG F Y, WANG Y F, CHEN Y Z, et al. Federated ecology: from federated data to federated intelligence[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(4): 305-311.
- [20] CARD S K, MORAN T P, NEWELL A. The psychology of human-computer interaction[M]. [S.l.]: Lawrence Erlbaum Associates, 1983.
- [21] CARBONELL J R. AI in CAI: an artificial-intelligence approach to computer-assisted instruction[J]. IEEE Transactions on Man-Machine Systems, 1970, 11(4): 190-202.
- [22] GANAPATHIRAJU A. Conversational artificial intelligence in production - challenges and advances[C]//Proceedings of 2021 IEEE Signal Processing in Medicine and Biology Symposium. Piscataway: IEEE Press, 2021: 1.
- [23] 郭小爽. 人机交互中的动态手势识别及应用研究[D]. 西安: 西安电子科技大学, 2014.
GUO X S. Research on dynamic gesture recognition and its application in human-computer interaction[D]. Xi’an: Xidian University, 2014.
- [24] YI X, YU C, XU W J, et al. COMPASS: rotational keyboard on non-touch smartwatches[C]//Proceedings of 2017 CHI Conference on Human Factors in Computing Systems. New York: ACM Press, 2017: 705-715.
- [25] YU C, SUN K, ZHONG M Y, et al. One-dimensional handwriting:

- inputting letters and words on smart glasses[C]//Proceedings of 2016 CHI Conference on Human Factors in Computing Systems. [S.l.:s.n.], 2016: 71-82.
- [26] 尤肖虎, 张川, 谈晓思, 等. 基于 AI 的 5G 技术: 研究方向与范例[J]. 中国科学: 信息科学, 2018, 48(12): 1589-1602.
- YOU X H, ZHANG C, TAN X S, et al. AI for 5G: research directions and paradigms[J]. Scientia Sinica (Informationis), 2018, 48(12): 1589-1602.
- [27] 张彤, 任奕璟, 闫实, 等. 人工智能驱动的 6G 网络: 智慧内生[J]. 电信科学, 2020, 36(9): 14-22.
- ZHANG T, REN Y J, YAN S, et al. Artificial intelligence driven 6G networks: endogenous intelligence[J]. Telecommunications Science, 2020, 36(9): 14-22.
- [28] 龙慧, 朱定局, 田娟. 深度学习在智能机器人中的应用研究综述[J]. 计算机科学, 2018, 45(S2): 43-47, 52.
- LONG H, ZHU D J, TIAN J. Research on deep learning used in intelligent robots[J]. Computer Science, 2018, 45(S2): 43-47, 52.
- [29] 喻祥尤. 基于深度学习的机器人场景识别研究[D]. 沈阳: 沈阳工业大学, 2017.
- YU X Y. Research on robot scene recognition based on depth learning[D]. Shenyang: Shenyang University of Technology, 2017.
- [30] 张小俊, 刘欢欢, 赵少魁, 等. 机器人智能化研究的关键技术与发展展望[J]. 机械设计, 2016, 33(8): 1-7.
- ZHANG X J, LIU H H, ZHAO S K, et al. Key technologies and development prospects of the intelligent robot study[J]. Journal of Machine Design, 2016, 33(8): 1-7.
- [31] 于华琛. 基于深度学习的挖掘机器人图像识别及铲斗目标跟踪研究[D]. 南京: 南京工业大学, 2016.
- YU H C. Research on image recognition and bucket target tracking of mining robot based on deep learning[D]. Nanjing: Nanjing University of Technology, 2016.
- [32] SINGH M, SRIVASTAVA R, FUENMAYOR E, et al. Applications of digital twin across industries: a review[J]. Applied Sciences, 2022, 12(11): 5727.
- [33] SUTTON R S, BARTO A G. Reinforcement learning: an introduction[J]. IEEE Transactions on Neural Networks, 1998, 9(5): 1054.
- [34] DABRE R, CHU C H, KUNCHUKUTTAN A. A survey of multilingual neural machine translation[J]. ACM Computing Surveys, 2020, 53(5): 1-38.
- [35] LEE L H, BRAUD T, ZHOU P Y, et al. All one needs to know about metaverse: a complete survey on technological singularity, virtual ecosystem, and research agenda[J]. arXiv preprint, 2021, arXiv:2110.05352.
- [36] 李小双, 王晓, 杨林瑶, 等. 元电网 MetaGrid: 基于平行电网的新一代智能电网的体系与架构[J]. 智能科学与技术学, 2021, 3(4): 387-398.
- LI X S, WANG X, YANG L Y, et al. MetaGrid: a parallel grids based approach for next generation smart power systems[J]. Chinese Journal of Intelligent Science and Technology, 2021, 3(4): 387-398.

[作者简介]



武强 (1985-), 男, 博士, 电子科技大学基础与前沿研究院复杂系统与社会计算研究中心博士后, 主要研究方向为复杂系统、复杂网络和人工智能。



季雪庭 (1998-), 女, 杭州师范大学阿里巴巴商学院硕士生, 主要研究方向为技术创新管理。



吕琳媛 (1984-), 女, 博士, 电子科技大学基础与前沿研究院复杂系统与社会计算研究中心教授, 主要研究方向为统计物理、复杂网络和网络信息挖掘。