

# 面向沉浸式元宇宙的显示、交互和应用

何泽浩, 曹良才\*

清华大学精密仪器系, 精密测试技术及仪器国家重点实验室, 北京 100084

**摘要** 概述了元宇宙涵盖的基本领域, 调研了当前面向元宇宙的沉浸式交互产业布局, 阐述了沉浸式体验对元宇宙的意义; 介绍了头戴显示、平视显示和大尺寸显示的技术原理, 展望了各类显示技术在元宇宙沉浸式体验中的前景; 分析了交互技术在元宇宙沉浸式体验中的重要性, 归纳了计算机图形学、人工智能和机器视觉等技术在沉浸式交互中的价值; 基于沉浸式体验特性, 对元宇宙技术的潜在应用领域进行了预测。

**关键词** 元宇宙; 沉浸式显示; 三维显示; 沉浸式交互

## 1 元宇宙和人机界面

1992年, 元宇宙 (metaverse) 一词首次出现在了科幻小说《雪崩》中<sup>[1]</sup>, 是元 (meta) 和宇宙 (universe) 的合成词, 指代的是一个带上了特殊眼镜就能进入的虚拟空间。其中, “元”来自于哲学中的形而上学 (metaphysics)。形而上学研究的是现象的原因和本源, 被亚里士多德称为“第一哲学”。结合了“本源”和“第一”等意象, “meta”便被赋予了“元”的含义。此外, 考虑到形而上学和物理学 (physics) 之间唯心和唯物的对应关系, “meta”又常

被赋予“虚拟”的含义, 故而元宇宙最初几乎等同于“虚拟世界”。

近年来, 伴随着行业的快速进步, 元宇宙的含义也得到了极大的扩展。从技术上讲, 元宇宙被认为是一个虚实交错的共享空间, 由数字增强的物理世界和数字构建的虚拟世界共同组成<sup>[2]</sup>。如图 1 所示, 元宇宙涉及的领域主要包括底层技术、环境感知、场景构建、人机界面、秩序管理和应用推广等方面, 其中人机界面是连接用户、真实世界和虚拟世界的重要节点, 涉及显示和交互等关键领域<sup>[3-5]</sup>。沉浸式体验是元宇宙对人机界面的终极要求, 具体

收稿日期: 2022-12-05; 修回日期: 2023-02-10

基金项目: 国家自然科学基金青年科学基金项目 (62205173); 国家自然科学基金重点基金项目 (62035003)

作者简介: 何泽浩, 博士后, 研究方向为全息三维显示和通信, 电子信箱: hezehao@mail.tsinghua.edu.cn; 曹良才 (通信作者), 教授, 研究方向为全息成像与全息显示, 电子信箱: clc@tsinghua.edu.cn

引用格式: 何泽浩, 曹良才. 面向沉浸式元宇宙的显示、交互和应用[J]. 科技导报, 2023, 41(5): 6-14; doi:10.3981/j.issn.1000-7857.2023.05.001

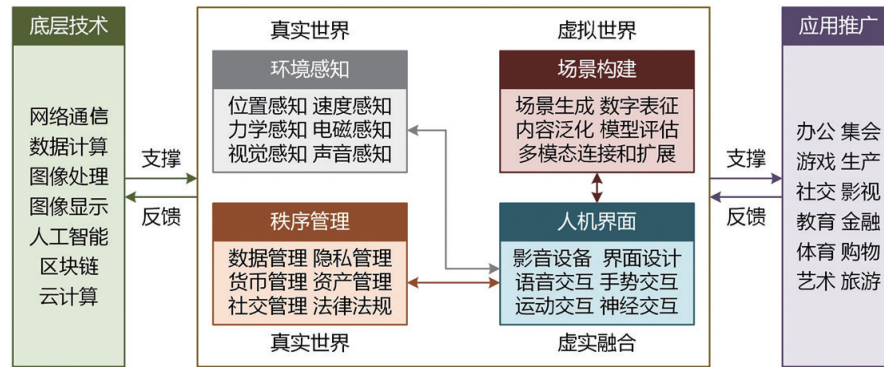


图1 元宇宙涉及的技术和人文领域

体现在显示三维(3D)化和交互真实化2个方面<sup>[6-9]</sup>。

当前,众多科技企业正致力于提升人机界面的沉浸感。Meta(原Facebook)公司先后收购虚拟现实(VR)设备厂商Oculus、智能眼镜厂商Ray-Ban Stories、应用开发公司Within和神经接口公司CTRL-Labs,掌握了相关显示和交互技术,是人机界面沉浸化的主要推动者;微软公司主要依托混合现实设备Hololens开展研究,在沉浸式人机界面领域开展了积极的布局;谷歌的母公司Alphabet不仅持续改进头戴显示(HMD)设备,还推出了Starline这样的大尺寸沉浸式会议系统<sup>[10]</sup>;字节跳动公司收购了VR设备厂商Pico,对HMD设备的结构进行了重新设计,显著提升了人机界面的视觉沉浸感。此外,Roblox这样的元宇宙主要参与者虽然不直接开发硬件设备,但也优化了HMD设备摄像头的控制方式,增加了第一和第三视角间的切换选项,显著

提升了人机界面的视觉沉浸感。

## 2 面向沉浸式元宇宙的显示技术

相比二维内容,用户通过头戴设备或裸眼设备观察3D内容时,视觉感受更接近观察真实世界,可以更直观地了解图像内容。3D显示是沉浸式元宇宙的必然要求。从科技企业的技术布局来看,当前人机界面的呈现主要依靠HMD设备,一般只能用于近眼显示,限制了元宇宙的使用场景。Alphabet的Starline系统实现了大尺寸沉浸式会议展示,为人机界面的未来发展提供了新的思路。根据观看距离和显示尺寸,未来人机界面的显示设备预计可分为3类,分别是HMD设备、平视显示(HUD)设备和大尺寸显示设备,如图2所示。

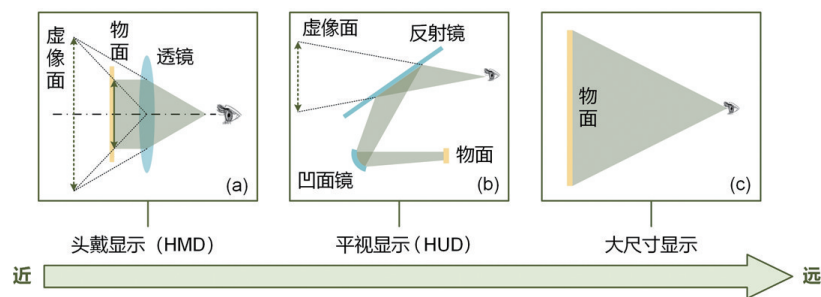


图2 元宇宙人机界面的显示设备

### 2.1 HMD设备

HMD设备既是元宇宙的最初呈现载体,也是当前最常见的呈现载体,其概念在小说《雪崩》中就

已出现。如图2(a)所示,HMD设备的光路中,物面与透镜的间距通常小于透镜的焦距,因此目标图像经透镜后成为正立、放大的虚像<sup>[11-12]</sup>。

对于HMD设备而言,产品重量是需要着重衡量的因素。考虑到用户的承受能力,当3D显示设备的重量超过特定数值时,就无法以HMD的形态存在。由于用户的重量承受能力具有显著的个体差异,业界针对HMD设备没有公认的重量阈值。作为参考,几款畅销产品,包括Oculus Quest 3、微软Hololens 2和字节跳动Pico 4等,其主机重量均在550 g左右。市售主流产品中,相对较重的HTC Vive Focus 3也未曾超过800 g。

当前,3D显示可以通过双目视觉、光场、体3D和全息等技术实现,各类技术的原理如图3所示。双目视觉技术中,显示设备向人眼投射的是两幅稍有差别的二维图像。在大脑的图像融合机制下,用户最终可以获得3D的视觉感知<sup>[13]</sup>。光场技术中,显示设备向空间中投射3D场景不同视点下的图像

信息。当人眼捕捉到2个以上的视点图像时,可以基于大脑融合机制获得3D感知<sup>[14-16]</sup>。体3D技术中,基于机械运动装置、多层显示屏或空气电离装置可以实现离散体像素点的点亮。显示设备针对显示3D图像的强度分布,按需点亮空间相应位置的体像素点,进而给用户以3D感知<sup>[17]</sup>。全息技术中,待显示3D图像首先需要被编码为衍射条纹分布。衍射条纹中的振幅和相位可以在相干照明光的照射下实现重建,进而提供用户以3D感知<sup>[18-20]</sup>。HMD设备可基于双目视觉、光场和全息等技术实现3D显示。考虑到体3D技术通常需要基于机械运动产生体像素,机械运动产生的振动对用户头部的舒适性影响极大,因此HMD设备通常不会基于体3D技术进行设计。

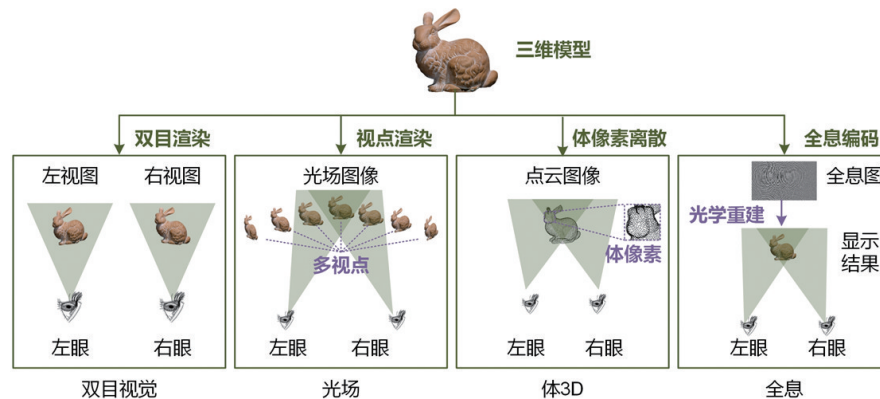


图3 不同的3D显示实现方式

## 2.2 HUD设备

HUD设备是元宇宙的又一呈现载体,填补了HMD设备和大尺寸设备之间的空白<sup>[21-23]</sup>。HUD设备在沉浸式元宇宙领域主要针对具有下列特征的场景:第一,佩戴HMD设备难以保证使用的安全性和功能性。第二,空间大小不足以安装可以直接观看的大尺寸3D显示设备。这两种特征在模拟训练、远程协作和汽车驾驶等应用场景中体现得尤为明显。因此,在新能源汽车行业智能座舱概念火热发展的当下,HUD在最近数年得到了广泛的关注。

如图2(b)所示,HUD设备的典型光路结构中,待显示图像一般通过投影光机出射,经过光学系统

放大,最后经由反射镜到达人眼。用户观看到的图像通常也是正立、放大的虚像。由于汽车和模拟训练机对装载设备的空间利用率要求较高,HUD设备通常使用凹面镜、自由曲面反射镜等元件构建光学放大系统,以实现转折、压缩光路的作用。

HUD设备的3D显示通常可基于双目视觉、光场和全息等技术实现。双目视觉方案中,光机投影系统和用户佩戴的眼镜需要物理分离,以提升使用的舒适性、安全性和功能性。其中,光机投影系统可以基于时分复用原理或偏振原理设计,用户佩戴远比HMD设备轻便的快门式或偏振式眼镜,即可获得3D视觉感知。不过,相比双目视觉方案,光场

和全息技术支撑的裸眼 3D 显示方案在 HUD 领域显然更具吸引力。体 3D 技术在 HUD 设备上存在应用的可能性,但考虑到这类设备体积较其他方案更加庞大,其应用前景预计较为受限。

### 2.3 大尺寸显示设备

当物理空间足够大时,直接观看的大尺寸显示设备能够提供理想的元宇宙体验,其典型光路结构如图 2(c)所示。这类设备可以基于双目视觉、光场、体 3D 和全息等技术设计,其中双目视觉技术仍然是当前的主流。由于对佩戴舒适性和空间体积的限制较少,大尺寸显示设备可以利用体 3D 技术实现 3D 显示,但需要重点解决体像素密度较低导致显示精细度不足的问题。在大尺寸显示领域,光场技术被认为是当下颇具商用价值的沉浸式元宇宙解决方案,在 3D 显示质量和可视角度等关键指标上具有良好表现,但需要重点改善多视点条件下显示分辨率下降的难题。全息 3D 显示是沉浸式元宇宙的终极解决方案,但当前受到器件像素尺寸精细度不足的限制,可视角度仍有待拓展。

特别指出的是,除了上述 3 种设备形态以外,

直视型中小尺寸显示设备(如智能手表、手机和显示器等)也能够用于元宇宙的内容呈现,并且可以通过双目视觉、光场和全息等技术实现 3D 显示。然而,这类设备的可显示尺寸较小,既难以完全隔绝真实世界,又难以与真实世界充分融合。因此,从沉浸感的角度来看,中小尺寸显示设备的性能较为有限,但也可以作为沉浸式元宇宙在特定应用场景下的补充呈现形式。

## 3 面向沉浸式元宇宙的交互技术

广义的元宇宙交互包含视觉、听觉、触觉和运动等方面的信息传递与反馈。本文主要关注与显示内容相关的交互,并归纳相关的底层支撑技术。

### 3.1 元宇宙的交互方式

用户与显示内容的交互,指的是元宇宙设备通过采集、处理不同的用户指令,进而实现显示画面渲染和呈现的过程,主要包括眼动、手动、运动和语音等交互方式(图 4)。

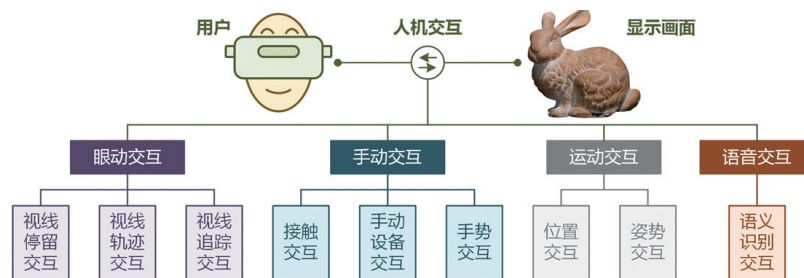


图 4 面向沉浸式元宇宙的交互技术

1) 眼动交互。眼动交互指的是人机界面里的显示图像随着眼睛注视位置和注视时间而变化的交互方式,主要包括视线停留、视线轨迹和视线追踪等<sup>[24-27]</sup>。类比当前手机触摸屏上的交互操作,视线停留类似手指点按,指的是眼球追踪装置检测到用户视线在特定位置停留超过一定时间后触发显示画面内容变化的交互机制;视线轨迹类似触摸轨迹识别,指的是眼球追踪装置识别到用户眼睛的扫视轨迹与特定图案近似时触发显示画面内容变化的交互机制;视线追踪类似于手指滑动,指的是显

示画面随着眼睛注视位置实时改变的交互机制。其中,视线停留和视线轨迹是准静态交互方式,其触发因素是一定积分时间内的特征信息,实现难度较低;视线追踪是动态交互方式,其触发因素是实时改变的特征信息,对响应时间和计算速度要求较高,当前实现难度偏高。

2) 手动交互。手动交互指的是人机界面里的显示图像随着双手位置、动作和姿势而变化的交互方式,主要包括接触式交互、手动设备交互和手势交互等<sup>[28-31]</sup>。接触式交互指的是直接在显示设备

上进行点按、滑动的交互机制;手动设备交互指的是通过在手柄、手环等辅助配件上进行移动和点按等操作触发显示画面内容变化的交互机制;手势交互指的是通过机器视觉配件识别手部动作和姿势、进而触发画面变化的交互机制。其中,接触式交互多使用在 HUD 设备和大尺寸设备中,并且要求可交互的画面内容必须位于物理屏幕上,适用范围较为受限;手动设备交互需要握持专用的交互配件,一定程度上影响了交互的便捷性,但可以弥补当前手势识别处理时间长、响应精度低的不足;手势交互是一种裸手交互方式,具有适用范围广和应用便捷性强的特点,但对机器视觉配件的响应速度和识别精度提出了较高的要求。

3) 运动交互。运动交互指的是人机界面里的显示图像随着身体位置、动作和姿势而变化的交互方式,主要包括位置交互和姿势交互<sup>[32-35]</sup>。位置交互指的是借助加速度传感器、方向传感器、陀螺仪传感器、重力传感器和气压传感器等元件判断用户的所处位置和运动状态、进而改变显示画面的交互机制;姿势交互指的是借助动捕装置和图像识别装置判断用户的身体姿势、进而改变显示画面的交互机制。相比于眼动交互和手动交互,运动交互在交互形式上更加自由,能够基于空间坐标、运动方向、行走速度和真实空间场景创造不同的交互方案。

4) 语音交互。语音交互指的是人机界面里的显示图像随着声音信号而变化的交互方式<sup>[36]</sup>。相比于眼动、手动和运动交互,语音交互处理的信息更为抽象。一般来说,用户的语音指令首先会被识别为文字信息,进而基于自然语言处理技术被转换为可以由计算机识别的语义信息。最后,基于语义信息进行图像渲染和显示,人机界面才能呈现出符合用户要求的画面。当前,语音交互的技术门槛主要体现在语义识别上。让设备充分理解用户的语音命令,是提升人机交互体验的重中之重。

### 3.2 元宇宙交互的支撑技术

元宇宙的沉浸式交互依赖于超快速的指令识别和高质量的图像生成与显示,需要计算机图形学、3D 成像、机器视觉、动作捕捉和人工智能等多种技术的支撑(图 5)。

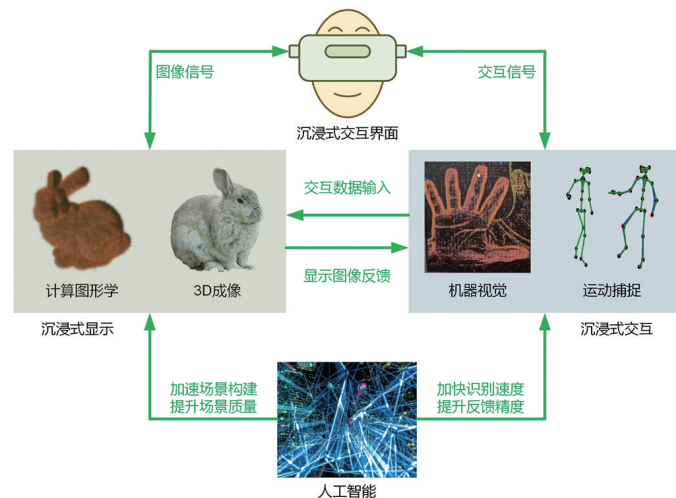


图 5 面向元宇宙交互的支撑技术

1) 计算图形学。沉浸式元宇宙的一大核心任务是将高质量的 3D 场景呈现给用户,而基于计算机图形学的高精度 3D 场景构建则是高质量 3D 呈现的重要基石。计算机图形学主要包含了建模、渲染和动画等方面<sup>[37-38]</sup>。其中,3D 建模指的是建立 3D 场景的几何模型表达,它既可以通过非均匀有理 B 样条法和细分曲面造型法等数学方法实现,也可以通过 Autodesk 3ds MAX 和 Dassault SolidWorks 等商业化软件完成;3D 渲染指的是在几何模型的基础上加入照明、光泽和材质等信息,提升 3D 场景的视觉真实感;3D 动画首先通过多幅静态 3D 模型的连续播放实现 3D 场景的动态化,进而通过高真实度的动态模拟技术提升动态场景的显示沉浸感。

2) 3D 成像。虚拟场景的构建主要依赖计算机图形学,而真实场景的数字化则主要依赖 3D 成像技术。用于 3D 成像的常用设备包括光场相机、激光扫描仪、飞行时间(TOF)相机和结构光相机等<sup>[39-42]</sup>。光场相机通过内部放置的透镜阵列采集 3D 场景各个视角下的二维图像,进而基于 3D 融合算法实现 3D 场景模型的建立;激光扫描仪通过成像设备的 360 度旋转捕捉 3D 场景的表面点云,进而通过点云绘制算法实现 3D 场景模型的建立;TOF 相机将光脉冲投射至 3D 目标,通过探测光脉冲飞行的往返时间确定 3D 目标的表面形貌;结构光相机将条纹图像按照时间序列依次投射在 3D 目标表面,通过采集条纹图像并进行条纹解码,进而实现高精度 3D 点云的重建。上述设备中,光场相

机是被动式照明成像设备,自身无需集成照明光源;激光扫描仪、TOF相机和结构光相机是主动式照明成像设备,设备内部通常需要安装有符合要求的照明光源。伴随着计算图形学和3D成像的不断发展,这2种技术在近年来呈现出了显著的协同效应,并催生和推进了数字人技术的快速发展。

3) 机器视觉。3D成像关注的主要是真实场景的数字化建模。但在部分沉浸式元宇宙应用中,人机交互设备不仅需要忠实地采集真实场景,还需要从场景图像中提取部分信息进行处理并加以理解,进而用于实际反馈和控制<sup>[43-45]</sup>。眼动交互、手势交互和姿势交互等技术都离不开机器视觉的支撑。典型的机器视觉系统通常包括六大模块,分别是光源投射模块、图像采集模块、图像数字化模块、数字处理模块、判断决策模块和信号反馈模块。机器视觉系统对信息进行判断决策后反馈给人机交互界面,人机交互界面中的显示画面将相应发生变化。

4) 运动捕捉。机器视觉技术可以基于图像信息实现目标对象的动作捕捉。但是,在很多沉浸式元宇宙应用中,用户的动作信息不完全依赖机器视觉进行采集。其他运动捕捉技术包括机械式捕捉、声学式捕捉、电磁式捕捉和惯导式捕捉等<sup>[46-48]</sup>。机械式捕捉技术将目标对象和机械结构相连,通过机械结构的运动状态推演目标对象的运动姿态;声学式捕捉技术使用多个超声探头捕捉超声波发射器发出的信号,通过计算声波的相位差确定探头的位置和方向,进而求解目标对象的运动姿态;电磁式捕捉技术使用电磁波发射器产生电磁场、使用多个

粘贴在目标对象关键位置的电磁波接收器探测电磁信号,通过电磁信号的反演实现目标对象运动姿态的重构;惯导式捕捉技术通过多种惯导传感器确定目标对象的运动加速度、方位和倾斜角等信息,进而基于数字信号处理实现目标对象运动状态的重构。运动捕捉在手动交互和运动交互等方面发挥着重要作用。

5) 人工智能。理想的人机交互设备应该能够及时响应用户指令并实现3D显示画面的高质量快速切换,这对建模、成像、机器视觉、运动捕捉和语义理解等技术的精细度和时效性都提出了极高的要求。人工智能技术为这些问题的解决提供了革命性的新思路<sup>[49-54]</sup>。例如,传统的3D成像技术受限于设备和环境因素,常常存在色彩失真、亮度不均和深度错乱的问题,且修复过程需要消耗大量算力,修复速度难以满足沉浸式交互的需求。基于人工智能的图像增强技术可以自动筛选并修复异常影像,显著提升了3D成像的采集速度和图像质量,预计可以为人机交互设备提供满足沉浸式交互需求的3D实时影像数据。

## 4 沉浸式元宇宙的应用

如图6所示,根据元宇宙显示画面的呈现方式,元宇宙应用可被分为两大类,即虚拟构建式应用和物理增强式应用。其中,虚拟构建式应用基于VR设备呈现,物理增强式应用基于增强现实(AR)、混合现实(MR)和扩展现实(XR)等设备呈

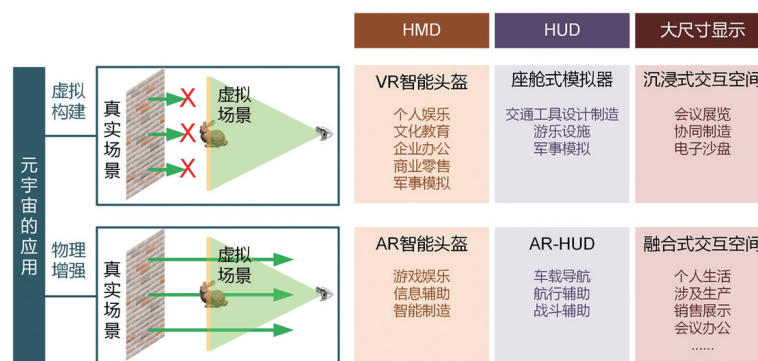


图6 沉浸式元宇宙的应用

现<sup>[55]</sup>。虚拟构建式应用较为封闭,用户除了目标图像外,通常无法看见真实世界中的景物,因此具有强烈的沉浸感;物理增强式应用允许用户同时观察目标图像和真实世界,但只有当二者充分融合时才能呈现理想的沉浸感。因此,物理增强式应用更为依赖环境感知器件,并且对机器视觉、图像处理等底层技术要求更高。相同成本下,物理增强式应用的视觉体验相比虚拟构建式应用往往存在较大差距。

当HMD、HUD和大尺寸显示各自与虚拟构建式的元宇宙相结合时,产生的典型应用形态分别是VR智能头盔、座舱式模拟器和沉浸式交互空间。其中,VR智能头盔是当前最为成熟的应用形态,Meta和字节跳动等大型科技企业已在相关领域进行了广泛的布局,可应用于个人娱乐、文化教育、企业办公、商业零售和军事模拟等众多领域;座舱式模拟器通过封闭的显示环境和定制的运动模拟装置,能够仿真汽车、飞机和轮船等交通工具的行驶状态,可应用于交通工具设计制造、游乐设施和军事模拟等领域;沉浸式交互空间通过在较大物理范围内构建封闭式的交互环境,能够同时为众多用户提供沉浸式的元宇宙交互,可应用于会议展览、协同制造和电子沙盘等领域。

相似地,当HMD、HUD和大尺寸显示各自与物理增强式的元宇宙相结合时,产生的典型应用形态分别是AR智能头盔、AR-HUD和融合式交互空间。其中,AR智能头盔当前得到了长足的发展,并得到了微软公司的重点关注,可用于游戏娱乐、信息辅助和智能制造等领域;AR-HUD是当前汽车行业的关注焦点,被认为是车载导航的未来形态,同时也可应用于航行辅助和战斗辅助等领域;融合式交互空间通过在较大物理范围内构建真实世界和虚拟世界相互交融的交互环境,有望在未来彻底突破真实与虚拟的边界,实现人与人、人与物的自由连接,应用于人类工作和生活的方方面面。

## 5 结论

简述了元宇宙概念的来源和发展,介绍了元宇宙涉及的主要技术领域,阐述了沉浸感对元宇宙的

意义,并初步调研了国内外科技企业在沉浸式元宇宙领域的技术布局。强调了3D显示技术对沉浸式元宇宙的重要作用,并以用户的观看距离为顺序,详述了HMD、HUD和大尺寸显示等技术方案的原理,以及不同方案中用到的3D显示技术,包括双目视觉显示、光场显示、体3D显示和全息显示。结合沉浸式元宇宙对人机界面高质量显示和超高速响应的要求,介绍了几类沉浸式元宇宙必备的底层交互技术,包括计算图形学、3D成像、机器视觉、运动捕捉和语音交互。最后,归纳了沉浸式元宇宙在虚拟构建式应用和物理增强式应用中的具体形态。

## 参考文献(References)

- [1] Stephenson N. Snow crash[M]. New York: Bantam Spectra, 1992.
- [2] Mystakidis S. Metaverse[J]. Encyclopedia, 2022, 2(1): 486-497.
- [3] Zhao Y H, Jiang J J, Chen Y, et al. Metaverse: Perspectives from graphics, interactions and visualization[J]. Visual Informatics, 2022, 6(1): 56-67.
- [4] Park S M, Kim Y G. A metaverse: Taxonomy, components, applications, and open challenges[J]. IEEE Access, 2022, 10: 4209-4251.
- [5] 王文喜,周芳,万月亮,等. 元宇宙技术综述[J]. 工程科学学报, 2022, 44(4): 744-756.
- [6] Fan Y C, Lee C M, Lee M Y, et al. 3D display and interactive technology in metaverse[J]. ITE Transactions on Media Technology and Applications, 2022, 10(4): 179-189.
- [7] Kanematsu H, Barry D M, Shirai T, et al. Virtual experiences of metaverse using mobile type head-mounted displays and their applicability[J]. Procedia Computer Science, 2022, 207: 4428-4433.
- [8] 曹良才,何泽浩,刘珂瑄,等. 元宇宙中的动态全息三维显示: 发展与挑战[J]. 红外与激光工程, 2022, 51(1): 267-281.
- [9] 张汉乐,邢妍,胡晓帅,等. 面向元宇宙的集成成像3D显示技术进展[J]. 指挥与控制学报, 2022, 8(3): 239-248.
- [10] Lawrence J, Goldman D, Achar S, et al. Project starline: A high-fidelity telepresence system[J]. ACM Transactions on Graphics, 2021, 40(6): 242.
- [11] Cheng D W, Wang Q W, Liu Y, et al. Design and manu-

- facture AR head-mounted displays: A review and outlook[J]. *Light: Advanced Manufacturing*, 2021, 2(3): 350–369.
- [12] Itoh Y, Langlotz T, Sutton J, et al. Towards indistinguishable augmented reality: A survey on optical see-through head-mounted displays[J]. *ACM Computing Surveys*, 2022, 54(6): 120.
- [13] Tian X, Liu R, Wang Z Y, et al. High quality 3D reconstruction based on fusion of polarization imaging and binocular stereo vision[J]. *Information Fusion*, 2022, 77: 19–28.
- [14] Watanabe H, Omura T, Okaichi N, et al. Full-parallax three-dimensional display based on light field reproduction[J]. *Optical Review*, 2022, 29(4): 366–374.
- [15] Wang Y D, Sang X Z, Xing S J, et al. Three-dimensional light-field display with enhanced horizontal viewing angle by introducing a new lenticular lens array[J]. *Optics Communications*, 2020, 477: 126327.
- [16] Ravishankar J, Sharma M, Gopalakrishnan P. A flexible coding scheme based on block Krylov subspace approximation for light field displays with stacked multiplicative layers[J]. *Sensors*, 2021, 21(13): 4574.
- [17] Smalley D E, Nygaard E, Squire K, et al. A photophoretic-trap volumetric display[J]. *Nature*, 2018, 553: 486–490.
- [18] Hua J Y, Yi D H, Qiao W, et al. Multiview holographic 3D display based on blazed Fresnel DOE[J]. *Optics Communications*, 2020, 472: 125829.
- [19] Blanche P A. Holography, and the future of 3D display[J]. *Light: Advanced Manufacturing*, 2021, 2(4): 446–459.
- [20] He Z Z, Li K X, Sui X M, et al. Holographic 3D display using depth maps generated by 2D-to-3D rendering approach[J]. *Applied Sciences*, 2021, 11(21): 9889.
- [21] Mu C T, Lin W T, Chen C H. Zoomable head-up display with the integration of holographic and geometrical imaging[J]. *Optics Express*, 2020, 28(24): 35716–35723.
- [22] Skirnewskaja J, Wilkinson T D. Automotive holographic head-up displays[J]. *Advanced Materials*, 2022, 34(19): 2110463.
- [23] Lv Z, Liu J, Yang Y. Dual-view and multi-content head-up display using a single picture generation unit and two-layer volume holographic grating[J]. *IEEE Photonics Journal*, 2022, 14(4): 7035208.
- [24] Valtakari N V, Hooge I T C, Viktorsson C, et al. Eye tracking in human interaction: Possibilities and limitations[J]. *Behavior Research Methods*, 2021, 53(4): 1592–1608.
- [25] Drakopoulos P, Koulieris G A, Mania K. Eye tracking interaction on unmodified mobile VR headsets using the selfie camera[J]. *ACM Transactions on Applied Perception*, 2021, 18(3): 11.
- [26] Pai Y S, Dingler T, Kunze K. Assessing hands-free interactions for VR using eye gaze and electromyography[J]. *Virtual Reality*, 2019, 23(2): 119–131.
- [27] 朱富丽, 杨磊, 申玉斌, 等. 基于眼动特征的视觉交互状态分类方法研究[J]. *航天医学与医学工程*, 2021, 34(6): 426–431.
- [28] Pavlovic V I, Sharma R, Huang T S. Visual interpretation of hand gestures for human-computer interaction: A review[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 677–695.
- [29] Li Y, Huang J, Tian F, et al. Gesture interaction in virtual reality[J]. *Virtual Reality & Intelligent Hardware*, 2019, 1(1): 84–112.
- [30] Cheok M J, Omar Z, Jaward M H. A review of hand gesture and sign language recognition techniques[J]. *International Journal of Machine Learning and Cybernetics*, 2019, 10(1): 131–153.
- [31] 林星雨, 邢妍, 张汉乐, 等. 基于 Leap Motion 手势识别的悬浮真 3D 显示实时交互系统[J]. *液晶与显示*, 2022, 37(5): 654–659.
- [32] Yen H Y. Smart wearable devices as a psychological intervention for healthy lifestyle and quality of life: A randomized controlled trial[J]. *Quality of Life Research*, 2021, 30(3): 791–802.
- [33] Yen H Y, Liao Y, Huang H Y. Smart wearable device users' behavior is essential for physical activity improvement[J]. *International Journal of Behavioral Medicine*, 2022, 29(3): 278–285.
- [34] Ringeval M, Wagner G, Denford J, et al. Fitbit-based interventions for healthy lifestyle outcomes: Systematic review and meta-analysis[J]. *Journal of medical Internet research*, 2020, 22(10): e23954.
- [35] 刘任高. 具有复杂人机交互功能的可穿戴产品设计研究[D]. 南京: 南京航空航天大学, 2021.
- [36] Liu J N, Qian H, Xiao A P, et al. Human-machine interaction based on voice[J]. *AASRI Procedia*, 2012, 3: 583–588.
- [37] Wellmann F, Caumon G. 3-D structural geological models: Concepts, methods, and uncertainties[J]. *Advances in Geophysics*, 2018, 59: 1–121.

- [38] Paulsen J, Ali T M L, Collas P. Computational 3D genome modeling using Chrom3D[J]. *Nature*, 2018, 13: 1137–1152.
- [39] Cao Xuan, Geng Z, Li T. Dictionary-based light field acquisition using sparse camera array[J]. *Optics Express*, 2014, 22(20): 24081–24095.
- [40] Serabyn E. Pupil segmentation in the light-field camera and its relation to 3D object positions and the reconstructed depth of field[J]. *Applied Optics*, 2019, 58(5): A273–A282.
- [41] He L, Liu K, He Z, et al. Three-dimensional holographic communication system for the metaverse[J]. *Optics Communications*, 2023, 526: 128894.
- [42] Yu P F, Yang S R, Chen S Y. Accuracy improvement of time-of-flight depth measurement by combination of a high-resolution color camera[J]. *Applied Optics*, 2020, 59(35): 11104–11111.
- [43] Coffey V C. Machine Vision: The eyes of industry 4.0[J]. *Optics & Photonics News*, 2018, 29(7): 42–49.
- [44] Ren Z, Fang F, Li Z, et al. Intelligent evaluation for lens optical performance based on machine vision[J]. *Optics Express*, 2022, 30(15): 26251–26265.
- [45] 颀永鹏. 基于机器视觉的工业机器人目标识别和定位研究[D]. 沈阳: 沈阳工业大学, 2022.
- [46] Bregler C. Motion capture technology for entertainment [J]. *IEEE Signal Processing Magazine*, 2007, 24(6): 158–160.
- [47] Chan J C P, Leung H, Tang J K T, et al. A virtual reality dance training system using motion capture technology [J]. *IEEE Transactions on Learning Technologies*, 2011, 4(2): 187–195.
- [48] Choe N, Zhao H Y, Qiu S, et al. A sensor-to-segment calibration method for motion capture system based on low cost MIMU[J]. *Measurement*, 2019, 131: 490–500.
- [49] Luck M, Aylett R. Applying artificial intelligence to virtual reality: Intelligent virtual environments[J]. *Applied Artificial Intelligence*, 2010, 14(1): 3–32.
- [50] Qin F W, Li L Y, Gao S M, et al. A deep learning approach to the classification of 3D CAD models[J]. *Journal of Zhejiang University Science C*, 2014, 15(2): 91–106.
- [51] Zhu Z T, Wang X G, Bai S, et al. Deep learning representation using autoencoder for 3D shape retrieval[J]. *Neurocomputing*, 2016, 204: 41–50.
- [52] Cheng F C, Zhang H, Fan W J, et al. Image recognition technology based on deep learning[J]. *Wireless Personal Communications*, 2018, 102(2): 1917–1933.
- [53] Luciano L, Hamza A B. A global geometric framework for 3D shape retrieval using deep learning[J]. *Computers & Graphics*, 2019, 79: 14–23.
- [54] 武强, 季雪庭, 吕琳媛. 元宇宙中的人工智能技术与应用[J]. *智能科学与技术学报*, 2022, 4(3): 324–334.
- [55] 何泽浩, 隋晓萌, 赵燕, 等. 基于全息光学的虚拟现实与增强现实技术进展[J]. *科技导报*, 2018, 36(9): 8–17.

## Display, interactions and applications of immersive metaverse: Progress and outlooks

HE Zehao, CAO Liangcai\*

State Key Laboratory of Precision Measurement Technology and Instrument, Department of Precision Instruments, Tsinghua University, Beijing 100084, China

**Abstract** Some critical aspects of metaverse are briefed. The significance of immersion to metaverse is highlighted. Efforts to improve the sense of immersion are investigated. The principles of some immersive display technologies including head-mounted display (HMD), head-up display (HUD) and large size display are introduced and the corresponding prospects are envisioned. The importance of immersive interaction on metaverse is analyzed. Values of computer graphics, artificial intelligence and machine vision for metaverse are induced. The potential applications of metaverse are predicted based on the mentioned features.

**Keywords** metaverse; immersive display; three-dimensional display; immersive interaction ●



(责任编辑 卫夏雯)