头戴式显示设备（HMD）及其交互技术发展调研

## 一．引言

人类与计算机之间的交互方式一直在不断演进，早期的人机交互主要依赖于键盘和鼠标进行输入和操作。随着时间的推移，触摸屏、语音识别、手势控制等新技术的出现使人机交互变得更加多样化和自然化。而虚拟现实、增强现实等新兴技术的崛起，头戴式显示设备（head-mounted displays，HMD）应运而生，它作为一种重要的人机交互工具推动着这一领域的发展。本报告将探讨头戴式显示设备在虚拟现实（Virtual Reality，VR）、增强现实（Augmented Reality，AR）、混合现实（Mixed Reality，MR）等领域的发展历程，以及它们对人机交互的影响。

## 二．虚拟现实和增强现实

头戴式显示设备的历史可以追溯到早期的虚拟现实和增强现实技术萌芽阶段。虚拟现实和增强现实是两种不同的技术，尽管在某些技术层面可能有一些重叠，但它们的核心原理是不同的。

虚拟现实是一种可以创建和体验虚拟世界的计算机系统，它由计算机生成，通过视、听、触、嗅觉等作用于用户，为用户产生身临其境的感觉的交互式视景仿真技术。沉浸感(immersion)、交互性(interaction)和构想性(imagination)是虚拟现实系统的三个基本特征，其交互性是在交互设备支持下能以简捷、自然的方式与计算机所生成的“虚拟”世界对象进行交互作用,通过用户与虚拟环境之间的双向感知建立起一个更为自然、和谐的人机环境来保证的[1]。

虚拟现实一方面需要感知用户的肌肉运动、姿势、语言和身体跟踪等多个感官通道的输入信息:另一方面可以从人类的视觉、听觉、触觉、嗅觉等多个感官通道。模拟逼真的真实世界的感觉;还可以根据用户的视点变化和不同输入,迅速为用户提供实时的眼、耳、身体所能感受到的感觉信息.从交互的观点看,虚拟现实是一种新的人机交互界面[2]。虚拟现实发展至今已经有近30年的历史，然而由于其通常需要使用价格昂贵的特定输入设备，直至最近十年来才逐步进入主流市场。

增强现实（AR）是将现实世界的环境与数字信息进行融合，创造出一个增强的、混合的现实体验。在增强现实中，虚拟信息与现实世界中的场景相叠加，用户通过头戴式显示设备或移动设备（如智能手机或平板电脑）观察和交互。与虚拟现实不同，增强现实并不将用户从现实世界中隔离，而是在现实世界的基础上添加了数字内容。

AR技术可以分为多种类型，包括基于标记的AR、基于位置的AR、基于识别的AR和基于投影的AR等。这些技术在不同的应用场景下，提供了各种各样的增强现实体验。AR技术的实现离不开一系列关键技术的支持，包括跟踪和定位技术、实时图像处理、用户交互技术和显示技术等[3]。这些技术共同作用，使得用户能够在现实世界中与虚拟内容进行交互，并享受到增强的视觉体验。

VR和AR技术在近些年来得到了长足的发展，在头戴式显示设备的助力下获得了广泛的应用，包括游戏娱乐、职业训练、设计创作、医疗保健等。例如，它们都为游戏提供了更加沉浸式和互动性的体验，为商业活动增添了新的营销手段，同时也为手术模拟、教育培训等领域带来了更高效、更精准的解决方案。

## 三．头显设备交互技术的发展：

#### 1. 视线追踪技术

视线追踪技术是指利用传感器或摄像头等设备跟踪用户的视线方向和注视点，从而实现对用户注视目标的检测和分析。这项技术在头显设备中扮演着关键角色，因为它可以实现更自然、更直观的用户交互体验。

早期的视线追踪技术采用机械传感器或追踪器，通过测量用户头部的移动来推断视线方向。这种技术存在精度低、反应速度慢等缺点，限制了用户的交互体验。红外线传感器在之后被引入了头显设备，用于追踪用户眼睛的位置和注视点。虽然红外线追踪在一定程度上提高了精度，但仍存在受环境光影响、距离限制等问题。

随着计算机视觉技术的发展，视线追踪逐渐转向了基于摄像头的方法，这些摄像头可以捕捉用户的眼睛运动并分析视线方向。这种方法使用摄像头捕捉用户眼睛的图像，并通过图像处理和计算机视觉算法来检测眼睛位置和注视点[4]。这种方法可以提供更高的精度和灵活性，但也面临着对光照条件、摄像头分辨率和算法效率的挑战。近年来，深度学习技术的发展使得视线追踪更加准确和稳定。深度学习算法可以训练模型来识别眼睛和注视点，从而提高了追踪的准确性和鲁棒性。

除了计算机视觉技术外，光学追踪技术也在视线追踪中发挥着越来越重要的作用。眼动追踪技术利用红外或激光光源照射用户眼睛，并通过检测反射光斑来确定注视点。这种技术保证了视线追踪的高精度和实时性。而虹膜追踪技术利用摄像头捕捉用户虹膜的图像，并通过比对已知的虹膜特征来确定注视点。虹膜追踪具有高度的个性化和安全性，但也需要更复杂的设备和算法支持[5]。

#### 2. 手势与姿势交互技术

在虚拟现实中,用户身体或肢体的运动可以作为一种重要的输入通道.通过跟踪器或计算机视觉的方法跟踪人体相关部位(如头、手、臂或腿),获得人在物理世界的运动姿态信息作为虚拟现实系统的输入,通过识别算法解释为手势(gesture)或姿势(posture)。这是目前头显设备最主要的输入方式之一。

计算机视觉算法可以通过观察手或手指姿势以及肢体或头部的视频图像,来识别出特定的姿势。微软发布的Kinect深度相机为基于视频的手势交互带来了革命性的进展。Kinect利用一个普通的RGB摄像头和一个检测深度的红外摄像头来获取人的RGB图像和深度图像，基于深度图像数据跟踪人体的骨骼、手、头等各个部位的精度、速度都有大幅的提升,为手势交互走向实用提供了基础。

雷达技术也在手势识别中发挥了作用，它能够通过持续发射和接受手部反射的电磁信号,测量精细、复杂的动作变化,而后转码分析、识别[6]。Google Project Soli是这一技术的典型代表,技术人员设置并分类了四种手指的隔空动作指令: 转盘手势、刻度手势、按手势和位移手势,用户可以通过搓一下手指来调手表的时间,捏一捏空气来浏览地图,不需要实体按键就可以隔空选音乐、调音量等。

## 3. 语音交互技术

语音输入是一种很自然的输入方式，它能将不同种类的输入技术(即多通道交互)结合起来形成一种更有连贯性和自然性的界面。如果功能适当,尤其是用户的两只手都被占用的时候,语音输入将成为虚拟现实用户界面中很有价值的工具。语音有许多理想的特点:它解放了用户的手;采用一个未被利用的输入通道:允许高效、精确地输入大量文本;是完全自然和熟悉的方式,在虚拟现实用户界面中、语音输入尤其适合非图形的命令交互和系统控制,即用户通过发布语音命令来请求系统执行特定的功能、更改交互模式或者系统状态。此外,语音输入也为虚拟现实中的符号输入提供了一种完整的手段。

语音交互主要有两种方式:语音识别和语音对话系统，在使用语音交互开发虚拟现实人机交互时,首先要考虑语音界面要执行的交互任务,交互任务将决定语音识别引擎的词汇量大小,即任务和它所运行的范围越复杂，越需要更多的词汇量。对于仅有少量功能的应用,采用简单的语音识别系统就足够了,高度复杂的应用则需要包含语音对话系统来保证理解语音输入的全部功能。语音交互的关键是语音识别引擎,目前有代表性的语音识别软件包括微软 Speech API,IBM ViaVoiceNuance 和国内的科大讯飞等。而随着语音识别技术的逐步开放和开源语音技术门槛逐渐降低，CMU-Sphinx,HTK-Cambridge,Julius和RWTHASR等语音开源交互平台也越来越受到关注。

## 四．头戴式显示设备的应用进展：

目前，VR/AR的头显设备已经进入了商业化的阶段，比较成功的产品有主要面向游戏市场的Oculus Quest，HTC Vive以及面向办公场景的Microsoft HoloLens。 Oculus Quest是由Meta公司推出的一款VR头显设备，它配备有高分辨率显示屏和精准的头部追踪技术，为用户提供沉浸式的虚拟现实体验。HTC Vive是由HTC与Valve合作推出的VR头显设备，采用了房间尺寸追踪技术，使用户能够在更大范围内自由移动，并提供了更真实的虚拟现实体验。另外，Microsoft HoloLens是微软公司推出的一款AR头显设备，它将虚拟图像与现实世界相融合，为用户提供增强的、混合的现实体验。它配备有高分辨率显示屏、深度摄像头和传感器，支持手势、语音和眼动等多种交互方式，广泛用于教育、医疗、设计等领域[7]。

除了面向个人用户外，头显设备也开拓了面向企业的市场。在医疗领域，Simulab公司开发了名为"VR Surgical Simulation"的虚拟手术模拟系统，该系统于2018年在美国医疗机构推出，通过虚拟现实技术，让医生和外科团队在实战模拟手术中练习和培训，模拟各种手术场景和复杂情况，为医疗行业带来了创新的培训解决方案。在军事领域，BAE Systems公司在英国军队推出了名为"Virtual Command and Control Centre"的虚拟指挥训练系统，该系统于2020年在英国陆军和空军进行试用[8]。这个虚拟指挥训练系统利用了头戴式AR设备和沉浸式仿真技术，让指挥官可以在虚拟指挥中心中模拟作战指挥和决策，远程协同训练和指挥部队，提高军队的作战效能和战斗力。

在娱乐行业，SM Entertainment在韩国推出了名为"SM Town VR"的虚拟现实音乐会和演出，该音乐会于2022年在首尔举办。SM Town VR利用VR技术，让粉丝可以通过头戴式VR设备参与虚拟现实音乐会，欣赏明星表演和与偶像互动。这种虚拟现实音乐会为粉丝提供了全新的音乐体验，超越了传统音乐演出的空间限制，让粉丝可以随时随地感受到偶像的魅力和热情。

## 五．混合现实与Apple Vision Pro

混合现实（Mixed Reality，MR）是一种新兴的交互式计算环境，它将虚拟世界与现实世界相结合，创造出一种新的混合体验。在混合现实中，虚拟对象和现实对象能够在同一个空间中共存，并且可以互相影响和交互，从而为用户提供丰富、沉浸式的体验。混合现实技术融合了虚拟现实（VR）和增强现实（AR）的特点，旨在打破虚拟与现实之间的界限，为用户创造全新的交互和体验模式。

混合现实的核心技术包括传感器、计算机视觉、姿态追踪、投影技术等。通过这些技术，混合现实设备能够实时感知用户的周围环境，并将虚拟对象与现实环境进行精准地对齐和交互。同时，混合现实设备还能够实现虚拟内容的空间感知和真实物体的虚拟增强，使得用户能够与虚拟世界进行更加直观和自然的交互[9]。

在混合现实中，虚拟内容可以呈现在用户的视野中的任何位置，并且可以根据用户的动作和行为进行实时调整和交互。用户可以通过头戴式显示设备、手持式设备或者其他交互工具来体验混合现实技术，从而与虚拟和现实世界进行沟通和互动。

2023年6月，苹果公司首次发布了Apple Vision Pro，作为一款MR头显，Vision Pro的出现具有着里程碑意义。Apple Vision Pro是一款革命性的空间计算设备，通过用户的眼睛、手和声音提供自然和直观的输入方式。Apple Vision Pro 搭载 visionOS，基于 macOS、iOS 和 iPadOS 数十年的工程创新，visionOS 提供了强大的空间体验，由于 visionOS 利用现有的开发人员框架，Apple Vision Pro 上有超过100万个熟悉的 iOS 和 iPadOS 应用程序可用，并且这些应用程序会自动与新的输入系统配合使用。它配备了一个全新的三维用户界面和输入系统，完全由用户的眼睛、手和声音控制，导航体验神奇。直观的手势让用户只需看着应用程序、轻触手指进行选择、轻挥手腕进行滚动，或使用虚拟键盘或语音输入进行打字[10]。

Vision Pro采用 Apple 硅片打造的突破性超高分辨率显示系统，结合微 OLED 技术，将 2300 万像素封装到两个邮票大小的显示屏中，并具有宽色域和高动态范围。它还配备了高性能的眼球追踪系统，该系统使用高速相机和一圈 LED，将不可见的光线投射到用户的眼睛上，以实现响应迅速、直观的输入。EyeSight功能实现了突破性的环境监测，当有人靠近佩戴 Vision Pro 的用户时，该设备看起来是透明的，让用户可以看到他们，同时还会显示用户的眼睛。当用户沉浸在环境中或使用应用程序时，EyeSight 会向其他人提供有关用户关注内容的视觉线索。

然而，即便作为目前业界最顶尖的MR头显设备，Apple Vision Pro还存在着许多显著的问题。比如外接电源的存在使得整个设备过重，佩戴不便；续航只有不到三小时，不能满足长时间办公或娱乐的场景需求；软件生态系统不够完善，大部分应用都是基于iPad应用的修改版，大多流媒体厂商都还在观望，用户在选择和使用应用程序时会受到限制，从而影响到整体的使用体验。

## 六．头戴式显示设备面临的挑战

我认为，目前的头戴式显示设备虽然进入了批量市场化的阶段，但其存在的问题仍然十分显著，如何解决这些问题，可能是头显设备下一步的发展方向。

由于HMD是一种视觉显示器，其相关参数以人眼为基础，眼罩的大小限制了人眼只能看到一定空间范围内的虚拟图像，而用于航空、制造和娱乐等行业的HMD设备迫切需要大的视野和更高的分辨率，而要同时实现大视场角和高分辨率并不容易，因此开发实现这一目标的方法成为一项重大技术挑战。HMD的微型显示器光谱带宽还必须与人眼的敏感光谱区域相匹配，而商业应用通常要求 HMD 具有全彩显示功能。这就产生了必须消除或校正色差的问题。此外，还要求微型显示器在人眼舒适范围内有较大的亮度调节范围，以便在人眼接收虚拟信息时，外部环境与虚拟信息之间形成足够的亮度对比[11]。

首先，由于头显设备通常需要集成复杂的传感器、显示屏和计算单元，因此目前它们的体积和重量往往较大，使得用户在长时间佩戴的情况下可能会导致不适。其次，由于头戴式设备直接接触于用户的头部和面部，因此其佩戴结构和材料非常重要。如何尽可能减轻硬件设备的重量，更人性化地设计设备的结构，研发更透气舒适的佩戴材料将成为头显设备面临的一大挑战。此外，设备的体积和重量也影响了其便携性，而电源、手部传感器等配件也增加了携带的繁琐程度，使得其用户在移动时可能需要额外的携带或支架来支撑设备。

在内容创作和应用场景方面，头显式VR和AR设备也面临挑战。首先，与传统的应用程序相比，虚拟和增强现实应用需要更加沉浸和交互性的体验，这意味着开发团队需要具备更多的技术和设计能力，来开发出适用于头戴式设备的内容。此外，由于头戴式设备的硬件限制（如处理能力、存储空间、电池寿命等），开发者需要在内容创作过程中考虑如何优化性能和资源利用率，以确保用户体验的流畅性和稳定性。其次，如何找到适合头戴式设备的实际应用场景，并将其与用户需求相匹配也是一项重要的论题。虚拟现实和增强现实技术的应用范围非常广泛，涵盖多个领域，然而，对于每个特定领域，开发者需要深入了解用户需求和行业特点，设计出切实可行且有价值的应用场景。同时，还需要解决与现有技术、法律法规、隐私安全等方面相关的问题，以确保应用的可持续性和合法性。

## 七．头戴式显示设备的发展建议与展望

我认为，头显设备在未来还有许多发展的可能性。随着材料科学的进步与人体工程学的运用，头部的穿戴设备将变得更加轻薄和柔软，使得佩戴者能够长时间舒适地使用。未来，我们可以着力研发采用更轻、更柔软的材料，如柔性电子材料和纳米材料，以及更人性化的设计，如可调节式头带和柔软填充物，以提高佩戴舒适性。此外，光学仪器精密性的进一步提升有望消除当前头显设备普遍存在的色差问题，提供更贴近人眼的真实视觉体验。未来的头显设备可能会采用更先进的光学元件和成像技术，如全息光学元件和超高清显示屏，以实现更真实、更清晰的图像效果。

未来，我们可以期待头显设备的分辨率进一步提升，从当前的高清甚至超高清，到更加细腻和逼真的8K甚至16K级别。同时，刷新率也将提高到更加流畅的水平，消除视觉上的闪烁和模糊感，更高的分辨率和更快的刷新率使得用户能够更加清晰地感知虚拟世界。此外，设备的视野范围也应进一步扩展，让用户更自然地感受到虚拟和现实世界的融合。未来的头显设备可能会采用更广角的光学镜头和更大尺寸的显示屏，以扩大用户的视野范围，并提供更加沉浸式的体验。

智能化功能也可能成为未来头显式显示设备的一大研究方向。设备可以配备更先进的传感器和处理器，提升算法的识别准确度与灵敏度，实现更智能的交互方式，例如更精准的语音和手势识别、全方位高分辨率的眼动追踪等。未来的头显设备可能还可以引入更先进的人工智能算法和机器学习技术，从而实现更智能、更个性化的交互体验。例如，设备可以根据用户的行为习惯和偏好，自动调整显示内容和交互方式，提供更加智能化的用户体验。此外，未来的头显设备还可能支持更多样化的交互方式，如脑机接口和心理识别技术，进一步提高设备的人机交互性和用户友好性。

此外，未来的头显式显示设备还有可能和其他技术融合，例如5G网络、、云计算等，从而在这些技术的助力下实现更强大的功能和更广泛的应用场景。例如，5G网络的高速和低延迟将为设备提供更加流畅的数据传输，支持更多复杂的虚拟和增强现实应用。云计算技术将为设备提供强大的计算能力，减少设备自身的负载，支持更复杂的虚拟现实场景和应用。

我认为，未来的头戴式显示设备有着广阔的发展空间，它有可能成为数字化生活的重要设备，为人们带来更加丰富、沉浸式的数字体验，推动智能化生活的进一步普及和发展。

### 【参考文献】

[1] 张凤军,戴国忠,彭晓兰.虚拟现实的人机交互综述[J].中国科学:信息科学,2016,46(12):1711-1736.

[2] Çöl B G, İmre M, Yıkmış S. Virtual reality and augmented reality technologies in gastronomy: A review[J]. Efood, 2023, 4(3): e84.

[3] Dargan S, Bansal S, Kumar M, et al. Augmented reality: A comprehensive review[J]. Archives of Computational Methods in Engineering, 2023, 30(2): 1057-1080.

[4] 侯守明,贾超兰,张明敏.用于虚拟现实系统的眼动交互技术综述[J].计算机应用,2022,42(11):3534-3543.

[5]卓林海.高精度视线跟踪算法研究及其在VR系统中的应用[D].华南理工大学,2019.DOI:10.27151/d.cnki.ghnlu.2019.001037.

[6] Mehrfard A, Fotouhi J, Taylor G, et al. A comparative analysis of virtual reality head-mounted display systems[J]. arXiv preprint arXiv:1912.02913, 2019.

[7]Jensen L, Konradsen F. A review of the use of virtual reality head-mounted displays in education and training[J]. Education and Information Technologies, 2018, 23: 1515-1529.

[8]邱睿,贺志伟,吉峰等.基于人机交互的虚拟现实技术在军事中的研究综述[J].软件,2021,42(06):123-125.

[9] Rokhsaritalemi S, Sadeghi-Niaraki A, Choi S M. A review on mixed reality: Current trends, challenges and prospects[J]. Applied Sciences, 2020, 10(2): 636.

[10] <https://www.apple.com/apple-vision-pro/>

[11]Cheng D, Wang Q, Liu Y, et al. Design and manufacture AR head-mounted displays: A review and outlook[J]. Light: Advanced Manufacturing, 2021, 2(3): 350-369.