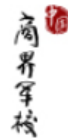




PHBS

北京大学汇丰商学院



元宇宙三十人论坛
Metaverse 30 Forum



安信证券
ESSENCE SECURITIES

元宇宙2023：硬件的“大”年

北京大学汇丰商学院商业模式研究中心

安信证券元宇宙研究院

元宇宙三十人论坛

2023年1月1日

魏炜、焦娟、王晶晶、冯静静、王利慧、易欢欢



研究
思路

01

2023版元宇宙研究框架、元宇宙的上半场与下半场

02

2023：硬件的“大”年

03

硬件之“沉浸感”、“交互范式”、“操作系统”

04

元宇宙之百舸争流

- 虚拟人：交互世界的新物种；NFT：数字世界的魔法棒
- Web 3：内核是“治理”，“去中心化”实则为“反抗有理”
- AI与后端基建：2023年迎来景气上行
- 场景：2024-2025年的“八面来风”

继2021年FB改名Meta是为“**穿云箭**”，我们对2022年全年的判断为“**蓄积的力量**”。2022年我们看到了PICO/META/HTC/SONY/TCL/联想/创维等的新一代**硬件**产品、看到了NVIDIA/UNITY的**底层架构**系列化发布、看到了**AI作画/AIGC（内容与场景）**的巨大争议，看到了国内外科技巨头伴随股价下跌而来的争议甚至是批评——2022年确是预期中的“**分歧期**”；

着眼于国内二级市场，我们预计7月底8月初会有新一轮元宇宙行情的开启，在去年**NFT、元宇宙、虚拟数字人**的基础上，我们认为细分方向有望新增**Web 3、人形机器人、XR硬件相关**等，各方向相关标的在去年基础上会有明显的发散，但核心仍然在于弹性。2022年国内A股的元宇宙行情与2021年有一定的呼应——2021年7、8月份NFT，9月元宇宙、11月虚拟数字人；2022年7月人形机器人、8月XR硬件、9月/10月调整、11月Web 3/AIGC——2021年行情的触发要素是Meta改名，2022年则是5部门联合发布的虚拟现实五年发展计划；

过去的2022年，我们兢兢业业且持续迭代研究框架，无论是元宇宙的年度预判，还是国内二级市场的节奏推演，唯一值得庆幸感恩的，是**我们敢真正去下前瞻判断，且得到了基本兑现的正向反馈**。在我们给出**2023年的年度前瞻判断**之际，我们将同时：1）迭代2021年建立的**研究框架**（六大投资版图）；2）新增元宇宙**下半场**的系统推演（《新硬件主义》，2022年5月出版）；3）植入112页**拆机报告**的独家研究结论（硬件入口的沉浸、交互、操作系统）；4）深化对**虚拟人与NFT**的理解与定位；5）提示对**Web 3**的警惕；6）展望将要在2024-2025年“八面来风”的**场景**的投资框架；7）**AIGC**；

我们非常期待2023年！2023年有可能是“**预期差**”最大的一年，我们认为它**将是“硬件的大年”**，“**硬件**”是指MR，即通用硬件设备，“**大年**”，则是综合考虑了“**周期、全局、结构**”。之所以会导致“预期差”最大，是市场尚未认知到“交互范式的定义”对硬件入口的重要性，此外，市场也尚未意识到，与硬件配套的软件部分，借力于AI的应用/场景/模式甚至是内容的创新（AIGC）速度极快，这是认知**结构**层面的“预期差”；META&MICROSOFT与APPLE，通用硬件设备的开源&封闭系统之争，犄角之势将领航2023年的**全局**；2023年又是多周期共振出清的节点，市值的底部有望迎来的并购潮、巨头市值下跌将加速其对“第二曲线”的“打出明牌”，是大**周期**轮动给2023年的大助力。



研究
思路

01

2023版元宇宙研究框架、元宇宙的上半场与下半场

02

2023：硬件的“大”年

03

硬件之“沉浸感”、“交互范式”、“操作系统”

04

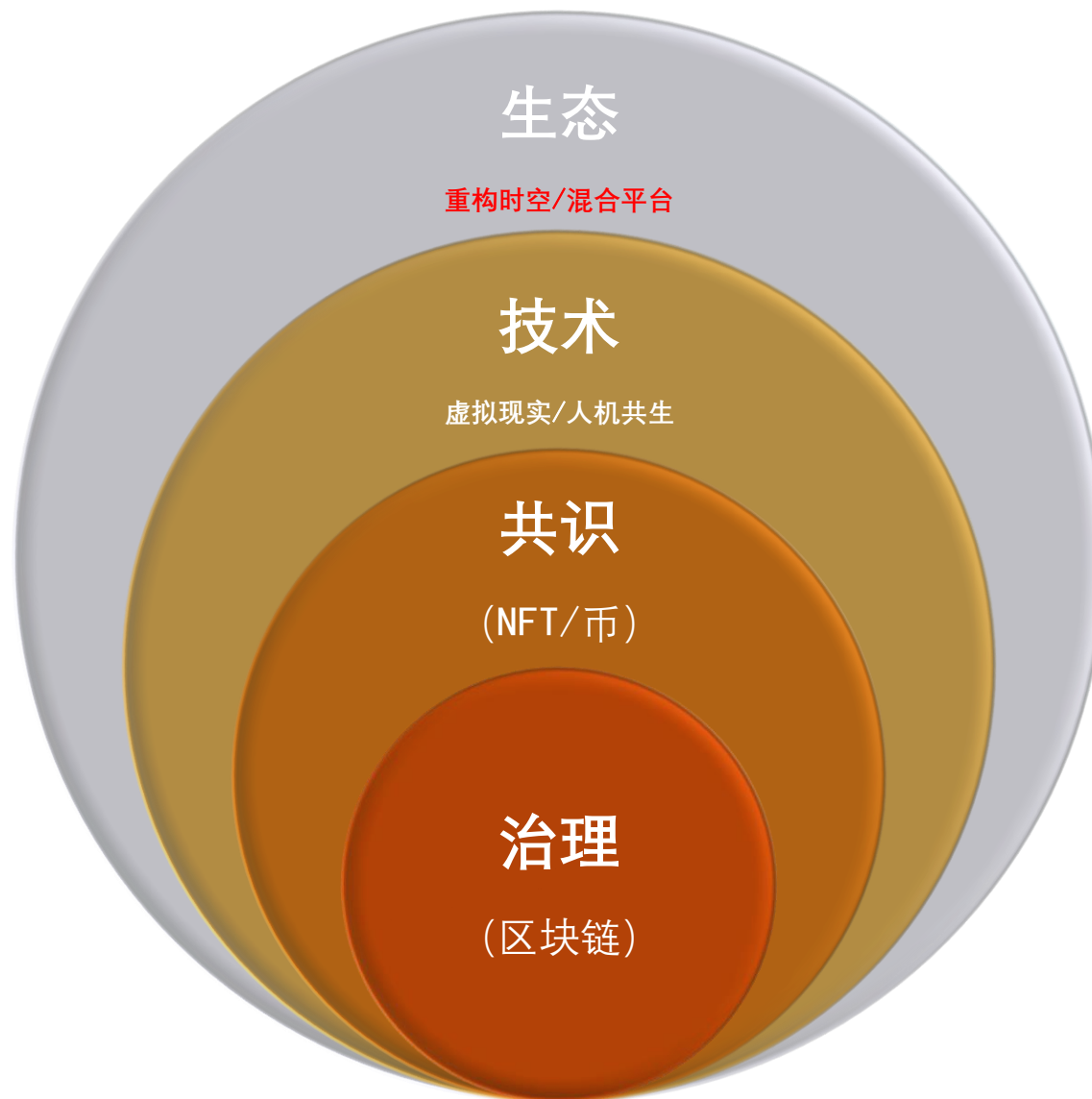
元宇宙之百舸争流

- 虚拟人：交互世界的新物种；NFT：数字世界的魔法棒
- Web 3：内核是“治理”，“去中心化”实则为“反抗有理”
- AI与后端基建：2023年迎来景气上行
- 场景：2024-2025年的“八面来风”

2023版元宇宙研究框架

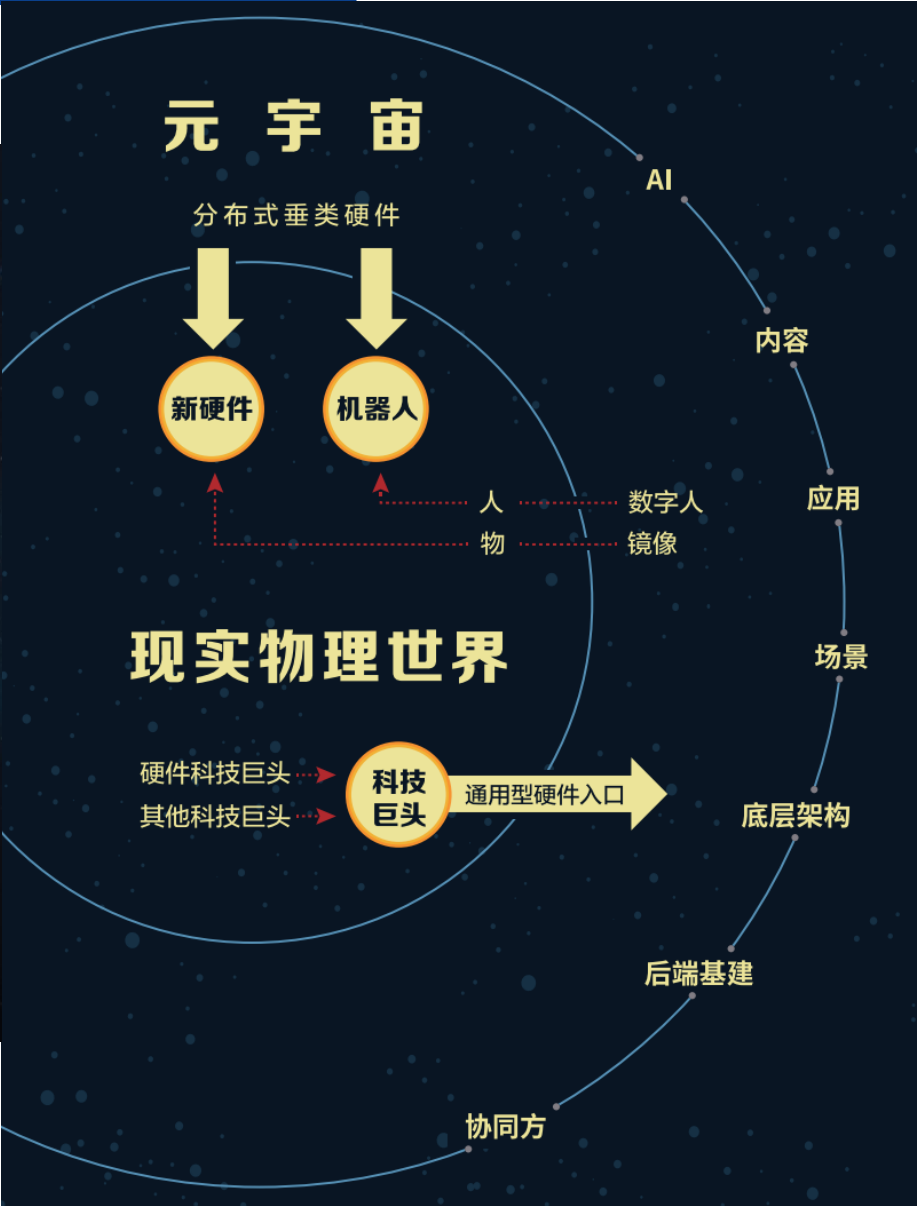


研究框架所配套的认知体系



所有繁杂概念都可归入的认知体系

元宇宙的上半场与下半场



资料来源：安信证券研究中心整理

2026年是上下半场的中间节点

我们更新了元宇宙的版图，以“智能”的真正实现为总纲，下括**元宇宙**与**混合平台**这两条肩负着真正实现虚拟现实的技术路径，其中：

重构时空的元宇宙，其构建需要六大板块（硬件与内容/场景；后端基建/人工智能；底层架构/协同方）；

混合平台包括人形机器人（以机器为载体）、脑机接口（以人为载体）。

元宇宙是实现了智能的“内在部分”；如果将人的身体看作“硬件”，智能的实现过程，也是**新智能交互硬件层出不穷**、**旧硬件迭代甚至是重塑**的过程。

2023年焦点为硬件（MR）、2024/2025年核心在应用/场景/模式等配套硬件的诸多新内容；

2026年是马斯克指向的人形机器人进入家庭端的节点年份、本质均为AI的人形机器人或将开启2027-2029年的垂类硬件的快速发展期；

2026年之前是元宇宙的上半场建设过程，2026年之后是元宇宙下半场囊括现实物理世界的过程，2030-2032年元宇宙或将已如移动互联网般深入人心。

由特斯拉人形机器人透视元宇宙下半场

1. 特斯拉的人形机器人之所以火爆的原因？
2. 人形机器人与智能电车、Neuralink有什么关系？
3. 以全局的眼光来理解马斯克的“庞大商业版图”背后的科技意图；
4. 新硬件之所以“新”，跟智能手机及之前的硬件逻辑已完全不同；
5. 新硬件的历史观、本质、全球24家相关科技巨头；
6. 新硬件与元宇宙有什么关系？AI的重要性被低估；
7. 引入古印度的“唯识学”，提供一个特别视角，去理解真正的“智能”。

由特斯拉人形机器人透视元宇宙下半场

特斯拉人形机器人“火”的背景

特斯拉以及其背后的马斯克，备受科技圈瞩目，马斯克在科技界的领袖风范堪比智能手机时代的乔布斯。也正因为特斯拉以“鲶鱼”的姿态进入中国，触发了2019年以来新能源电车浩浩荡荡的大行情，堪比2013年左右开始的移动互联网资本浪潮；

2021年开始，马斯克对特斯拉的期许，越来越转移至机器人“擎天柱”身上，2022年4月更是直言“擎天柱”是特斯拉的“第二曲线”、“这比车重要（“这”指特斯拉推出人形机器人的原型机）”；

当下仍处于新能源电车流畅且热烈的行情中，出于对科技前沿的敏锐、更出于对复制“新能源电车”火热行情的期待，国内资本市场对“擎天柱”关注，期待“机器人”板块性行情；

很多人开始感受到了这种“火热”，但绝大多数人仍一头雾水，“机器人”由来已久，“擎天柱”的特别之处在哪里？

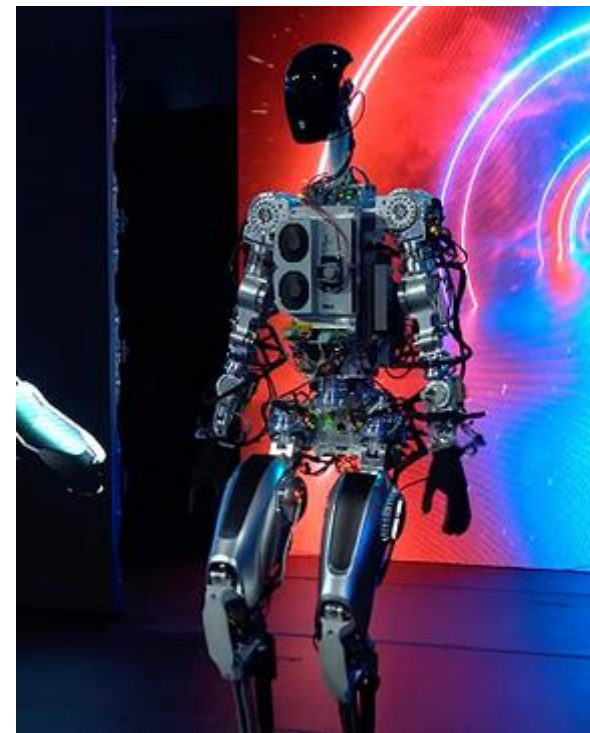


由特斯拉人形机器人透视元宇宙下半场

“擎天柱”这个单点，确实不简单

本质上与特斯拉的智能电车相同，均是建立在“输入——计算平台——输出”模型上的智能交互硬件。也就是说，完全不同于工业界逐渐在普及的“机器臂”，“机器臂”是一种生产力，而“擎天柱”是新一代智能交互硬件，你甚至可以把它看做一个独立“个体”，能独立应付复杂的环境、消解较为含混的“指令”，比如指令为“照顾好我奶奶”，背后是人工智能走向了认知与决策阶段；

这一次之所以资本市场反应非常快，是因为马斯克已经给出了“擎天柱”落地、投产、广泛应用的时间表：2022年出原型机、2023年出有用的呈现（场景）、最快两年能做到小规模应用，此后，机器人应用逐年增长、成本降低、产量规模扩大；且已经给出了“擎天柱”规模化推广及应用的节奏：1）第一批机器人服务于危险、无聊、重复的工作或人们不想做的工作，主要是to B的应用；2）第二批为大规模使用的机器人，让机器人在现实世界中导航，没有明确的指令下做有用的事情；3）第三批是机器人进入家庭端。

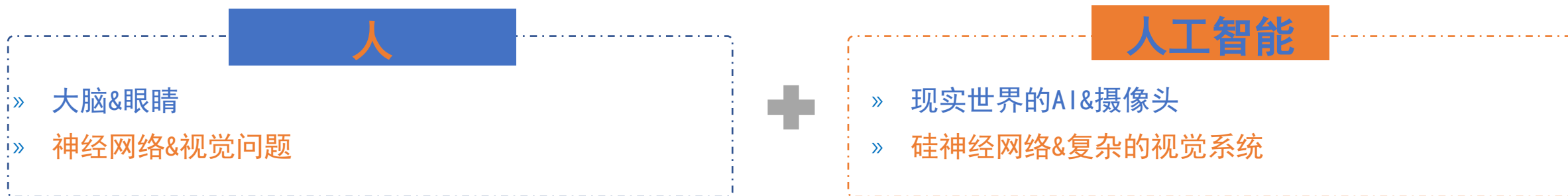


自动驾驶与人形机器人

马斯克认为“擎天柱”比自动驾驶容易（对速度的要求、速度高低的风险敞口完全不同），即使刚开始价格高（产量低），但规模化生产后，成本比一辆车便宜（汽车的复杂性和成本都高于类人机器人）；

我们先接受一个前提假设：移动互联网已是“旧周期”，新一代的计算平台是“元宇宙”，未来的所有智能交互硬件，都会架构在新一代计算平台上，比如你的可穿戴设备——智能手表，只会跟你的智能手机连接，而不可能再架构在你的PC电脑上，更不用说架构于家中客厅里的那台电视机；

从智能的角度，未来新的智能交互硬件绝对不止人形机器人这一种，我们认为新硬件有两类：通用型硬件入口、分布式垂类硬件，其中分布式垂类硬件又包括三小类：基于情感需求投射的垂类硬件（对应元宇宙的内容/应用/场景）、机器人（对应元宇宙中的数字人）、未来将被重塑的现实世界中的“物”。



人形机器人将共享特斯拉的自动驾驶能力、规模化生产的能力；
传感器与执行器当下已有解决方案运行得不错，目前缺少的是足够的智能（AI通用平台）。



特斯拉的 自动驾驶能力

1. 已建立了高质量统一的向量空间；
2. 8个摄像头绑一起拍摄环绕图，加入时间维度，生成环绕立体视频；
3. 环绕的视频必须被自动标记，特斯拉编写了自己的标签工具，创造了自动标签系统；人类标签只是辅助，对已经生成的视频标签做一些细微的修改，然后将这些修改反馈给未来的订单标签机；最终你得到了一个“飞轮”，在这里汽车标签能够接收大量的视频，高精度地自动标签视频汽车、车道线、驾驶空间；
4. 需要对时间和空间进行存储，存储不能是无限的，会消耗大量内存。



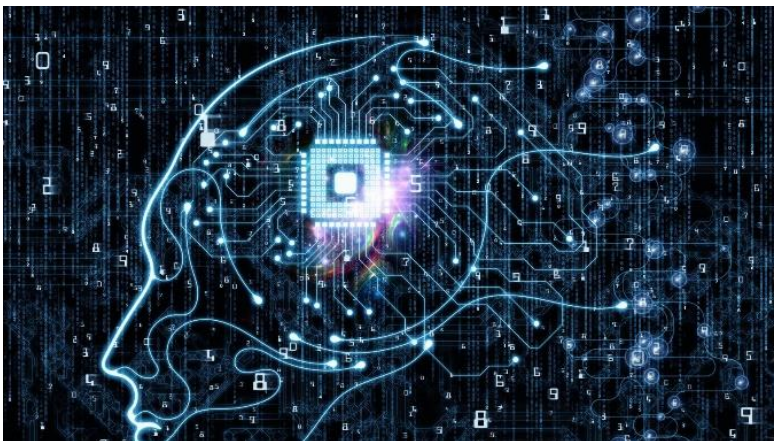
特斯拉的 自动驾驶能力

1. 计算机（AI、摄像头）替代人（大脑、眼睛），实际上解决的是现实世界的AI和视觉问题；
2. 根据特斯拉在无人驾驶上具备的能力（自动驾驶、规模化生产），类人机器人上，特斯拉只需要设计专门的驱动器以及类人机器人所需要的传感器；
3. 安全设施方面，需要在新硬件上安装一个本地化的ROM芯片，且无法在线更新；
4. 设立人工智能监管机构，这对公共安全很重要；
5. 机器人的风险在于人工智能或数字超级智能与人类的集体意志脱钩。

Neuralink：与人形机器人并行的工程方案

混合平台是智能的一条技术路径，在这条技术路径上，脑机接口与人形机器人是不同的工程方案，即智能的实现分别以人、机器人为载体：

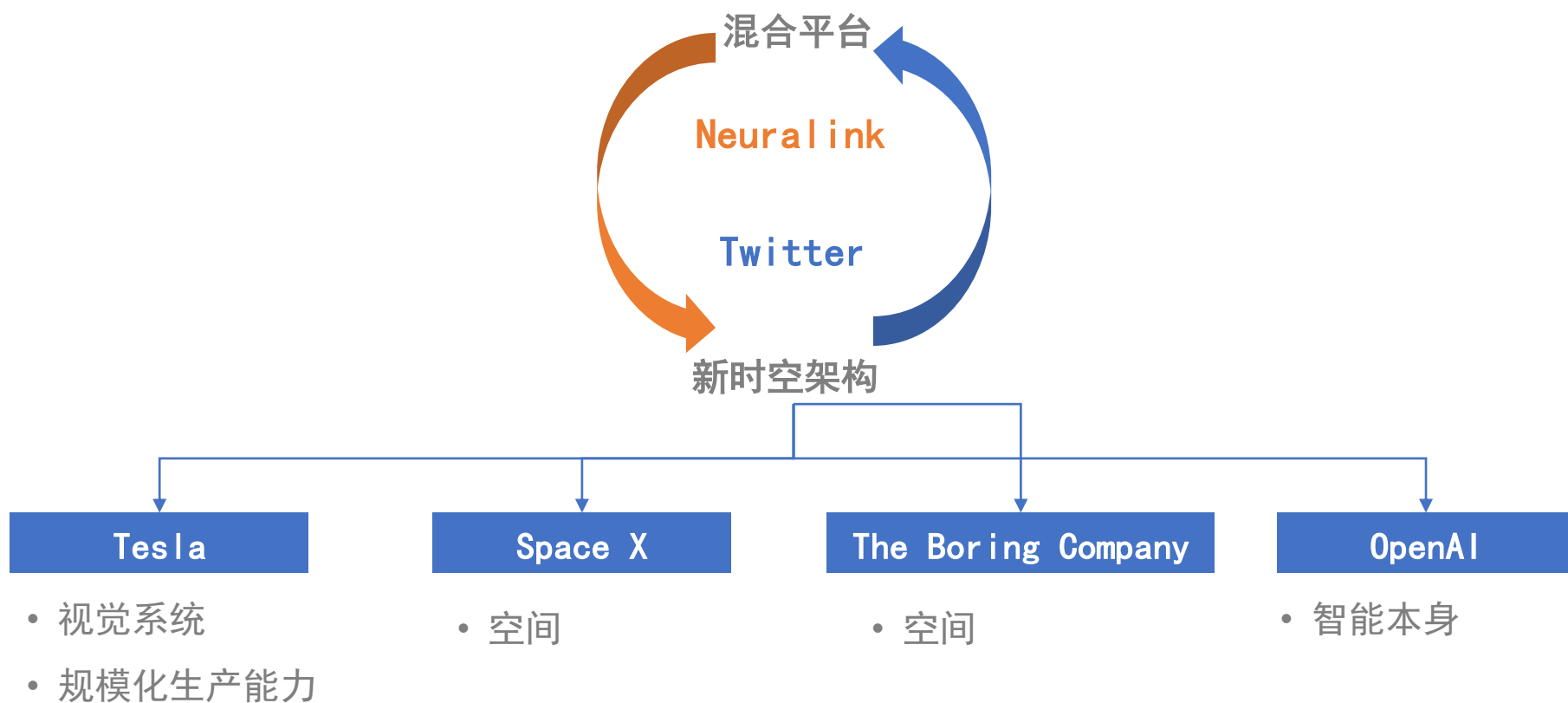
1. Neuralink背后的想法：为了更紧密地结合人类集体世界；
2. 数字超智，将人类的集体世界转化为数字智能；
3. 同时解决很多脑损伤和脊髓损伤。



1. 我们已是终端的一部分——计算机是我们自身的延伸。
2. 是什么抑制了人机共生？是数据速率；
3. 读取神经元的基本原理，用微小的电极在神经元上进行读写并非新概念；
4. 但当下没有一种产品能很好地运转且带宽高、安全，均为实验室测试款；
5. Neuralink装置的方式，像Fitbit或苹果手表，数字智能与生物智能紧密联系在一起；
6. 需要很多年做到近似于高带宽神经接口的东西，让人工智能与人类共生；
7. 长期是来应对人工智能带来的文明风险。

马斯克的布局版图

海外对AI的重视程度很高，真正的“智能”有嵌套的四个层次，若要实现真正的智能：1) 混合平台，如人机接口；2) 重新建构一个环境，元宇宙即新的时空架构。



自哥伦布大航海时代以来，关于地理空间的探索已阶段性到头，“上天”“入地”“元宇宙”均是空间探索。

软硬一体/高度集成/门槛提高/竞争加剧

1. 全局角度，未来新的智能交互硬件绝对不止人形机器人这一种，我们认为新硬件有两类：通用型硬件入口、分布式垂类硬件，其中分布式垂类硬件又包括三小类：基于情感需求投射的垂类硬件（对应元宇宙的内容/应用/场景）、机器人（对应元宇宙中的数字人）、未来将被重塑的现实世界中的“物”。
2. 擎天柱藏着两个秘密：1）未来的智能交互硬件，本质都是一样的，能做人形机器人，自然能做出更多新的智能交互硬件；2）“擎天柱”身上藏着一个马斯克尚未实现的秘密——特斯拉尚未实现真正的“智能驾驶”，这也是特斯拉将2022年的“人工智能日”推迟到9月30日的原因——一旦攻克了“基于现实世界的“智能”问题”，自动驾驶与人形机器人将同时成功，且马斯克未来能毫无难度的推出更多的新硬件，那将产生特斯拉的“第三曲线”、“第四曲线”

以“擎天柱”为代表的新一代智能交互硬件，将与智能手机及之前的交互硬件有本质性的区别：

- 1) 软硬一体将大势所趋；
- 2) AI的重要性更加凸显；
- 3) 入局的玩家将前所未有的广泛。

软硬一体/高度集成/门槛提高/竞争加剧

以“擎天柱”为代表的新硬件（硬件入口、垂类硬件）已走出软硬一体的趋势，**软件、硬件、算法等诸多核心部位在一家公司内高度集成**（**苹果模式**，但有别于过往硬件的主流模式——开放平台结合标准化的硬件与软件，再整合），竞争门槛急剧提高，但在这个方向能成功的未来“王者”，将是真正的“集大成者”，会让竞争对手真正的“望其项背”；

AI在元宇宙全貌中是递进式的发展脉络：1）元宇宙上半场中，AI是元宇宙的六大组成部分之一，是元宇宙中的关键生产要素；2）元宇宙中场，AI以供给的方式辅助于人的决策，挖掘人的更多需求/影响人的认知；3）元宇宙下半场中，AI将以新硬件/垂类硬件的方式，叠加入现实物理世界，以更好的满足人的需求/影响人的认知，由未来回溯当下，AI的重要性严重被低估，AI是新硬件的内核。我们在《新硬件主义》一书中，花了不少篇幅，着重讲“供需”、“AI”，就是基于未来人的认知，大概率将要被“AI”的供给所深度影响，大家要注意，“AI”在元宇宙中是内容、应用或场景，但“AI”也将显现为“各式垂类硬件”出现在现实物理世界中，以完成元宇宙对现实物理世界的囊括与重塑，这一幕很震撼，终将到来时也会让人“身不由己”——围绕一个人的24个小时，哪哪都是AI；

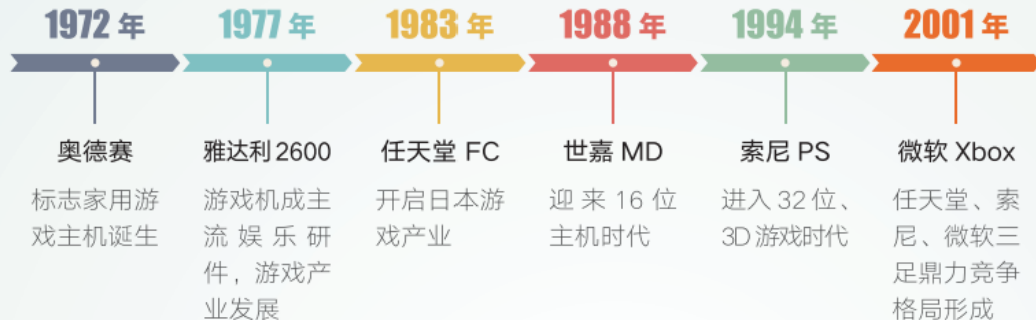
“王者”、难以“望其项背”的诱惑太大了，**新硬件将前所未有地吸引最广泛的玩家：**1）竞争的胜负手由芯片、硬件入口至少扩展至芯片、通用型入口、AI、分布式垂类，入局方将包括但不限于XR厂商、智能手机厂商、智能电车厂商、AI与芯片巨头、垂类厂商（游戏主机、智能电视）、互联网巨头、诸多细分子行业的龙头、初创型公司等；2）新硬件的竞争会更残酷——元宇宙的前半场是二维移动互联网升维至元宇宙，难度系数高；但元宇宙的后半场是三维元宇宙反向叠加现实物理世界，难度系数确定性会降低，全球科技巨头这一总决战，酣战至中场，利益格局的划分有可能就已接近于终局。

● 聚焦元宇宙下半场，描述元宇宙怎样囊括现实物理世界

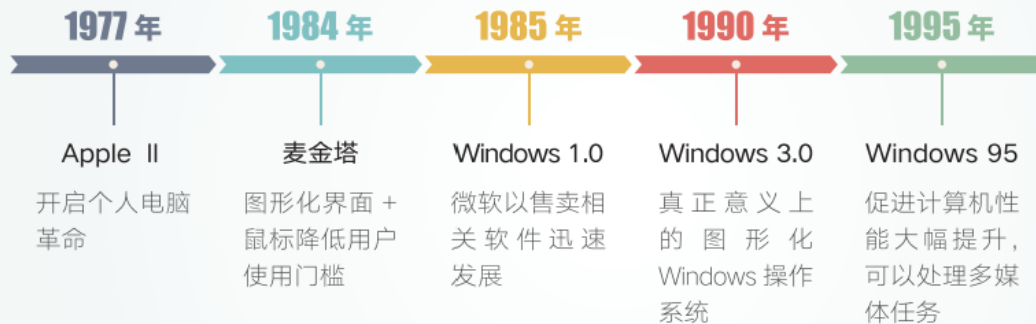
1. 《元宇宙大投资》专注于元宇宙上半场建设过程中的研究与投资框架；
2. 《新硬件主义》聚焦于元宇宙下半场，刻画元宇宙怎样囊括现实物理世界的认知体系；
3. 两本书的成书时间相隔半年，全球化时代背景扭转为全国统一大市场，《元宇宙大投资》作为一种分析框架，立场较为中性，《新硬件主义》作为时代背景扭转之后的一种认知体系，立场有主观色彩——时代背景变了，思路与打法需要彻底扭转，蹚一条自己的路，在此之前，则需要判断出正确的方向——新硬件；
4. 中国科技在元宇宙硬件入口方向上的资源投入，转一部分投向分布式垂类硬件，这一方向胜率与赔率更高，更重要的是，这是一条自己蹚出来的路，而非跟随策略。简单来说，硬件入口不允许有任何短板，分布式垂类硬件可以充分利用长板效应。

新硬件历史观：交互硬件50年历史梳理

垂直计算硬件
游戏主机



通用计算硬件
个人电脑与 windows 系统



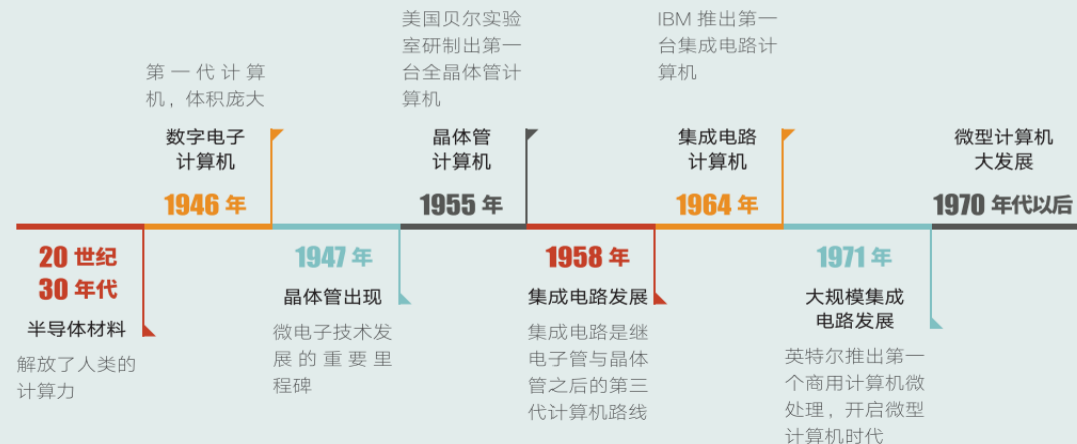
小型化计算硬件
掌上游戏机与智能手机



● 新硬件

1. 通用型硬件入口；
2. 分布式垂类硬件：基于情感需求投射的新硬件、机器人、当下现实物理世界中部分被重塑的“物”。

硬件集成了各时代最先进的生产力



新硬件本质：合之为—、衍之为万

● 软硬一体：内核为AI，外显为垂类硬件

- » 芯片是通用型硬件入口的命门，AI则是分步垂类硬件创新的命门，故全球科技巨头在元宇宙这盘大棋中的胜负手，上半场主要是硬件入口、芯片，下半场则增加了垂类硬件、AI；硬件入口与垂类硬件的发展并非绝对的先后顺序，垂类硬件中的机器人、当下现实物理世界中部分“物”的重塑（如冬奥会中各类黑科技硬件）已启动；
- » 垂类硬件中最重要的当属“基于情感需求投射的新硬件”，为什么？当下已启动的机器人、物理世界中部分“物”的重塑，仍处于“感知”阶段（或正走向“认知”阶段），只有当元宇宙走完上半场，AI作为新的生产要素，以它的供给，充分挖掘出人在元宇宙中的更多、新情感需求，同时在科技的需求下，AI反向映射入现实物理世界，显现为“基于情感需求投射的新硬件”时，各式垂类新硬件已由“认知”阶段走向“决策”阶段，真正完成了元宇宙对现实物理世界的重塑及囊括。此外，还有基于“利益”需求的新硬件。



新硬件本质：合之为—、衍之为万

● 软硬一体：驱动力是“人与科技的需求”

- » 我们将新硬件划分为两大类，通用型入口与分布式垂类，其中入口只负责“承接”，交接用户的是元宇宙内的万千存在；
- » 此外，AI映射入现实物理世界显现为分布式垂类硬件，这背后均有底层架构、后端基建、内容与场景、协同方的合围力量；
- » AI能真正的独立产品化，是未来的可能性，也是当下的发展方向；
- » 新硬件本质上就是AI，这也回答了此前部分读者的一个疑问：元宇宙是软硬一体化发展吗？我们认为元宇宙一定是软硬一体化，甚至更进一步——新硬件即是AI，AI以显现为新硬件的方式更契合未来AI的进化需求。

元 宇 宙

分布式垂类硬件

新硬件

机器人

人
物

数字人
镜像

现实物理世界

硬件科技巨头

其他科技巨头

科技
巨头

通用型硬件入口

底层架构

后端基建

协同方

AI

内容

应用

场景

底层架构

后端基建

协同方

元宇宙全貌 《元宇宙大投资》 《新硬件主义》

● 两本书给出了元宇宙的研究框架、认知体系，持续迭代以服务于未来十年的元宇宙真实发展

1. 元宇宙：囊括现实物理世界、数字化everything的虚拟集合。数字化人的更多感官体验，由移动互联网的二维升至三维；
2. 新硬件主义：中国科技的打法必须要变；因为思路与打法的转变，需要推敲正确的方向。元宇宙叠加入现实物理世界，速度会很快。

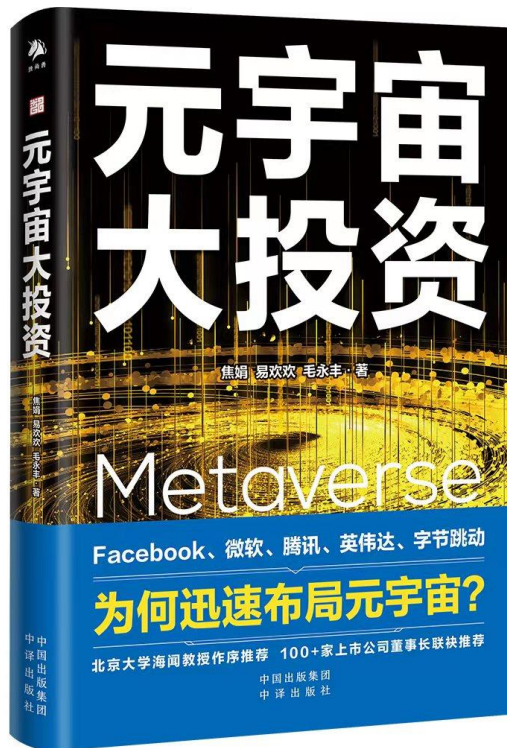
《元宇宙大投资》 2021年11月出版

- » 推演元宇宙上半场，元宇宙的建设阶段
- » 定义、元宇宙的本质/历史观/终局、六大组成部分、全球科技巨头布局元宇宙的资源禀赋
- » 元宇宙长什么样子？
- » 中美优势比较
- » 中国企业：受益于全球化时代背景的惯性思维与打法

《新硬件主义》 2022年5月出版

- » 刻画元宇宙下半场，成型的元宇宙去囊括现实物理世界
- » 怎样囊括？为什么要囊括？
- » 交互硬件50年的历史梳理；新硬件包括元宇宙上半场的通用型硬件入口（包括但不限于VR/AR、触觉手套）、分布式垂类硬件（基于情感需求投射的垂类硬件、机器人、现实物理世界中未来会被重塑的“物”）
- » 垂类硬件与AI：合之为一、衍之为万
- » 全国统一大市场这一时代背景下，需要切换思路与打法

元宇宙全貌 《元宇宙大投资》 《新硬件主义》



元宇宙：囊括现实物理世界、数字化everything的虚拟集合
数字化人的更多感官体验，由移动互联网的二维升至三维

背后的驱动力量：人的需求、科技的需求

新硬件主义：中国科技的打法必须要变
因为思路与打法的转变，需要推敲正确的方向
元宇宙叠加入现实物理世界，速度会很快



新硬件主义：全球24家科技巨头

		资源禀赋 & 正在布局				未来走向				显现
全球科技巨头		通用型 硬件入口	AI			分布式 垂类硬件	AI			
			内容	应用	场景		内容	应用	场景	
VR AR	苹果	✓	✓	✓	✓	★	✓	✓	✓	
	Meta	✓	✓	✓	✓	★	✓	✓	✓	
	字节跳动	✓	✓	✓	✓	★	✓	✓	✓	
	腾讯	★	✓	✓	✓	★	✓	✓	✓	
	HTC	✓				★				
智能手机	三星	✓				★				
	小米	✓				★				
	OPPO	✓				★				
	华为	✓	✓	✓	✓	★	✓	✓	✓	
游戏机	任天堂	✓								
	索尼	✓								
	微软	✓	✓	✓	✓	★	✓	✓	✓	
智能电车	特斯拉	✓	✓	✓	✓	★	✓	✓	✓	
	蔚、小、理	✓								
AI	阿里巴巴		✓	✓	✓		✓	✓	✓	
	亚马逊		✓	✓	✓		✓	✓	✓	
	谷歌	✓	✓	✓	✓	★	✓	✓	✓	
	百度	✓	✓	✓	✓	★	✓	✓	✓	
芯片	英伟达		✓	✓	✓		✓	✓	✓	
	高通		✓	✓	✓		✓	✓	✓	
	英特尔		✓	✓	✓		✓	✓	✓	
	联发科		✓	✓	✓		✓	✓	✓	

现实物理世界
映射入元宇宙

元宇宙反向
映射入现实物理世界

注：✓表示已有布局，★表示未来有希望布局

1. 元宇宙的入局方可以是六大框架中的每一个企业，目前均处于起跑线附近；
2. 美国与中国各有优势，通用型硬件入口非常关键，元宇宙的后半场——以AI为内核、以新硬件为显现的分布式垂类硬件，国内的全面赶超，在路径上，可以有更多确定性的选项，硬件入口不能有短板，但分布式垂类硬件天生比拼的是长板应用；
3. 我们信奉“德不孤必有邻”，但海外科技巨头的强势思维叠加资本市场的杠杆效应，其高举高打更为流利顺畅，从微观角度，A收购一家优质标的，作为竞争对手的B就少了一个落子之处且多了一个竞争对手；
4. 元宇宙的前半场是二维移动互联网升维至元宇宙，难度系数高；但元宇宙的后半场是三维元宇宙反向叠加现实物理世界，难度系数确定性会降低，全球科技巨头这一总决战，酣战至中场，利益格局的划分有可能就已接近于终局。

分布式垂类硬件，中国胜率与赔率更大

元宇宙上半场，中国优势在“内容与场景”、“协同方”

- 美国在硬件与操作系统、底层架构、后端基建、人工智能方面优势突出
- 中国在“内容与场景”、“协同方”两部位的优势突出
- 中国未来在后端基建、人工智能两部位，2024年起预计将有市值公司跑出来
 - » 现实物理世界并非一定需要人工智能，但元宇宙一定需要
 - » 后端基建与人工智能，必须等底层架构的升级

硬件入口

硬件入口包括但不限于VR/AR、触觉手套等：

- » 中国在智能交互硬件方面，智能手机时代是跟随策略，受益于全球化大背景
- » 硬件入口的命门在交互（沉浸式的迭代架构已相对清晰），当下元宇宙的硬件入口方向上，中国胜率及赔率一般
- » 垂类硬件的命门在AI（人工智能），中国头部科技公司可以在此方向走自己的路，大无畏、有担当；在全国统一大市场时代背景下，更能团结合作
- » 未来的垂类硬件，合之为—（AI）、衍之为万（诸多分布式垂类硬件）
- » 垂类硬件的衍生，基于元宇宙中的诸多内容、应用、场景，这是中国擅长的

防疫机器人



递送机器人



送餐机器人



咖啡机器人



调酒机器人



炒菜机器人



AI在元宇宙全貌中的递进式发展脉络



AI的
递进

01

人工智能在元宇宙上半场

元宇宙的六大组成部分之一，是元宇宙中的关键生产要素
存在于硬件、内容等诸多环节，由感知逐渐走向认知

2023年下半年乃至2024年才有AI在元宇宙领域的明显进展（2021年底的预判）

02

人工智能在元宇宙建设成型时（中场）

AI在元宇宙中，以供给的方式辅助于人的决策，挖掘人的更多需求/影响人的认知
移动互联网时代的“供给决定需求”在此阶段，将发挥得淋漓尽致

关键节点：由认知走向决策、AI通用技术平台成型、AI真正独立成产品

03

人工智能在元宇宙下半场

叠加入现实物理世界，以更好的满足人的需求/影响人的认知
背后是人的需求、科技的需求，共同驱动的

AI在物理世界显现为各式垂类硬件，包括新智能交互硬件、机器人、旧“物”改造
机器人包括虚拟数字人在现实世界“显现”（虚拟数字人原本即是AI驱动）

古印度的“唯识学”能提供一个特别视角

真正的智能严格来讲要有自我意识

目前人工智能大部分的工作只是在建构它捕捉外部世界的能力——不管是一台机器，一个人形机器人还是一个机器盒子，业内做的大量工作都是帮助它们构建捕捉外部世界的能力，严格来讲这只能算是对一个复杂问题有求解的能力

科技的发展在进入第三个1000年时发生了本质性的变化：第一次调头、针对人自己，我们用科技不再对外而是对自己的，这是人类历史上、科技史上、技术史上从没有过的——表现出的最大特点就是意识问题

唯识学认为，一个智能有四个层面，且四个层面是相互嵌套的（一个是另一个的基础，层层嵌套），经典人工智能，是从这四个层面的“最表层”开始去探索，并未整体去考虑智能本身（更意识不到智能有四层、这四层是互相嵌套生成的），而是过于工具化的去想怎么技术上去解决，即经典人工智能认为“智能”就是人的大脑、是对数据/符号的抽象化处理

从哲学角度，人工智能最棘手的问题，是它看上去充满未来感和高科技，但实际上是建立在早已过时了至少两百年的哲学体系上



能实现自学习、自编程，才能叫做人工智能

感知

- 人工智能最大的标志是自我意识，而自我意识这件事情是东方最擅长的，不管是禅定这种最古老的方式还是“唯识学”细密的、反身对自己意识的理解、剖析
- 感知不是第一代人工智能讲的纯粹的、被动的接收器，真正的感知是很复杂的，有意识混合的
- 西方对智能的理解，在哲学以及实践上其实已经开始知道要有智能、意识、有能力处理符号系统、人工神经网络，还是要有“身体”。我们的智能不在大脑里，而是遍布全身的，如果没有身体、对外的感知、运动，是没有办法做智能的

身体

- 马斯克的人形机器人目前看错解了身体，智能结构里的身体不是有四肢就可以解决的，它的身体背后一定有意愿跟它混合，不是单纯的人形的、有接收的感应器的硬件，不是一个物理装置

智能最底层

- 阿赖耶识，智能本身就是一股创造世界的力量而已；对阿赖耶识来说，以机器为平台来运行算法还是以人为平台来运行算法都无所谓



研究
思路

01

2023版元宇宙研究框架、元宇宙的上半场与下半场

02

2023：硬件的“大”年

03

硬件之“沉浸感”、“交互范式”、“操作系统”

04

元宇宙之百舸争流

- 虚拟人：交互世界的新物种；NFT：数字世界的魔法棒
- Web 3：内核是“治理”，“去中心化”实则为“反抗有理”
- AI与后端基建：2023年迎来景气上行
- 场景：2024-2025年的“八面来风”

为何2023年是“大年”



判断2023年为硬件的“大”年，综合了周期、全局、结构的考量

结构



认知**结构**上的“**预期差**”

- ① “交互范式的定义”对硬件入口的重要性
- ② 通用硬件设备的交互范式定义出来后，AI加持下的场景/应用/模式甚至是内容的配套会极快

全局



领航全局的“**犄角之势**”

- ① Meta & Microsoft VS Apple，开源与封闭体系
- ② 其他厂商则跟随策略

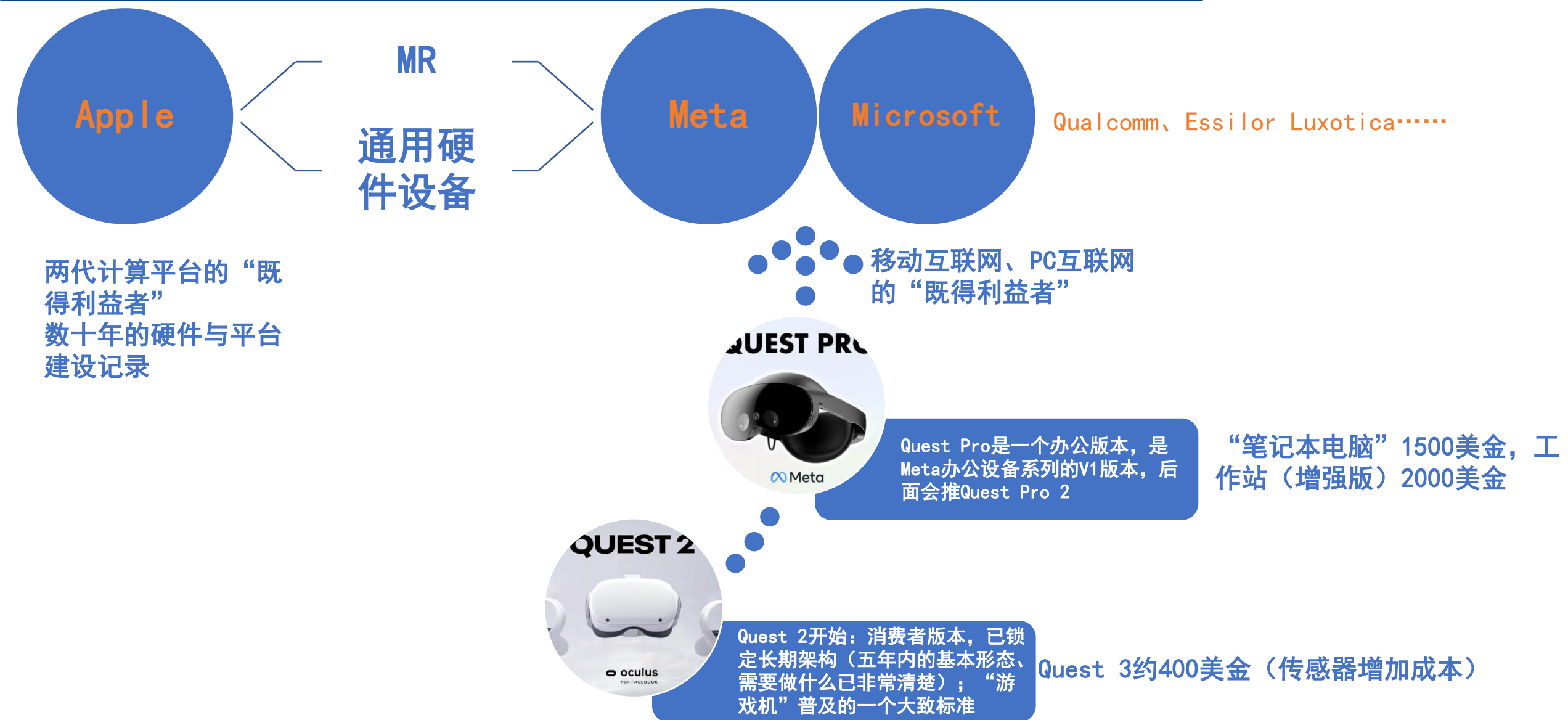
周期



大周期的**助力**

- ① 多周期共振出清，底部并购潮，以补足资源禀赋上的短板
- ② 巨头补跌，“第二曲线”的明牌化

为何2023年是“大年”



为何2023年是“大年”

正式对打
通用设备

Pro是通用计算设备的V1版；与微软一起构建开放的元宇宙和下一代平台的开放版本，且确保开放生态系统在**下一轮计算中获胜**；通用计算设备，需要一个接一个地开辟用例，游戏、社交、健身、办公；2023年，通用计算平台中的一个重大问题是：**输入是什么？**神经接口这样的输入方式（神经界面、肌电图腕带等）

交互/跨屏

用户将在一个有着一大堆这些平台（电脑手机VR）的世界里生活，**交互时如何顺利跨屏将是23年的亮点之一？**

形成标准
商业模式

Teams**本质上就像一个浏览器**；Teams、Microsoft 365、Intune Device Management、Azure Active Directory、Windows Streaming、Xbox；HoloLens构建的应用将会带到包括Quest这样的所有平台之中（强大的ISV（独立软件供应商））

Horizon更像是**元宇宙的Web端**（而不是应用商店），更像一个独立、本地应用程序

谁在投资？谁拥有任何计算形态中**最大的装机量？**

每一个主要的计算平台都有一个注重合作的开放生态系统，以及一个更加封闭和集成的生态系统：PC的Windows与Mac；手机的Android与iOS

入局且**真正形成标准**需要提前3-5年出发；最好的硬件，基本上以收支平衡点销售硬件甚至可以略有亏损，发展生态系统，商业模式通过软件和服务获得收入

为何2023年是“大年”

	硬件入口	后端基建	底层架构	人工智能	内容与场景	协同方
Meta	✓		✓	✓	✓	
字节跳动	✓		✓		✓	
Unity		✓	✓		✓	
Roblox		✓	✓		✓	
Epic Games			✓		✓	
Nvidia	✓	✓	✓			
高通	✓		✓		✓	
Sony	✓				✓	
网易					✓	✓
腾讯		✓	✓		✓	✓
Amazon		✓			✓	
阿里巴巴		✓	✓		✓	
Microsoft	✓	✓	✓		✓	
Google	✓	✓		✓		
百度		✓		✓	✓	
Apple	✓		✓	✓	✓	
华为	✓	✓	✓			
小米	✓				✓	
HTC	✓				✓	
Valve	✓				✓	
以太坊			✓		✓	



元宇宙是一种技术路径，各家则在实施各自的**工程方案**，背后：

- 1) 既有的资源禀赋、收购以拼图；
- 2) 23年极为需要“元宇宙”这第二曲线。

围绕沉浸式的各工程方案，未来五年的迭代框架已相对清晰，23年将围绕“交互范式的定义”，定义交互是海外巨头的强项（有成功经验、所需资源禀赋具足，即便不足也能大手笔收购以补足），也将开启硬件与内容的正反馈；

美股23年或将是多周期共振出清的节点，与2001-2002年相比是全新的水域全新的风险，大巨头若补跌，当下的业务结构撑不起持续快速增长的预期，需要“第二曲线”对抗周期动荡的持续性与连锁反应；且底部利于并购重组：

- 1) 美股是强势思维，杠杆更流畅（估值）；
- 2) 元宇宙的前半场是二维移动互联网升维至元宇宙，难度系数高；但后半场是三维元宇宙反向叠加现实物理世界，难度系数降低，中场的利益格局的划分有可能就已接近于终局。

国内：虚拟现实22-26年五年规划

- 最大的意义是政策层面对元宇宙的技术本质（虚拟现实）“盖章”认可
- 未来5年的指导思想和方针，对技术要求、场景应用进行全面规划

3500亿产业规模，2500万硬件销量：

- 1) 硬件700亿左右
- 2) 工业生产500亿（to B与to G）
- 3) 内容场景（包括NFT、游戏）700亿
- 4) 融合媒体（VR直播、线上直播等）300-400亿
- 5) K12应用700亿
- 6) 体育赛事&健康100亿左右
- 7) 商贸创意（其他细分应用）在200亿
- 8) 其他100亿（智慧城市）

关键点：

- 产业发展战略窗口期已然形成
- 培育100家具有较强创新能力和行业影响力的骨干企业
- 10个具有区域影响力、引领虚拟现实生态发展的集聚区
- 推进关键技术融合创新：围绕近眼显示、渲染处理、感知交互、网络传输、内容生产、压缩编码、安全可信等关键细分领域
- 工业和信息化部、教育部、文化和旅游部、国家广播电视总局、国家体育总局：所覆盖的所有场景

价值链拆分：3500亿/2500万台，1.4万/台的价值链

3000元的单价，**2500元是硬件成本**
(镜片、芯片、摄像头等供应商)

1.1万是内容，内容、服务（应用场景）、模式（会员、道具），内容领域空间巨大

Pico: 市占率70%

华为: VR与AR, 平台及软件工具

OPPO & VIVO: AR眼镜

小米、TCL、创维、联想等





研究
思路

01

2023版元宇宙研究框架、元宇宙的上半场与下半场

02

2023：硬件的“大”年

03

硬件之“沉浸感”、“交互范式”、“操作系统”

04

元宇宙之百舸争流

- 虚拟人：交互世界的新物种；NFT：数字世界的魔法棒
- Web 3：内核是“治理”，“去中心化”实则为“反抗有理”
- AI与后端基建：2023年迎来景气上行
- 场景：2024-2025年的“八面来风”

＜ PART 01 ＞

硬件之“沉浸感”

围绕沉浸感的技术架构，国内外均已成型

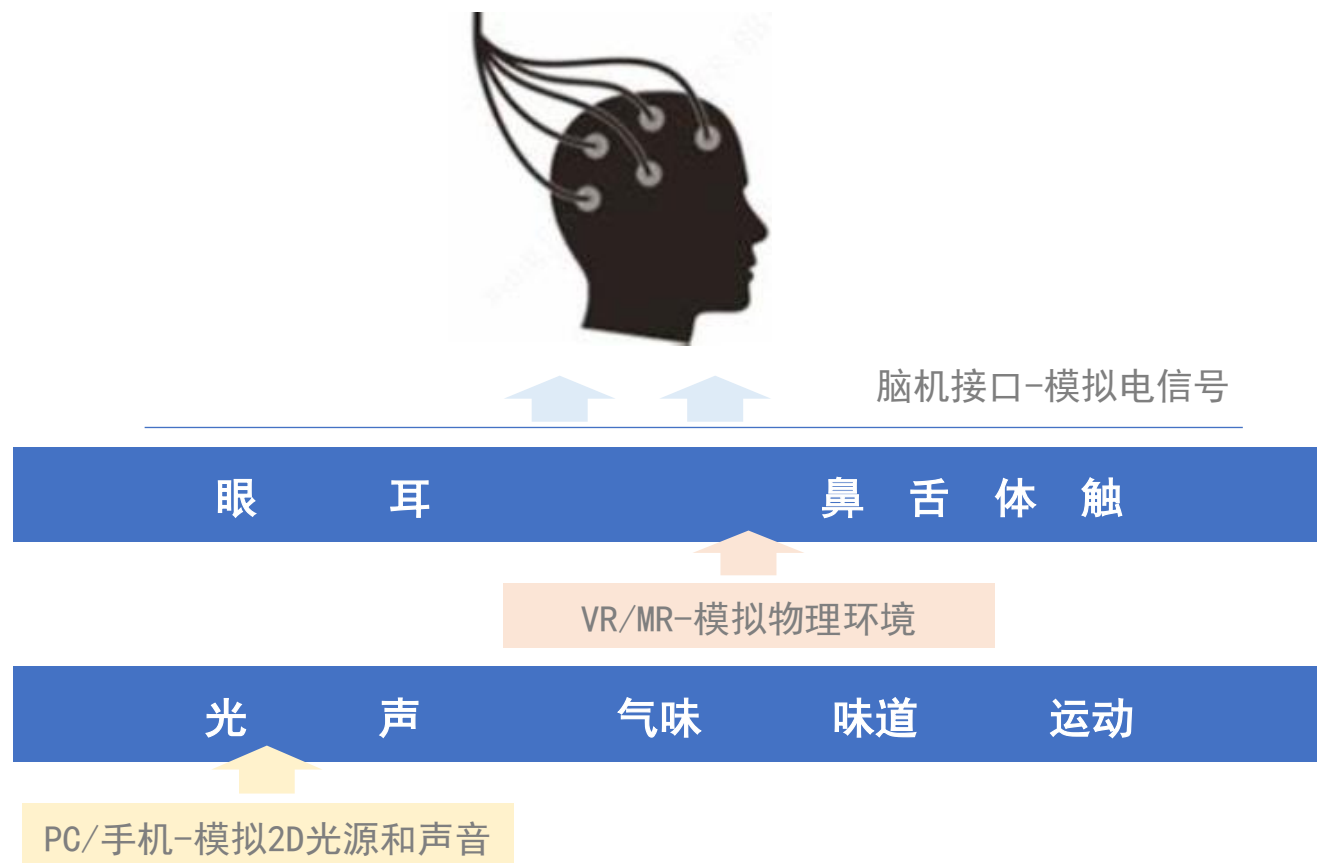
围绕沉浸感的各技术指标

相对VR的沉浸感，AR核心在于融合感

产业链关键环节——光学、显示、交互、芯片

沉浸感是区分互联网与元宇宙的最直观体验

- VR通过模拟物理环境——光、声音、气味、味道、运动，并以模拟出的物理环境来影响人的眼、耳、鼻、舌、体感、触觉，由此产生化学激素和电信号来影响到人。这也就是VR让人们产生“沉浸感”的心理基础。
- 手机和PC之所以无法产生“沉浸感”就是因为手机和PC仅模拟了光源和声音，无法像VR一样模拟整个物理环境。



元宇宙/全真互联网

- ✓ 以VR/AR/MR为载体，手势、语音交互
- ✓ 实时性、所见即所得、全息投影
- ✓ 沉浸感高

移动互联网

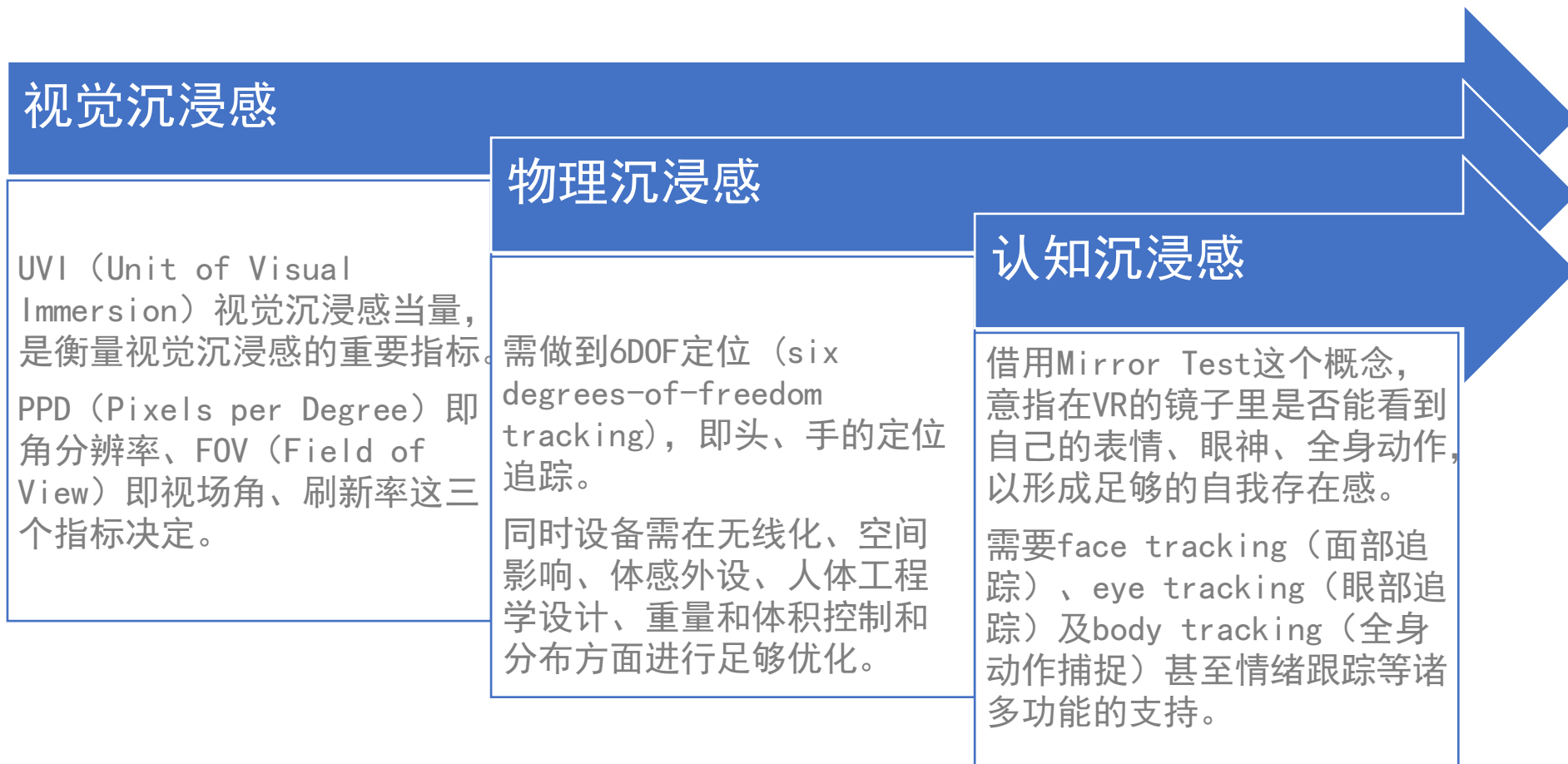
- ✓ 以智能手机为载体，触控、语音交互
- ✓ 便携性、低实时性、小屏幕
- ✓ 沉浸感提升

PC互联网

- ✓ 以PC为载体，键鼠交互
- ✓ 固定性、匿名性、非实时性、大屏幕
- ✓ 沉浸感低

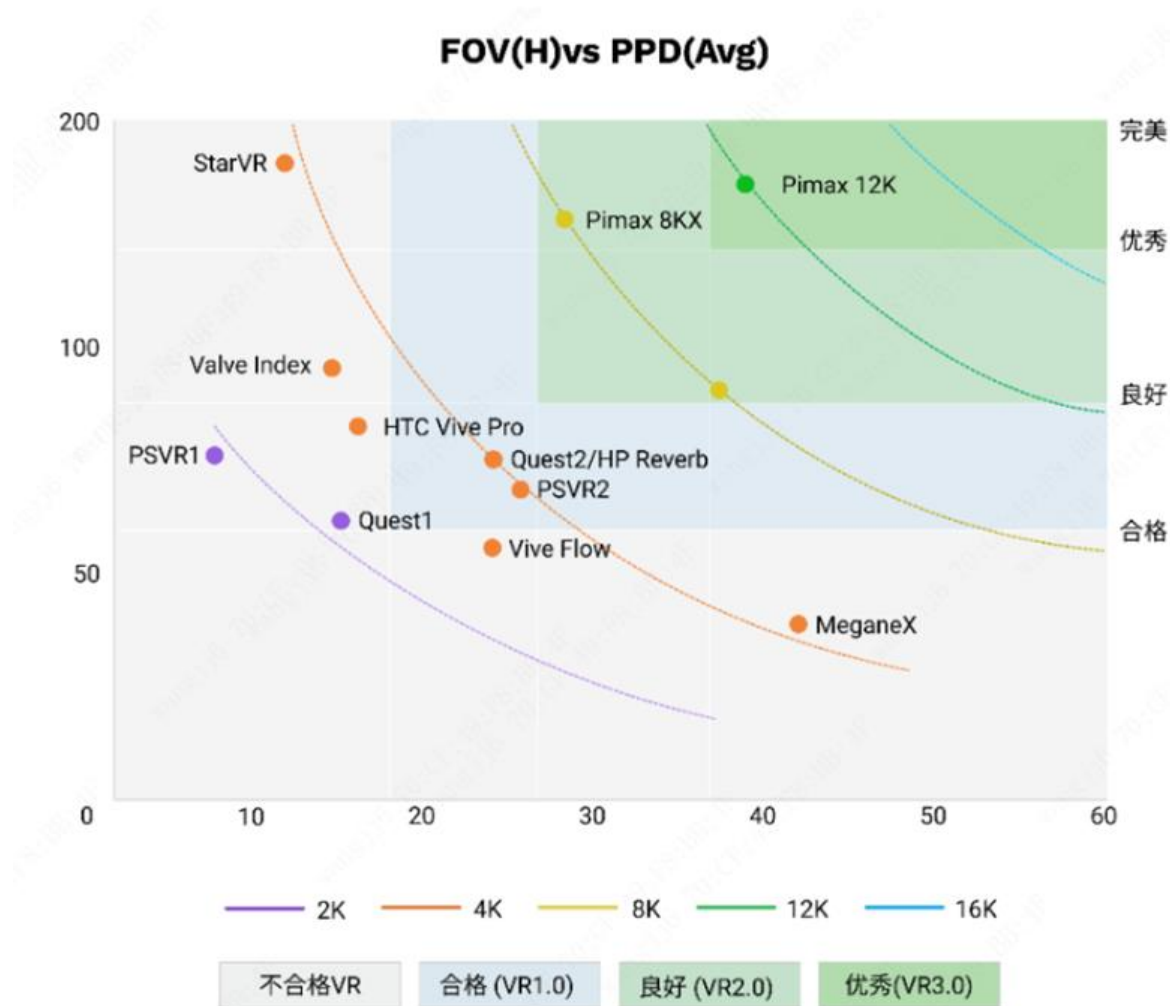
沉浸感的三个维度：视觉、物理、认知

- 沉浸感是支撑元宇宙概念成立的核心要素之一，可以借助高性能VR硬件获得。
- 国内VR厂商小派科技Pimax将这一系列技术提升的目标归纳为VR 3.0的技术标准，即在视觉沉浸感、物理沉浸感、认知沉浸感三方面技术达标。



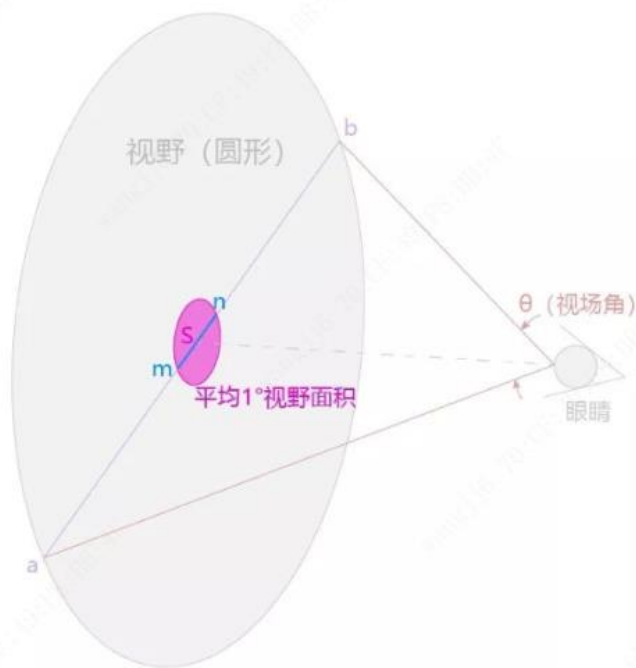
视觉沉浸感：UVI当量

- UVI (Unit of Visual Immersion) 即视觉沉浸感当量，是衡量视觉沉浸感的重要指标。通常认为PPD (Pixels per Degree) 即**角分辨率**、FOV (Field of View) 即**视场角**、**刷新率**这三个指标都分别存在绝对的天花板水平：PPD=60代表了视网膜水平；FOV=220度代表了人眼自然视场角；刷新率大于180Hz人眼基本无法分辨。
- 任何一款VR头显的这三个指标分别除以该指标的天花板水平（也就是100%水平）得到一个百分比，就代表该头显在该项指标上的表现。这个三个百分比的加权调和平均数就是UVI。
- 我们认为这个模型能够很好地量度视觉沉浸感水平。100%的UVI就意味着人眼已经完全无法分辨看到的是头显中的图像还是真实的世界。这个标准听上去很科幻，但是实际上离我们不再遥远。



UVI 量化指标——PPD

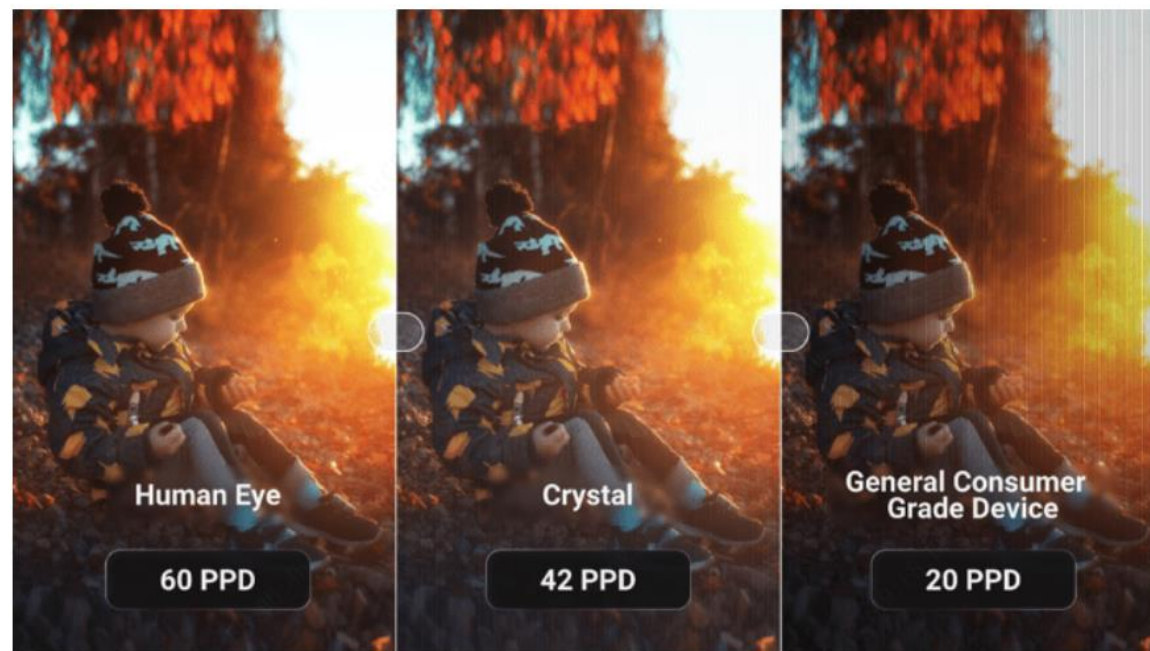
- PPD (Pixels Per Degree) 角分辨率，指视场角中的平均每 1° 夹角内填充的像素点的数量。
- 计算公式—— $PPD = \text{单眼成像外接圆的直径像素点数量} \div \text{单眼成像的对角线的视场角 FOV}$
- 人眼正常视力下的分辨能力极限为50-60PPD。当PPD小于50时，人眼正常视力下的能够感受到头显产品的像素晶格感，PPD越小，晶格感越强烈，即“纱窗效应”。头显产品的PPD越接近60，产品的成像清晰度就越接近人眼的分辨极限，人眼看图像就觉得越清晰。



$\angle \theta$: 视场角 FOV
线段ab : 视野直径填充的像素点数
粉色圆S : 平均 1° 视野面积
线段mn : PPD (平均 1° 视野直径像素点数)

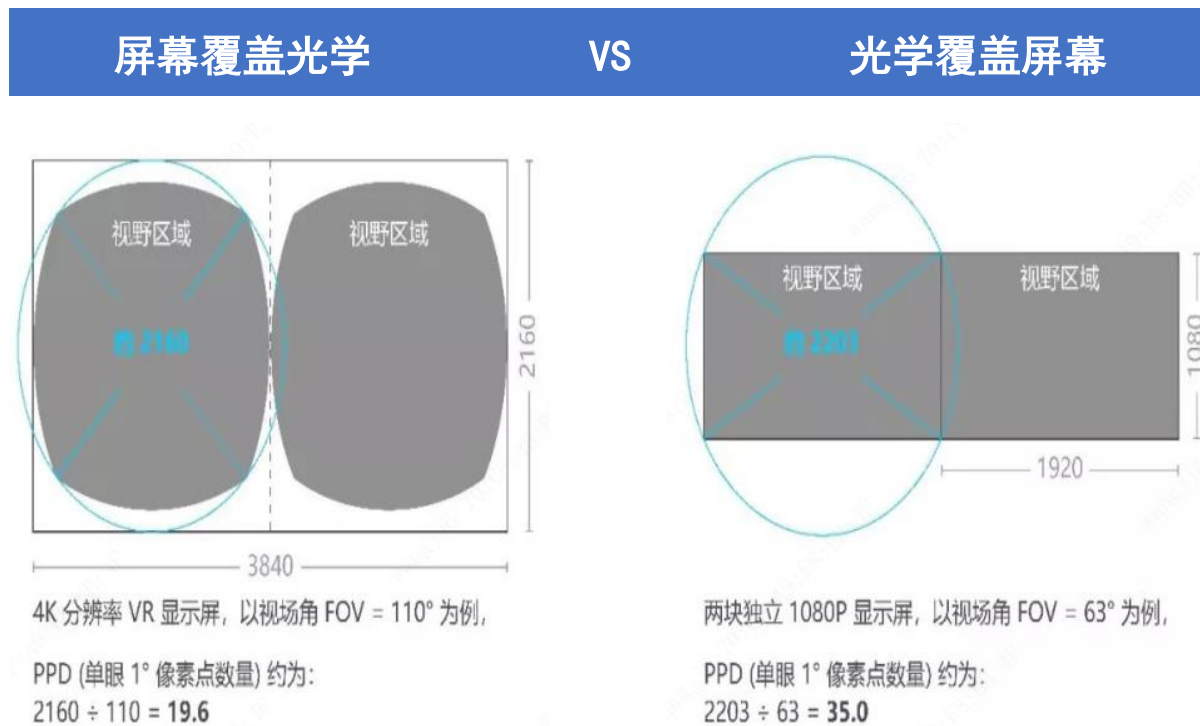
$$PPD = \text{线段ab} \div \angle \theta$$

$$S = \pi (PPD \div 2)^2$$



屏幕分辨率与光学系统共同作用影响 PPD

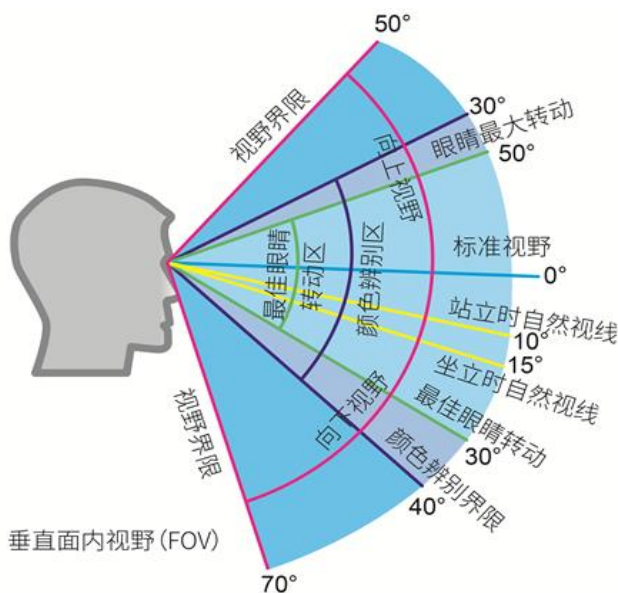
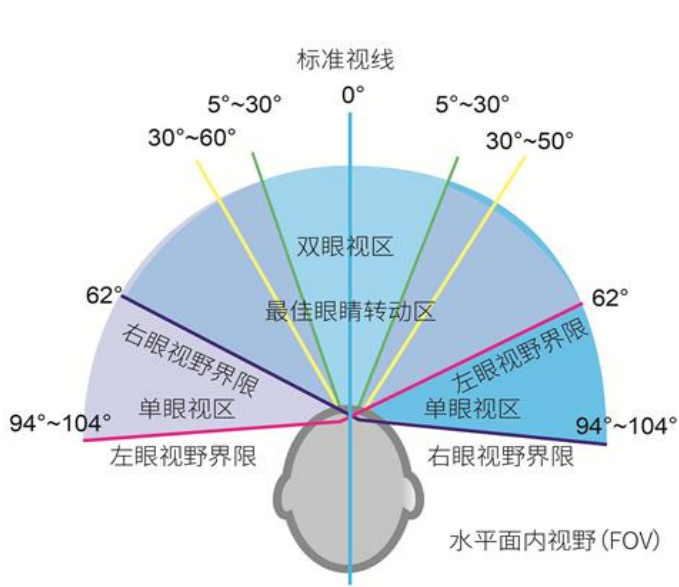
- 由PPD的计算公式可知，PPD由总像素数和视场角共同决定。其中，屏幕分辨率决定像素数，是清晰度的决定因素之一；光学系统决定视场角，对放大图像的清晰度影响很关键。
- 头显产品有两种光学成像系统：屏幕覆盖光学、光学覆盖屏幕。屏幕覆盖光学成像系统的视场角可以做得更大，但是会浪费了一些“屏幕边角料”，也就是产生“无效像素点”，实际上被有效利用的像素点数量比屏幕分辨率少一些。光学覆盖屏幕类型的头显产品，相对来说视场角有所减小，但能够实现屏幕像素点的全部利用。



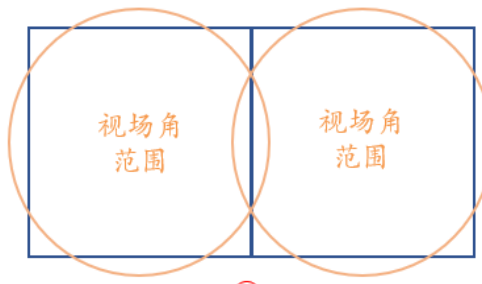
- ✓ 屏幕覆盖光学的头显：3840*2160 分辨率（4K 单屏）、视场角 FOV = 110°，——计算单眼PPD为19.6
- ✓ 光学覆盖屏幕的头显：1920*1080 分辨率 *2（1080P 双屏）、视场角 FOV = 63°，——计算PPD为35.0
- ✓ 屏幕分辨率相似的情况下，光学系统对PPD有直接影响

UVI 量化指标——FOV

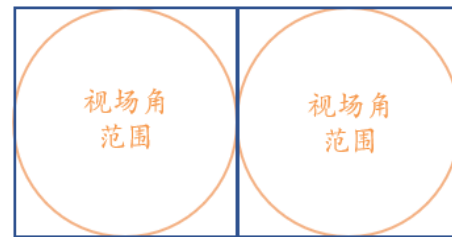
- 视场角（FOV, Field of View），即可见区域或可见角度，代表了人眼所能看到的空间区域范围。在计算机相关显示系统中，视场角就是显示器边缘与观察点（眼睛）连线的夹角。
- 视场角在VR中的作用主要是与沉浸感密切相关，一般来说，视场角越大、越不容易产生眩晕感，沉浸感也就会越强；视场角小意味着看到的影像小，对分辨率的要求低。对于VR设备而言，其水平视场角至少要达到90度左右，才能够保证用户获得较高的沉浸感。目前市场上主流VR设备的视场角范围在90度-110度之间。
- 视场角并非越大越好，以屏幕大小为例，合适的视场角要搭配合适尺寸的屏幕才能展现出最佳的显示效果。



①



②

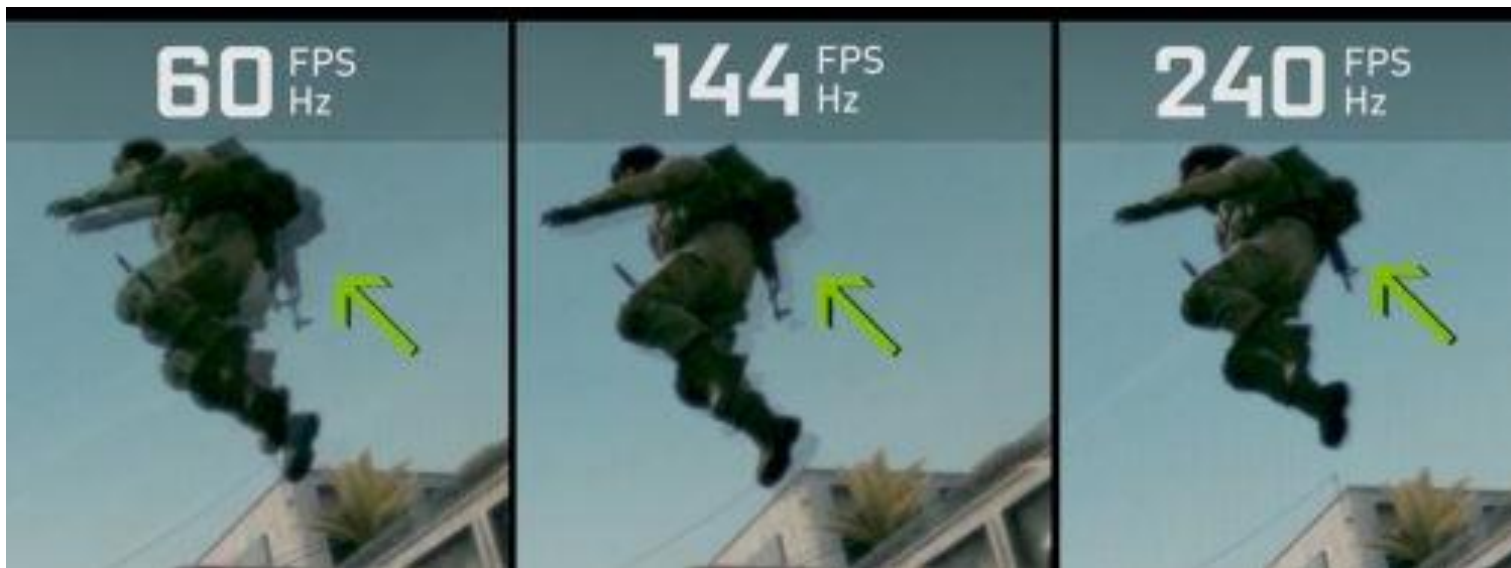


③

假设屏幕尺寸固定，在不同的视场角下：
①若视场角过小，导致看不全屏幕
②若视场角过大，会看到屏幕之外的黑边
③视场角正合适

UVI 量化指标——刷新率和帧数

- 刷新率，是指显卡将显示信号输出刷新的速度，60Hz就是每秒钟显卡向显示器输出60次信号。刷新率是由显示器决定的。
- 帧数，是指画面改变的速度，60fps就是每秒钟显卡生成60张画面图片（理论上，每一帧都是不同的画面）。帧数FPS是由显卡所决定的。
- **刷新率与帧数的关系：**在VR体验中，离开帧数只谈提高刷新率没有意义。假设帧数是刷新率的1/2，意味着显卡每两次向显示器输出的画面是用一幅画面。相反，如果帧数是刷新率的2倍，画面每改变两次，其中只有1次是被显卡发送并在显示器上显示的。即高于刷新率的帧数都是无效帧数，对画面效果没有任何提升，反而可能导致画面异常。
- **刷新率和帧数必须匹配，才可以获得流畅的VR体验。**当前主流的VR设备基本可以做到90HZ的刷新率，但是帧数受限于芯片性能和网络带宽基本在60fps左右。



物理沉浸感：定位、追踪、重量、体积

- 以VR硬件为例，足够好的物理沉浸感需做到头+手 6自由度定位，其中Inside-out定位方案是主流的发展方向，精确度足够高、价格便宜、使用便捷。
- 同时设备需在无线化、空间影响、体感外设、人体工程学设计、重量和体积控制和分布方面足够优秀。物理沉浸感是影响VR设备的用户时长与佩戴舒适度的重要因素。
- 此外，see-through功能已经形成全视野AR功能的雏形，为VR过渡到MR打下基础。
- 以旗舰VR设备PICO4、Quest pro为例，头手6DOF+Inside-out定位方案已成为主流，重量、体积相比上一代产品也显著优化。



- ✓ 头6DOF+手6DOF
- ✓ Inside-out定位方案，Pro版本支持眼动追踪
- ✓ 重量586g（外壳重量295g），重量减轻26.2%，厚度减少38.8%
- ✓ 支持全彩透视



- ✓ 头6DOF+手6DOF
- ✓ Inside-out定位方案，支持眼动/面部追踪
- ✓ 重量722g
- ✓ 支持全彩透视

Quest Pro

定位&追踪： Inside-Out降低部署成本

- Outside-in：使用外部定位模块+头部追踪器，来标定头盔位置。Outside-in方案的最大弊端在于，头盔的位移空间受制于外部定位器的放置位置，并且每组定位器都有其极限捕捉范围，并且易产生校准偏差，校准过程繁琐。典型产品是HTC Vive、Oculus Rift。
- Inside-out：使用头盔内的摄像头，来标定头盔位置。Inside-out 方案解决了空间限制，让用户可以更加专注在VR本身的体验上。典型产品是Oculus Quest 2。

	Outside-in外向内追踪技术	Inside-out内向追踪技术
原理	由外而内的定位：需要事先在环境中布置定位器，一般两个以上，实现从外到内的位置计算	由内而外的定位：不需要额外布置空间定位设备，借助VR设备自身的传感器进行环境的感知与事实位置计算
追踪精度	准确度较高	目前的准确度略低（算法依赖）
延迟	延迟相对低	延迟相对高
可移动范围	仅限于传感器监测范围	无空间限制，活动自由
成本	较高	较低
对算力的需求	较低	较高
事先环境部署	需要	不需要
外部传感器	需要	不需要
主流设备应用	HTV VIVE pro、Oculus Rift、Varjo VR等	Oculus Quest 2、HTC VIVE Cosmos、HTC VIVE Focus、Pico Neo等
总结	成熟但昂贵	主流方案

定位&追踪：头手6DOF是多元交互的基石

- 6DOF，就是指物体除了具备在X、Y、Z 三轴上旋转的能力之外，也具备在X、Y、Z三轴上移动的能力。6DOF的VR设备几乎可以模拟所有的头部动态。
- 头部交互：6DOF中，头盔增强了许多主动性的行为，如微动作（歪头、缩脖子、半身前倾等）、主体移动（可以走、跳、蹲甚至躺等）。用户的头部以及身体能够依靠与现实世界认知通用的反应来躲避对手的攻击。
- 手部交互：手柄往往是标配，支持移动的定位方式，6DOF中，常规手柄键位设计极为相似，主要划分为三大区域——拇指区域（摇杆+按键）的方式，食指区域（扳机），后三指区域（侧键）。

头部

+

手部

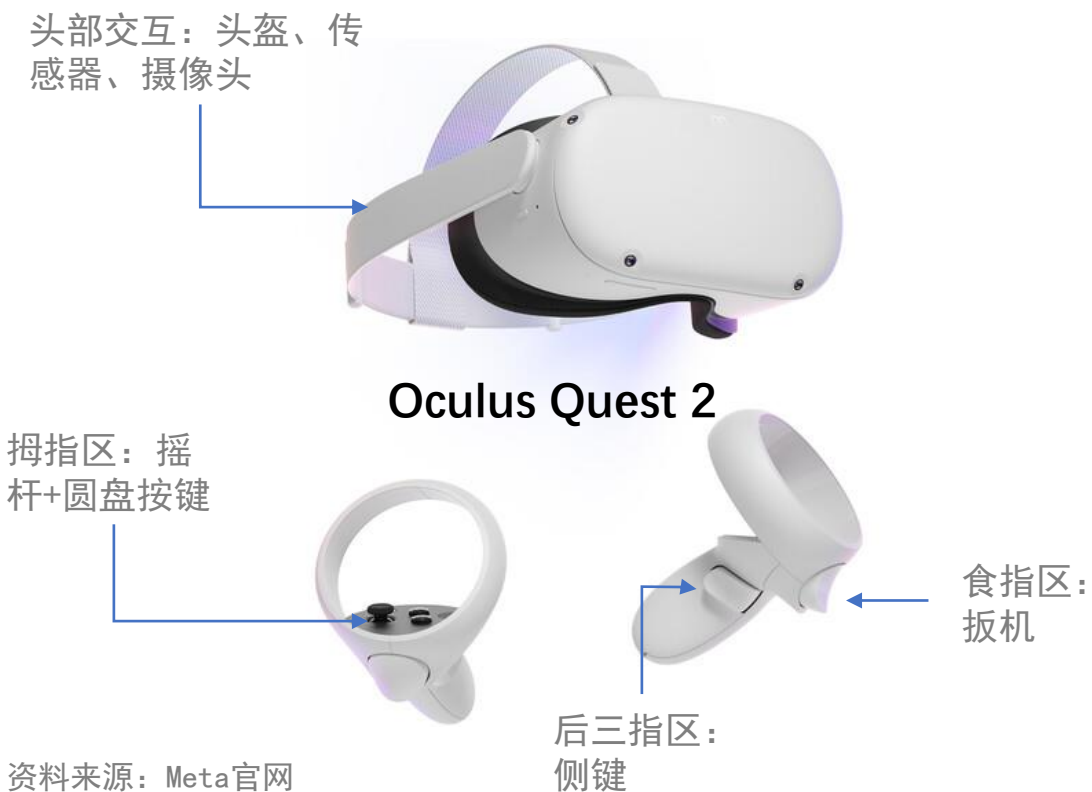
6DOF

交互

认知更匹配：通过标定高度、位置、微动作等空间移动信息，为用户提供更加真实的观察角度，便于用户将虚拟环境和自我感知绑定，增强沉浸感。

自由度更高：可移动的定位方式可以满足更多用户个性化需求，极大地丰富了探索VR世界的可能性。

交互更多元：头手位移象限的解锁，使传统交互中的浏览、选择、点击再也无法满足用户的探索欲望，握、拽、弹、砸，更多肢体语言被解放出来，人的潜意识行为更值得被关注。



重量&体积： Pancake光学方案助推VR轻薄化

- VR光学方案发展大致经过非球面透镜—菲涅尔透镜—折叠光路（Pancake）三大阶段，目前菲涅尔透镜是市场主流方案，Pancake方案是光学系统的重大创新，在VR头显轻薄化、小型化等方面起到重要推动作用，将成为VR光学的主要升级方向。

	非球面透镜	菲涅尔透镜	折叠光路Pancake
重量	600g	400-500g	200-300g
模组厚度	40-50mm	40-50mm	15-20mm
优势	- 成像质量较高 - 光路简单，光损较小 - 制造工艺成熟，制造成本低	- 相比非球面透镜更轻薄 - 视场角扩大 - 光路简单，光损较小 - 制造工艺成熟，制造成本低	- 焦距短，显著轻薄化 - 多片式透镜，支持屈光度调节 - 成像质量好，畸变小
劣势	- 焦距长，模组厚 - 不支持调节屈光度	- 焦距长，模组厚 - 不支持调节屈光度	- 制造工艺不完全成熟，制造成本高 - 多次折返，光损较大 - 易出现画面重影
代表产品	3Glasses D1(2014)，Oculus Rift(2015)，PSVR(2016)，爱奇艺奇遇2Pro(2019)，奇遇Dream(2021)	HTC Vive(2015)，DPVR P1(2018)，Pimax 5K XR(2018)，Oculus Quest 2(2020)，Pico Neo 3(2022)	华为VR Glass(2019)，3Glasses X1S(2020)，HTC Vive Flow(2021)，arpara 5K(2021)，创维PANCake 1(2022)，PICO 4(2022)，Quest Pro(2022)，苹果MR(2023)，PSVR2(2023)
单组价格	¥5-10	¥15-20	¥150-200

重量&体积： Pancake光学方案助推VR轻薄化

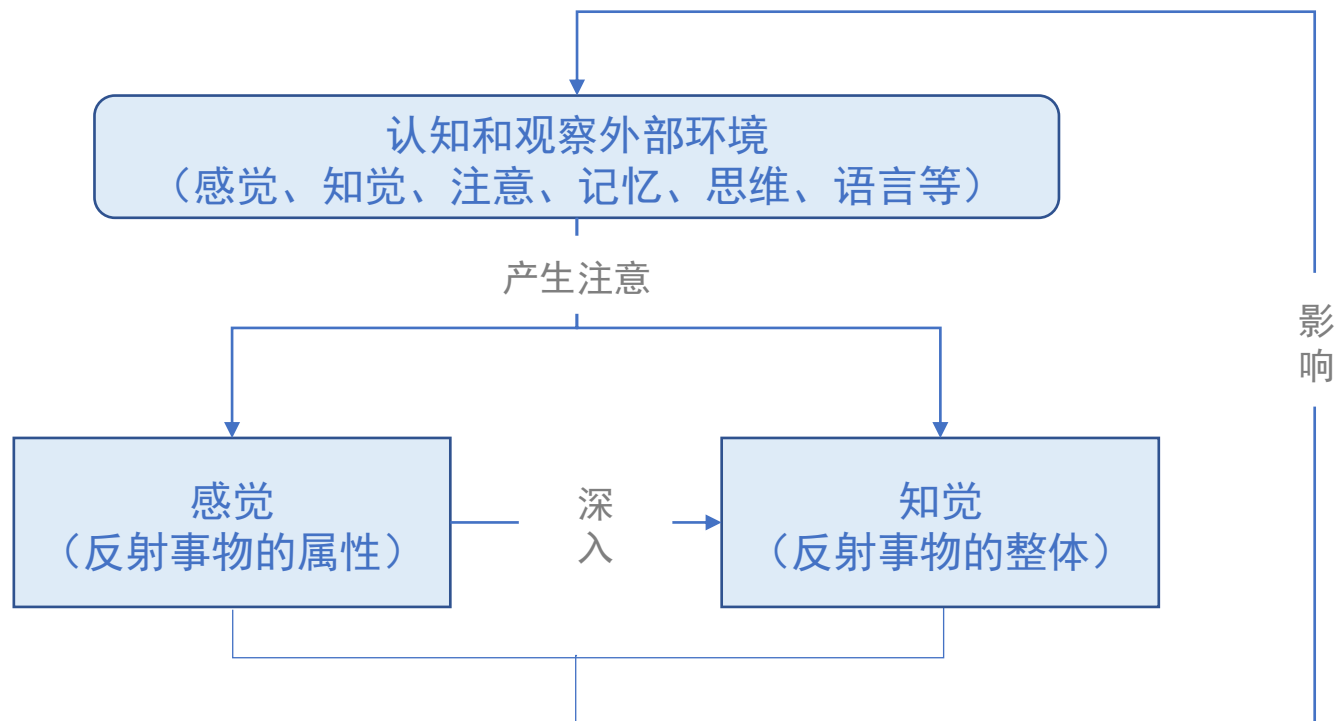
- Pancake方案在初期存在制造成本高、光学效率低等问题，仅在部分分体式VR 设备中采用。伴随供应链及众多终端品牌的协同技术调整，目前生产良率和使用效果已得到大幅改善，有望逐步替代菲涅尔方案。
- 近两年Pancake 方案已被arpara、HTC、YVR、创维等品牌应用于一体式VR，HTC Flow系列、PICO VR Glasses、3Glasses X系列、创维S6及Pancake 1、arpara、Huawei、Shifitall等头显设备均已采用Pancake方案，还包括行业龙头的新一代产品，如近期发布的PICO 4、Meta Quest Pro，后续将发布的苹果MR、索尼PS VR2等。



资料来源：WellSenn XR，安信证券研究中心

认知沉浸感：眼球追踪、面部追踪、虚拟化身

- 借用Mirror Test概念，意指在VR的镜子里是否能看到自己的表情、眼神、全身动作，以形成足够的自我存在感——认知。
- 只有在充分的认知沉浸感的支持下，才有可能在VR中形成真正有效的社交场景。认知沉浸感需要除了物理沉浸感要求的头手6DOF的跟踪定位之外，还需要face tracking（面部追踪）、eye tracking（眼部追踪）以及body tracking（全身动作捕捉）甚至情绪跟踪等诸多功能的支持。



眼球追踪



面部识别

眼球追踪：可应用范围广泛，逐渐成为标配

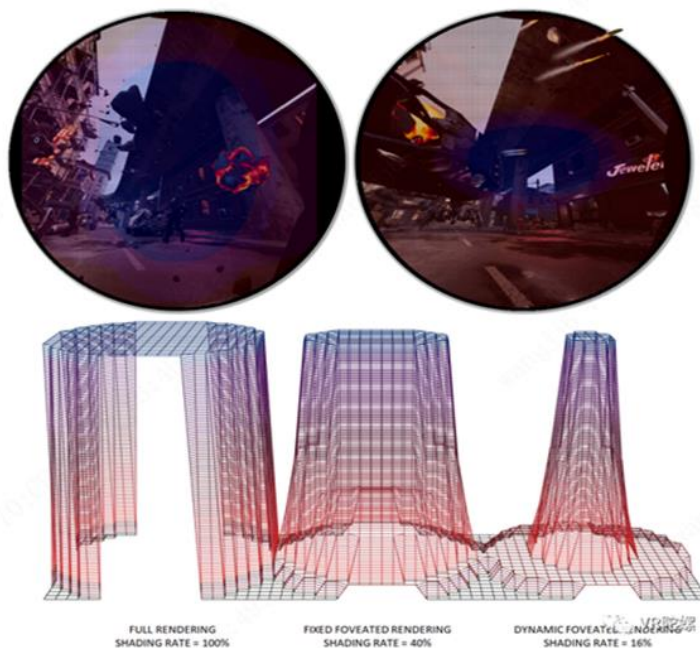
- 眼球追踪就是通过图像采集设备对眼球运动的信息进行捕捉，实现对眼球的跟踪和分析，从而服务于具体的场景当中。眼球追踪可应用于光学显示、交互控制、目标识别、身份验证、健康监测、社交和多人协作等方面，几乎覆盖AR/VR的所有环节。
- 从眼睛生理结构的特性出发，目前全球主流的智能眼镜公司的眼动追踪方法有五种：**瞳孔角膜反射法、视网膜影像、对眼睛建模后计算视觉中心、视网膜反射光强度、角膜的反射光强度**。其中第一种和第二种都是通过计算机对眼睛的图像进行处理、特征点提取，进而获得的眼睛视觉中心。第三种通过摄像头（红外相机、深度相机）对眼睛重建为三维模型再计算视觉中心。第四种和第五种仅仅可以通过一个或者几个光敏传感器元件捕获眼睛反射光的强度获取眼睛视觉中心，反射光可能来自角膜中心或者视网膜。

分类	技术原理	设备/器件	代表企业
瞳孔角膜反射法	通过角膜中心和瞳孔中心的连线进行眼动追踪。	红外摄像头（2+）、红外光源，如LED红外传感器（2+）	Magic Leap one 、Digilens、Tobii 、SMI、Apple、索尼。
视网膜影像	因为视网膜上不规则的毛血管、中央凹等生理结构形成的图案，通过计算视网膜图像变化来进行眼动追踪。	光波导光学器件、全息波导	Magic leap（第二代或者第三代）、lumuc、三星。
对眼睛建模后计算视觉中心	1. 利用结构光对眼睛建模；2. 利用光场相机对眼睛建模	结构化的发光器、光场摄像头（2+）	Facebook。
视网膜反射光强度	激光透过角膜、瞳孔、晶状体最终打在视网膜上，视网膜会对入射的激光反射，外部光敏传感器通过检测视网膜反射光强度确定眼球运动的方向。	光敏传感器、红外光发射器、透镜	Facebook、 Lumuc。
角膜的反射光强度	当激光刚好射在角膜中心时(角膜中心是最高点)，激光的反射强度最大，因此通过光敏传感器检测角膜反射光的强度确定眼球运动的方向。	光敏传感器、红外光发射器、MEMS反射镜	微软、华为、 Adhawk。

资料来源：《AR/VR行业兵家必争之地——眼动追踪技术大全》，天蝎科技，映维网，安信证券研究中心

眼球追踪带来两大关键突破， 注视点渲染&自动变焦

- **注视点渲染：**通过眼球追踪来识别用户在VR中的注视点，从而动态调节VR屏幕的清晰度，注视点区域最清晰，而余光部分则降低清晰度。注视点渲染方案有助于降低硬件压力，同时提升VR视觉清晰度。



动态注视点渲染相比固定注视点渲染，等降低2倍左右的GPU着色负载，能更有效地维持高帧率，优化计算资源。

- **自动变焦：**VR光学变焦主要有机械式变焦和电子液晶透镜变焦两种，机械式变焦通过在Pancake模组加入移动式机械变焦系统（音圈致动器和弯曲铰链阵列），配合眼球追踪来移动透镜的位置实现变焦实现无极可变焦调节。



机械式可变焦显示器采用精密齿轮传动模组来调节VR/AR眼睛的两组透镜组件的间距，直至与使用者的瞳距相匹配。

面部追踪：现阶段与Avatar化身配合使用

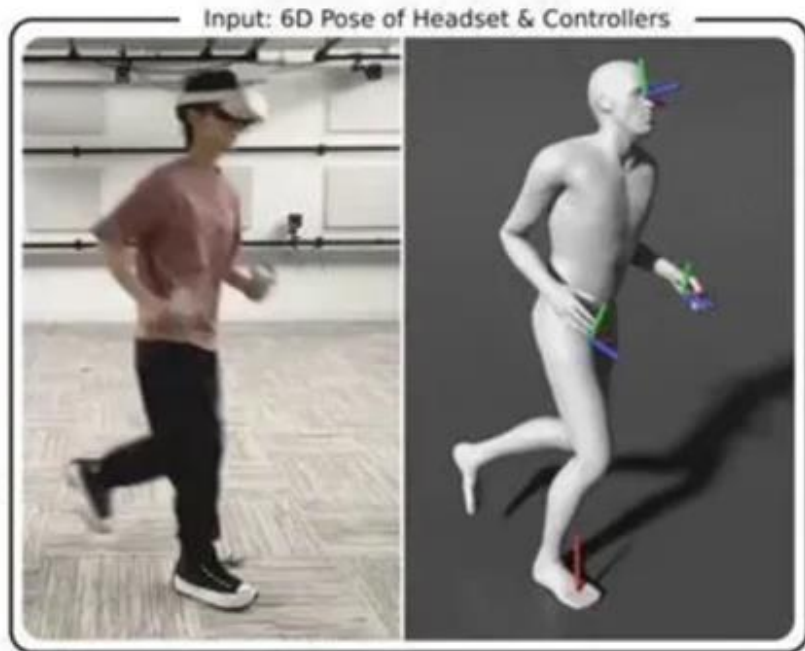
- 面部追踪是动作捕捉技术的一部分，指使用机械装置、相机等设备记录人类面部表情和动作，将之转换为一系列参数数据的过程。与捕捉由关节点构成、较为稳定的人体动作相比，面部表情更为细微复杂，因此对数据精度要求更高。
- Meta Quest Pro搭载了全新的面部追踪系统，该系统能通过内部摄像头感知用户眼睛以及脸部某些部位的运动情况，结合免校准的机器学习模型，头显能将摄像头所看到的内容实时驱动Avatar的面部动作。



- ✓ 据新浪VR，Meta 已发布了 Aura 面部追踪示例供开发人员进行试验，以及与“运动 SDK”相关的开发人员示例，这些示例为开发人员创建具有眼睛、面部和身体追踪的富有表现力的化身奠定了基础。
- ✓ 在短期内，面部追踪系统将与 Meta 现有的Avatar化身一起使用。

全身动捕：高度依赖AI算法，普及尚有距离

- 常见的全身追踪技术主要包括光学动作捕捉技术、惯性动作捕捉技术以及视觉动作捕捉技术三种。**光学动作捕捉技术**：操作的时候会直接在人的身体上进行简单的标记，标记点会直接反射到提前设定好的摄像机上，然后再通过反射的不同位置的成像信息来预算标记点的空间运动信息，最终将信息进行简单地定位以及输出。**惯性动作捕捉技术**：会直接在人的身上佩戴陀螺仪，人在运动的时候，陀螺仪也会跟着进行旋转。此时，直接通过感知陀螺仪的旋转信息将人的运动推算出，然后实现动作捕捉。**视觉动作捕捉技术**：在操作的时候是不需要标记和佩戴设备的，只要在人的活动范围内通过普通的摄像头进行动作的录制，将人体关键信息进行识别，然后采用特殊AI算法实现动作捕捉。
- 全身追踪在VR中普及还有一段距离，由于技术的成本和使用门槛对新手并不算友好。



- ✓ 据映维网，Meta提出了一种仅通过Quest实现全身动捕的解决方案。对于上半身追踪，通过在AI训练过程中获得的经验，仅需来自现实世界的少量输入就足以将双手准确地转换到虚拟世界。

虚拟化身：元宇宙社交与互动的关键一步

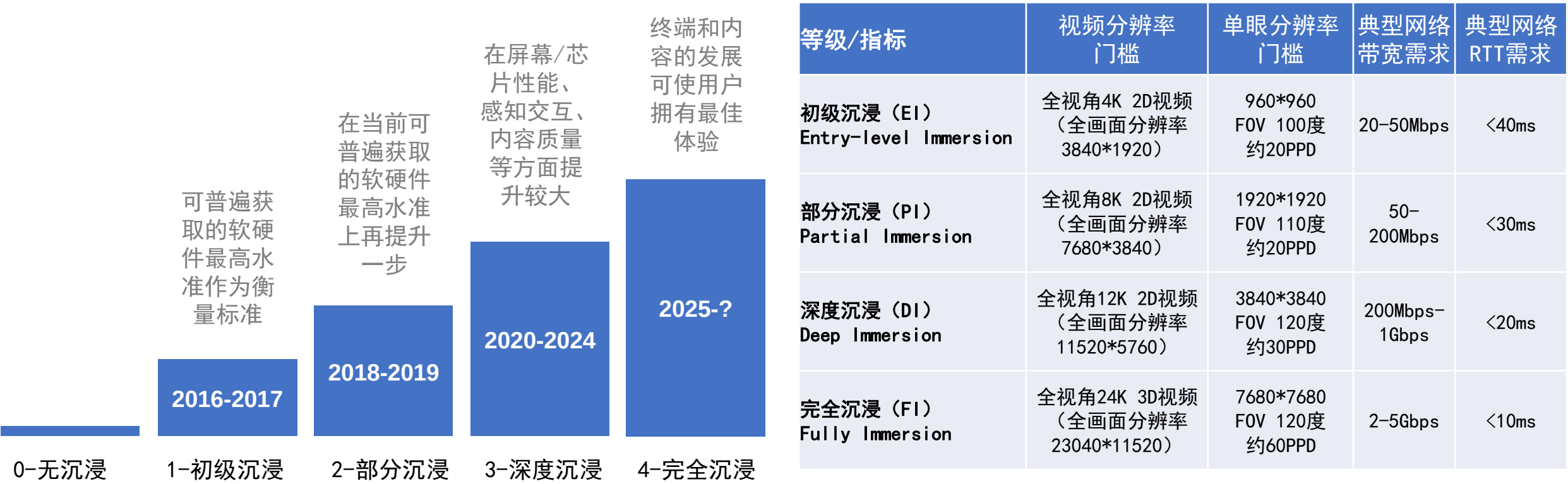
- 虚拟化身，是打造元宇宙沉浸感不可或缺的一部分。虚拟化身在元宇宙生活，成为模糊现实和虚拟世界界限的重要突破口。
- 区分于社交头像，虚拟化身决定了人类个体在虚拟世界的独一无二，且能够实现脸部五官、情绪表情、手势姿态的变化，提升互动感和真实感。
- Meta一直致力于打造元宇宙社交体系，包括社交平台Horizon Worlds、社交形象Meta Avatar等。眼动追踪与面部追踪技术的普及有利于元宇宙社交的推进。



- ✓ 据青亭网，Meta Avatar旨在提升VR头像的表现力、个性化功能（号称有超过1018种搭配），并且支持更多VR应用。
- ✓ 将采用基于机器学习的动作预测算法，进一步提升虚拟头像运动的流畅性，可自然模拟挥手、微表情、注视点、语音驱动的面部表情等效果。
- ✓ Meta Avatar全身版的进展很慢，目前只能在第三人称视角中为Avatar配备虚拟腿。

当前处于部分沉浸向深度沉浸的过渡阶段

- 沉浸体验的提升有赖于相关技术的突破和进步，是分阶段演进的过程，不同发展阶段对应不同体验需求。
- 部分沉浸：在当前可普遍获取的软硬件最高水准上再提升一步；深度沉浸：在屏幕/芯片性能、感知交互、内容质量等方面提升较大。
- 当前虚拟（增强）现实的发展阶段已走过部分沉浸期，处于深度沉浸期。



资料来源：中国信通院《虚拟（增强）现实白皮书》，安信证券研究中心

注：我们认为进入完全沉浸时间点由2023年后移至2025年

当前处于部分沉浸向深度沉浸的过渡阶段

- 虚拟（增强）现实的底层技术主要包括近眼显示、网络传输、感知交互、内容制作等，按照各技术所达参数基本可判断元宇宙当前处于部分沉浸到深度沉浸的过渡阶段。

体验层级		初级沉浸	部分沉浸	深度沉浸	完全沉浸
 近眼显示	分辨率	接近1K	1. 5K-2K	3K-4K	≥8K
	角分辨率PPD	≤15	15-20	30左右	60左右
	视场角FOV	90-100°	100-120°	140°	200°
	刷新率	60HZ	75HZ	90HZ	≥120HZ
 网络传输	码率（Mbps）--弱交互	≥40	≥90	≥290/≥160	≥1090/≥580
	码率（Mbps）--强交互	≥40	≥90	≥360	≥440
	MTP时延（ms）	20	20	20	20
 感知交互	追踪定位	Outside-in	Inside-out	Inside-out	Inside-out
	眼动追踪	无	无	有	有
	面部识别	无	无	有	有
	全身动捕	无	无	有	有
	触觉反馈	无	无	有	有
 内容制作	渲染计算	2K/60 FPS	4K/90 FPS	8K/120 FPS	16K/240 FPS
	渲染优化	--	--	注视点渲染	注视点渲染
	虚拟化身	无	无	虚拟化身	精细化虚拟化身

沉浸感的理想状态——高度模拟现实物理世界

- 早期沉浸：VR游戏《半衰期：爱莉克斯》中，玩家需要在场景中翻箱倒柜寻找关键道具，可与游戏内几乎任何物体进行交互。
- 深度沉浸：《头号玩家》中，玩家通过VR及可穿戴式设备进入游戏，体验真实互动。通过体感衣，玩家可以感受到身体所受攻击的痛感；通过全自动触觉椅，玩家可以体验到游戏中的坠落飞行等体感，通过多款设备采集玩家信息并向玩家实时输出反馈信息，玩家在虚拟空间中的映射感更真实，从而获得身临其境式体验。
- 终极形态：通过脑机接口技术，实现嗅觉、味觉等感知体验，同时与虚拟世界自由交互，显著提升拟真体验与沉浸感。

VR游戏《半衰期：爱莉克斯》



仅手部交互

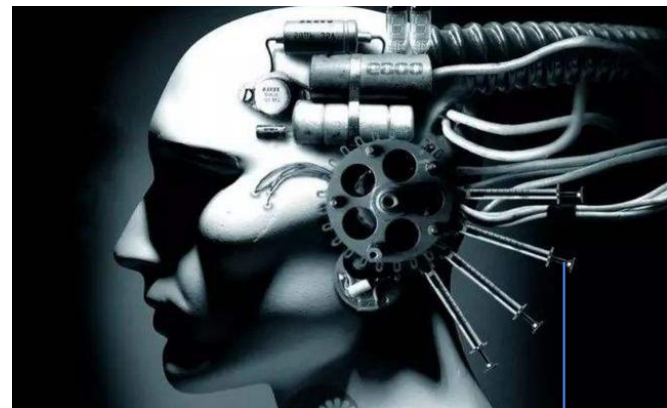
VR世界《头号玩家》



头戴式设备

体感衣

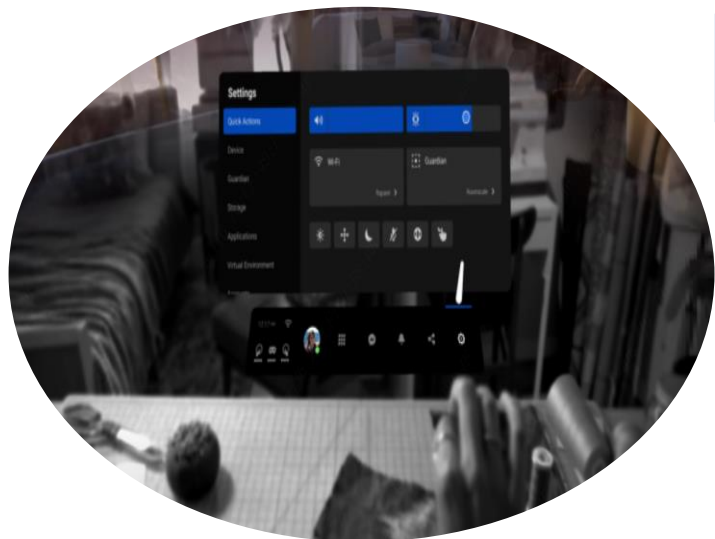
终极形态：脑机接口



五感
电信号，化学信号

全彩透视正打破物理壁垒，真正“虚实融合”

- VR设备上支持的多数为黑白透视。黑白透视只需要利用空间定位摄像头以及一些简单的算法，就可以实现不错的透视效果。
- 全彩透视相比黑白透视，背后则又需要考虑多个技术的融合。全彩透视主要可以分为：透视、空间锚点（放置AR物体）、三维重建、环境理解及对象识别。

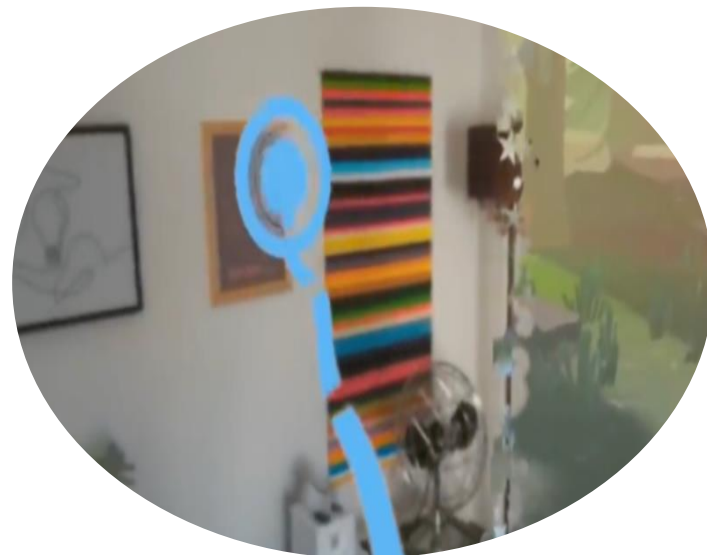


黑白透视场景

全彩透视对 VR 设备的硬件提出更高要求，技术难点主要在于摄像头和算法。

✓ 摄像头：黑白透视下，搭配四广角黑白摄像头方案；全彩透视下，将采用 3 颗全彩广角摄像头（2 颗全局快门摄像头，1 颗普通全局变焦镜头）+1 颗 TOF 深度摄像头，。

✓ 算法：全彩透视对画面的分辨率、清晰度、即时性要求提高，需要一套能够实时处理外界 RGB 全彩信息的算法。



全彩透视场景

与VR沉浸感→MR并行的，是AR的贴合感

- VR (Virtual Reality)，虚拟现实（VR），又称为计算机模拟现实。它是一种通过人机交互和计算机生成的三维模拟所创造的体验。我们可以通过使用虚拟现实设备，例如，耳机、控制器，在环境中进行交互。
- AR (Augmented Reality)，增强现实（AR），是一种实时的、直接或间接的对现实世界物理环境的观察。它将真实环境中看到的内容与计算机软件生成的数字内容相融合，以某种方式增强真实环境。
- MR (Mixed Reality)，混合现实（MR），是现实世界与虚拟世界的融合，以产生新的环境和可视化。MR根据环境渲染3D物品并融入所见的环境中。这些3D内容往往会遵从现实世界的大小，并且遵守物理规律，仿佛是真实存在的物件。
- 沉浸程度排序：VR>MR>AR；与现实世界的交互程度：AR>MR>VR

VR



MR



AR



沉浸感

现实世界

贴合感的理想状态：实时贴合于时空与节奏

- AR，增强现实，其目标是通过AR技术，参与者可以与虚拟对象进行实时互动，从而获得不一样的视觉体验，且它能够突破空间、时间以及其它客观限制。

AR将广泛应用于各种场景

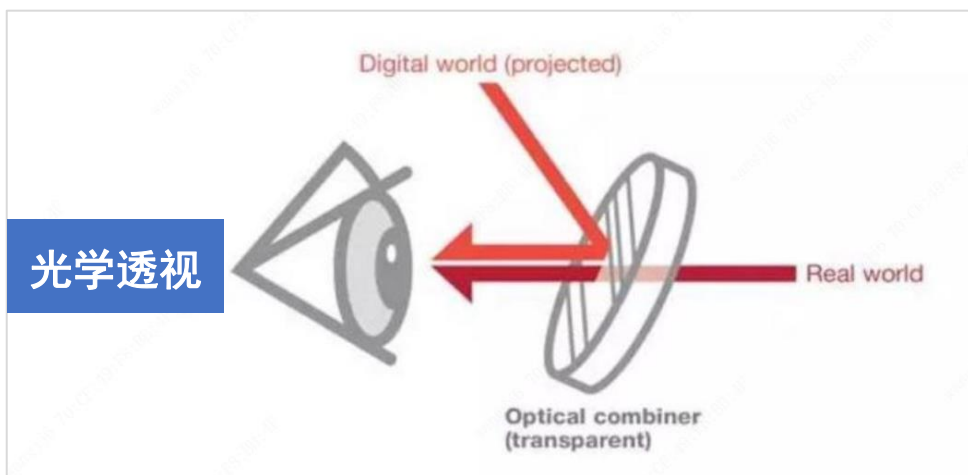
1. 场景是争夺用户时长与可支配收入的产品与服务的组合，to C的场景如此，to B/G的场景亦然；
 2. 从供给的角度，技术（虚拟现实还是人机协同/共生）决定操作系统，模式（to C还是to B/G）决定操作路径，要素与场景则以产品或服务形式共同作用于人性（顺人性还是逆人性）；
 3. 基于技术与模式的任何产品与服务，底层逻辑均是作用于用户的“人性”，“顺人性”与“逆人性”是此消彼长的两大方向。顺人性更多的是情感需求，逆人性更多是利益需求；
 4. 元宇宙是重构时空，场景是如何切割新时空；
 5. 场景的本质之一，是各类新的时空，包括时间与空间两个维度；但从用户体验的角度，趋势上，时间是征服空间；
 6. 基于抢用户时长、可支配收入的考量，场景一定是有自身独特节奏的，节奏是一种基于体验的抽象概念，具体分析起来，作用于人的时空观——折叠空间（也包括反向折叠）、给时间加杠杆（也包括去杠杆）。
- AR 技术作用于场景，即通过扫描现实场景，让虚拟物体能够自然地 and 现实场景的平面贴合；
 - 贴合感的理想状态，即实时的，贴合于时空与节奏。

AR在2022年有一定发展，但变化幅度不大

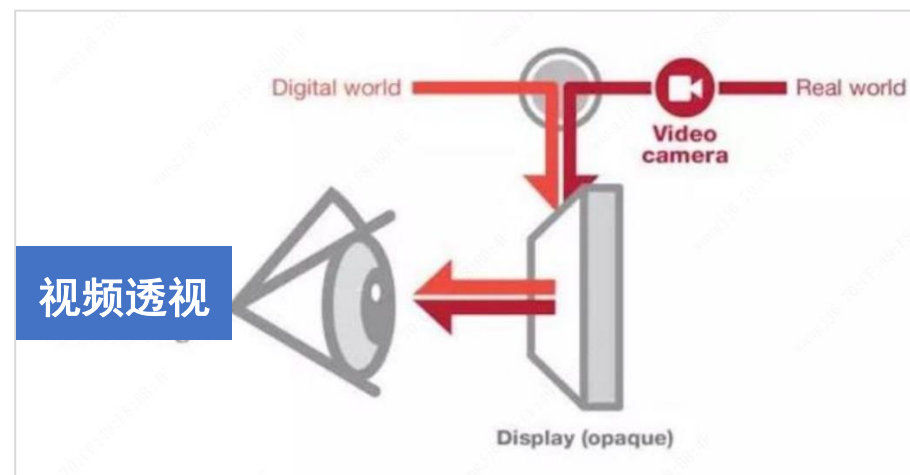
- VR与AR两者诸多技术互通。VR硬件目前基本达到消费级硬件标准，下一步聚焦于升级硬件性能及丰富内容生态。AR光学系统不同于VR，尚存技术难点待攻克，且轻量化要求与性能矛盾更大，目前仍处于硬件发展早期阶段。
- VR：VR产业链核心环节趋于成熟，硬件产品体验大幅提升。光学显示器件Fast-LCD + Pancake是目前主流方案，芯片/处理器由高通芯片占据统治地位，追踪定位环节Inside-Out+头手6Dof功能日趋完善，未来向手势、眼球识别方案发展。VR内容及应用逐步发力，其中游戏内容生态已形成爆款游戏驱动用户增长—用户反哺游戏内容丰富的良性循环。2022-2025年为VR硬件性能爬升期，VR头显将增加多样化功能并增强性能以提升用户体验，2025年有望达到硬件成熟期。预计2023年，随着Meta、苹果、索尼等众多重磅VR/MR头显的发布，推动全行业技术升级和出货量持续提升。
- AR：AR发展阶段晚于VR2-3年，光学与显示部分器件量产仍存难点，尚未推出相对成熟能大规模放量的C端产品。当前最佳技术路径已锁定Micro LED+光波导，Micro LED目前技术难点在于巨量转移技术，高性能光波导的量产尚未形成高性价比方案，未来海兹定律有望推动成本持续下探。我们预计2025年前后，苹果和Meta预计将推出真正的AR眼镜，新一代硬件面世后有望带动产业链快速成长。
- 但AR具备相比VR更大的市场潜力。AR具备虚实融合、赋能现实的特性，可适用于大多数 B 端和 C 端场景，AR眼镜相比VR一体机更加轻量化，可定位为未来的生产力工具和计算平台（与智能手机同一等级）。AR眼镜硬件发展初期预计将以手机配件形式发行，可类比 TWS 耳机和智能手表等可穿戴设备；未来一体机成熟后，预计将取代手机，拥有十亿级出货量的广阔市场空间。

区分AR与MR：“现实”的来源不同

- MR与AR看起来都是在真实环境中进行，但现实环境的来源是不同的。AR（增强现实）的“现实”指的是真实的物理环境。MR（混合现实）的“现实”是通过摄像头捕捉来的画面信息。
- 相应地，See-Through实现方式也主要有两种：光学透视（Optical See-Through, OST）和视频透视（Video See-Through, VST）。光学透视(OST)：真实世界是通过放置在用户眼前的半透明光学合成器看到的。光学合成器也被用来将计算机生成的图像反射到用户的眼睛里，从而将真实世界和虚拟世界结合起来。视频透视(VST)：真实世界是通过相机捕捉到实时视图，然后与计算机图像技术结合在一起，呈现在不透明的显示器上。使用视频透视，算法可以完全控制视觉集成，允许虚拟和真实物体之间的完全遮挡，甚至可以对真实物体进行更高级别的修改。



VS



XR均是基于视觉这一技术路径的不同工程方案

- 视觉只是技术路径之一，这一技术路径上有不同的工作方案，走向是充分开放的。



资料来源: 小派科技《VR/AR为基础的MR将成为下一代计算平台?》

2022年VR跟踪之PICO



硬件产品



PICO 4



Pico Neo 3/3 Pro



Pico G2 4K

软件产品

游戏串流助手

- 在VR一体机串流PC玩SteamVR游戏

PICO VR助手

- 辅助PICO VR的手机App

PICO OS 5.0

- 基于Android定制开发，专为VR一体机设计的操作系统

远程播放助手

- 通过网络将PC视频传输至VR一体机

内容生态

VR游戏

VR视频

VR直播

VR社交

VR健身

...

PICO 4产品亮点

自研Pancake光学透镜

- ✓ 前端重量减少26.2%，厚度减少38.8%

4K+超视感屏幕

- ✓ 搭载2块2.56英寸的Fast-LCD屏幕，整体分辨率4320X2160，1200PPI，20.6PPD，90HZ刷新率



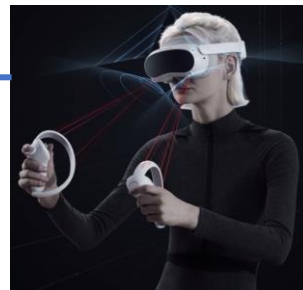
F0V105°

- ✓ 105° 超广视野，上一代F0V为98°



亚毫米级定位精度、毫秒级追踪速度

- ✓ 搭载自研高精度四目环境追踪和红外光学定位系统，光学传感器进一步提升，以实现更好的追踪定位效果。



无级电动瞳距调节

- ✓ 支持62-70毫米无级瞳距调节，瞳距适配更精细



PICO 影视：VR+电影破局，引入三体、灵笼重磅IP

- 目前，国内VR影视内容的制作产能体系尚未成熟，PICO的影视内容多采用外部合作方式。
- **引入3D电影大片：**PICO同迪士尼、索尼影视、环球、派拉蒙及狮门联手打造“PICO 3D大片重燃计划”，引入漫威《复仇者联盟》系列、《X战警》系列、《星球大战》系列以及《神偷奶爸》、《加勒比海盗》、《银翼杀手2049》等100多部3D经典影片。
- **试水VR短片：**PICO与国内拥有最多精品VR内容版权代理权之一的沙核科技达成深度合作，引入多部优质短片。PICO还开启了“超沉浸观影计划”，《苍穹》、《纸鹤》、《水母》等6DoF作品已在PICO视频上线。
- 据PICO发布会，三体版权方三体宇宙已与PICO达成合作，将共同制作首个VR版本《三体》互动叙事作品，预计将在2023年上线；PICO也成为头部国创动画《灵笼》的VR动画番外篇的联合出品，预计将于2023年5月上线。

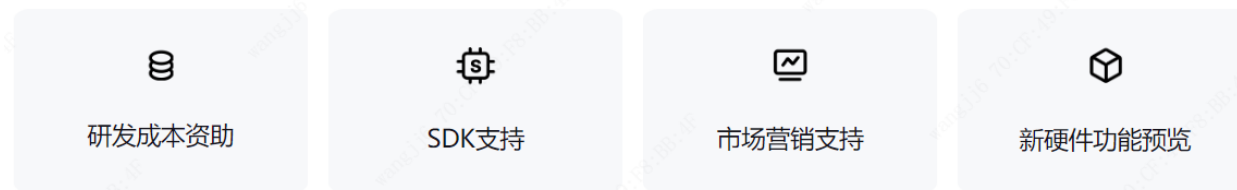


PICO 游戏：不乏TOP级VR大作，开发者生态尚未建成

- **引入TOP级VR大作：**据IDC统计，截至2021年Q2，Pico Store已有类似《亚利桑那阳光（Arizona Sunshine®）》、《MultiBrush》等优质6DoF游戏及应用160余款，其中《SUPERHOT VR》《幻音骑士（Synth Riders）》《雇佣战士（CONTRACTORS）》等TOP级VR大作，目前国内一体机平台均仅上线Pico。
- **本地化运营：**Pico还充分发挥本地优势，积极为海外引进的优质游戏进行本地化适配，针对《雇佣战士（CONTRACTORS）》《超能军团（Hyper Dash）》《乒乓：致胜11分（Eleven Table Tennis）》等多人联机类游戏，Pico会为其在国内自建服务器以保证流畅运行，并进行本地化玩家社群运营。
- **开发者生态：**目前尚未看到Pico公开层面的内容筛选策略或规则，相比Quest和开发者共建生态，Pico在开发者生态仍然处于起步期。



PICO 开发者计划



PICO 直播：明星、虚拟偶像、才艺主播三箭齐发

- PICO的直播内容形式主要包括明星演唱会、虚拟直播和秀场直播。
- 演唱会-借用明星效应出圈引流：2022年至今，PICO共举办了王晰、郑钧、汪峰3场明星直播演唱会，1场A-SOUL虚拟偶像演唱会，VR线上演唱会这一形式已走进大众视野。
- 虚拟主播-探索虚拟偶像技术边界：区别于其他VR直播形式，虚拟直播的主播形象均为虚拟建模，主播通过动作捕捉技术实现互动。
- 才艺主播-套用抖音成熟模式：抖音早期直播依靠才艺直播起家，积累丰富主播资源及运营经验。
- 赛事转播：PICO全程直播2022卡塔尔世界杯比赛。用户可通过超清巨幕观看世界杯赛事直播，还能邀请好友进入互动场实时语音畅聊，享受真正沉浸的观赛体验。



PICO 社交：从0到1阶段，主打虚拟形象+PICO轻世界

安信证券
ESSENCE SECURITIES

- PICO作为国内VR社交的开创者，已成立专门的VR社交中心。据36氪，波粒子科技创始人、前小米VR负责人马杰思任职Pico社交中心负责人。北京波粒子科技有限公司专注于二次元虚拟社交，2022年被字节跳动收购后原有团队50余人也整体并入Pico社交中心。
- PICO虚拟数字人+PICO轻世界，对标Meta Horizon Worlds+Avatar。PICO的VR社交布局当前处于从0到1的阶段，主打虚拟数字人+PICO轻世界，围绕用户在VR世界当中的形象与活动空间发力，路径类似Meta的Horizon Worlds+Avatar。



三款团队协作游戏

- ✓ 射箭、投篮、网球
- ✓ 娱乐性强，增加互动



定制虚拟形象

- ✓ 官方形象模板
- ✓ 个性化定制形象
- ✓ 可以使用语音、手势、动作进行交互



创造模式

- ✓ 玩家可以使用创造编辑工具，创造自己的世界
- ✓ 轻世界的“灵魂”

PICO 健身：内容差异化拓展，兼顾趣味性与专业度

- 对比Meta，PICO当前虽然硬件设备在中国市场已占据最高的市占率，但还未构建起可良性循环的完整生态。Meta的内容生态高度依赖重度游戏，但PICO想要实现弯道超车，更有可能借力字节的视频内容、营销生态等资源优势走出差异化。VR+健身就是其中一个突破点。
- PICO的VR+健身，兼顾趣味性、专业度、社交。
- 趣味性：《多合一运动》《OHSHAPE》《实况钓鱼》等趣味性运动游戏，比较适合轻度健身爱好者；
- 专业度：与超级猩猩合作的《超燃一刻》，配备专业教练与课程，内设奖励与闯关体系，运动强度较大的《莱美搏击》，比较适合中重度健身爱好者；
- 社交：PICO加入自研CalSense体能算法。用户可以在运动中心内查看卡路里燃烧量，运动数据可视化。同时，运动排行榜功能可与好友进行PK，增加社交属性。



2022年VR跟踪之META——从Rift到Quest

- 三款里程碑式产品



Rift DK1



Quest2



Quest Pro

- **DK1——VR产品范式的定义者**:Oculus Rift DK1, 简称Oculus一代。它的硬件部分是一个头戴式的显示设备, 通过HDMI或DVI输入, 可以将计算机渲染的画面显示在你面前的小屏幕里。这种全封闭的设计, 虽然看起来有些笨重, 但是却可以带给人以全方位的沉浸体验。
- **Quest2——现象级消费级VR一体机**:Quest2是Facebook于2020年9月在Facebook Connect上正式发布的VR一体机。它是既能独立工作又能连接到游戏PC的虚拟现实头盔; 作为VR头部企业的年度旗舰VR一体机, Quest2一上市便获得了极大的关注。其配置与技术, 乃至价格与外观设计等都或多或少的影响着日后VR一体机的走向。
- **Quest Pro——场景拓展, 娱乐-办公**:Quest Pro是Meta高端产品线上的首款头显, 为合作与创作而生, 旨在扩展VR的可能性, 其添加的一系列新技术能让创作者和用户在元宇宙中做到更多的事情。Meta希望Quest Pro能够为建筑师、工程师、设计师等工作者所用, 帮助他们改进工作体验、提升创造力。

VR产品系列——Quest系列

● Meta产品路线

机型	Oculus Rift DK1	Oculus Rift	Oculus Go	Oculus Rift S	Quest	Quest2	Quest Pro
外观							
发布时间	2012. 8	2015. 5	2017. 10	2019. 3	2018. 9	2020. 9	2022. 10
头显类型	PCVR	PCVR	一体机	PCVR	一体机	一体机	一体机
售价	300美元	599美元	199美元	399美元	399美元	299/399美元	1499美元

Quest 2产品亮点



硬件方面：Quest 2配备了卓越的Qualcomm Snapdragon XR2平台，有助于提供更好的AI功能，在性能翻倍的同时还带来5G接入能力。



电池续航方面：与一代基本一样，游戏时间为2至2.5小时（测试时间为2至3小时）。

瞳距调节方面：Quest 2摒弃了此前的底部滑轨设计（58Hz-72Hz），改用了屏幕侧滑三档调节（58mm、63mm和68mm）。



内存方面：从原先的4G升级至6G。

显示方面：改进后的显示屏单眼具有1832 x 1920像素。为了获得更身临其境的体验，Quest 2还配备了90Hz显示屏。



控制器：新版Oculus Touch控制器具有先进的手势追踪功能，可以随心操控，享受更真实的VR体验。



Quest pro产品亮点



面部跟踪

Quest Pro有五个面向面部的传感器，可以捕捉眼睛和面部，从而创建各种表情。

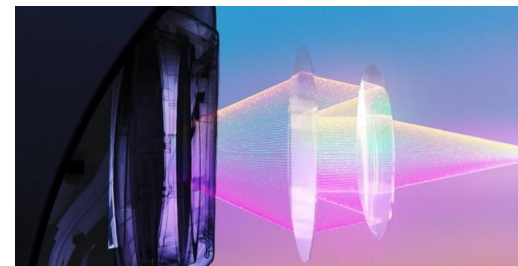


更强大的MR体验

全新的传感器架构，配备高分辨率的对外摄影机，进而打造高分辨率、全彩的MR体验。



作为Meta首款高端头显设备，Pro机型支持全彩MR混合现实体验的高分辨率传感器、高清LCD显示屏、以及眼动追踪和面部追踪功能。



全新光学组件

采用Pancake透镜，全新搭载的两个LCD显示器还使用了高端桌面显示器才会使用的局部调光和量子点技术，在颜色方面的表现优秀。



自动追踪手柄

每个手柄内置了三个传感器，能在不借助头显的情况下追踪其自身在3D空间里的位置，为用户提供全方位的360度追踪。

Quest 内容生态

- 2019年Meta建立Oculus Rift和Quest内容平台，并在2021年2月推出APP Lab平台，允许开发者以相对宽松的方式发布自己的游戏。2021年10月，Meta宣布1,000万美元的创作者基金，鼓励更多的内容创作者进行VR内容创作。
- 目前，Quest商店有超过400款内容，其中1/3的作品创造了超过100万美元的收入，有33款作品的总收入突破1000万美元。自2021年以来，收入超过500万美元的作品数量翻了一番，达到55款。
- 在 META Connect2022大会期间，Meta宣布现在正在与十几个合作伙伴合作，为Quest生态系统带来新的工具。其中微软是重要的合作伙伴，将为Quest平台带来丰富的微软工具，包括办公套件、Xbox云游戏和 Windows等。



Quest 游戏

- 据VR陀螺，截至2022年3月，Quest Store应用总数超400款，截至2022年8月，Quest Store排名前十游戏有《Beat Saber》、《Blade&Sorcery: Nomad》、《Warplanes: Battles over Pacific》等，涉及射击、音乐、健身、战斗沙盒等题材；其中《Beat Saber》是目前最畅销的VR游戏之一。

收购Beat Games
收购Play Giga
收购Sanzaru Games
收购Ready At Town
收购Downpour
收购Unit2 Games
收购Bigbox VR
收购Within



- ✓ 截至2021年年底，Quest商店交易规模已超过10亿美元；
- ✓ 截至2022年4月，Oculus游戏平台VR内容达1738个。



Horizon Worlds

可自定义形象与环境并邀请
好友进入做客



Horizon Home

推出Messenger支持；
Instagram、Facebook等服
务成为Home的2D面板应用



Horizon Workrooms

跨距离虚拟会议室，以虚拟
形象参加会议，或通过视频
电话从计算机拨号进入虚拟
房间

PICO VS Quest：硬件旗鼓相当，差异在于内容

Pico Neo 3 VS Quest 2

产品	Quest2	Pico Neo3
CPU	高通XR2	高通XR2
运行内存RAM	6GB	6GB
存储ROM	128/256GB	128/256GB
基础版价格	399美元	2499RMB
是否支持扩展存储	不支持	不支持
材质	LCD	LCD
分辨率	3664*1920	3664*1920
刷新率	120HZ	120HZ
电池续航	3640mAh	5300mAh
视场角FOV	88度	98度
交互自由度	6dof	6dof
是否支持Steam	支持	支持
重量	450g	395g



从硬件参数看，Pico Neo3可与Quest2一较高下，甚至在视场角、电池容量等方面优于Quest2。然而它们的核心差距还是在于其VR内容生态，Quest内容生态游戏、娱乐、社交等更加丰富。

PICO VS Quest：硬件旗鼓相当，差异在于内容

● PICO 4 Pro VS Quest Pro

产品	Pico4 Pro	Quest Pro
处理器	骁龙XR2	XR2+Gen1
内存	8GB	12GB
存储	512GB	256GB
视场角	105度	106度
分辨率	单眼2160*2160 LCD(两块)	单眼1800*1920 LCD(两块)
刷新率	72hz/90hz	72hz/90hz
光学模组	Pancake	Pancake
透视	全彩	全彩
电池容量	5300mAh	5000mAh
表情追踪（眼睛和面容）	支持	支持
IPD调整范围	62-72mm	55-75mm
重量	586g	722g
零售价	3799RMB	1500美元



综合硬件指标来看，PICO4 Pro、Quest Pro两款设备均属同代产品中的旗舰产品，Quest Pro已着力于切入办公场景，PICO4仍着眼于娱乐与体验。

PICO VS Quest：硬件旗鼓相当，差异在于内容

- PICO与Quest的比拼，实际上就是字节与Meta的比拼。Quest和Pico的硬件体验感差异不大，但Quest内容强于游戏，Pico内容强于视频，PICO有望借助字节的短视频内容生态弯道超车。
- PICO当前在内容方向上的发力路径有两条，一是游戏，主要依靠串流Steam平台、海外VR大作的移植以及中小开发者的扶持；而是泛娱乐，包括视频、全景式沉浸视频、互动剧、直播等等。



2022年VR跟踪之Pimax系列

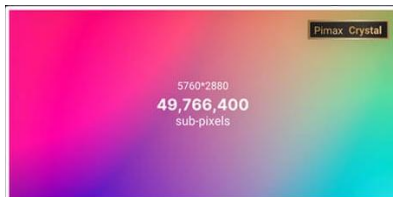


- 小派科技（Pimax）创立于2015年，创始团队由国内第一批对虚拟现实技术有狂热爱好的技术工程师组成，已经先后参与全球顶尖的VR头显项目研发与生产，团队成员大多来自虚拟现实、智能手机、可穿戴设备、互联网、游戏、医疗等专业领域。



Pimax Crystal

屏幕方面



OLED+Mini LED屏幕

160Hz硬核刷新率

5760x2880分辨率

硬件方面：

搭载了高通XR2和Pimax定制的PCVR引擎双处理双芯片。



眼动追踪功能&Auto-IPD：

快速自动调节瞳距；
完美眼追、虚拟社交、无限真实。



突破VR镜片设计限制：

通过可更换透镜个性化调整用户所需的PPD和FOV，可达消费级VR目前已知的最高PPD。

体验方面：

WiGig：支持60GHz频率无线传输PC数据；
6000mAh电池；
四摄像头、双手柄、6DOF。

6DOF自由度
4 Cameras



- 在2022年6月的Pimax Crystal发布会上，小派就预告了自己的应用商城Pimax Store即将上线。在选取上架内容时，Pimax Store将兼顾用户体验和内容与硬件的匹配度，不仅会注重内容品质，也愿意与独特的、硬核的甚至小众的VR内容达成合作，扶持上架。
- 小派应用商城首批上架阵容将涉及游戏，VR工具软件，VR视频播放器，VR社交软件等。首发的PCVR游戏阵容将以能展现小派头显特性的游戏为主，其中就包含2021年大热僵尸FPS VR游戏Surviv3，该游戏也曾荣登2021年度Steam大赏VR游戏黄金榜单。
- 此外，小派已经与业内王牌拟真赛车游戏厂商iRacing，建立深度合作关系，进一步巩固小派产品在拟真游戏应用上的绝对首选地位。



2022年VR跟踪之HTC系列

HTC VIVE系列产品



VIVE Pro系列

VIVE Pro系列：专业级PC-VR设备。VIVE Pro为满足专业级的VR用户需求而设计，是简单易用的PC-VR系统，满足各种业务需求进行扩展



VIVE Cosmos系列

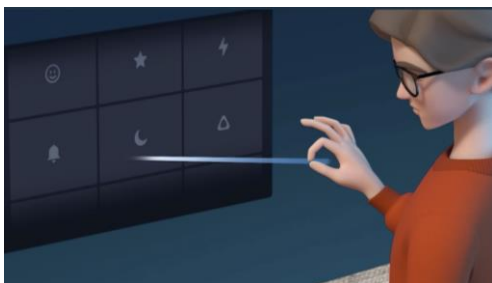
VIVE Cosmos系列：定制化PC-VR设备。采用全新定位系统，六个摄像头传感器与最新的软件优化功能相结合，可实现精确的 inside-out 追踪。



VIVE Focus系列

VIVE Focus系列：多模式6dof高端VR一体机。VIVE Focus3是HTC VIVE最新推出的VR一体机，搭载最新高通骁龙XR2处理器，以5K分辨率、120°视场角和更平滑的90Hz刷新率，带来沉浸式视觉效果。

HTC Vive Focus3



显示方面：

5K分辨率；
90Hz刷新率；
120度视角。



高精度手部追踪功能：

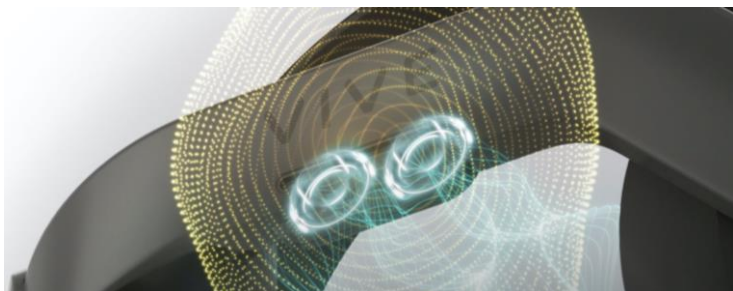
用户可以随时放下操控手柄，以激活26点手部追踪，确保用户在虚拟环境中交互的精准度达到新水准。



VIVE Focus3

远程协作：

在不同地理位置远距离与团队合作，就像面对面一样，使用强大的交互式VR工具进行互动。



沉浸式3D空间音效：

开放式双驱动扬声器可提供3D空间音效。



2022年VR跟踪之PS VR系列

- 2016年，索尼的第一代VR设备PS VR正式推出，正是VR行业的第一次兴盛期。
- 2022年，索尼宣布其第二代VR产品PS VR2即将在2023年2月发布，并且直接公布了产品的参数信息。
- PS VR，看起来更像是PS游戏机的附属品。从官宣信息来看，PS VR2仍然是从PS游戏机配件的角度做VR，需要搭配PS5使用。

项目	PS VR	PS VR2
分辨率	1920*1080	4000*2040
显示屏	OLED	OLED HDR
刷新率	90Hz 120Hz	90Hz 120Hz
视角	100度	110度
跟踪	通过PS MOVE进行LED跟踪	眼动追踪
触觉反馈	仅手柄	头显和手柄
声音	3.5mm耳机接口	3.5mm耳机接口
连接接口	USB-A和HDMI	Type-C



4KHDR显示器

通过最高120fps的4K高动态范围视觉效果，体验清晰的虚拟现实环境。两个2000*2040的OLED显示器将提供原始PS VR头戴设备四倍以上的分辨率。



PlayStation VR2



内向外追踪

通过嵌于头戴设备中的四个摄像头追踪用户与控制器，用户的动作与视线方向即会反映在游戏内。

头戴振动反馈

眼动追踪

头戴设备会侦测眼球的移动，让用户在线与其他玩家相遇时能以逼真的全新方式进行互动，感受到的情绪反应和表达细腻度都会因而提升。



PlayStation VR2 Sense控制器

通过触觉反馈、自适应扳机、精准跟踪及手指触摸侦测，以全新方式与虚拟世界互动。这全都整合于一个以人体工学设计的独特外型中。

- 索尼持续投入优质独占游戏，利用自身游戏平台的优势，借助经典热门IP吸引用户流量。包括《生化危机7》《蝙蝠侠》《钢铁侠VR》《Hitman》等。
- 索尼这几年悄然加快了在VR内容上的投入，公司计划未来通过积极投资建设全球工作室，加强游戏自研能力，打造优质独占游戏，维持自身在内容领域的优势。



生化危机7
2016



蝙蝠侠：阿卡姆
2017



漫威钢铁侠VR
2020



Hitman3
2021

2022年AR跟踪之Rokid系列

- Rokid创立于2014年，目前致力于AR眼镜等软硬件产品的研发及以YodaOS操作系统为载体的生态构建。公司通过语音识别、自然语言处理、计算机视觉、光学显示、芯片平台、硬件设计等多领域研究，将前沿的AI和AR技术与行业应用相结合，为不同垂直领域的客户提供全栈式解决方案，其AI、AR产品已在全球七十余个国家和地区投入使用。



Rokid Air

显示方面

4K显示屏
75Hz刷新率
1800nit亮度
视场角43度



双高品质定向扬声器

环绕立体声震撼音效

拓展性强

可连接手机、PC、平板；
兼容所有主流视频平台app。



Rokid Station

Rokid Station为Rokid Air智能眼镜的专配智能终端，也可化身为遥控器，当我们进行菜单选择或打游戏时，都可通过Station上的按键进行操作。



屈光度可调

支持单眼500度以内屈光度调节

记忆金属鼻托

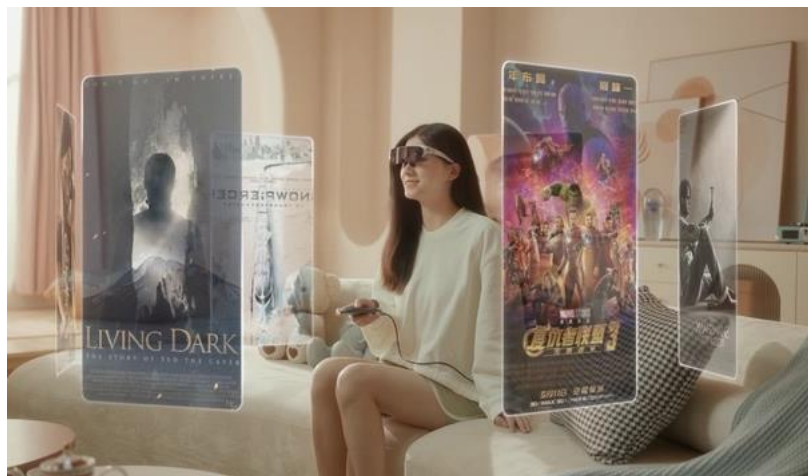
各类鼻梁均使用

轻巧便携

仅重83g，机身可折叠易携带

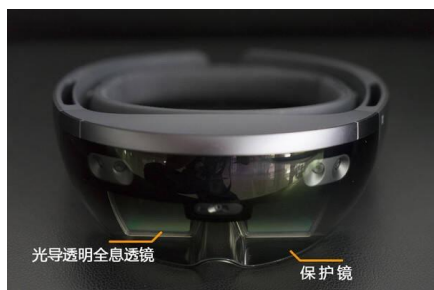


- 在内容生态层面，目前阿里、b站、腾讯、字节等头部大厂都给予高度认可并与Rokid达成合作。
- 引入了bilibili、抖音大屏版、葫芦视频、南瓜视频、CCTV新视听，在应用商店提供下载。全网影视、综艺、赛事、直播，应有尽有。还引入了QQ音乐、全民K歌等音乐视频内容，提供更丰富的泛娱乐应用。
- 在游戏领域，Rokid Station引入上万款本地经典休闲游戏，同时还支持原神、最终幻想、仙剑奇侠传、真三国无双等数百款知名3A大作，并和中国移动、中国电信等运营商云游戏平台达成了合作，共同探索便携沉浸的巨幕云游戏体验。
- 在办公领域，Rokid Station创新引入阿里无影云电脑。连接Rokid Air，配合蓝牙键鼠，成为一套小巧、便携的移动办公套件，随时随地满足工作、学习需求。



2022年MR跟踪之HoloLens

- HoloLens MR头显由Microsoft公司于北京时间2015年1月22日发布。
- MR的意思是Mixed Reality混合现实，HoloLens运用的MR技术具备环境学习能力，能够实现全息影像和真实环境的融合。



功能介绍



模拟游戏



观看视频



辅助3D建模

模拟登陆火星场景

产业链关键环节——光学、显示、交互、芯片

相较于智能手机，我们认为XR硬件产业链的关键环节体现在光学、显示、交互、芯片。

- 1) 光学：对于VR而言，Pancake助力VR头显进一步轻薄化，消费级VR头显的商用进程正在加快，视场角、分辨率、PPD等参数同样在进步提升。对于AR而言，光波导技术持续突破，从光学效果、外观形态来看具备较高的发展潜力，有望探索更广阔的的消费前景。
- 2) 显示：Fast-LCD由于成本优势当前仍然占据主流，但Micro OLED正快速导入应用，作为Pancake方案的配合加速提升显示性能，Micro LED技术也在逐渐成熟，有望接力Micro OLED驱动近眼显示再升级。
- 3) 交互：向接近人自然的交互不断进化。眼动追踪、面部识别、自动IPD调节、可变焦显示等技术越来越多地应用在新一代VR当中，手势识别、Inside-out、6DoF等技术也不断成熟，与人交互的范围也越来越广。
- 4) 芯片：高通芯片持续迭代，从VR2到AR1，且定制化芯片成为最新趋势，手机芯片商/AIOT专业芯片商/品牌厂商也纷纷参与开发，希望抢占XR产业链核心部位。

VR光学： Pancake方案带来里程碑式体验提升

- VR光学方案沿着传统透镜-菲涅尔透镜-Pancake超短焦方案持续迭代。Pancake方案使光线在镜片、延迟片、反射式偏振片中多次折返，将光路压缩至2-3片偏振片的狭窄空间内，打破菲涅尔透镜对焦距的要求，帮助头显重量降至200g以内，厚度缩减至传统终端的三分之一，大幅提升佩戴舒适性。

产品名称	Oculus Quest 2	Arpara A10 5K VR一体机	Arpara 5K VR头显	华为VR Glass
产品形态	一体式	分体式	分体式，无计算模块和电池	分体式

产品示意图



选择方案	菲涅尔透镜+Fast LCD	Pancake+Micro OLED	Pancake+Micro OLED	超短焦+Fast LCD
主机厚度	142.5mm	—	30mm	26.6mm
整机重量	503g	380g	200g	166g

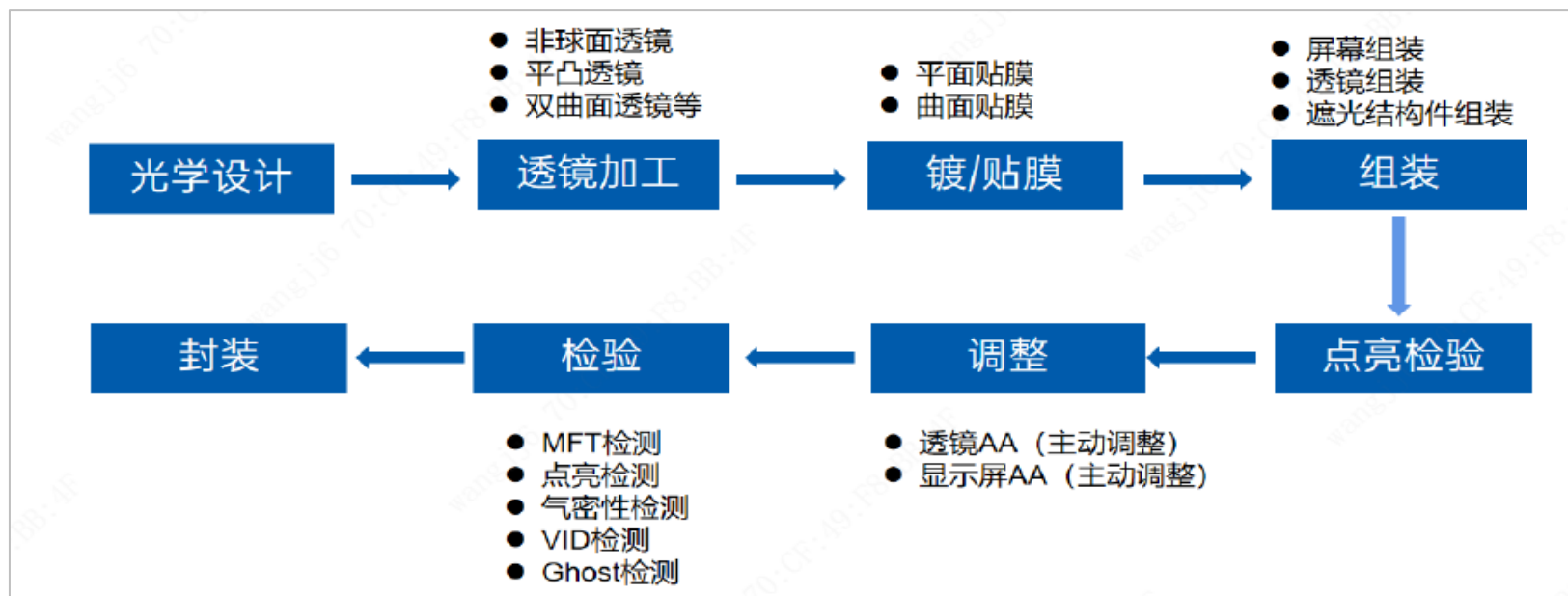
Pancake商用趋势确定，存在透光率/鬼影等痛点

- **Pancake方案光效较低，理论最高25%，通常仅约10%。**受光学原理限制，光线在Pancake模组中每次经过偏振/半反射环节，光效损失50%。以两片式Pancake 简易模组为例，光线从屏幕发出后至少经过一次圆偏振镜，两次半透半反镜，光效折损已到12.5%，考虑光线传播中不可避免的其他损失，通常Pancake模组光效仅约10%。因此，Pancake光学方案通常对屏幕亮度要求更高，光学与显示方案需配套迭代。
- **Pancake方案的鬼影问题比传统透镜方案更严重。**在光学成像系统中，透镜界面多次反射、透镜缺陷散射、物理结构散射等因素造成的杂散光，在画面中的某个位置形成的“像”被称为“鬼影”（ghost）。鬼影现象会直接导致图像质量的降低。Pancake方案因为光线多次折返，鬼影问题相比常规非球面/菲涅尔方案更为严重，一般通过改善透镜材料、形状等方式优化。
- **当前的Pancake方案能实现的视场角较小。**尽管Pancake方案视场角理论上限较高，但目前可实现的视场角（90° 左右）明显低于菲涅尔透镜方案（100° 以上）。曲面贴膜工艺能够扩大FOV，改善Pancake视场角小的缺点，但是工艺难度较大。
- **以上提及的Pancake方案的三大痛点中，光效低需通过配套高亮度显示方案解决，但鬼影现象、视场角小均可通过改进Pancake模组加工工艺解决/优化。**

Pancake工艺难点：膜材料质量/贴膜工艺/组装调整

- **膜材料质量：**手握Pancake核心材料的公司美国3M，其反射偏振模是目前Pancake模组厂商难以绕开的关键材料。反射偏振片和1/4相位延迟片工作的准确性与稳定性对于Pancake光学成像质量构成关键影响，膜材质量的标准和对应要求较高，目前供应商以海外厂商3M、旭化成等，3M也联合华硕推出了基于Pancake方案的VR参考设计。
- **贴膜工艺：**光学贴膜的方式可以分为平面贴膜和曲面贴膜两种。平面贴膜技术难度较低，但会牺牲部分光学性能和成像质量。曲面贴膜可以带来更大的FOV和更优质的成像质量，但是工艺难度较大，容易边缘褶皱和翘起。
- **组装调整：**Pancake方案光路设计复杂，细微差异便会导致光学模组整体的光路变化，因此对组装调整环节的要求较高。高精度AA设备在Pancake组装过程中起到关键的效率和良率提升作用。

图：Pancake模组的加工流程



AR光学：光波导发展趋势清晰

- 与VR的不同之处在于，AR眼镜需要透视（see-through），既要看到真实的外部世界，也要看到虚拟信息，所以成像系统不能挡在实现前方，这就需要多加一个或一组光学组合，通过层叠的方式，将虚拟信息和真实场景融为一体。
- 目前比较成熟的AR光学方案主要包括棱镜、BirdBath自由曲面、阵列光波导、光栅光波导。其中，棱镜方案畸变小、体积小但FOV小、镜片太厚难以普及；BirdBath自由曲面FOV较大、体积适中但是有较严重的重影问题。光波导方案可分为阵列光波导与衍射光波导，阵列光波导FOV较大显示效果好，但光效低、量产难度高，衍射光波导还存在较大技术难点，工艺上难以做到均匀性。

技术路径	优点	缺点	代表公司
棱镜	畸变小、体积小、设计难度小	FOV小、遮挡视线 镜片较厚、外形难被大众接受	RealWear、Google Glass
BB自由曲面	FOV较大 体积适中、设计难度小	透光度低（<25%） 重影、镜片厚（>10mm）	Epson、耐德佳、Nreal、影创、Meta
阵列光波导	FOV较大、显示效果佳 透光度高（>80%），镜片薄（1-2mm）	光效率相对低 加工难度高	珑璟、Lumus、灵犀、理湃
衍射光波导	FOV、Eyebox可以设计得很大 半导体工艺极限成本较低	均匀性差，存在鬼影和倒影现象 材料难找，设备贵，生产工艺难	Microsoft、Vuzix、Magic Leap、Waveoptics、Digilens

AR光学：自由曲面、BirdBath为过渡方案

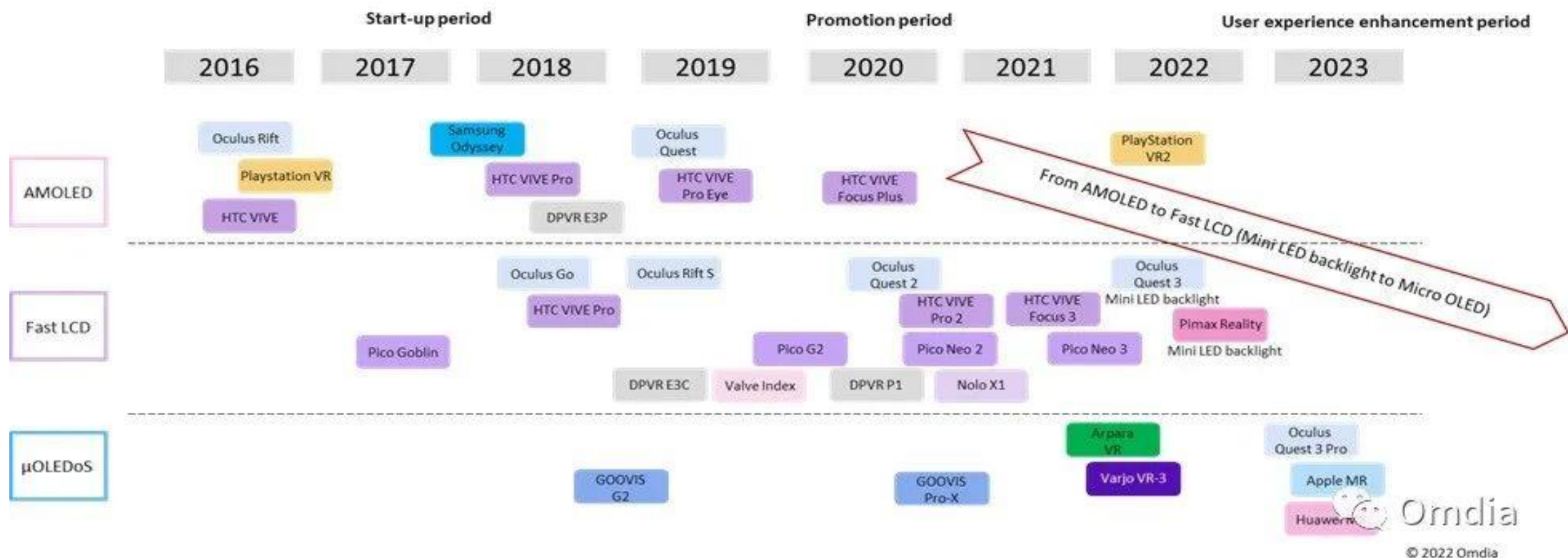
- 自由曲面和 BirdBath 小幅改善镜片厚度、其他性能良好、量产制造成熟，成为近几年的过渡方案。
- 自由曲面方案由表面形状不能被连续加工、具有传统加工成型的任意性曲面担当反射镜，对显示屏光线进行准直和成像，因此成像质量较高，色彩饱和度和光学效率表现优秀。但自由曲面结构局部精度低，带来低分辨率和画面扭曲，使得现实世界和虚拟世界光线传递时存在畸变现象。
- BirdBath 方案下，显示屏光线经 45 度角的分光镜反射至曲面镜弹射入眼，而现实光线透过曲面镜和分光镜入眼。光学结构简单，光效高、视场角大；但眼动范围受限，同时透射入眼面临图像畸变、光线透过率低的缺点。自由曲面和 BirdBath 光学结构相对简单，一方面光效高，显示屏选择灵活，另一方面制造难度低，可以较低成本规模量产，成为目前中低端或消费级 AR眼镜的主要光学方案。但其他性能一般，存在畸变等问题，致命的是，为实现可用视场角，镜片厚度压缩极限为 8mm，无法做到日常眼镜般的轻薄机身。

AR光学：光波导为最终发展方向

- 光波导由于轻薄性及显示效果优势，一旦突破量产瓶颈后有望实现快速渗透。
- 光波导可以分为两类技术方案，几何光波导与衍射光波导。几何光波导成像效果好但量产难度高，衍射光波导具备基础量产条件有望成为主流技术方案。
- 几何光波导由一系列半透半反镜面组成，镜面是嵌入到玻璃基底里面并且与传输光线形成一个特定角度的表面，每一个镜面会将部分光线反射出波导进入人眼。几何光波导的概念最先由以色列公司Lumus提出并一直致力于优化迭代，市场上还未出现大规模的量产眼镜产品。
- 衍射光波导可进一步细分为全息光波导与表面浮雕光栅波导，主要原理为通过衍射光栅替代传统几何光学器件，采用镜片表面的光栅结构实现光束的扩展和耦出。通过合理的设计光栅结构，光栅波导技术可以实现出瞳的二维扩展。
- 全息体光栅方案厂商需要具有IDM能力，提供从材料到量产完整的解决方案，代表公司包括苹果公司收购的Akonia、Digilens、Sony等。
- 表面浮雕光栅波导已具备成熟批量生产条件，代表公司包括微软HoloLens 2, Magic Leap One等。

显示：LCD&OLED持续演进，硅基OLED路径明确

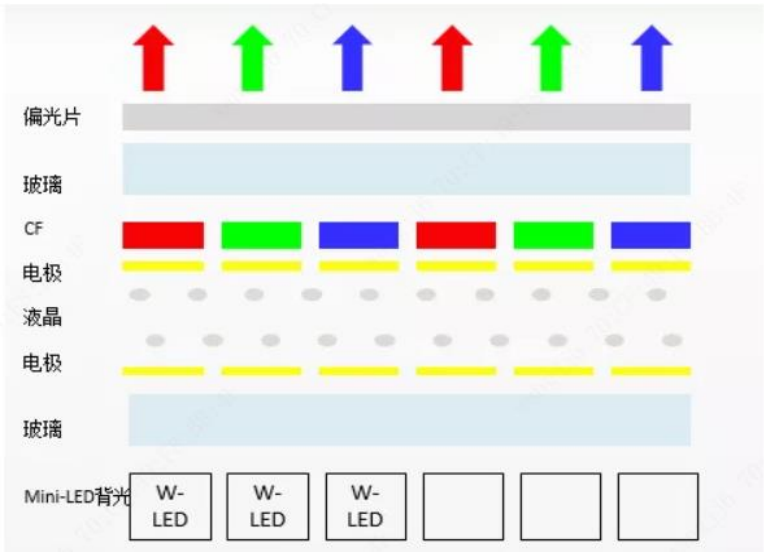
- AMOLED-Fast LCD (Mini LED背光)-Micro OLED. 早期VR头戴式设备大部分采用普通LCD，而后逐渐被OLED取代。2016年前后，OLED技术尝试用于VR设备中，并一度成为VR设备生产商的首选。2018年以后，Fast-LCD技术的出现让LCD逐渐成为主流选择。
- 自2022年起，显示方案进一步迭代，行业龙头开始使用带有Mini LED背光的Fast LCD显示面板，或使用硅基OLED (OLEDoS)，来提升显示性能，比如Apple MR将使用硅基OLED显示方案，Quest Pro将使用Mini LED背光的Fast LCD显示面板等。



显示：趋势1——Mini-LED背光LCD

- Mini LED (Mini Light-Emitting Diode, 次毫米发光二极管), 指100~300微米大小的LED芯片, 芯片间距在0.1~1mm之间, 采用SMD、COB或IMD封装形式的微型LED器件模块, 往往应用于RGB显示或者LCD背光。Mini LED的应用主要分为作为使用Mini LED芯片+LCD的背光方案与直接使用Mini RGB显示屏的自发光方案。当前Mini LED背光方案已经进入爆发期, 苹果、三星等多家品牌厂商都已开始推出相关产品。
- Fast-LCD与Mini LED相结合, 不仅可以很好的解决漏光难题, 也能够进一步提升Fast-LCD在高对比度、高刷新率、高亮度等方面的性能, 辅以HDR功能, 将更好地发挥VR产品近眼超清显示的画质要求。
- 截至目前, 已发布的应用Mini LED的VR头显共有四款, 为Varjo Aero、Pimax Reality 12KQLED、Pimax Crystal、创维Pancake 1pro。

图：Mini LED背光结构图



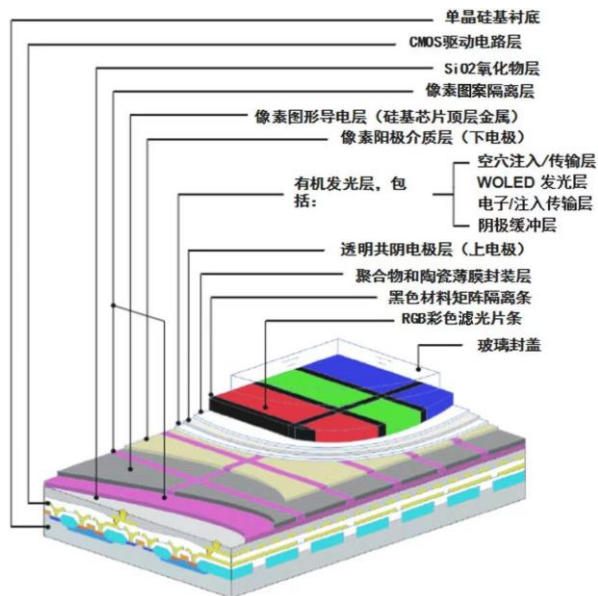
图：全球首款Mini LED VR-Varjo Aero

Varjo Aero参数	
分辨率	单目2880x2720 (7.8MP), mini-LED
刷新率	90Hz
透镜	非球面
官方FOV	对角线134°, 水平115° (出瞳距12毫米时)
IPD调节	电动调节, 范围57-73毫米
接口	USB-C, 通过接线盒链接USB-A 3.0、DP 1.4
连接线长度	5米
定位	SteamVR 1.0或2.0
内置摄像头	双眼球追踪摄像头
交互方式	支持SteamVR手柄
音频	3.5毫米接口
麦克风	无 (支持外接)
透视模式	不支持
重量	机身487g, 头带和配重230g

显示：趋势2——Micro-OLED（硅基OLED）

- Micro OLED（又称硅基OLED，OLED on silicon），是半导体技术与OLED技术的结合，Micro-OLED光源模块是通过使用气相沉积将OLED沉积到衬底上而产生的。Micro OLED（硅基）包含了夹在两个电极之间的有机发光材料，当电流流过时二极管发光。随后通过滤光片生成所需的颜色。硅基OLED创新性地结合了半导体与OLED，显示器件采用单晶硅芯片基底。硅基OLED都是在6英寸、8英寸的晶圆上小面积蒸镀，大大减小了生产OLED时要克服的蒸镀均匀性难题。
- Micro-OLED微显示器件具有OLED自发光、薄、轻、视角大、响应时间短、发光效率高等优异特性，而且更容易实现高 PPI（像素密度）、体积小、易于携带、功耗低等应用效果，特别适合应用于近眼显示设备。

图：硅基OLED原理图



图：全球首款Micro OLED VR-arpara 5K VR

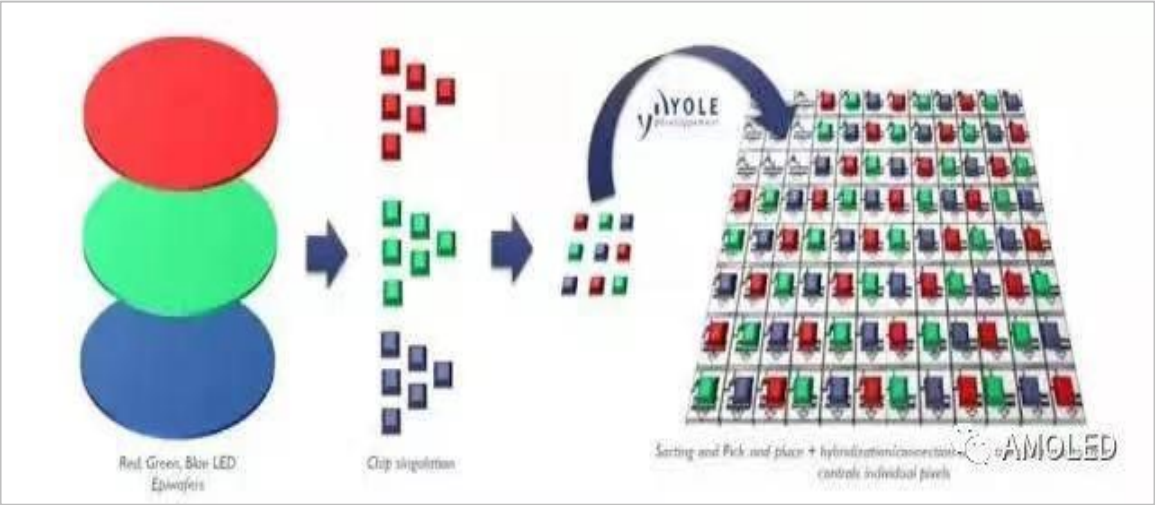
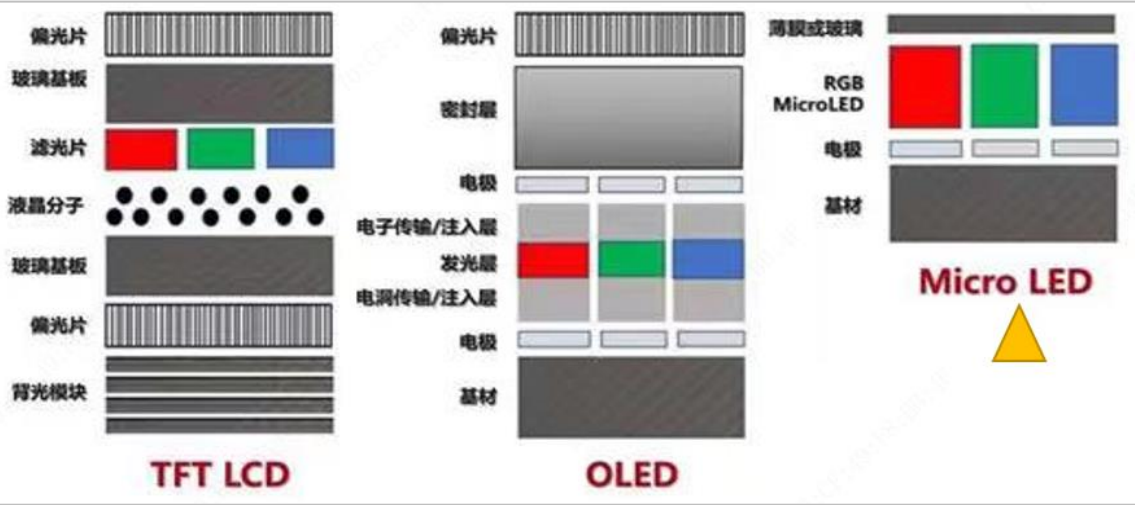


arpara 5K参数	
分辨率	单目2560x2560，Micro-OLED
刷新率	120Hz
透镜	超短焦Pancake
官方FOV	95°
IPD调节	56mm~72mm
屈光度	1D~5D
定位	主机3DOF+配件实现头手Outsidein 6DOF
定位设备	arpara Tracker
传感器	内置9轴传感器 G-Sensor 陀螺仪 距离传感器 电子罗盘
重量	200g

显示：趋势3——Micro LED

- Micro LED（Micro Light-Emitting Diode，微型发光二极管），指在一个芯片上集成的高密度微小尺寸的LED阵列，如LED显示屏每一个像素可定址、单独驱动点亮，可看成是户外LED显示屏的微缩版，将像素点距离从毫米级降低至微米级。
- 通过巨量转移技术，将微米等级的RGB三色Micro LED搬到基板上，形成各种尺寸的Micro LED显示器。Micro LED的芯片到了肉眼难以分辨的等级，可以直接将R、G、B三原色的芯片拼成一个像素点，变成“一个像素”的概念，不再需要滤光片和液晶层。
- Micro LED用到的自主发光器件，即LED半导体灯珠，是一个正负极性明确、体积形状固定的固体小颗粒灯珠，再通过“巨量转移”技术平面化集成大量的微小LED灯珠器件。“巨量转移”技术是Micro LED产业独有技术，难以从其它相关产业获得巨大的技术帮助。目前，Micro LED巨量转移技术依然处在实验室攻关阶段。

图：LCD、OLED、Micro LED结构图对比：Micro LED明显更薄 图：巨量转移技术原理图



显示：硅基OLED有望成为行业主流

- 从性能参数角度看，相比Fast-LCD和Micro LED，硅基OLED色彩表现要更强，沉浸体验更好。对比LCD屏幕来说，由于其拥有黑矩阵的工艺，因而限制了LCD的PPI提高，即便像素密度提高到1000ppi以上也难以完全消除纱窗效应；对比Micro LED来说，尚存在工艺和产业化问题，成熟度不够高，而且也可能存在像素点亮度不一致的问题。
- 从工艺制造角度看，其技术工艺主要分为单晶硅光刻的基底驱动层技术和OLED蒸镀技术。对于单晶硅光刻的基底驱动层技术而言，硅基OLED采用的是28nm、55nm、或者180nm成熟的CMOS工艺，其工艺难度要远低于目前的顶级光刻单晶硅技术。

显示技术	传统LCD	Fast-LCD	Mini-LED	OLED	Micro-OLED	Micro-LED
技术类型	背光/LED	背光/LED	背光/LED	自发光	自发光	自发光
对比率	1, 000: 1	5, 000: 1	10, 000: 1	无上限	无上限	无上限
亮度	500	500	1000+	500	1000+	5000+
发光效率	低	低	中等	中等	高	高
反应时间	毫秒 (ms)	毫秒 (ms)	纳秒 (nm)	微秒 (μm)	微秒 (μm)	纳秒 (nm)
对比度	低	低	中等	高	高	高
寿命(小时)	60k	60k	80-100k	20-30k	30k	100k
柔性显示	难	难	难	容易	容易	难
成本	低	低	较高	中等	较高	高
功耗	高	高	约LCD的40%-50%	约LCD的60%-80%	约LCD的30%-40%	约LCD的10%
运作温度	-40℃-100℃	-40℃-100℃	-1000-120℃	-300-85℃	-500-80℃	-100℃-120℃
产业化进展	已大规模量产	已大规模量产	已量产	已规模量产	初步规模量产	研究阶段
产业成熟度	高	高	较高	高	较低	低
主要机型	早期VR盒子	Quest 2、创维Pancake 1、PICO Neo3	创维Pancake 1 pro、预计Quest Pro	Quest 1、HTC Vive、预计索尼PS VR2	预计苹果MR	—

VR交互：需传感器、芯片、算法等多方共同参与

- VR的感知交互交互需要传感器、芯片、算法等多方参与共同完成整体流程。
- 传感器：利用含摄像头在内的传感器精准实时捕捉用户行为；
- 芯片：多传感器融合和校准后的图像经芯片算力、算法打造多维感知效果；
- 屏幕：最后从设备呈现给用户。



AR交互：复用VR交互，但需优化算法和传感器

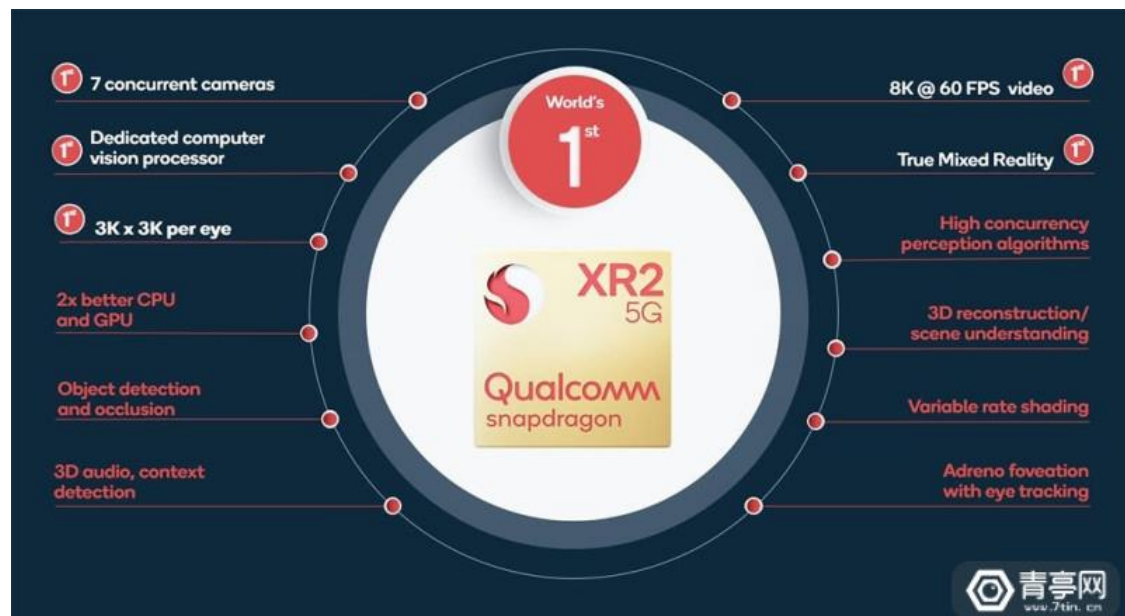
- AR交互的整体思路是复用相对成熟的VR感知交互技术，但受轻量化、功耗和成本限制，现搭载功能有限，空间交互、手势识别将最先应用。AR 面临的难点是在使用较少数量传感器的情况下，保持高自由度和高精度，现阶段通过提升算法、传感器和软硬适配进行效果优化。
- 目前AR多用的交互方式包括控制器、手势交互、语音交互等。部分B端眼镜可像VR般搭载大量传感器，交互精度较高，如微软 HoloLens 分别搭载 8 和 9 颗摄像头，附加 IMU 等传感器。但这同时带来设备体积、重量和价格上升，难推广至消费端。
- 空间定位是AR的必要交互功能。AR区分于VR的目标是实现虚实融合，因此需要感知空间和分辨场景。
- 手部交互具备进步空间。AR一般不搭配手柄使用，未来作为生产力工具使用需要进一步精进手部交互，最终目标是裸手交互。

AR交互：SLAM实现虚拟信息和现实世界的结合

- SLAM (Simultaneous Localization and Mapping)，同步定位与地图构建，指在运动过程中通过重复观测到的环境特征定位自身位置和姿态，再根据自身位置构建周围环境的增量式地图，从而达到同时定位和地图构建的目的。
- SLAM 在AR 中的主要应用是1) 现实物体与虚拟物体的有效交互，2) 实现语义理解，优化智能辅助功能。
- 实现虚拟世界和现实世界之间坐标叠加、实现几何物理信息交互。与电脑、平板、手机的3D 显示不同，AR 更注重虚拟信息与真实信息的无缝融合，即图像出现的平面位置与景深准确、带来沉浸感的良好体验。这就需要利用SLAM 算法，准确叠加虚拟坐标系和真实坐标系。同时，真实环境中高低起伏、有障碍物、有遮挡关系，AR 可以让虚拟信息跟这些真实环境中的物理信息进行交互。
- 实现语义理解，优化智能辅助功能。随着机器学习和深度学习的发展，虚拟信息可以“理解”真实世界，让二者的融合更趋于自然。当前计算机已经可以已经可以认出图片上的内容，但没有理解内容之间的关系，当前的一项研究方向是，应用SLAM+AI 技术，通过特征提取，实现机器的语义理解，优化AR系统的辅助功能。

芯片：从XR2到AR2，高通XR芯片持续迭代

- 早期，高通主要基于手机芯片骁龙820、835、845研发VR开发套件，2018年发布XR专用芯片XR1，2019年推出全球首个支持5G的XR专用芯片XR2，相比上代CPU和GPU性能提升2倍、视频带宽提升4倍、单眼分辨率提升6倍、AI性能提升11倍，并多方面定制化升级。
- 据青亭网，2022年推出全球首个AR专用芯片AR2。骁龙AR2 Gen1 平台，以彻底改变头戴式眼镜的外形，并为现实世界 / 虚拟世界的组合开创一个空间计算体验的新时代。



芯片：定制化开发成为最新发展趋势

- 目前VR/AR平台中，高通芯片一家独大，国产芯片正奋力追赶。多方竞相开发，有望持续推动XR芯片向高性能、定制化方向升级。
- 手机芯片平台商：高通：与Meta已签署协议，为Meta的Quest生产定制芯片组。
- AIOT专业芯片商：瑞芯微16年推出RK3399，算力2.4TOPS，21年发布新一代RK3388，算力提升至6TOPS；全志科技16年发布H6VR入门芯片，17年推出VR9，支持4K@60fps解码，主要用于中低端VR设备。
- 品牌厂商：华为海思针对XR领域开发3796V300芯片，集成8核64位Cortex-A73 CPU和Mali-G52 MC6GPU，算力达到9TOPS；三星推出Exynos 芯片，配置Mali-G71 MP20 GPU；苹果预计将推出定制化M2 Staten芯片与协处理器Bora 芯片以运用在MR新品。

硬件与内容/应用/场景互为增长飞轮

- 硬件与内容/应用/场景相互螺旋式渐进。
- 沉浸式体验只是表象，各种前沿的技术也只是辅助手段，内容/应用/场景才是VR有别于其他娱乐方式的内核。



硬件对内容/应用/场景的促进：

- ✓ 硬件参数提升到合格的水准，内容/硬件/场景的体验才具备合格的底层；
- ✓ 硬件达到优秀的水平（性能优秀、价格大众化）为内容/应用/场景的繁荣创造条件；
- ✓ 硬件的进步与升级能够刺激为内容/应用/场景的创新。

内容/应用/场景对硬件的反向促进：

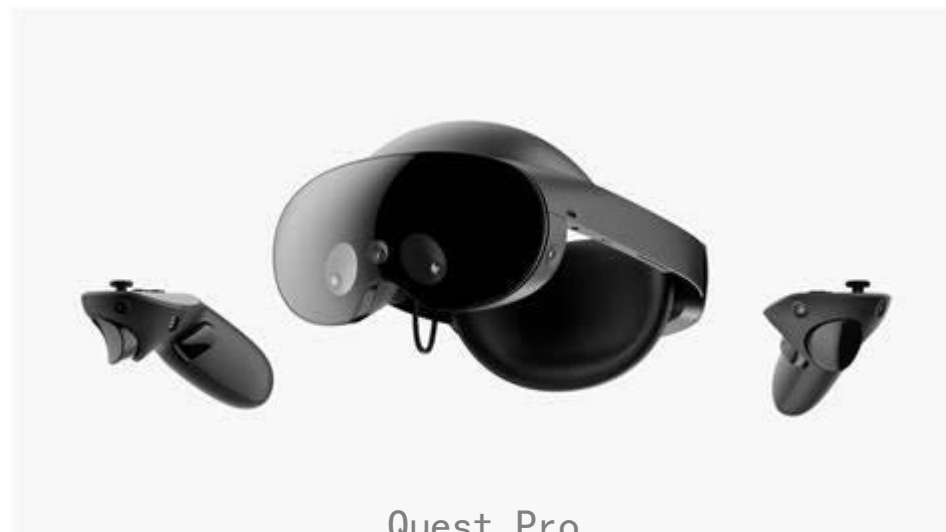
- ✓ 杀手级内容/应用/场景是推动硬件普及的最核心力量；
- ✓ 内容/应用/场景的创新倒逼硬件升级。

国内外对XR的两种认知

- 国内停留在娱乐与体验，则对应的硬件出货量的天花板为游戏主机。
- 海外认知已上升至生产力工具，则对应的硬件出货量的天花板至少为个人PC。
- 国内外对XR的两种认知会随着国内VR硬件的起量与内容的丰富逐渐追平。
- PICO 4 VS Quest Pro就是这两种认知的最佳体现。



PICO 4



Quest Pro

- 游戏将是元宇宙沉浸式体验的先行者。游戏相比其他场景在行业规律、用户属性及技术准备方面有着最高的综合准备度。游戏与Metaverse重叠的需求点众多，包括虚拟身份、真实沉浸的社交系统及独立开放的经济系统。
- 同时，受疫情影响，许多线下音乐节、演唱会和各种展会被迫取消，在线娱乐方式激增。AR/VR技术的普及和5G网络的高速发展，促进了虚拟演出与展会市场的火热，沉浸式体验越来越受到大众接受与青睐。



iBox元宇宙演唱会



ROBLOX: 全球最大的互动社区之一及大型多人游戏创作平台



虚拟沉浸式影院



第一视角沉浸式射击游戏

元宇宙游戏特性

平台性

整个游戏就是一个开放性平台，不仅包含了游戏，还包含大量的非游戏内容，比如社区活动，音乐节和生日聚会等。

用户创作

作为一个开放性平台，用户将具有极高的自由度，可以创作大量内容。

虚实结合

元宇宙不是空中楼阁，大多数现实场景也会同步映射到元宇宙中。

持久身份

同现实世界一样，身处元宇宙中的用户也是唯一的，因此与ta关联的一切资产也拥有唯一恒久归属感。

区块链加成

区块链的不可修改，唯一归属属性恰好适配上一条持久身份，因此区块链与元宇宙的联系也会在元宇宙游戏中体现得更为深刻。

- 作为元宇宙的关键要素，社交一直被视为“短期内构建元宇宙的现实入口”。目前元宇宙的发展处于初级阶段，至少在未来10年，元宇宙的概念将依旧集中于社交网络、游戏、内容等领域。
- 元宇宙对传统社交的颠覆在于“数字拟真世界”，在连接范围更广、沟通效率更高、沟通方式更拟真的优势下，改变以往文字或视频的沟通交流模式，双方可以用更真实的数字形态进行交流互动。
- 虚拟社交与元宇宙社交的区别就在于社交形式的改变，虚拟社交突破了时间和空间的限制，拉近了社交的距离，而元宇宙社交在突破时间和空间的基础上又增加了沉浸感。

元宇宙社交的核心要素



社交角色

角色塑造是元宇宙社交的基础特点。在线上交流中，身体作为社交“媒介”的具象感被弱化了，而元宇宙社交将通过虚拟化身重新拾起这一部分。



社交空间

不同于被局限在屏幕之中的互联网交互，元宇宙增加了空间和时间维度，提供了一个全景式、开放式、全天候的社交空间。



社交场景

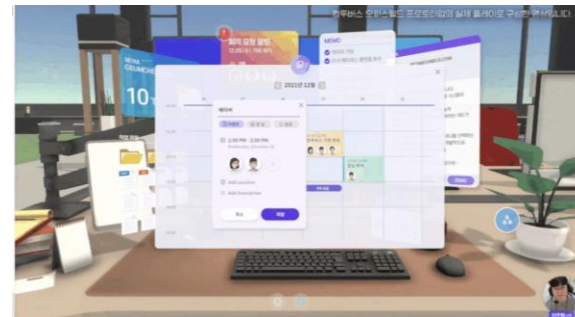
场景的搭建是实现元宇宙愿景的重要组成部分，拟真社交功能和可供创造的虚拟空间带来了与人交互相关的更多场景，也意味着在服务业、商业和工业领域落地的可能性。



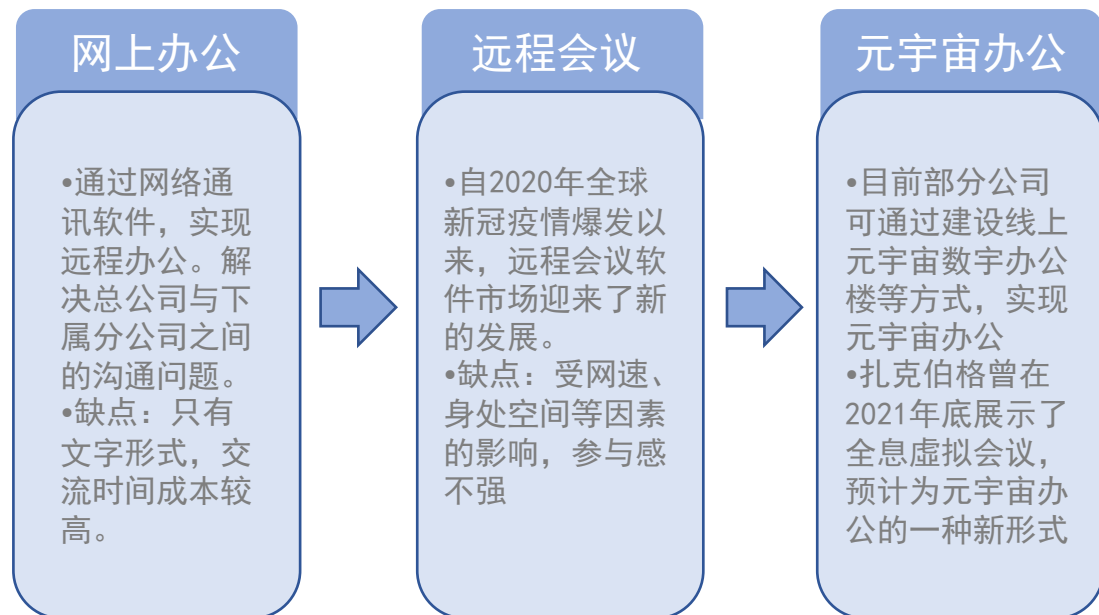
资料来源：Horizon Worlds，青亭网，映维网

办公元宇宙

- 工作场所将再次面临新的“进化”，对比我们当下的远程办公，元宇宙技术对于办公的最大改变就是增加了临场感和沉浸感。
- 与传统办公相比，元宇宙办公中所配备的虚拟化技术，让企业员工通过各种端口随时随地登录并访问元宇宙办公平台；不断提升的运行效率以及交互、协同能力，使得其可以完美地适用于各类复杂的办公场景，进一步提升办公效率。



在Office World办公中使用日程表管理行程



元宇宙办公发展历程



韩国初创公司Zigbang 推出了一款虚拟办公室产品“Soma World”，在这个虚拟世界中，有30层高的办公大楼、可容纳 3,000 人的会议中心、网络中心等和现实世界相仿的办公设施。



Com2uS号称首个实现“元宇宙上班”的公司，依托Com2uS去年底发布的《Com2Verse》平台，公司员工的打卡、开会、演讲均可以在元宇宙中进行。

- 现阶段元宇宙办公还不能完全替代实体办公，除了技术上的局限，长时间的虚拟办公员工的心理健康状况、人与人之间的联系和社群的建立都是未知的，很难衡量。

PART 02

硬件之“交互范式”

人机交互范式的变与不变

回溯互联网的交互——键鼠、触控、手势、语音

前瞻的交互方向：裸手交互、脑机接口、情景感知等

由虚实交互展望下一代硬件

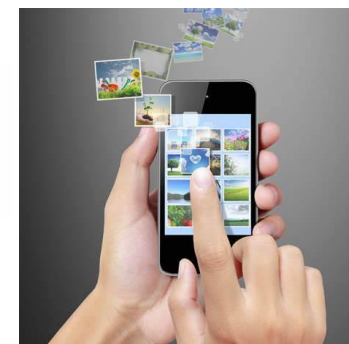
新硬件关键在全新体验，体验关键在交互

- 我们认为，下一代XR硬件不等于一系列硬件参数的升级，核心在于重新定义（产品端）和体验升级（用户端）。
- 重新定义（产品端）：以iPhone为例，初代iPhone完全抛弃键盘，而是采用全触摸屏幕，配合独特的外观设计、拟物化界面与重力感应等。
- 体验升级（用户端）：功能体验上，iPhone可看做是iPod+手机+Internet浏览器的融合产物，将手机从通讯工具升维为信息处理工具；交互方式上，由键盘转为触控，更符合人的操作习惯。
- 新的交互会是什么样？无法事前预测，需要被逐渐定义出来。

✓ iPhone重新定义智能手机

产品端：重新定义智能手机

用户端：更丰富的功能，更自然的交互



当前的VR/AR还无法达到“新硬件”的标准

- 下一代新硬件的核心在于重新定义（产品端）和体验升级（用户端）。
- VR: Quest 2已经是一款成功的VR产品（销量超千万台），但是距离“新硬件”（重新定义+体验升级）尚远。Quest 2是迄今为止销量最高的VR产品，双手柄的交互设计也基本奠定了当前的VR产品形态，但VR仍然存在过重、晕眩、适用范围有限等问题，消费端渗透率短期难以提高到手机的量级。
- AR: AR的应用范围更加广阔，长期看市场潜力更大，但AR产业链成熟度相比VR更低，光学、交互等环节尚存技术难点，目前2B/2G的AR成熟度更高，而消费级AR的功能主要局限于作为手机配件/扩展屏使用。



Quest 2

- ✓ 极高的性价比，产品体验较均衡；
- ✓ 渗透率与普及率仍然不足，局限于娱乐体验。

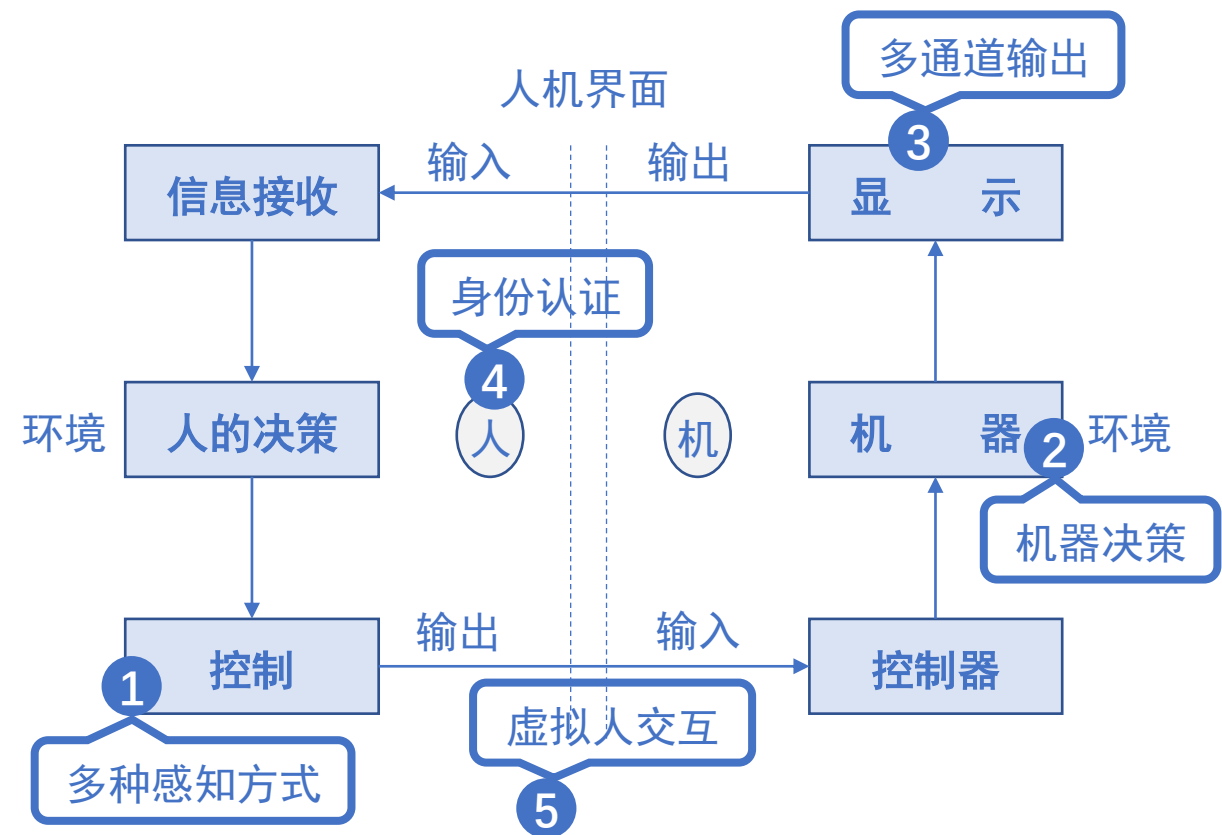


HoloLens

- ✓ 造价昂贵，多用于B端辅助生产、G端军队辅助训练；
- ✓ 语音、手势交互，不够自然。

交互的本质：信息感知-决策-输出

- 新技术背景下人机交互的信息感知与输出、信息处理和交互对象正在发生变化。
- 交互的本质仍然是信息感知-决策-输出。
 - 1. 信息感知：机器已经可以感知语音、手势等指令；
 - 2. 信息处理：和人的决策相对应，机器开始借助人工智能技术进行信息处理和决策；
 - 3. 信息输出：机器逐步普及语音输出通道，并出现简单的触觉信息输出；
 - 4. 身份识别：在人机交互的过程中对人进行身份认证；
 - 5. 交互对象：虚拟人开始作为交互主体出现在人机交互过程中。



PC互联网时代的交互——键盘+鼠标

- 个人电脑的结构基于冯·诺依曼体系结构的图灵计算机，包括运算器、控制器、存储器、输入设备以及输出设备。其中，输入和输出设备是人机交互的方式，输入设备包括鼠标、键盘、手柄，输出设备包括显示器、音箱、耳机。
- 关键产品的出现一举奠定个人PC的产品范式：1) 1981年，IBM公司推出第一台IBM PC，正式名称为IBM 5150型，成为第一台被业界广泛采用的PC，从此点燃了个人电脑市场的快速增长，并彻底改变了商业计算。2) 1983年，Apple Lisa 作为世界首款图形用户界面电脑问世，Lisa 的出现引领了操作系统的发展方向。苹果的Lisa机鼠标（A9M0050）是第一个完全商业化的鼠标，并投入到市场。



IBM PC，1981年发布，商业意义上第一款PC。



Apple Lisa，1983年发布，图形用户界面，键盘+鼠标输入。



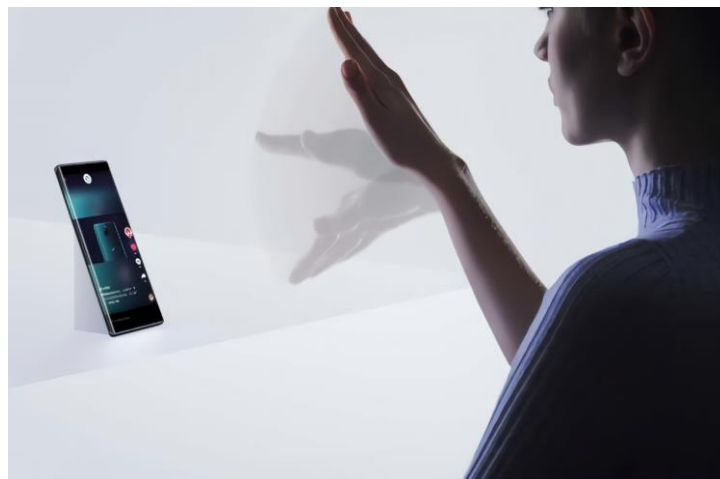
1995年，微软发布Windows95，微软操作系统外观基本定型。图形用户界面极大降低了用户门槛。

移动互联网时代的交互——触控、手势、语音

- 触控交互、手势交互和语音交互是当下智能终端的三种主要交互方式。其中，触控交互最为精简高效，已经成为最基本的交互方式，语音交互在IoT领域应用广泛，手势交互也越来越多地走进大众视野（这里的手势交互是指通过计算机视觉识别出的手势动作，无需和实体进行接触）。
- 交互效率：触控交互>语音交互>手势交互。触控交互最符合人的操作直觉，操作路径最简洁。语音交互中唤醒设备+清楚表达指令的过程通常会花费更多的时间。手势交互的指令性内容的表达同样需要更多时间。
- 学习成本：手势交互>触控交互>语音交互。语音交互最自然，触控交互已经普及，手势交互的学习成本最高。
- 指令准确性：触控交互>语音交互>手势交互。触控交互是三种交互方式当中唯一与载体发生接触的，准确度更高。



- ✓ 触控交互技术不断迭代，精度与灵敏度越来越高，由单点触控进化为多点触控



- ✓ 华为Mate30提供AI手势交互功能，作为核心卖点之一
- ✓ 手势交互的学习成本太高，偏离用户直觉，短期难以普及

触控交互效率最高，学习成本较低

- 一般来讲，当前的触控交互需要满足三项：1）触摸式屏幕，2）单指或多指直接接触该屏幕（或近距离接触，如果显示器具有邻近感应传感器并支持悬停检测），3）触摸接触时移动（或不移动，具体取决于时间阈值）。
- 相比鼠标、键盘、触笔交互，触控交互是更高级别的交互方式，有更高的交互效率和较低的学习成本。

因子	触控交互	鼠标、键盘、笔/触笔交互
Precision	指尖的接触区域大于单个 x-y 坐标，增加无意中激活命令的几率	鼠标和笔/触笔提供了精确的 x-y 坐标。
	接触区域的形状在整个移动过程中不断变化。	鼠标移动和笔/触笔笔划都提供精确的 x-y 坐标。 键盘焦点非常清晰。
	没有鼠标光标来帮助确定目标。	鼠标光标、笔/触笔光标以及键盘焦点都可以帮助确定目标。
人体解剖学	指尖移动并不精确，因为使用一个或多个手指沿直线移动非常困难。 这是由于手关节的曲率和运动涉及的关节数量导致的。	使用鼠标或笔/触笔进行直线移动就很容易，因为控制它们的手所移动的物理距离要比光标在屏幕上移动的物理距离短。
	由于手指的姿势以及用户对设备的控制，显示设备触摸表面上的某些区域可能很难接触到。	鼠标和笔/触笔可以达到屏幕的任何一个部分，同时任何控件都应该可以通过键盘按照 Tab 键顺序进行访问。
	一个或多个指尖或用户的手可能会遮住对象。 这称为封闭。	间接输入设备不会造成封闭。
对象状态	触摸使用两个状态的模型：显示设备的触摸表面为已触摸（打开）或未触摸（关闭）。 没有可以触发其他视觉反馈的悬停状态。	鼠标、笔/触笔以及键盘全都显示三种状态的模型：向上（关闭）、向下（打开）以及悬停（聚焦）。
		悬停允许用户通过与 UI 元素关联的工具提示来了解信息。 悬停和聚焦效果都可以传达哪些对象是交互对象，并且还可以帮助确定目标。
丰富交互	支持多点触摸：触摸表面上的多个输入点（指尖）。	支持单一输入点。
	通过手势（点击、拖动、滑动、收缩和旋转）支持的直接操作。	不支持直接操作，因为鼠标、笔/触笔以及键盘为间接输入设备。

手势交互适配非接触式用户界面TUI

- 用户界面的进化方向是从文本用户界面到图形用户界面 (GUI) 再到自然用户界面 (NUI)，非接触式用户界面 (TUI) 是自然用户界面的一个分支。
- 图形用户界面限制了用户需要使用键盘、鼠标等输入设备或触摸屏幕，而非接触式用户界面 (TUI) 是通过身体动作和手势命令计算，无需触摸键盘、鼠标或屏幕的过程，可以更自然的传递信息或与环境互动。

手势的分类	手势识别的技术实现
指示类：用于指示物体或方向的手势	技术本质：手势识别通过数学算法理解人的手势。大多使用传感器、相机、雷达等设备采集用户手势信息。
信号类：传达一些意义的手势或动作	传感器：典型案例是可穿戴设备
情景演绎类：无对象执行的特定任务演示	相机：深度感知相机，由于它们的近距离能力，可以有效地检测手势
标志类：用于传达特定对象或实体店信息的手势	雷达：微型雷达捕捉细微的手部动作，利用无线电波的反射成像
操控类：手势可以在较短的反馈循环中引导运动	

语音交互是最自然的交互方式之一

- 在触控交互、手势交互和语音交互三种交互方式当中，语音交互1) 是最自然的交互方式，学习成本几乎为零，2) 语音交互也是非接触交互，不需要直接接触载体，3) 语音交互中AI的参与程度非常高。因此，我们认为语音交互会是下一个时代广泛应用的交互方式之一。
- 一次完整的语音交互需要经历ASR→NLP→Skill→TTS的流程。ASR，用于将声学语音进行分析，并得到对应的文字或拼音信息，一般分训练和解码两阶段。NLP，用于将用户的指令转换为结构化的、机器可以理解的语言。Skill即AI时代的APP，作用是处理NLP界定的用户意图，做出符合用户预期的反馈。TTS，语音合成，从文本转换成语音，让机器说话。



XR时代的交互——接近自然的交互

- 人机交互演进的一个原则是，不断接近人类现实世界信息获取和流动的方式。
- 区分于真实世界的2D内容/场景，元宇宙当中的3D内容/场景是一种格式上的升维，在时间和空间两大维度重构。
- 无论是PC还是智能手机，交互集中于人体的上半身，聚焦于视觉听觉的所听所感，但元宇宙强调高度的沉浸感，这就要求除了上半身的人体肢体也参与到交互当中，VR/AR/MR/智能汽车等新硬件就必须承载更多肢体参与、更接近自然状况的交互形态。
- 3D界面的交互：3D界面有着区别于2D界面的特殊属性，包括但不限于1) 交互空间：空间大小、平面类型、遮挡关系，2) 输入方式：语音、手势、控制器、外部传感器，3) 呈现信息类型：3D模型、图片、2D界面，3) 音效：空间音效、界面反馈音效，4) 信息与环境融合方式：无中生有、真实物体与虚拟信息叠加、遮挡真实环境。（5点要素由 8ninths 团队整理而成）

2D界面

- ✓ PC桌面操作：键盘、鼠标、点击Click
- ✓ 智能手机移动操作：单/多点触控、滑动、语音
- ✓ 视觉、听觉、触觉（震动反馈）

VS

3D界面

- ✓ Outside-in追踪，6DoF
- ✓ 空间大小、类型、遮挡
- ✓ 手柄输入
- ✓ 视觉、听觉、触觉（震动反馈）、手势、眼球、面部

XR时代的交互——接近自然的交互

- 按照器官及输入方式的不同，交互方式大致可以划分为头部运动追踪、眼动追踪、面部/表情识别、控制手柄、手势识别、全身动捕、脑机接口等。

器官	输入方式	定义	应用情况
头部	头部运动追踪	对用户头部方向和位置的定位	是VR/AR中最基本的输入手段，也是一种自然的输入方式
	眼球追踪	一般由眼动摄像机、光源和算法共同完成	新一代VR/AR设备逐步加入眼动追踪
	面部/表情识别	分析用户的面部肌肉运动	新一代VR设备逐步加入眼动追踪
	声音输入	语音聊天、VR音效、语音控制等	AR很多采用语音输入，部分VR设备搭载语音识别功能
	脑波控制	将大脑电波作为输入信号	仍处于实验室阶段
手部	控制手柄	通过对手柄的位置追踪，可以实现更自然、更具沉浸感的互动体验	Oculus、HTC、索尼等都有自己的体感手柄
	手势识别	直接识别手部骨骼与节点	代表性产品之一Leap Motion，但还不精确
	力反馈手套	通过空气的膨胀来为手部提供虚拟现实的感官触觉	尚未广泛应用
	数据手套	手套上安装有多个感应点，因而可以更精确地识别手部动作	相当于手势识别的另一种形态
	触摸板	通过滑动和点击位于头显侧面的触摸板来控制VR中的行为	三星Gear VR和PICO 1采用了这种输入方式
肢体	全身动作捕捉	通过在身体各个运动关键部位穿戴感应器，可以实现对用户全身动作的检测识别，并将动作信息在虚拟世界中表现出来	主要应用于B端游戏和影视制作中
	跑步机	能解决用户在VR场景中较大范围走动及跑动的需求，也能通过“所动即所感”来解决晕动症问题	小范围出货
体感	蛋椅/赛车/飞行模拟仓	好的体感设备可同时实现VR的输入和输出功能，能够给用户提供最好的体验	主要应用于VR线下体验使用
综合	位置追踪/可穿戴设备	结合以上各种交互方式形成的新的交互设备	

颠覆式创新总是诞生于人机交互的变革

- 颠覆式创新 VS 渐进式创新：渐变式创新着眼于技术，改良现有产品；而颠覆式创新着眼于产品，直接面向需求创造出新的更好的产品。——李善友
- 复盘数字经济百年创新史，堪称颠覆式创新的产品屈指可数，但几乎每次颠覆式创新的源头都是人机交互的变革。

第三次颠覆式创新：

谁能创造出更好的产品、提供更佳体验？

第二次颠覆式创新：

乔布斯在手机上再次突破边界，基于触屏交互技术革新性推出iPhone，成为即个人电脑后电子消费品领域第二次颠覆式创新。

第一次颠覆式创新：

个人电脑的早期是字符交互，乔布斯和比尔盖茨基于图形用户界面开发的麦金塔和Windows是电子消费品领域的第一次颠覆式创新。

VR典型的交互范式：一体机+手柄

- **头盔：从分体式VR到VR一体机。**以Oculus Quest为首的6DoF VR一体机推出以来，整个VR行业的产品形态逐渐转为以一体机为主。据VR陀螺统计，2020年开始，VR一体机发布数量陡增，第一次超越分体机；2021年，VR一体机数量几乎为分体机的一倍。
- **手柄：从手指追踪到手势追踪。**VR手柄最开始主要为棒状，代表是HTC Vive以及PSVR的手柄，后有Oculus与Valve进一步优化，推出Oculus Touch和Valve Index的Knuckles，向更优秀的手部追踪迭代进步。



Oculus Touch

- ✓ 支持空间定位和体感追踪功能。
- ✓ 类似HTC Vive的环形设计，不遮挡LED。
- ✓ 具备“电容传感电路”以及“接近感控制”两大黑科技，能够识别用户大拇指、食指和中指的不同姿态。



Valve Index Knuckles

- ✓ Index手柄的代号为Knuckles。
- ✓ 配备宽大的塑料跟踪条、操纵杆、小型触控板、几个面部按钮和一个可挤压的辅助输入。
- ✓ 内置87个传感器，能更精确、更全方位地触摸、运动，以及捕获压力和光学数据。

VR典型的交互范式：一体机+Portal

- 2022年11月，小派科技推出了一款创新型VR设备——Pimax Portal。据介绍，Pimax Portal 是一款4合1通用产品，可在掌上游戏设备、面对主流人群的VR头显、便携式PC游戏系统以及客厅娱乐中心之间切换。
- Pimax Portal是VR终端设备的创新与尝试，有助于进一步扩大用户群体，但本质上还是没有脱离VR的范畴，属于“老技术、新玩法”。



✓ 游戏掌机：Pimax Portal 的核心为一个搭载高通骁龙 XR2 的安卓游戏掌机，最高配备 4K 144Hz QLED 屏。

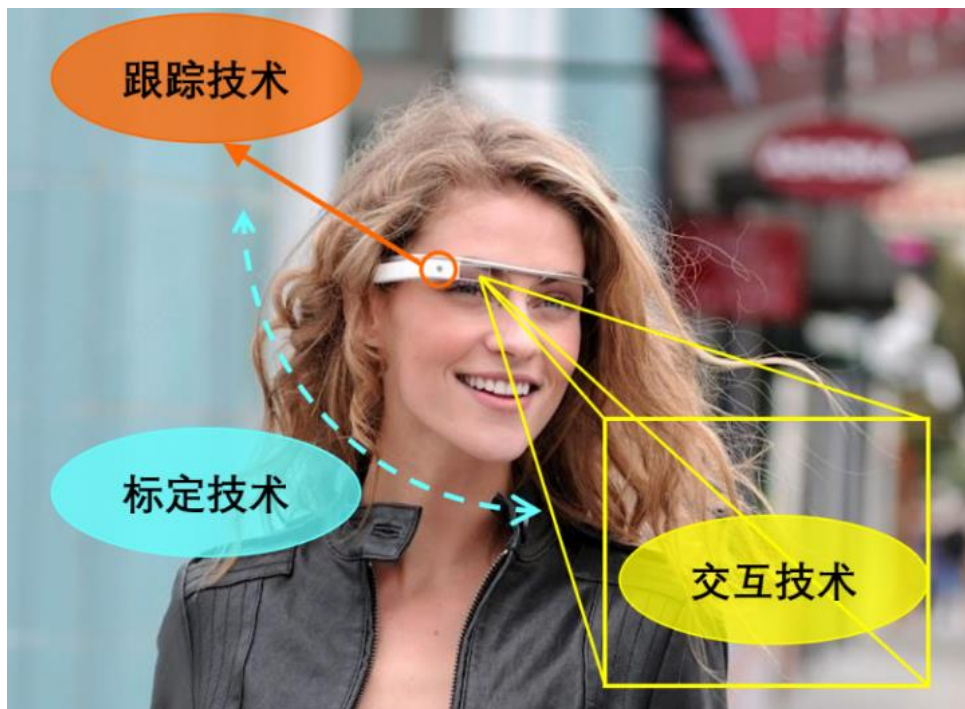
✓ VR形态：屏幕和手柄部分可拆开，分别放入 VR 头显和 VR 手柄中使用。屏幕部分配备 6DoF 追踪，直接变成 VR 设备，VR 头显部分还可以更换不同 PPD 和视角的版本。

✓ 便携式PC游戏系统：Pimax Portal XL 模式下，掌机能连接一个搭载 AMD R7 6800U 的 Mini Station，该硬件配有 32GB + 1TB 存储，可通过 WiGig 无线传输 PC 游戏到掌机游玩。

✓ 客厅娱乐中心场景：Pimax Portal 掌机可插入一个底座中，从而实现体感游戏的游玩，还能当作机顶盒观看视频。

AR典型的交互范式：标定、跟踪、交互

- AR眼镜中涉及到的技术包括：1) 标定技术，帮助用户看得准；2) 跟踪技术，帮助用户走得稳；3) 交互技术，实现信息多交换。三类技术彼此相关、共同协作。



标定技术：鉴于光学透视型的AR眼镜的特殊原理，我们往往不能直接获得人眼在目标空间中的坐标位置，而需要对人眼和AR眼镜的跟踪系统进行标定，实现“物理环境”-“跟踪系统”-“人眼成像”这条通路，保证人眼能看到准确的虚实融合效果。

跟踪技术：AR眼镜想要实现物理环境与虚拟元素的融合，必须要做到的就是能够跟踪虚拟环境。主流的视觉跟踪技术包括基于图像标记点的跟踪（Image-based Tracking）、基于物体的跟踪（Object-based Tracking）和同步定位与地图构建（SLAM）。

交互技术：AR眼镜比较方便的交互方式是语音、手势和遥控器。语音交互属于人工智能的重要研究领域，其中科大讯飞等已经做出了比较出色的语音交互产品，这对于AR眼镜的交互具有重要意义。手势交互因为不需要引入额外的交互设备，只需要用户做出手势就可以输入指令，也在一定程度上促进了AR眼镜交互技术的发展。遥控器作为传统的输入手段，虽然简单但是却能保证准确高效的输入，很多AR眼镜产品还是配了遥控器或者触控板作为交互设备。

AR典型的交互范式：眼镜+语音+手势

- 目前市面上流行的光学透视型的AR智能眼镜主要有2种：1) 一体式智能眼镜。这类眼镜的计算单元、电源、显示单元都集成在一起，技术难度高，价格昂贵。一体式智能眼镜的典型代表是HoloLens；2) 分体式智能眼镜。这类眼镜通常将电源和计算单元独立为一个类似手机的处理单元，但是将显示单元做成眼镜。分体式智能眼镜的典型产品是Magic Leap One、NED+ Glass X2、Nreal Light等。
- 语音交互即通过语音完成对AR眼镜的操控和输入。用户与AR眼镜间通过语音可以完成从简单到复杂的指令，复杂一点的语音指令需要人工智能，包括NLP的介入。
- 手势交互的实现，需要AR设备配备多种深度传感器来感知环境与用户的手势，用户直接通过手势控制虚拟的AR数字内容。手势交互一定程度上摆脱了控制器的束缚，让交互体验更加自然。

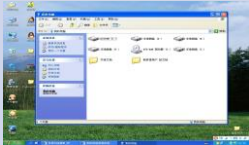


HoloLens：手势型交互

- ✓ 手势交互的定位机制是凝视。凝视和空中点击(Air tap)结合起来构成了凝视与行动这种交互方式。
- ✓ HoloLens目前能够识别两种核心手势——空中点击和Home手势。
- ✓ 复合手势是通过将点击、保持、释放和手的位移进行结合，HoloLens可以表现出含义更复杂的复合手势。

WMIP四大元件的更迭：PC-Phone-AR-VR

- WIMP是由“视窗”（Window）、“图标”（Icon）、“选单”（Menu）以及“指标”（Pointer）所组成的缩写，其命名方式也指明了它所倚赖的四大互动元件。

	PC	Phone	AR(HoloLens)	VR
W (Windows)				
I (Icon)				
M (Menu)				
P (Pionter)		N/A		
(Pointer 控制)				

资料来源：CSDN，《2D 到 3D 的跨越，探索最佳 AR 交互设计之道》

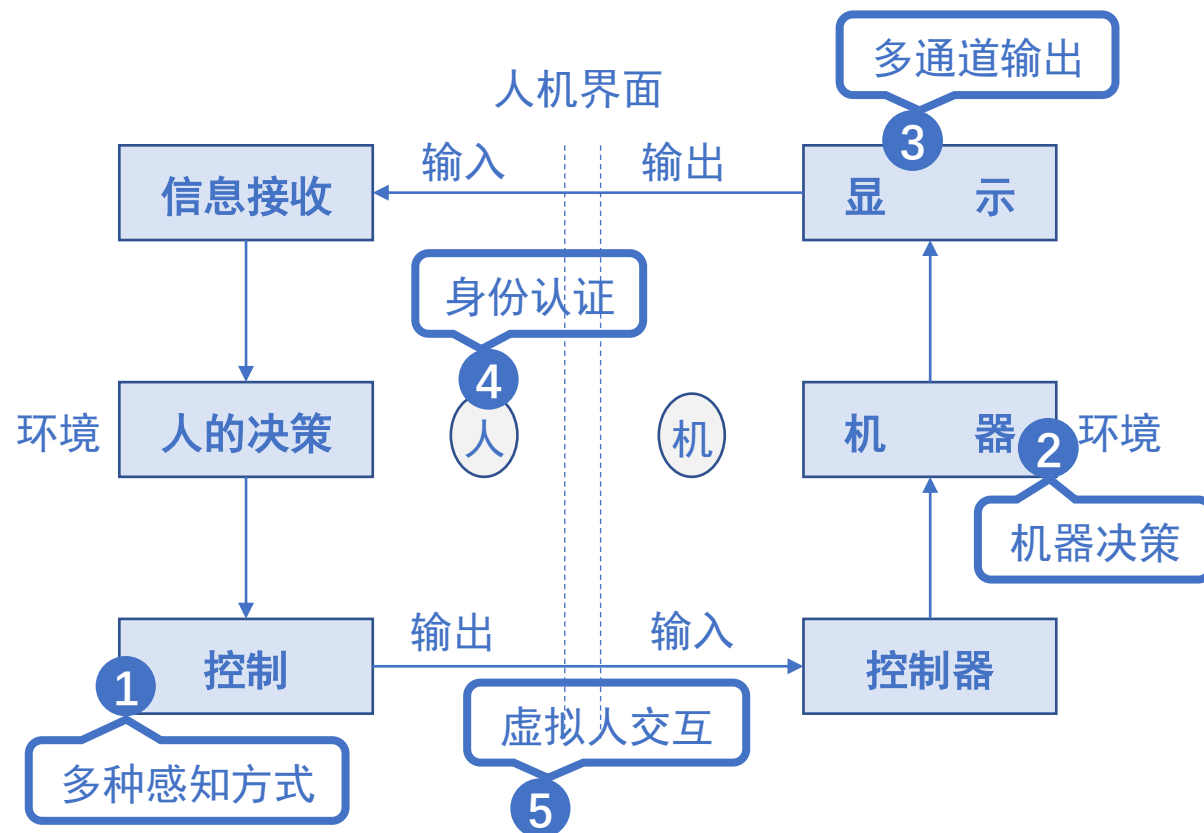
人机交互的变与不变：变的是范式，不变的是本质

- 新技术背景下人机交互的信息感知与输出、信息处理和交互对象正在发生变化。

- 交互的本质仍然是信息感知-决策-输出。

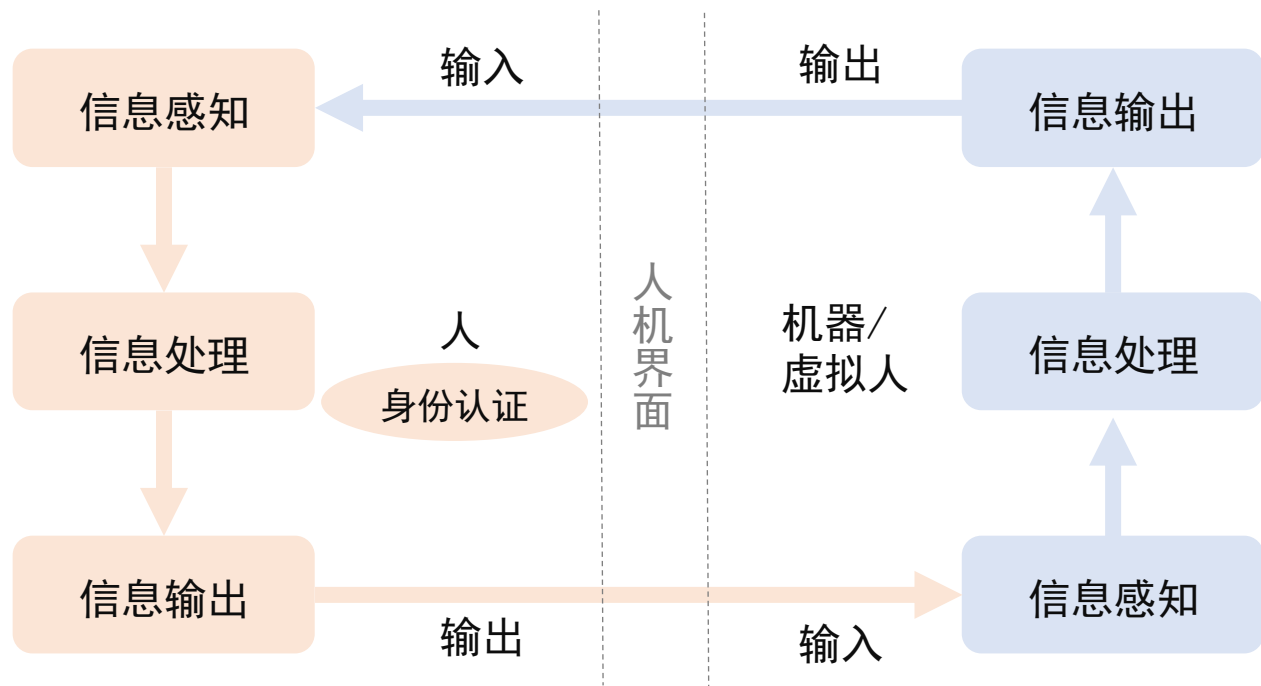
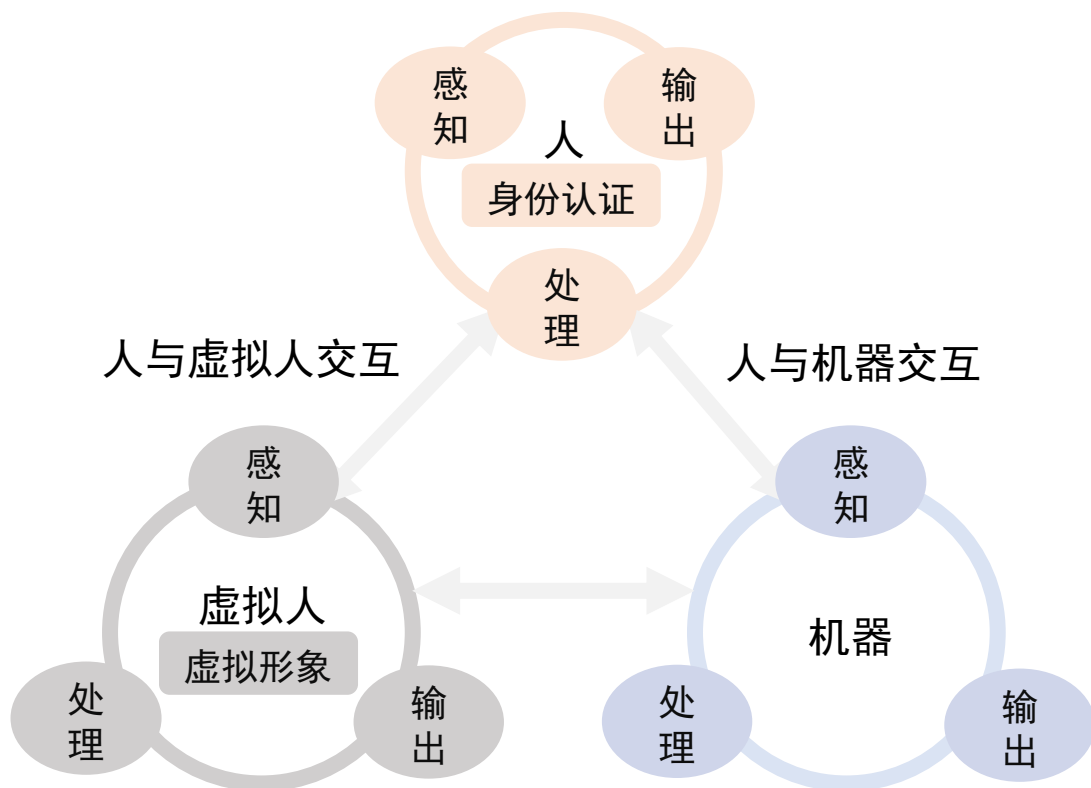
- ✓ 1. 信息感知：机器已经可以感知语音、手势等指令；
- ✓ 2. 信息处理：和人的决策相对应，机器开始借助人工智能技术进行信息处理和决策；
- ✓ 3. 信息输出：机器逐步普及语音输出通道，并出现简单的触觉信息输出；
- ✓ 4. 身份识别：在人机交互的过程中对人进行身份认证；
- ✓ 5. 交互对象：虚拟人开始作为交互主体出现在人机交互过程中。

图：传统的人机交互流程



未来人机交互的主体将拓展为人-机器-虚拟人三方

- 未来人机间将建立起人-机器-虚拟人三方互动的交互模型。
- 目前人与机器交互正在经历转变，人与虚拟人交互处于起步阶段。
- 每种交互模式均遵循信息感知-决策-信息输出的交互流程。



人与机器的交互：Tesla Bot人形机器人

- Tesla AI Day 2022会上，Tesla Bot发布。我们认为，Tesla Bot之于机器人行业与汽车行业都是一个颠覆式创新的角色。
- Tesla选择人形双足双手方案，是因为其定位是生产力机器人。



人与机器的交互：拟人化是为了“通用”

- 随着人机交互技术的发展与演进，机器人与人机交互体验均需要重构。
- 对于机器人而言，人工智能机器人是社会发展的第三代机器人，第三代机器人将具有识别、推理、规划和学习等智能机制，可以把感知和行动智能化结合起来，因此能在非特定的环境下作业，故称之为智能机器人。
- 对于人机交互体验而言，拟人化的机器人能够主动交流、情感交互、协同智能。拟人化存在一定的伦理风险，但真正的目标是为了实现机器人的“通用”。



小米CyberOne

- ✓ 身高177cm
- ✓ 体重52kg
- ✓ 自由度21
- ✓ 有双手，无十指
- ✓ 视觉Misense视觉空间系统
- ✓ 人工智能：视觉传感器、听觉传感器、音频算法、语义识别、情绪识别



Optimus

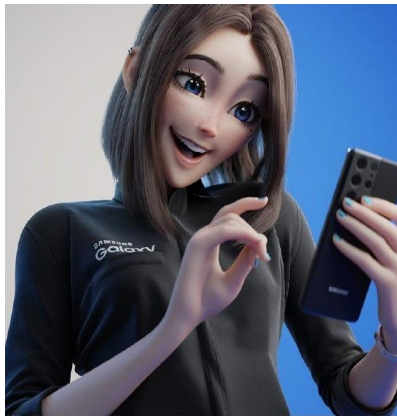
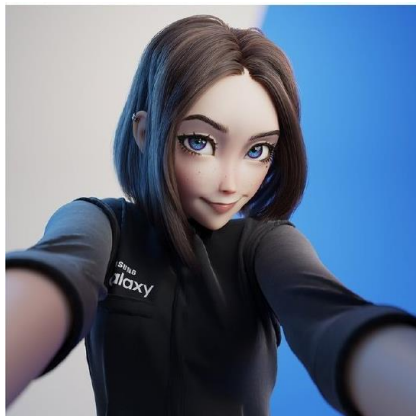
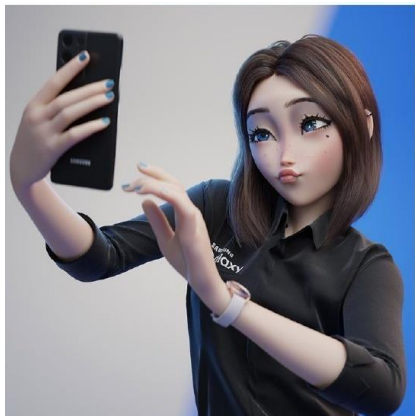
- ✓ 身高172cm
- ✓ 体重56.7kg
- ✓ 自由度40
- ✓ 有双手，有十指
- ✓ Autopilot cameras
- ✓ 人工智能：自动驾驶FSD系统，搭配多摄像头视觉架构的深度神经网络架构

拟人化：

- 1、外观像人，双足双手
- 2、具有通用智能

人与虚拟人的交互：AI加持下趋于无界

- “多模态交互”：融合了人的视觉、听觉、触觉等多种感官，计算机利用多种通信通道响应输入，并充分模拟人与人之间的交互方式。现被广泛运用到各种智能产品的交互设计中，人类通过语音、手势、表情等模态输入信息，计算机通过计算机视觉、听觉等通道响应就是典型的多模态交互。
- 进化之前：语音助手Siri、小爱同学。
- 进化之后：多模态交互的智能助手，具象化的三星虚拟助手Sam。

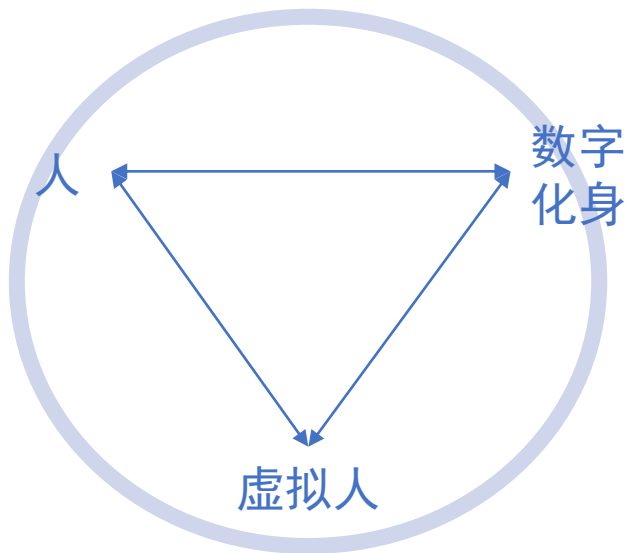


三星虚拟助手Sam

- ✓ 由三星旗下营销公司Cheil Worldwide，与巴西视觉艺术制作公司Lightfarm共同打造。
- ✓ Sam与一般语音助手最明显的区别，在于它（或者她）首次有了自己专属的虚拟形象。

人与虚拟人的交互：人设是关键和灵魂

- 人有人格，虚拟人也有人设。语音助手是没有人设的，人与语音助手的交互是单向的指令；虚拟人有人设，人与虚拟人（具有人设）的交互是双向的“交流”。
- 人与虚拟人的交互按照对象的不同可以分为三类：人与人的数字化身，人与虚拟人，数字化身与虚拟人。
- 虚拟人的制作可以拆分为两端，一端是技术端，打造虚拟人的模型与形象，另一端是运营端，树立虚拟人的人设和行为。



- ✓ A-Soul虚拟偶像团体，字节跳动负责开发，乐华娱乐负责运营。
- ✓ 五位成员嘉然、贝拉、向晚、珈乐、乃琳拥有各自的人设。

VR与AR在交互领域相通，目前处于拓荒期

- 长期来看，VR和AR设备在交互领域是相通的，目前处于拓荒期。操作界面正在从二维平面转移到三维空间，但是还未形成类似PC时代下鼠标键盘、移动时代屏幕触控等公认的交互范式。
- 据华为VR领域技术专家张梦晗，将VR/AR的交互范畴划分为以下三类：
 - ✓ 一类是由人主动发起的交互动作，比如手部交互、体感交互等等，这类交互往往发出的都是确定性的指令性信息，要求准确达成某些特定的目的，也会耗费最多的精力和体力。
 - ✓ 还有一类则是由机器主动发起的交互动作，比如眼球跟踪、脑电机电、头部位置跟踪、表情识别等等，这类交互是机器来感知使用者的状态来进行交互，很多时候使用者自身不需要牵扯太多精力。
 - ✓ 再有一类是机器对环境的感知交互，比如三维重建、情景感知、物体分割识别等等，很多时候尤其在AR中，只有先做好对外在环境的感知，人与机器之间的交互才能顺畅。

交互探索方向——裸手识别

- 裸手识别就是不需要任何介质的手部交互，具体案例包括Leap Motion（基于计算机视觉），Google Soli（基于毫米波雷达），此外还有一些基于超声、红外的解决方案。当前基于计算机视觉的裸手交互发展得最成熟，效果也最好。
- VR中非裸手交互更有优势，AR中反之。VR中3DoF或6DoF手柄已有广泛应用，非裸手具备优势，VR中人和虚拟物体和环境交互，通过某个媒介能提高交互的真实感。而在AR里，人不仅仅要与虚拟物体交互同时也能看到现实场景和物体，裸手情况在AR中更多。

计算机视觉方案



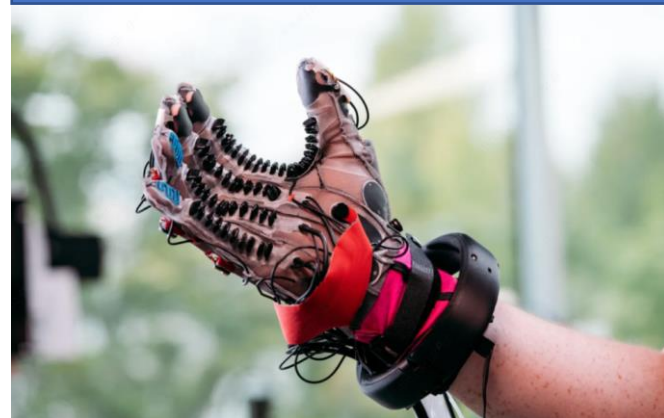
- Leap Motion，微软HL2的手部追踪，Oculus Quest的手部追踪，代表计算机视觉手部追踪的最高水准。

肌电控制方案



- 肌电信号反映的是肌肉收缩所产生的生物电，用户个体肌肉生理差异大。该技术领域的代表公司为CTRL—Labs，已被Meta收购。

数据手套方案



- Meta Reality Lab公布一款数据手套，触觉交互位置被布置在贴合佩戴者手掌、手指下侧和指尖等位置上。

交互探索方向——体感交互

- 体感交互是通过3D体感摄影机（体感互动专用的智能投影机）的采集、追踪、控制完成的。首先，3D体感摄影机发射红外线，从而对整个房间进行立体定位，摄像头则可以借助红外线来识别人体的运动。然后，动作感应器能追踪你全身的动作，并且会根据数据建立用户的数位骨架，它可以对人体的20个部位进行实时追踪。此外，体感互动系统还能利用即时动态捕捉、影像辨识、麦克风输入、语音辨识等功能丰富操作模式。
- 区别于传统的界面交互（WIMP），体感交互强调利用肢体动作、手势、语音等现实生活中已有的知识和技能进行人与产品的交互，通过看得见、摸得着的实体交互设计帮助用户与产品、服务以及系统进行交流。

体感游戏



智能“魔镜”



智能驾驶控制



资料来源：微软Xbox官网，新浪VR

交互探索方向——表情识别

- 表情识别的技术底层包括面部捕捉、眼球追踪以及人工智能。
- 实时面部捕捉存在技术难题。实时面部捕捉的运行过程大致如下，人的表情先被传感器追踪到，经数据处理后，通过高速宽带网络传输至社交对象的VR设备中并被在其虚拟环境中重现。追踪脸部表情并在虚拟画面中重现是同步的，所以其技术实现难度较大，主要在于：追踪与重现（映射到虚拟化身）的延迟、设备数据处理能力、追踪传感器的敏感精度、社交网络的带宽等等。
- 此外，表情识别对于提升虚拟世界社交的真实感很关键。人类在社交中的信息传达只有30%是通过语言实现的，剩余70%则是通过表情和肢体动作实现。



——MindMaze社最新面部识别VR组件《MASK》

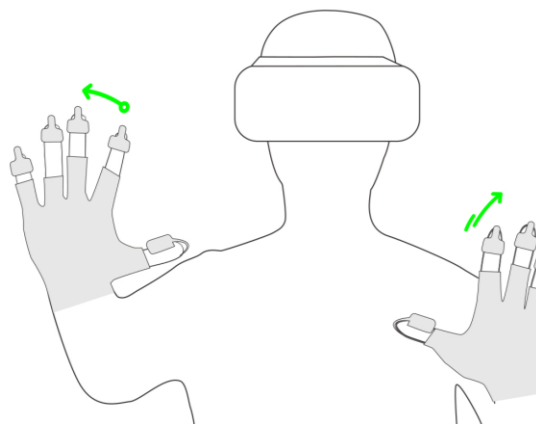
- ✓ 据映维网，MindMaze社提供的最新解决方案是“在VR眼镜周围安置高灵敏电极，通过捕捉玩家的面部肌肉动作产生的电气信号来还原识别面部表情。”
- ✓ 传统的表情识别解决方案是通过摄像机拍摄用户的面部然后去捕捉识别。这一方案的缺陷在于永远有延迟，难以同步。

交互探索方向——数据手套

- 数据手套是指通过在手套内置传感器，采集手部运动数据的硬件设备。
- 使用硬件设备操作虚拟信息：鼠标+键盘，数据手套。
- 数据手套相比手柄，与真实世界的互动更加真实。有别于传统的VR控制手柄，数据手套通过与VR设备相连接运用手套上的弯曲感应传感器实时记录并模拟手部动作，让佩戴者通过手部活动实现与虚拟物品的真实互动。
- 力反馈设备模拟真实物品感受。而由于力反馈设备的加入使得数据手套在现在能够让佩戴者充分感受到虚拟物品的材质以及触感。

手势识别

更加便捷的操作方式，给用户全新的VR体验。



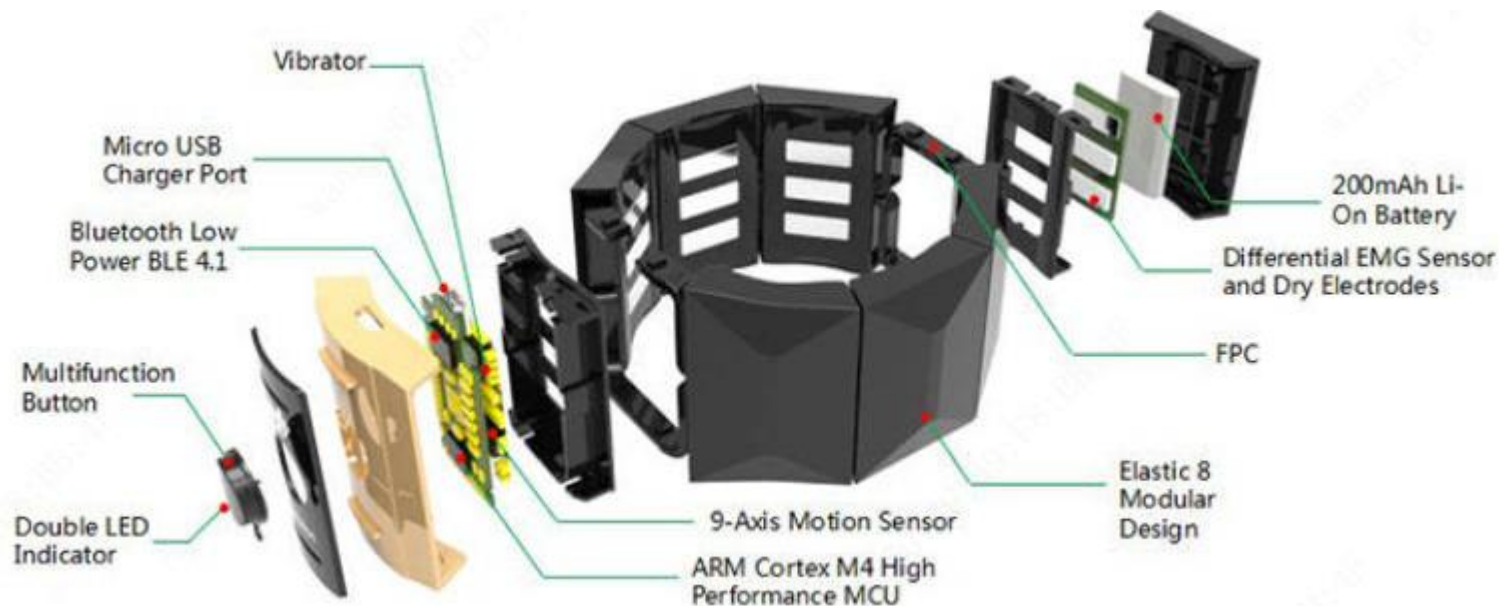
触控反馈技术

结合游戏内容，提供震动反馈，提升用户沉浸感。



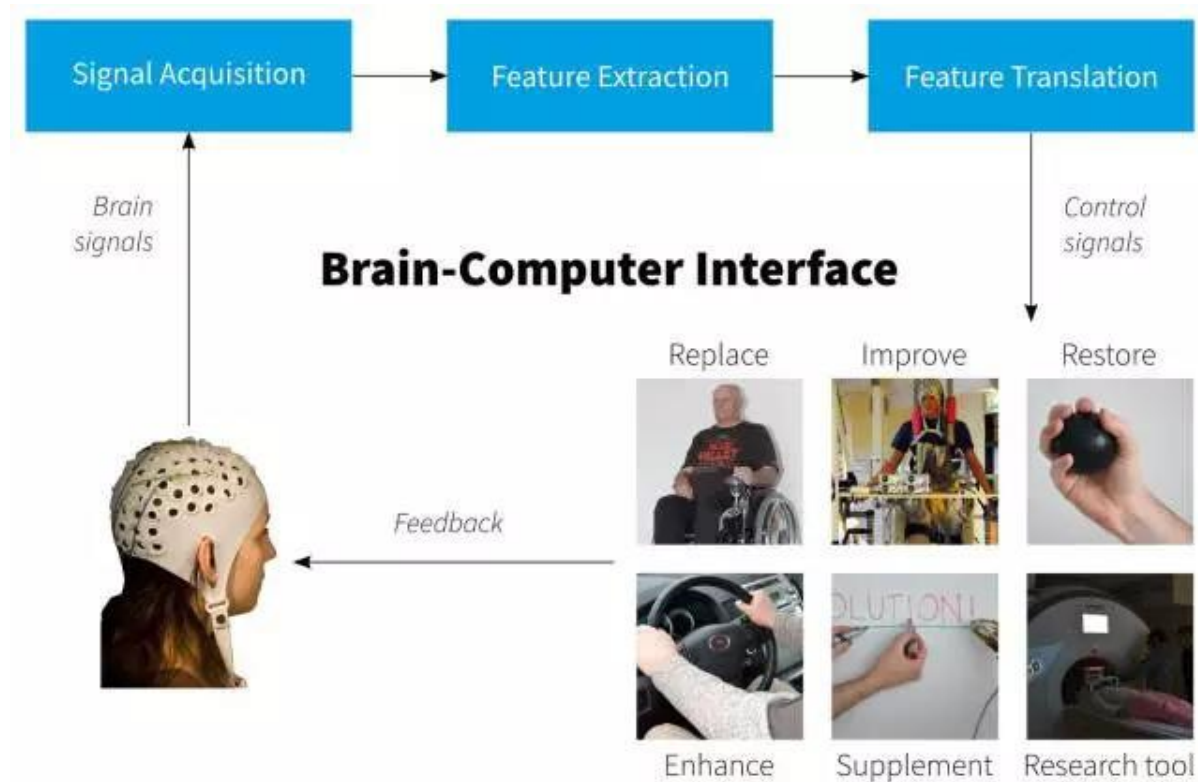
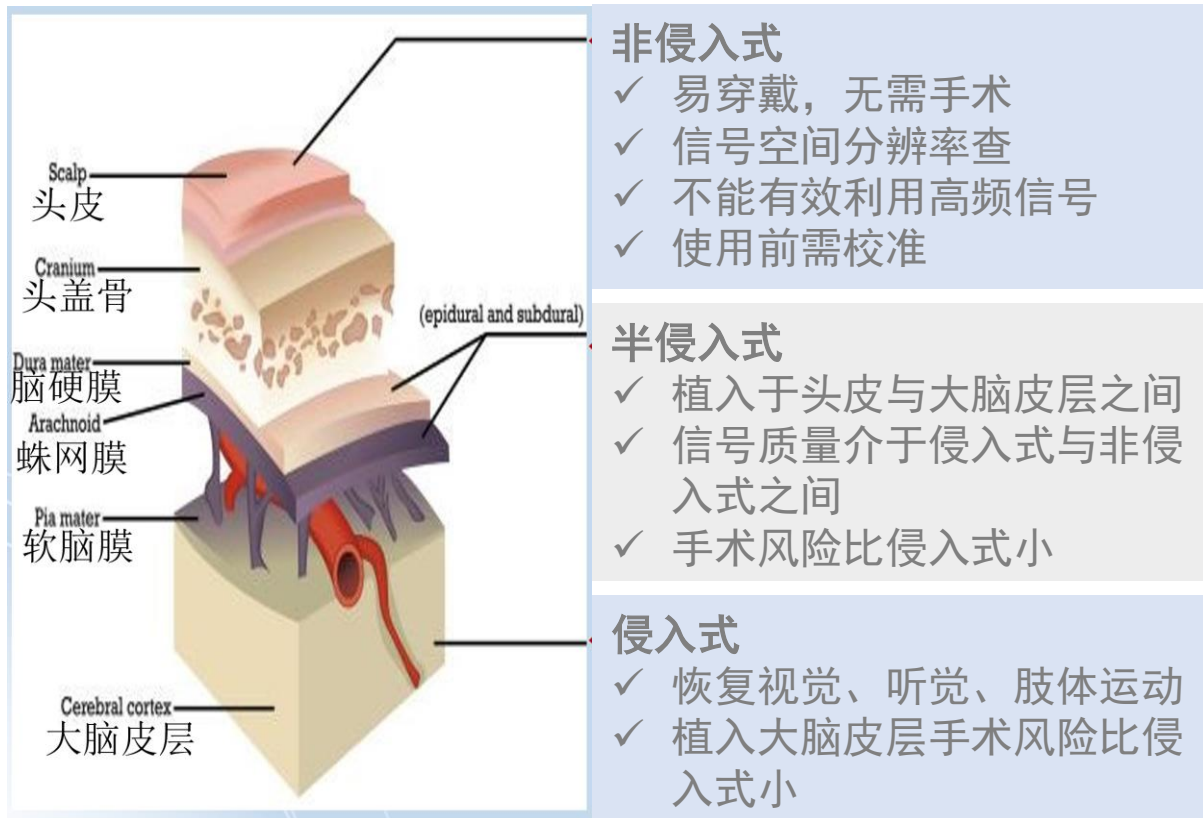
交互探索方向——肌电信号

- 肌电信号，据维基百科给出的解释，当肌肉细胞被生物电或者神经激活时产生的电能。将单个或者多个肌肉细胞或者部分肌肉组织活动时产生的生物电变化，经电极引导，放大、记录和显示所获得的电压变化的一维时间序列信号图形，就是肌电图（electromyogram，简称EMG）。
- 肌电手环是使用肌电信号控制设备的仪器。通常的肌电手环从获取信号到变成功能，需要经历：获取肌电信息——处理信息（降噪、放大等）——翻译处理器能够识别的信息——传输到终端——翻译成具体操作——输出等复杂流程，不仅每一步都需要大量运算，跨平台数据传输，而且在过程中还可能产生误差。



交互探索方向——脑机接口

- “脑机接口”的定义=“脑”+机“+”接口”。即，在人或动物脑（或者脑细胞的培养物）与外部设备间创建的用于信息交换的连接通路。“脑”意指有机生命形式的脑或神经系统，而并非仅仅是“mind”（抽象的心智）。“机”意指任何处理或计算的设备，其形式可以从简单电路到硅芯片到外部设备和轮椅。“接口”=“用于信息交换的中介物”。
- 脑机接口同样符合交互的本质，包括信息的感知-决策-输出。



- 情景感知的构建过程是：传感器提供有关现实世界中活动和事件的数据。感知算法理解这些刺激并将其在情境中予以分类。根据观察到的情境，系统的动作将被触发。
- 情景感知当前应用较为成熟的两大领域是智能家居与可穿戴设备。未来的情境感知将更加智能化，设备会主动与周边的世界进行交互，成为用户与周边世界互动的门户。通过分析用户所处环境、状态甚至情感的信息，运用眼动追踪、触觉反馈等技术看懂和识别身体语言或手势，并更有预见性的做出相应准备。

情景感知+智能家居

- ✓ 情景感知应用环境中智能设备的情境感知概念是指“智能化家居主要处于一种具备流动性、开放性、动态性的智能化空间环境当中”。
- ✓ 基于用户家庭环境改变情况，智能家居需要合理地对待侦搜的情境信息进行针对性的理、创设。
- ✓ 智能家居除了具备学习、模仿功能外，还能够通过分析环境及相关因素变化，对用户的真实需要进行预测，从而实现自主工作。

情景感知+可穿戴设备

- ✓ 智能手机和平板电脑将最终支持情景感知功能，但可穿戴设备更有优势。
- ✓ 可穿戴设备中的感应器突破智能手机感应器监测功能，将成为情景感知平台最为重要的数据来源之一。
- ✓ 情景感知的个性化具体指：给消费者他们想要的东西，知道他们何时、何地需要这种东西，它将虚拟在线世界和现实世界连接起来。

由虚实交互展望下一代硬件

- 进化的路径与阶段：虚拟建设-由虚向实-虚实交互（虚实融合）。三个阶段中，硬件载体相应变化。

虚拟建设

虚拟建设（现已起步，VR为主，AR尝试）：云计算、AI运算、网络等基础建设高速发展，游戏、影视娱乐、社交等内容不断丰富，硬件载体VR设备率先落地，提升沉浸感，由此吸引更多用户加入元宇宙，虚拟世界加速建设。

根据Mob研究院数据，2020年人均使用手机时长5.72小时/天（非睡眠时间占比约36%），预计未来将进一步提升，其中部分转化为虚拟世界沉浸时间。这个阶段硬件设备以VR设备为主，AR硬件及初级应用开始出现。

初期涉及领域：游戏、影视娱乐、社交、广告等领域率先落地，硬件载体以VR为主，AR尝试，计算方式以端为主，云为辅。

由虚向实

从虚拟到现实（五年维度，MR融合，AR落地）：虚拟世界建设丰富，元宇宙的发展不再局限于以娱乐和社交为主的虚拟世界，而是反馈现实世界，提高生产效率。

实现了虚拟现实与现实的初步融合，进而拓展了轻办公等丰富的应用场景。虚拟现实设备也继续增强沉浸感，市场需求不断增加。虚拟现实/MR设备的用户规模和使用时间将显著提高用户粘性。AR在VR的基础上发展，真正落地（社交、健康、时尚、互动等）。随着基础设施和技术水平的不断提高，AR可以实现更直接、更自然的虚实融合。

中期涉及领域：拓展电商、教育、旅游等轻度办公场景，硬件载体以MR融合，AR落地，计算方式以云为主，端为辅。

虚实交互（虚实融合）

虚实融合（十年维度，VR沉浸，AR常在）：用户未来在虚实融合世界的时间占比有望超过80%，为提升真实世界生产效率，AR/VR并行发展。VR成为专家场景，作为生产力工具以拓展更多应用，更强调沉浸式应用场景；AR成为主流场景，覆盖生活各个角落，专注强交互场景。

我们判断，未来AR设备有望成为类似手机、手表的主流强粘性终端（作为眼镜长时佩戴）；而VR设备继续发力强沉浸场景，设备使用频率和粘性进一步提升。

远期涉及领域：拓展医疗、工业、设计等重度办公场景，硬件载体为VR沉浸，AR常带，计算方式为云。

由虚实交互展望下一代硬件

虚拟世界

真实世界

万物互联



PICO 4&4 Pro：力求走通消费级VR叙事

产品名称	PICO Neo 3	PICO 4	PICO 4 Pro
产品示意图			
重量	395g	295g	
电池容量	5300mAh	5300mAh	5300mAh
分辨率	单眼 1832*1920	单眼2160*2160	单眼2160*2160
PPI	773	1200	1200
视场角	可视角98度	可视角105度	可视角105度
无级瞳距调节	不支持	不支持	支持
镜片方案	菲涅尔镜片	Pancake	Pancake
感知/交互	4*鱼眼单色摄像头	1600万像素RGB全彩摄像头、 4*SLAM灰度跟踪摄像头	1600万像素RGB全彩摄像头、 4*SLAM灰度跟踪摄像头
手柄	6DoF，红外光学	6DoF、宽频线性马达、红外传感器、全 手掌握把、真实触觉/振动反馈	6DoF、宽频线性马达、红外传感器、全手 掌握把、真实触觉/振动反馈
裸手识别	支持	支持	支持
面部识别	/	/	支持
眼球追踪	/	/	支持
彩色透视	/	16MP全彩透视	16MP全彩透视
价格	1999元	2499元（128G）/2999元（256G）	/

Quest Pro：强调MR，升维切入办公场景

- Quest Pro与Quest2相比，光机厚度减少40%，PPI增加37%，对比度增加75%。1) 芯片：采用Snapdragon XR2+处理器，这是第一款与高通深度合作开发的芯片，算力提升50%且散热更好。2) 光学：用Pancake折叠光路技术，光机厚度减少40%。3) 显示：采用具有局部高光功能的LCD类型面板（MiniLED），还包含一个量子点层以改善 LCD 的色域。4) 交互：手柄中新的传感器能自追踪手柄位置不需要头显介入，可以360°转动使用，手柄还搭载新的触觉触控，能够给更精准的反馈。5) 全彩透视、眼动追踪、面部捕捉：Quest Pro搭载高分辨率摄像头传感器，获取4倍于Quest2分辨率的透视影像；有一颗红外深度传感器切换Passthrough和全彩模式。头显内部有5枚摄像头，用于眼动追踪和面部追踪。
- Quest Pro的产品定位是办公场景，其设计重点考虑协作与生产力。Meta拟深度合作微软加速未来办公，计划引入微软多生产力工具，包括Mesh for Microsoft Teams、Microsoft 365、Microsoft Intune和Azure Active Directory等。定价1500美元。

图：Quest Pro示意图及核心参数



Meta Quest Pro参数

分辨率	相比Quest2，PPI增加37%、PPD增加10%，中心视图清晰度提升25%，色域范围扩大1.3倍
光学	Pancake折叠光路方案，光学深度减少40%
显示	两个局部高光功能的LCD类型面板（MiniLED），还包含一个量子点层
芯片	采用Snapdragon XR2+处理器，算力持续性提升50%且散热更好
全彩透视	搭载高分辨率摄像头传感器，获取4倍于Quest2分辨率的透视影像
眼球追踪	支持
面部捕捉	支持
定价	1499.99美元，10月25日出货

图：Microsoft Teams Meeting in VR



PSVR2：成熟内容生态是优势也是限制

- **硬件方面**，采用 4K 120Hz HDR 显示屏、110° 视场角和注视点渲染技术，其 OLED 显示屏将提供每眼 2000×2040 的分辨率和 90/120hz 刷新率，内置眼动追踪功能、重新设计的手柄、触觉反馈以及各种其他改变游戏规则的增强功能。此外，PSVR2支持透视视图。索尼的PSVR2透视视图可以让玩家保持佩戴头显的同时观察周围环境，这要归功于头显前部嵌入的摄像头。索尼高级产品经理Yasuo Takahashi解释说：“当用户想要轻松检查PSVR2 Sense控制器在房间内的位置时，它会派上用场，而无需摘下头显。”
- **内容方面**，PSVR2的护航游戏作品将达30多款，包括《地平线：山之呼唤》、《黑相集：Switchback VR》、《生化危机8》VR模式、《莫斯1&2》和《穿越火线：Sierra Squad》等。
- PSVR的最大优势来源于成熟而丰富的内容生态，但这也一定程度上限制了PSVR的场景局限于娱乐与体验，难以破圈。



《The Dark Pictures: Switchback VR》
《Crossfire: Sierra Squad》
《The Light Brigade》
《Cities VR》增强版
《Cosmonious High》
《Hello Neighbor: Search and Rescue》
《Jurassic World Aftermath》系列
《Pistol Whip VR》

HoloLens 2：主攻B端场景，搭载Windows系统

- HoloLens的交互方式简单自然。HoloLens以真实自然的方式实现全息图的触摸、抓握和移动。使用 Windows Hello，只需使用虹膜信息即可立即安全登录 HoloLens 2。通过智能麦克风和自然语言语音处理，它甚至可以在嘈杂的工业环境中执行语音命令。
- HoloLens在在艺术、工业、汽车、娱乐等方面均有广泛应用。其中，HoloLens目前最大的应用场景是与工业的结合，比如远程协助、工业维修等。



- ✓ 光学：透明全息透镜(波导)
- ✓ 分辨率：2k 3:2 光引擎
- ✓ 全息密度：>2.5k 辐射点(每个弧度的光点)
- ✓ 基于眼睛位置的呈现：基于眼睛位置的3D显示优化
- ✓ 头部追踪：4台可见光摄像机
- ✓ 眼动追踪：2台红外摄像机
- ✓ 深度：1-MP飞行时间(ToF)深度传感器
- ✓ IMU：加速度计、陀螺仪、磁强计
- ✓ HoloLens 2搭载了特殊版Windows，并引入了新的手势交互系统，支持在空气中点击、缩放，操作系统，并且 HoloLens 2 还加入了人眼追踪功能，用户可以通过注视来完成一些操作。

苹果MR：最有希望重新定义交互

- 苹果的首款MR设备或将1) 自有操作系统realityOS（据App Store上传日志和苹果开源代码）；2) 显示器的配置将包括两个微型OLED显示器和一个AMOLED面板（据显示器分析师Ross Young）；3) 芯片将采用代号Staten的M2芯片衍生版本，外加一个协处理器Bora芯片（据芯片供应链专家@Mobile chip Expert）；4) 将配备14个摄像头，可以追逐头、手、甚至腿部（Quest pro 10个摄像头，PIC04 5个摄像头）；5) 价格被预测为3000美元级别，需要额外的月度订阅（据TrendForce），Quest pro 1500美元。
- 苹果最有希望重新定义交互：据 The Information，预计苹果的设备将支持虹膜识别。这样用户只需将设备戴在头上，就可以快速登录自己的账户。该功能将有助于苹果多用户系统的出现，哪怕好几个人使用同一设备也不会产生影响。

图：苹果MR产品概念图



图：虹膜识别是一种生物特征提取



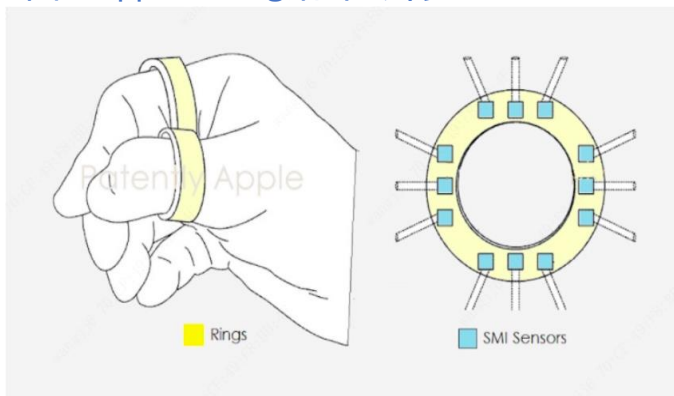
新硬件的交互范式亟待重构，猜想苹果MR的三种交互

- 第一种，基于Apple Watch的交互。Apple Watch的Assistive Touch功能可通过检测手的握紧和捏紧以及基于加速度计的控制来实现对手表的单手控制。该系统看起来与Apple和Facebook在XR输入上所做的工作非常相似。
- 第二种，基于Apple Ring的交互。自2019年以来，苹果已经为可戴在手指上并用于控制VR或AR眼镜的设备申请了多项专利。Apple Ring描述了一种带有SMI传感器的指环。它们发出的激光被物体反射并干扰传感器中的光。这些干扰信号可用于测量到附近物理物体的距离。SMI Sensors传感器可以在空间上定位并检测附近的物体。因此，VR或AR眼镜可以准确地知道手指在哪里以及它如何在空间中移动。Apple Rings与SMI传感器，相比目前的控制手柄，重量和体积都明显减小。
- 第三种，基于手套的交互。苹果的VR手套能够测量手指和拇指骨骼的运动，这款由智能织物技术制成的VR手套可用于许多应用，包括游戏、教育和军事训练。在苹果这项被授予的专利中，可以使用一个或多个电子设备来模拟VR / AR / MR技术。一种电子设备可以是VR头戴式设备，用户可以在其中使用VR头戴式设备查看模拟的虚拟环境。另一电子设备可以包括VR手套，VR手套可以戴在用户的手上，并且可以允许用户实时触摸、感知和握持虚拟对象，这可能需要能够准确地检测用户手和手指的位置以及运动。

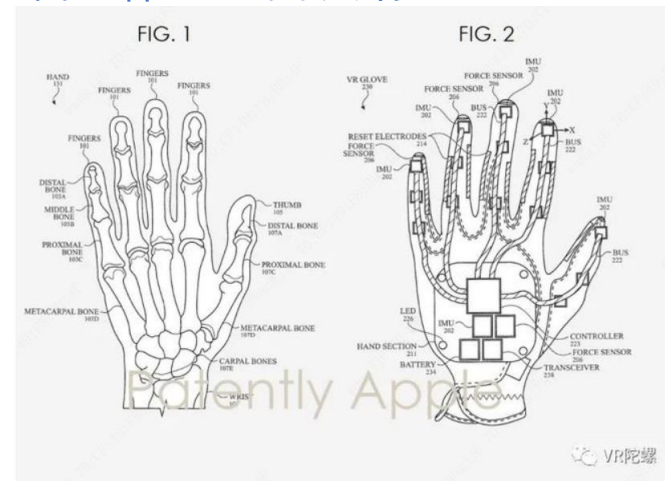
图：Apple Watch支持腕带式输入



图：Apple Ring指环专利



图：Apple VR手套专利



裸手交互正成为VR/AR的一个重要交互方式，尤其是基于腕带的裸手交互

Meta: 根据肌电图信号来实现裸手输入的腕带式设备；

Apple: 两只手或两根手指接触的方法和系统——“微手势”来控制设备的方法，如将拇指贴在食指的指根关节，并将其作为一种虚拟摇杆来进行方向选择，这与Meta的腕带方案接近；苹果的方案在腕带式设备的基础上增加手套，手套搭载的四个感测电极350A-350D可布置在手掌的大约四个角处，并可通过相应的感测电路感测，启用多方向滑动手势，例如向上滑动、向下滑动、向左滑动、向右滑动和对角滑动等从而提升感知精度。

交互范式的定义之所以重要，**还有一个原因**：“后面的”场景应用模式内容将快速得以呈现——未来AI基本上触及了一切，用户只需要描述出来就可以构建整个世界，相关技术将允许一个有想法的普通用户实现自己的愿景。

基于新一代硬件/计算平台，开发无处不在的输入技术

- 1) 可用于一天中遇到的各种情况；
- 2) 这个系统需要负责任地构建，从一开始就考虑隐私和安全，为人们提供有意义的方式来个性化和控制AR体验；
- 3) 界面同时需要直观、始终可用、不突兀和易于使用；
- 4) 理想情况下，它同时需要支持丰富的、高带宽的控制，适用于从操纵虚拟对象到编辑电子文档的所有操作；
- 5) 另外，它同时需要一个足够舒适的形状参数，可以穿一整天，并且足够节能，可以长时间使用。

基于上述5点，有两条思路变得非常清楚

- 1) 当时没有任何接近于所述标准的现有产品；
- 2) 任何最终出现的解决方案都必须穿戴于手腕。

为什么是手腕？

其他同样有用的输入源，声音十分直观，但由于背景噪音的影响，它在公共领域不够私密或不够可靠；诸如智能手机或游戏控制器这样的独立设备则需要存储在口袋里，从而在你和你的环境之间增加了一层摩擦。

当探索潜在的可能性时，在手腕放置一个输入设备成为了一个明确的答案：

- 1) 手腕常用于佩戴手表，已经对这一点已经非常**习惯和熟悉**，这意味着它可以合理地适应日常生活和社会环境；
- 2) 这是一个舒适的位置，**适合全天穿戴**；
- 3) 它就在你用来与世界交互的主要工具的周围：**手**，这种**接近度**将允许我们把丰富的AR控制能力带给你的双手，并实现直观，强大，令人满意的交互；
- 4) 可以轻松作为计算、电池和天线的**平台**，同时支持广泛的传感器阵列。

缺失的一块拼图是**寻找一条通向丰富输入的清晰路径**，而一个潜在的理想解决方案是肌电图：

- 1) 肌电图（EMG）利用传感器将通过手腕传递到前掌的电机神经信号转换成数字命令，并允许你用来控制设备的功能，所述信号可以允许你将清晰的一位元命令传达给设备；
- 2) 这是一种高度个性化的控制，可以适应多种情况；
- 3) 通过手腕的信号非常清晰，肌电图可以理解手指仅一毫米的运动，这意味着输入可以不费吹灰之力，就像点击一个虚拟的、随时可用的按钮一样轻松，而且最终甚至可以感觉到大脑移动手指的意图。

利用肌电图直接在桌面或膝盖高速打字

一开始，肌电图只提供一位元到两位元控制，我们称之为“点击”，它相当于点击一个按钮。这是基于动作的手势，比如拇指和食指的捏和放。走路、说话或就坐、双手放在身体两侧或口袋里，无论你在哪里或在做什么，捏放拇指和食指都非常容易执行。这样的操作永远可行，无需口言一字，从而使之成为第一个无处不在的超低摩擦AR交互方式；

但这只是第一步，肌电图最终将发展成更丰富的控制，最终，非常有可能能利用肌电图直接在桌面或膝盖高速打字，而且速度甚至可能比键盘的速度更快。

利用神经界面来支持你直接控制机器

利用神经界面来支持你直接控制机器，利用负责手和手指肌肉的周围神经系统的输出，特别是脑外神经。这样理解：你拍摄了大量的照片，但选择只分享其中的一部分；同样地，你会有大量的想法，但你只选择执行其中的一部分。当这种情况发生时，大脑会向手和手指发送信号，告诉它们以特定的方式运动，以便执行诸如打字和刷卡之类的动作。重点是在手腕位置解码所述信号（即你已经决定要执行的动作），并将它们转换为设备的数字命令；

与在手机屏幕点击选择歌曲、单击鼠标或在键盘上键入内容相比，这种根据意图来执行操作的方式会更为迅速；

神经界面的目标是改变人机交互的长久历史，并使之成为人类对机器具有更多控制权的开始。

过往的交互外设也会改变

如键盘：它可以随着时间的推移学习并适应你独特的打字风格，包括打字错误等；结果是键盘将慢慢地向你靠拢，与你合二为一，而不是与世间相同的物理键盘；这将比任何机械式打字界面都快，而且始终可用，因为你就是键盘；虚拟打字和点击等控件的优点在于，我们已经能够熟练地使用它们。

同时需要一种方法来闭环反馈回路：触觉

允许系统与用户进行反馈，让虚拟对象感觉有形，这就是触觉发挥作用的地方——用手和手指执行各种操作，然后你的手和手指会获得各种感觉；我们已经进化到利用这种触觉信号来理解世界——触觉允许我们使用工具和实现精细控制、触觉同时可以传达不同的情绪；在正确的语境中，不同类型的触觉反馈可以对应流行的表情emoji，这可能是一种可以实现更佳社交的新奇又好玩的方式；大脑会将视觉、听觉和触觉刺激结合起来，并赋予虚拟体验全新的维度；基于触觉维度的增加，一个全新的空间刚刚开始向我们开放，而许多可能都是从手腕丰富的触觉系统开始。

PART 03

硬件之“操作系统”

PC操作系统的演进——Windows、Mac OS

智能手机操作系统的演进——Symbian、Android、iOS

由此推演XR操作系统的竞争格局？

——寡头垄断格局，Apple、微软、谷歌、Meta？

PC操作系统发展Milestone

- 个人计算机面世之初，人机交互最早的形式是命令行。
- Apple II问世。在70年代末开创了个人计算机的产业。
- IBM兼容机+DOS操作系统。每次操作计算机，都要从记忆中抓取命令字符，准确输入，才能准确输出。
- Apple Macintosh问世，图形用户界面（GUI）开始成为主流。图形用户界面彻底改变了计算机信息呈现和交互。鼠标来回轻点，随便拖曳窗口，极大地降低了操作难度。
- Macintosh使用独立的Mac OS操作系统，微软研发的Windows操作系统发力兼容机市场，二者合力推动了个人电脑的大规模运用，也改变了人类的生产和生活方式。

PC操作系统发展初期由硬件厂商主导

- 90年代初期的个人电脑市场主要由三部分构成：1) 来自IBM的IBM PC以及OEM厂家的IBM PC克隆机：运行MS-DOS、Windows3.0或是OS/2操作系统；2) 苹果Macintosh：运行苹果System 6.0/7.0操作系统；3) 其他PC厂家（Atari、Commodore、Amiga、NeXT）。
- 此时操作系统主要由硬件厂家主导，硬件厂家可以一手包办硬件设备与软件操作系统。

IBM PC

- IBM PC与其克隆机占据了主要的市场份额，这部分电脑运行微软MS-DOS操作系统。1985年开始，微软与IBM合作研发OS/2操作系统，旨在代替MS-DOS。
- 1990年双方合作破裂，微软发布Windows 3.0。IBM独立推进OS/2。
- Windows成为IBM PC与其克隆机的主要操作系统。

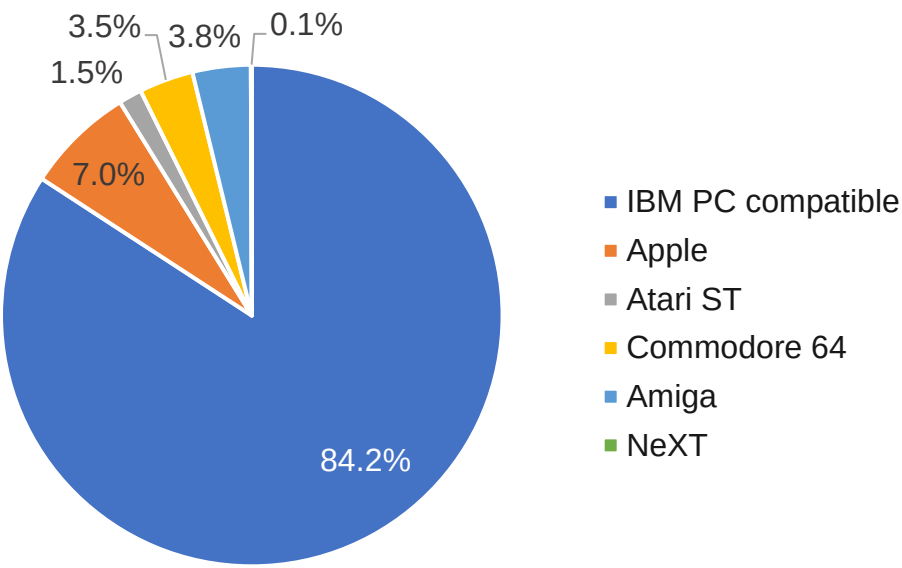
APPLE

- 90年代初期，苹果处于经营困难时期。
- 1996年，乔布斯回归苹果，主持发布了Mac OS 7.6，开创Mac OS 品牌。

其他PC厂家

- 任务逐渐复杂，需要越来越多软件与驱动支持，中小厂商的自研系统逐渐淘汰。
- 1991年，李纳斯托瓦兹在类UNIX系统的基础上开发出Linux。开源自由免费的Linux迅速代替UNIX占领了专业应用市场。

图：1990年个人电脑市场份额



资料来源：Statista

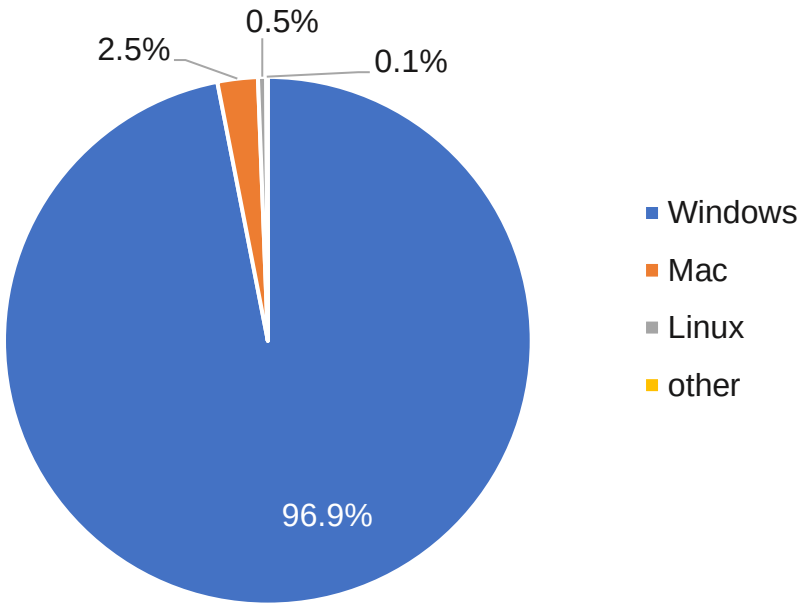
一超（Windows）两强（Mac OS、Linux）格局形成

- 微软、苹果双双发力。微软发布Windows 95, Windows 98, Windows Me, Windows XP, 巩固市场统治地位。微软的Windows和包括Word、Excel和PowerPoint在内的Office平台的巨大成功，令其占据主导地位。苹果发布第一代iMac与Mac OS 9，随后推出Mac OS X，令苹果在台式机市场重新占据一席之地。。
- 个人电脑操作系统“一超（Windows）两强（Mac OS、Linux）”的格局初步形成，其他厂商的操作系统基本被淘汰。

表：1990-2001年微软与苹果主要个人电脑操作系统发布情况

系列	名称	发布年份	备注
Windows消费者用户	Windows 3.0	1990	
	Windows 3.2	1994	首个中文版Windows
	Windows 95	1995	NT4.0内核，搭载IE1
	Windows 98	1998	
	Windows ME	2000	NT 4.9内核
	Windows XP	2001	NT 5.1内核
Apple消费者用户	System 7.0	1991	
	Mac OS 7.6	1997	乔布斯回归，更名
	Mac OS 8.0	1998	
	Mac OS 8.5	1999	
	Mac OS 9.0	1999	iMac G3发布
	Mac OS X	2001	

图：2000年个人电脑操作系统份额



资料来源：OneStat

一超（Windows）两强（Mac OS、Linux）格局稳固

Windows

- 2007年，微软发布Windows Vista，激进的技术路线尝试阻碍了Windows Vista的推广。
- 2009年，微软发布Windows 7。这是迄今为止最成功的个人电脑操作系统。截止2021年，仍然有18.3%的电脑在运行Windows7。

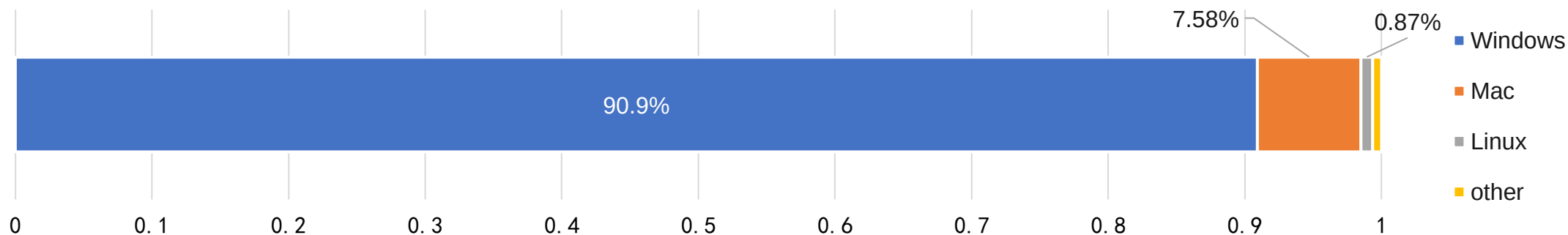
Mac OS X

- 2008年，苹果发布Mac OS X 10.6，该版本系统只支持x86架构。至此，苹果放弃了使用了近20年的PowerPC芯片，x86架构统一个人电脑市场。
- 同时，iPod、iPhone等产品的走红成功带动了苹果电脑的销售，Mac OS X的市场份额从2004年的2.6%增加到2012年的7.6%。

Linux

- 2004年，Linux Ubuntu发布。同年，中国企业也推出了Linux Deepin。
- 成熟的商业化Linux系统已经可以满足基本的办公需求。
- Linux系统可以更好的兼容RISC指令集的芯片，打破一些国家对x86架构芯片的垄断，从而吸引许多对安全性有要求的政企用户。

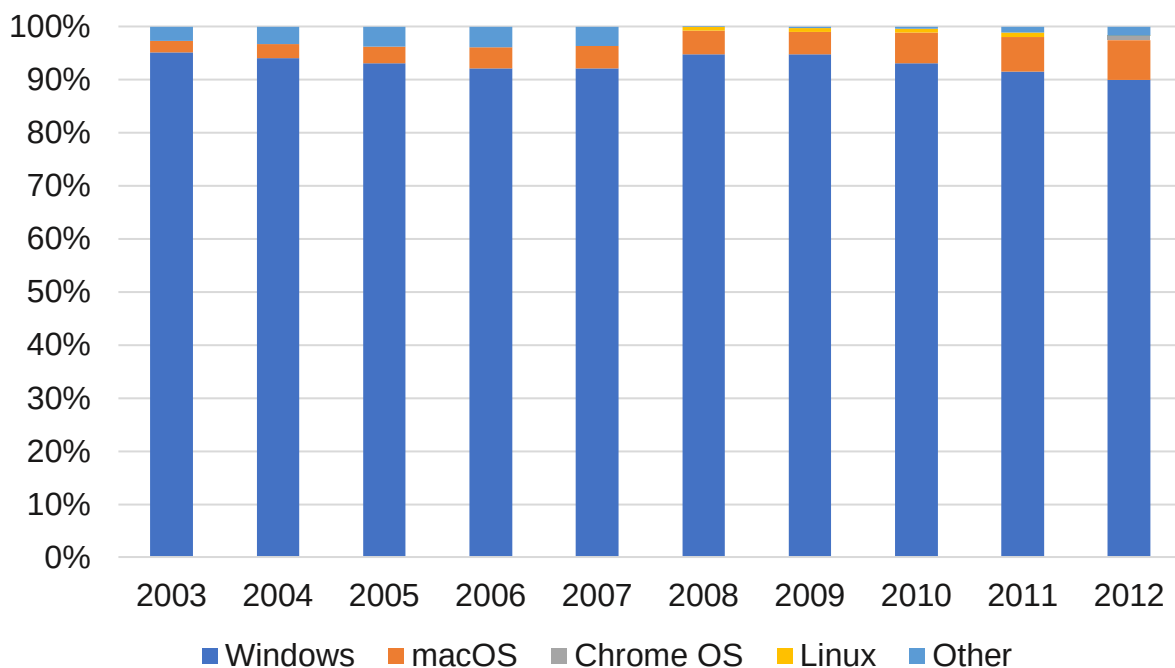
图：2012年个人电脑操作系统份额



一超（Windows）两强（Mac OS、Linux）格局稳固

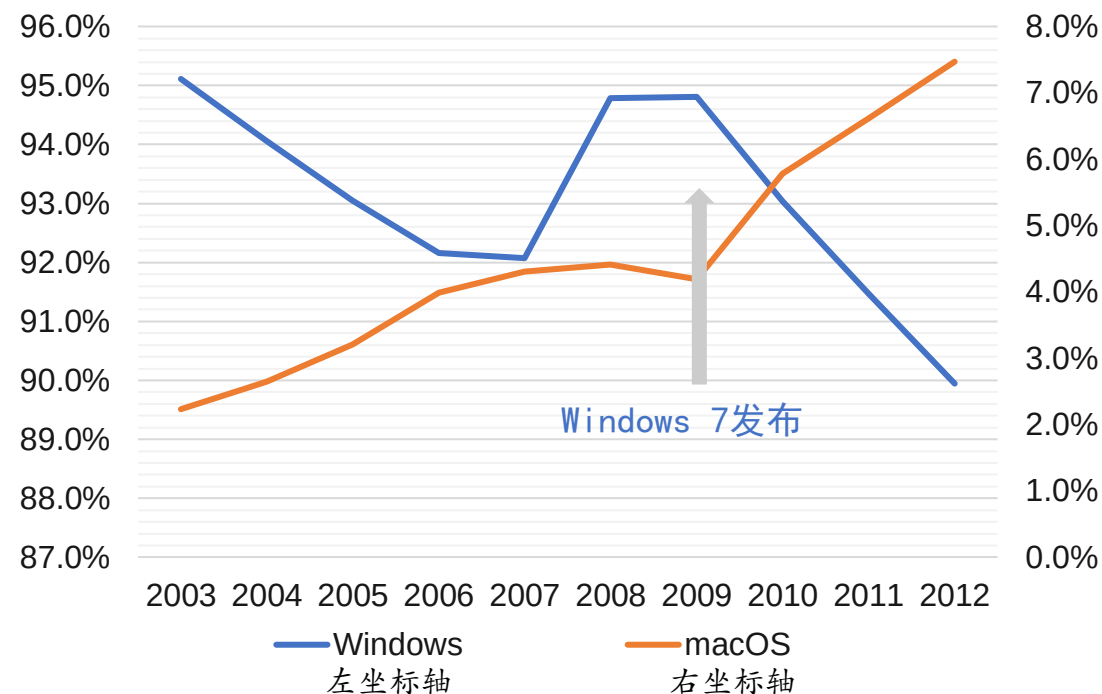
- 一超：Windows的统治地位稳固，Windows7及以后版本连续成功，常年占据90%以上的市场份额。
- 一强：苹果手机、电脑产品热销，助力Mac OS仍占据一席之地。
- 一强（逐渐式微）：Linux专业性较强，适用于特定场景。

图：2003-2012年个人电脑操作系统份额



资料来源：Statista、w3schools、statcounter

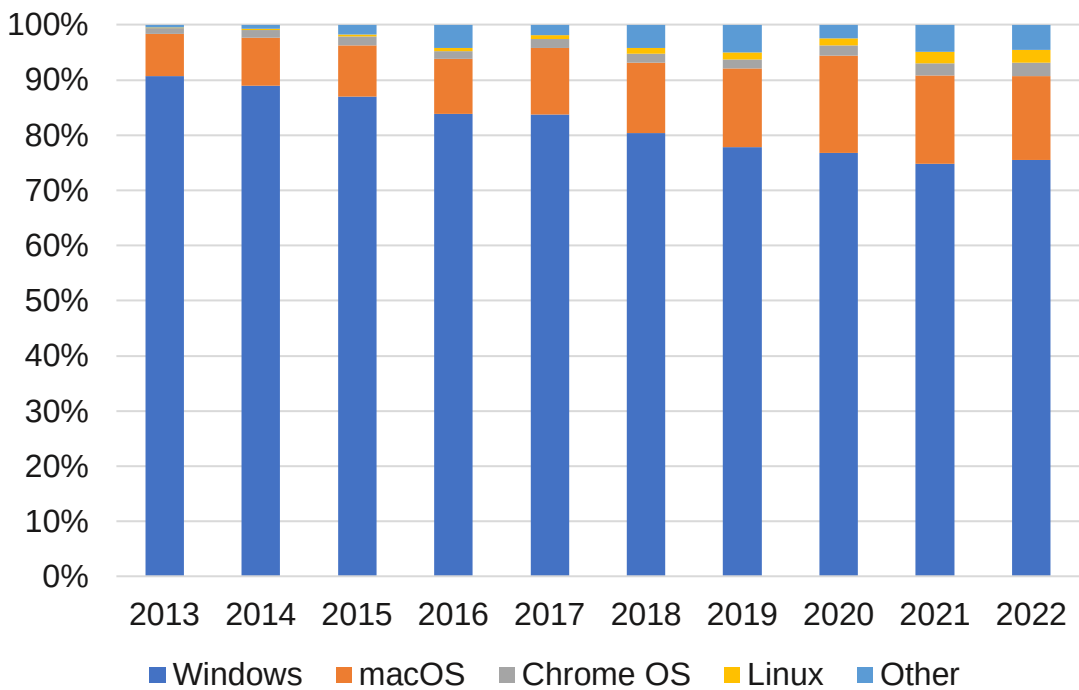
图：2003-2012年全球Windows与Mac OS市场份额变化



面向未来继续发展，多屏融合的尝试

- 随着技术的不断发展，手机和平板已经逐步成为了比个人电脑更重要的屏幕。更好的连接手机、平板和电脑，弱化设备之间的隔阂。从用户体验上来说，这一阶段操作系统的发展趋势主要有如下特点：
 1. 平板与电脑的统一：从Windows 8采用Modern UI，提供更好的触摸支持。平板电脑出现窗口化多任务。
 2. 移动设备与电脑的互联：2013年OS X 10.9升级实现了AirDrop功能。
 3. 应用通用：2021年发布的macOS12操作系统可以直接运行手机端iOS应用程序。同年推出的Windows 11中可以直接运行手机端安卓应用程序。

图：2013-2022年个人电脑操作系统份额



资料来源： Statista、statcounter

表：2001-2022年微软与苹果主要个人电脑操作系统发布情况

系列	名称	发布年份	备注
Windows消费者用户	Windows Vista	2006	支持64位芯片
	Windows 7	2009	
	Windows 8	2012	Modern UI
	Windows RT	2013	支持ARM架构
	Windows 8.1	2013	
	Windows 10	2015	
	Windows 11	2021	支持安卓应用
Apple消费者用户	Mac OS X 10.4	2005	开始支持x86架构
	Mac OS X 10.6	2008	结束支持Power 架构
	OS X 10.9	2013	扁平化设计，AirDrop
	macOS 11	2020	支持ARM架构（M1）
	macOS 13	2021	支持iOS应用

面向未来继续发展，多屏融合的趋势

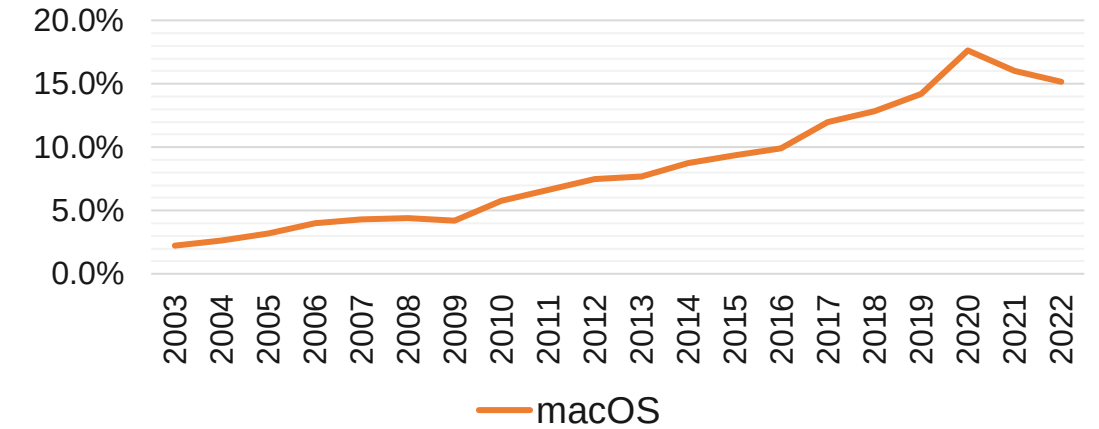
■ 轻量级开源操作系统出现

- 搭载Chrome OS的电脑，即ChromeBook依靠网页端应用来实现各种功能，不需要在本地进行复杂的数据处理。
- Chrome OS的电脑相较于普通电脑硬件成本较低，同时更加轻便，续航时间更长。

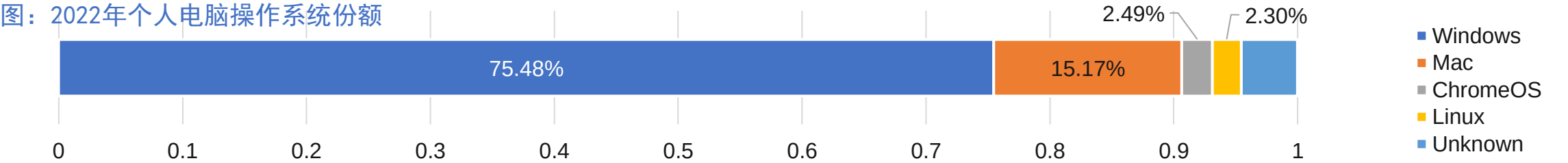
■ 不同设备之间的互联互通愈加重要

- 消费者会更倾向于购买同一生态链内的产品。
- 封闭生态系统的苹果将获得巨大的优势。苹果操作系统的市场占有率从2012年的8%上涨到2022年的15%。

图：2003-2022年全球Mac OS市场份额变化



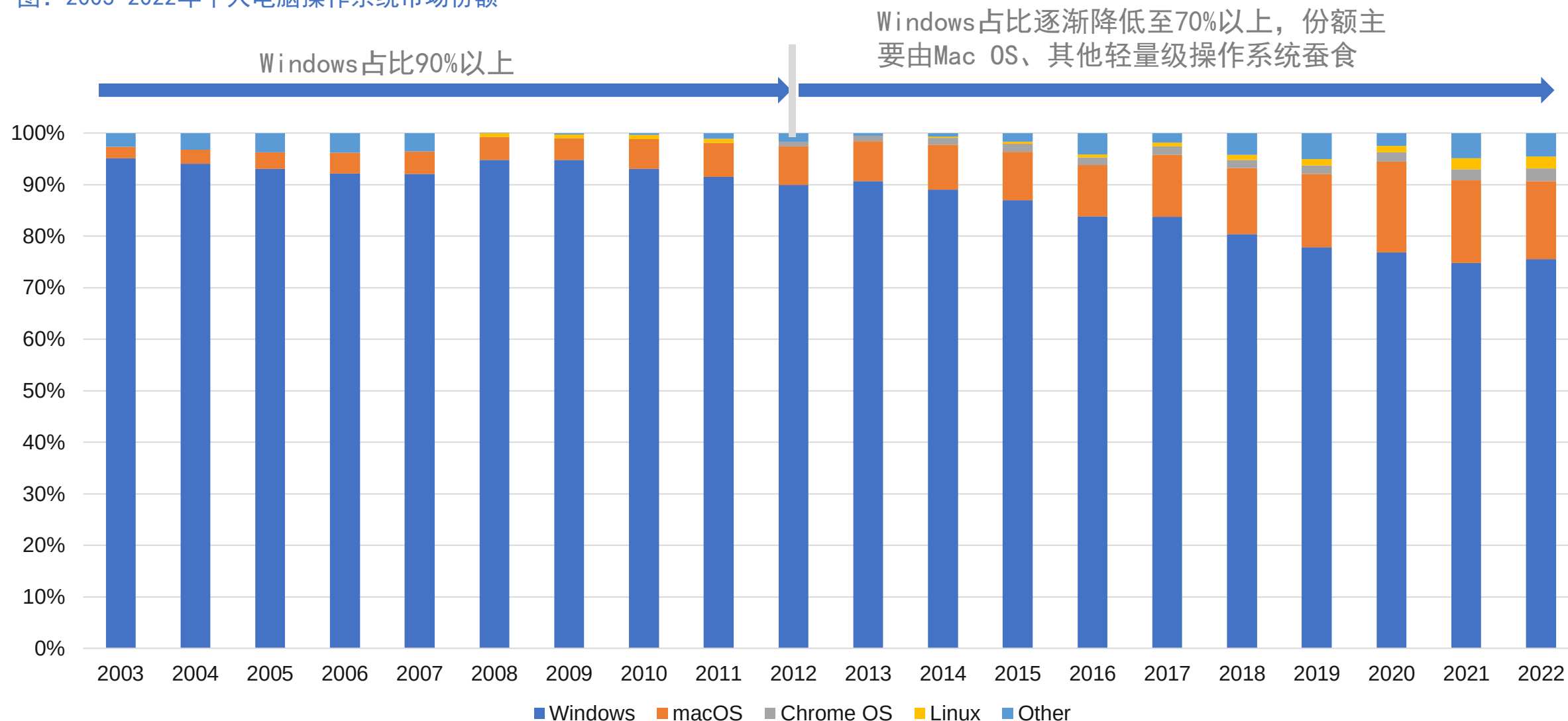
图：2022年个人电脑操作系统份额



资料来源： Statista、statcounter

PC操作系统市场份额变化

图：2003-2022年个人电脑操作系统市场份额



智能手机操作系统发展Milestone

- Symbian与诺基亚一同成就了功能机的巅峰，也一同消弭于智能机的浪潮。
- iPhone重新定义了智能手机，触摸取代键盘成为人机交互的主流。iPhone所搭载的iOS操作系统只能为苹果硬件服务，市场上其他的硬件厂商必须选择搭载其他操作系统来满足终端用户的需求。
- 谷歌组建OHA联盟，Android操作系统成为适配多元硬件的开放生态中最大的赢家。
- 智能手机一超一强格局延续至今，基本是Android（70%），iOS（20%+），其他（10%-）。但近几年，Android份额下降而iOS份额提升，主要是苹果在高端智能手机市场的优势进一步拉大。

智能手机操作系统鏖战，Android、iOS突出重围

Symbian OS

- 2007年，Symbian历史上最为成功的产品之一诺基亚N95正式发布，标志着Symbian巅峰时代。
- 2009年始，LG、索尼爱立信、三星等众多厂商逐步宣布退出Symbian平台，转投Android阵营。
- 2010年以后，由于Symbian先天性的设计缺陷（为键盘机而设计），导致其在触屏手机上的运行始终显得笨重而臃肿，不能和Android以及iOS系统抗衡，无法扭转颓势。
- 2013年，诺基亚正式宣布808 Pure View成为最后一款搭载Symbian系统的手机。意味着Symbian系统正式退出竞争舞台。

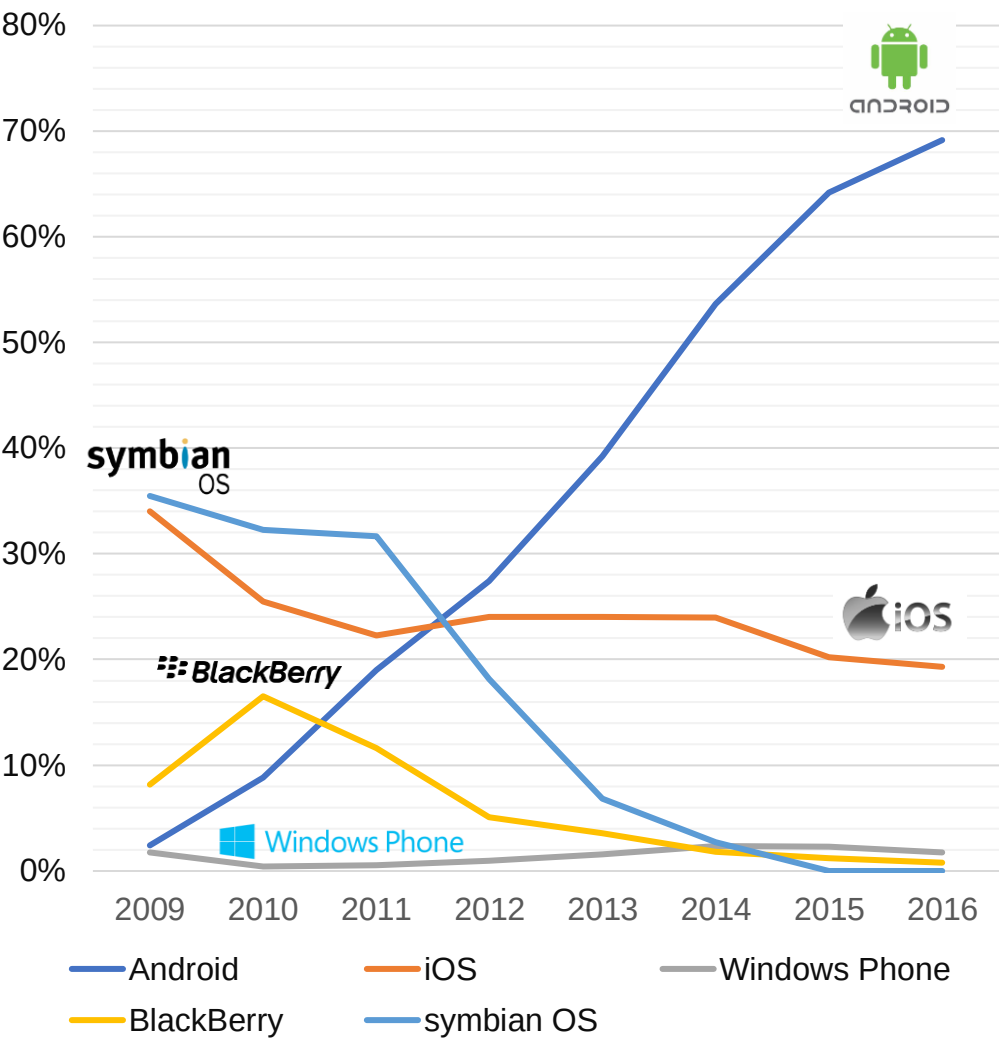
iOS

- 2007年，iPhone问世，重新定义了智能手机。2012年6月，iOS 6亮相WWDC 2012，提供了超过200项新功能。
- 2013年6月，iOS 7亮相WWDC 2013，重绘了所有的系统APP，去掉了所有的仿实物化，整体设计风格转为扁平化设计。

Android

- 2007年11月5日，Google公司正式推出了其基于Linux 2.6标准内核的开源手机操作系统，命名为Android。
- Google以Apache免费开源许可证的授权方式，发布了Android的源代码。开源的设计取得了硬件厂商的支持与响应，Android系统的市场份额快速放大。
- 目前销量最好的手机厂商如华为、中兴、小米、阿里、OPPO、Vivo、联想、魅族、360、TCL、酷派等所使用的操作系统，均基于Android进行定制改造。

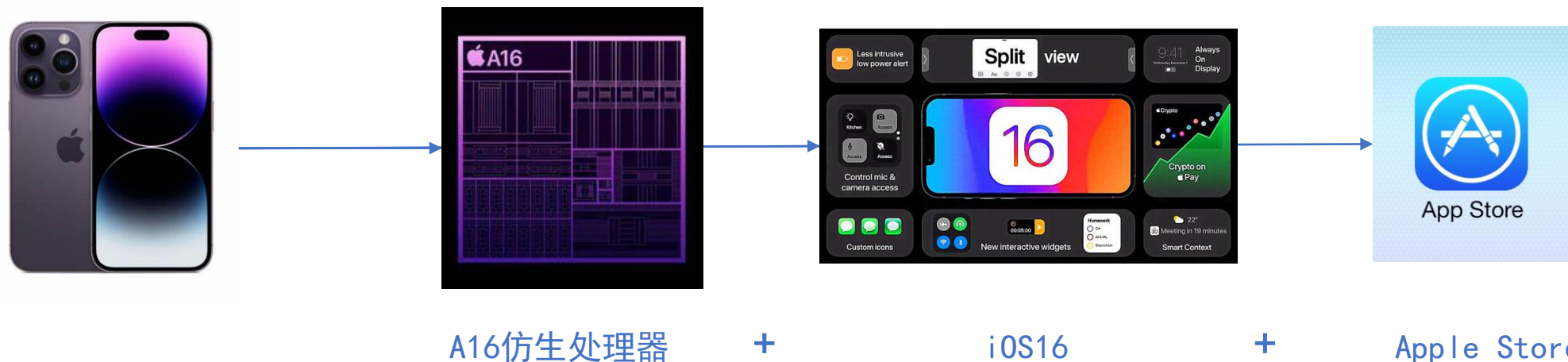
图：2009-2016年智能手机操作系统市场份额



资料来源：statcounter

iPhone的闭环生态：自研芯片+iOS+App Store

- 苹果iPhone以“自研芯片+iOS操作系统+App Store”构建软硬件一体的闭环生态。闭环生态的突出优势在于软硬件整合程度高、可升级、安全性高。
- **软硬件整合程度高：**iOS操作系统的软硬件整合度高，实现从“自研芯片+iOS操作系统+应用软件”的完美适配，避免兼容性差异等常态化问题，大幅提升系统的稳定性。与之对照，Android系统的分化程度较高，碎片化程度较高。
- **系统可升级：**通过FOTA（Firmware Over the Air，固件空中升级）技术，苹果用户摆脱了过去仅允许连接iTunes才能进行系统更新的限制，使其第一时间在联网状态下就能完成iOS系统的在线升级，以获得更强的手机性能。
- **安全性高：**在闭环生态下，苹果能够快速捕捉到系统中存在的安全漏洞，并通过FOTA方式及时推送给用户以实现在线修补。同时，应用软件开发商仅能通过苹果提供的工具包进行应用开发，并在其严格的审核机制下才能推出应用功能。



Android的联盟生态：开放硬件+开源软件

- 由于自身软硬件能力的缺乏，Android采取了与苹果不同的开源策略——**建立联盟生态**：联合手机制造商、软件开发商、芯片制造商等多方势力，共同成立OHA（Open Handset Alliance，开放手持设备）联盟，基于Android源代码构建包含操作系统、中间件、应用程序在内的开源软件框架，以实现Android系统与不同手机终端及芯片间的兼容与适配，形成从“芯片+操作系统+应用+终端”的Android联盟生态。
- OHA联盟中成员可分为五类：通信运营商、手机制造商、半导体公司（芯片厂商）、软件应用开发商、商业公司。

开放硬件

- ✓ Pros: 实现了Android系统的快速普及，实现了应用软件在不同终端中的良好兼容。
- ✓ Cons: 手机厂商为保持对Android系统的兼容性，而抹杀了自身的差异性，只能比拼硬件。

Open Handset Alliance(OHA)

Mobile Operators



Handset Manufacturers



Semiconductor Companies



Software Companies



Commercialization Companies



开源软件

- ✓ Pros: 基于开放性原则兼容了更多的应用软件开发商，应用矩阵也更丰富。
- ✓ Cons: 对用户而言，通过SOTA方式即能实现应用软件在不同终端以及不同渠道的获取（渠道可以是官方应用商城，也可以是腾讯应用宝、360等第三方应用平台）。

塞班的没落与微软的失利

symbian
OS

- ✓ 一是用户体验极差，塞班系统为了防止恶意的程序进入降低系统流畅度，就推出了一个签名机制，下载一个软件必须要使用签名机制，过程复杂，体验感差。
- ✓ 二是生态缺乏，开始时索爱 and 三星手机都可使用塞班系统，后来诺基亚将授权收回，其他的手机品牌都不能够使用塞班的系统，导致生态进一步衰落。

 **Windows Phone**

- ✓ 一是微软已经失去移动市场的先发优势，在大局已定的情况才进入，此时Android与iOS已经占据市场，微软下拉式菜单的使用习惯也给客户习惯切换带来难度；
- ✓ 二是无法形成一个完整的手机生态体系。与Android、iOS相比，Windows 10系统的生态系统过于匮乏，微软也曾通过基金池的形式鼓励开发者加入Windows平台，但效果不佳。

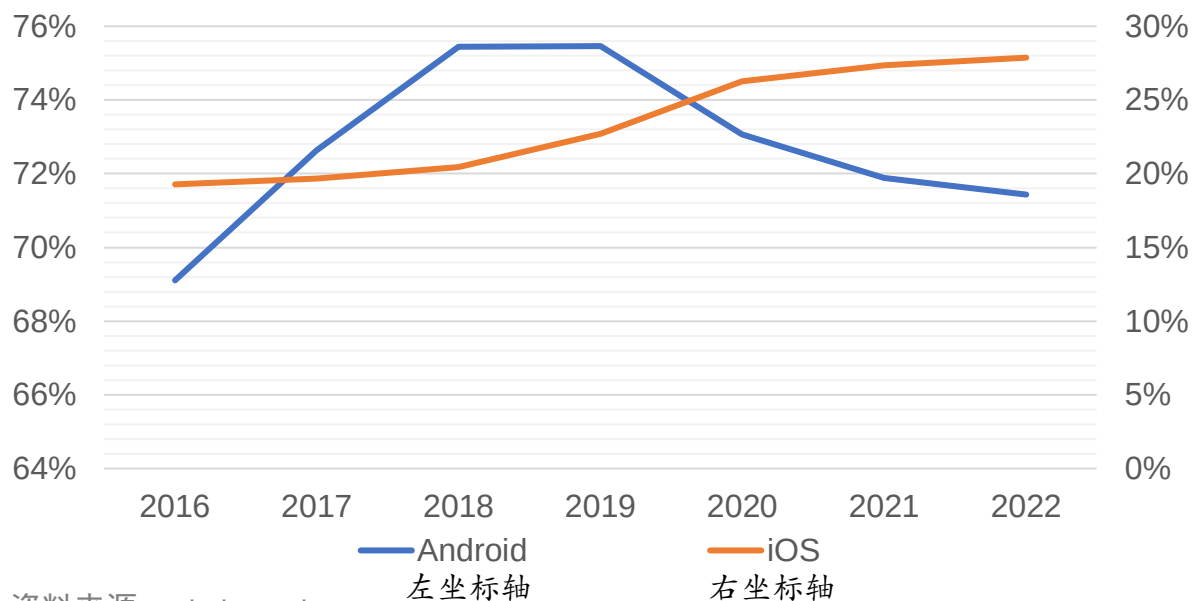
TIZEN

- ✓ 一是三星Tizen起步较晚，开发者Android、iOS已形成粘性，故APP在这个系统上的适配问题不断显现。
- ✓ 二是三星与其他品牌是竞争关系，很难说服手机厂商使用自己的系统，而谷歌并不会威胁到其它手机厂商的业务。

近五年Android统治地位有松动迹象

- Android市场份额由2018年的75.45%下滑至2022年的71.44%，降低4个百分点；同期，iOS市场份额则由20.47%增长至27.88%，增加7个百分点，与苹果手机全球市占率的提升相匹配。
- 外因：华为1) 受到制裁后芯片短缺，已无力参与高端手机市场的竞争，这部分用户转向苹果手机；2) 华为手机自研鸿蒙OS，所有搭载华为麒麟710芯片及其以上的机型均可升级。
- 内因：OHA成员需要保持对Android的兼容性，导致安卓手机同质化较严重、用户体验待优化，难以攻克高端机市场份额。

图：2016-2022年全球Android与iOS市场份额变化

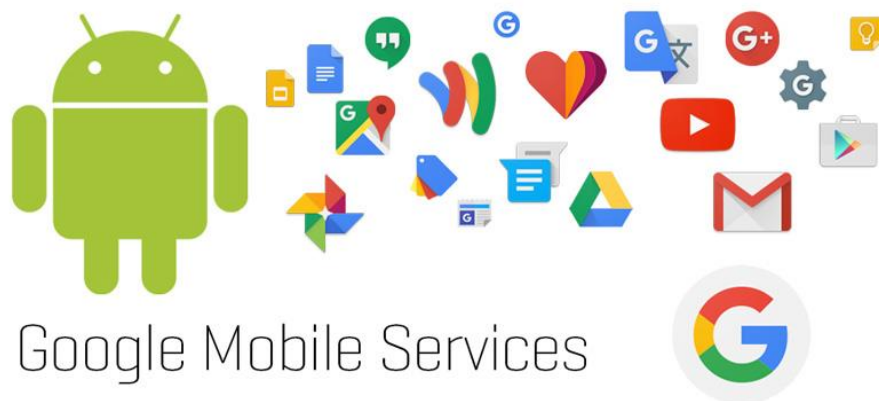


资料来源：statcounter



OHA联盟出现内部利益分化，谷歌VS手机厂商

- OHA联盟中成员可分为五类：通信运营商、手机制造商、半导体公司（芯片厂商）、软件应用开发商、商业公司。其中，各个环节内部互有竞争，手机厂商之间、芯片厂商之间、软件开发商之间，但仅手机厂商的发展与谷歌普及Android存在潜在冲突。
- 一方面，Android的兼容性制约了手机厂商的差异化。手机厂商只能通过销售硬件获得收入，同质化竞争严重、利润率低。
- 另一方面，手机厂商尝试多平台策略以增强竞争力。谷歌收购摩托罗拉给其他手机厂商带来危机感，HTC、三星等推出一两款机型转向Windows Phone操作系统。



VS



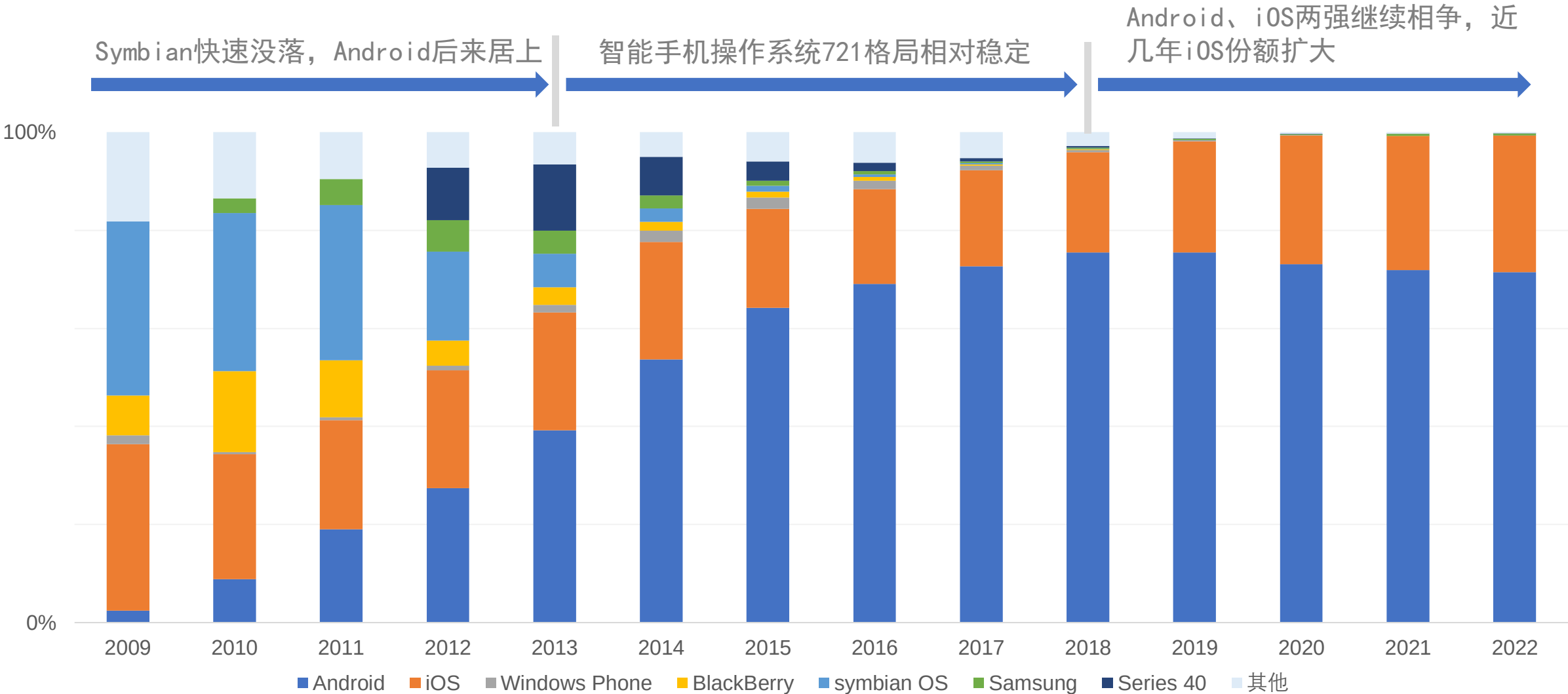
闭环生态日臻纯熟，iOS与Mac OS全面融合

- **统一化、平台化**是苹果闭环生态一直以来追求的目标。苹果的生态策略是希望将手机、平板、电脑三款不同的硬件平台融合，对应到执行层面是硬件、软件（系统）两大维度齐齐发力。
- **硬件：推出基于ARM架构的M1、M2芯片。**ARM为移动端而生，M系列芯片的综合性能及部分性能已经不落后甚至超过PC端主流的intel和AMD芯片。未来，随着苹果研发的跟进，iPad甚至iPhone都有可能搭载苹果M系列芯片。
- **软件（系统）：苹果体系内三大系统平台iOS、iPad OS、Mac OS已实现全面融合。**首先，iOS和iPadOS的软件兼容自诞生以来就可以。其次，2019年iPadOS独立，标志着苹果移动端和PC端系统融合正式开始。最后，2020年Mac迁移到自研芯片，开发者可以在不经修改的情况下让iOS、iPadOS App在Mac中可用。



智能手机操作系统市场份额变化

图：2009-2022年智能手机操作系统市场份额



资料来源：statcounter

XR操作系统：VR安卓占据主流，AR则乾坤未定

VR设备：安卓占据主流

- ✓ VR一体机目前主要使用安卓系统。
- ✓ 早期的VR一体机基本沿袭手机端的计算芯片与操作系统，如Oculus Go采用高通骁龙821芯片，Vive Focus采用高通骁龙835芯片，操作系统则是在安卓系统的基础上进行优化及定制。
- ✓ 类似地，分体式VR包括PC VR、PS VR以及华为VR Glass、创维V601等超短焦手机VR，其运行的操作系统仍以连接的主机为主，包括微软WMR、索尼PS及安卓系统。

AR设备：生态制高点，乾坤未定

- ✓ AR操作系统生态占领制高点。
- ✓ 目前市场上的创业型公司AR眼镜自身缺乏系统研发能力，当前仍以多以安卓系统为基础进行优化与定制。
- ✓ 科技巨头将AR操作系统视为战略制高点，如微软HoloLens采用以Windows NT为基础的Windows 10 Holographic系统，HoloLens 2则采用全新多平台操作系统Windows Core OS；Magic Leap自研专为空间计算设计的Lumin OS。
- ✓ 国内企业虹宇科技2020年发布自研VR/AR 3D多任务系统Iris OS，可以呈现2D、3D的窗口及全景应用，支持多人多设备协作，兼容各类芯片平台及光学模组。

Android

WMR

PS

Windows 10 Holographic

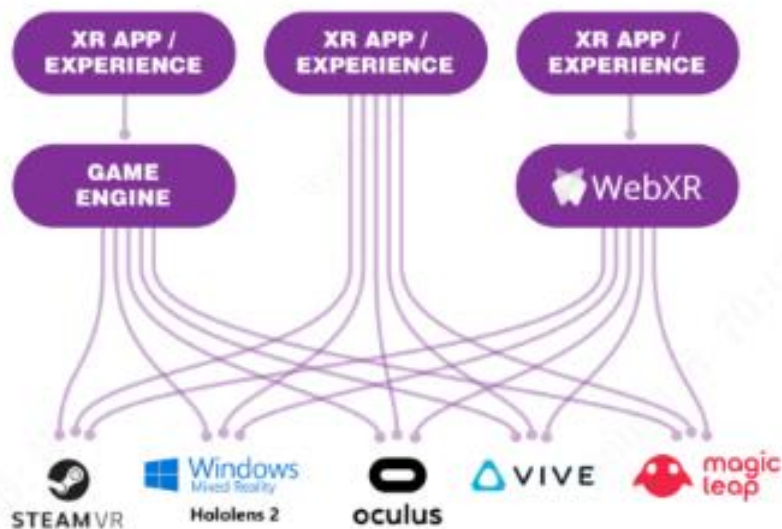
Windows Core OS

Lumin OS

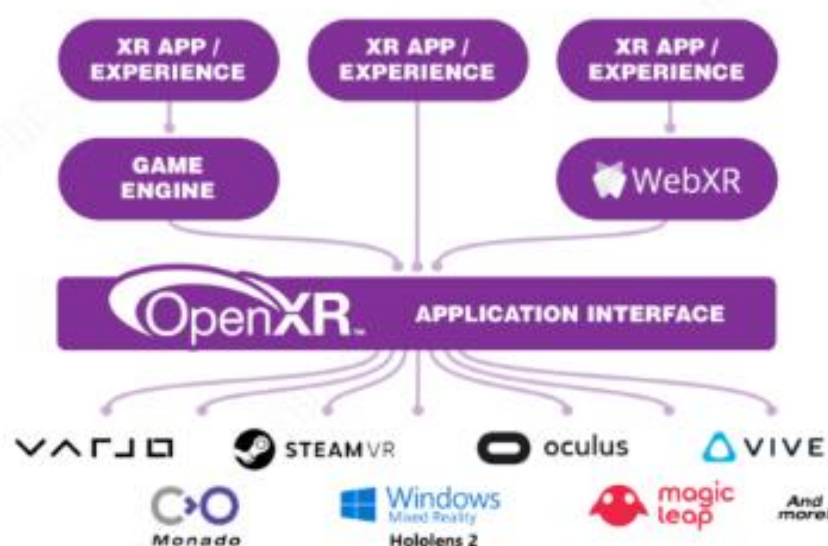
Iris OS

OpenXR行业标准已经出现

- OpenXR是一个由Khronos组织联盟开发的开放式、无版权费用的XR行业标准规范，旨在简化VR/AR软件开发，打通游戏引擎及内容底层连接，塑造具备互通性的开放生态。OpenXR最大的意义是，游戏开发者使用一个API接口就能让游戏在不同品牌的VR/AR中运行；硬件厂商则可利用现有OpenXR内容降低市场进入门槛，为消费者提供更丰富的内容体验。
- 目前微软在HoloLens、WMR头显均提供OpenXR支持；Unity推出OpenXR支持预览版；Epic宣布虚幻引擎5不再支持SteamVR、Oculus等平台，转而支持且仅支持OpenXR标准；Oculus推荐游戏引擎使用OpenXR等。国内厂商中华为、兆芯、Pico等也加入OpenXR联盟参与标准的讨论与制定工作。



Before OpenXR: Applications and engines needed separate proprietary code for each device on the market.



OpenXR provides a single cross-platform, high-performance API between applications and all conformant devices.

开端：XR区分于PC、智能手机的最本质差异

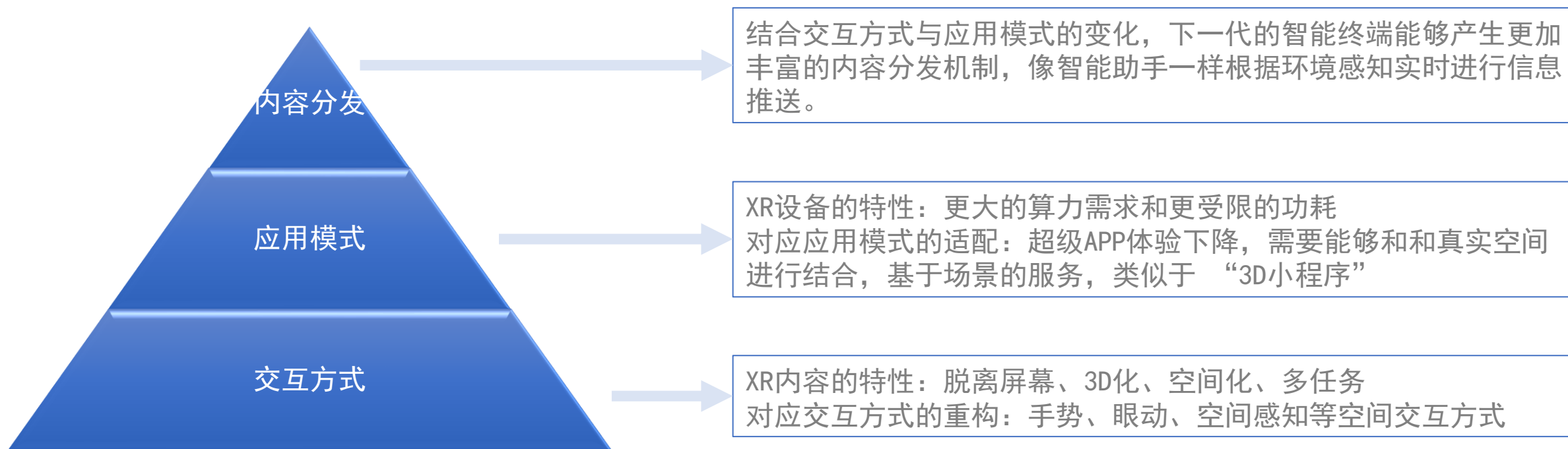
- VR（Virtual Reality 虚拟现实）是指，由计算机创建、渲染一个仿真三维空间并呈现信息；AR（Augmented Reality增强现实）是指直接以真实的三维空间为背景将信息投射其中。VST VR，它以VR相对成熟的硬件为基础，添加能够捕捉真实世界的摄像头，将真实世界重构于用户眼前，达到一种类似AR的效果，也称之为MR。
- 传统VR类似于游戏机，而VST VR（MR）更像是PC的延伸，AR更像是手机的延伸。
- 相比上一代移动互联网，VR/AR带来的最本质的变化是，用户能够在三维空间中与立体、逼真的信息进行互动。



	PC/笔记本	智能手机	XR终端
通讯技术	Internet	3G/4G	5G
交互形式	鼠标	多点触控	空间交互3D，手柄
应用模式	“Windows”	App	云应用Cloud
内容分发	浏览器	应用商店	环境感知、AI

中场：XR时代需要怎样的操作系统？

- XR需要一套全新的原生的操作系统。专门基于XR特性设计的包含交互方式、应用模式、内容分发的操作系统。
- XR设备、内容有着区分于PC或智能手机的显著差异，3D化、空间化、多任务等，对算力和功耗的要求也更苛刻。基于现有Android系统改造的XR操作系统长期难以适应XR的发展需求。



终局：XR操作系统之争，谁的胜率更大？

- 四大巨头角逐操作系统霸主。复盘PC及移动操作系统历史，寡头垄断可能是最终格局。争夺XR操作系统霸权的战争已经开始。下一代硬件仍在演进，但四家重要的参与者——苹果、微软、谷歌和Meta针对VR、AR头显提供操作系统的竞争已经非常激烈。



- ✓ 参与过两场操作系统之争，Mac OS、iOS 至今占据稳固份额
- ✓ 强大的产品能力始终是压舱石
- ✓ 独有的闭环生态效率与安全性双高，但同时也是普及的限制因素
- ✓ Mac OS
- ✓ iOS
- ✓ ARKit
- ✓ 苹果MR



- ✓ 胜于PC操作系统之争，落后于移动操作系统之争
- ✓ 基于Windows的VR/AR操作系统已有应用
- ✓ 积极寻求第三方合作，如Meta
- ✓ Windows
- ✓ HoloLens
- ✓ Windows Core OS



- ✓ 基于Android系统改写的操作系统占据VR主流
- ✓ 未直接研发VR/AR产品（Google Glass已宣告破产），具备再建开源生态的条件
- ✓ Android
- ✓ 无硬件
- ✓ 开源生态



- ✓ 从硬件来看，是VR/AR设备的先行者，VR全球市占率第一
- ✓ 新秀，未参与过操作系统之争
- ✓ 积极寻求第三方合作，如微软
- ✓ VR产品线
- ✓ AR产品线（在研）
- ✓ 策略最激进



研究
思路

01

2023版元宇宙研究框架、元宇宙的上半场与下半场

02

2023：硬件的“大”年

03

硬件之“沉浸感”、“交互范式”、“操作系统”

04

元宇宙之百舸争流

- 虚拟人：交互世界的新物种；NFT：数字世界的魔法棒
- Web 3：内核是“治理”，“去中心化”实则为“反抗有理”
- AI与后端基建：2023年迎来景气上行
- 场景：2024-2025年的“八面来风”

PART 01 >

虚拟人：交互世界的新物种

NFT的魔法棒：从数字内容到数字资产

虚拟人最重大意义之所在

NFT的魅力在于孵化与孕育未来的各类“明星”

虚拟人与NFT对当下的重构，亦会重塑企业的三张报表

虚拟人与NFT的交叉之处，是新一轮的IP孵化与商业化

虚拟人：交互世界的新物种

根据安信最新的逻辑架构图，元宇宙是实现真正智能的路径之一，即重构一套新的时空，架构在真正的智能实现上、元宇宙上的虚拟人，它是新物种，“新物种”是相对于“人”而言，从这个角度，新物种不仅包括虚拟人（人的数字人、虚拟数字人），也包括机器人（现实物理世界中的机器人，在元宇宙中一定会有虚拟形象）。

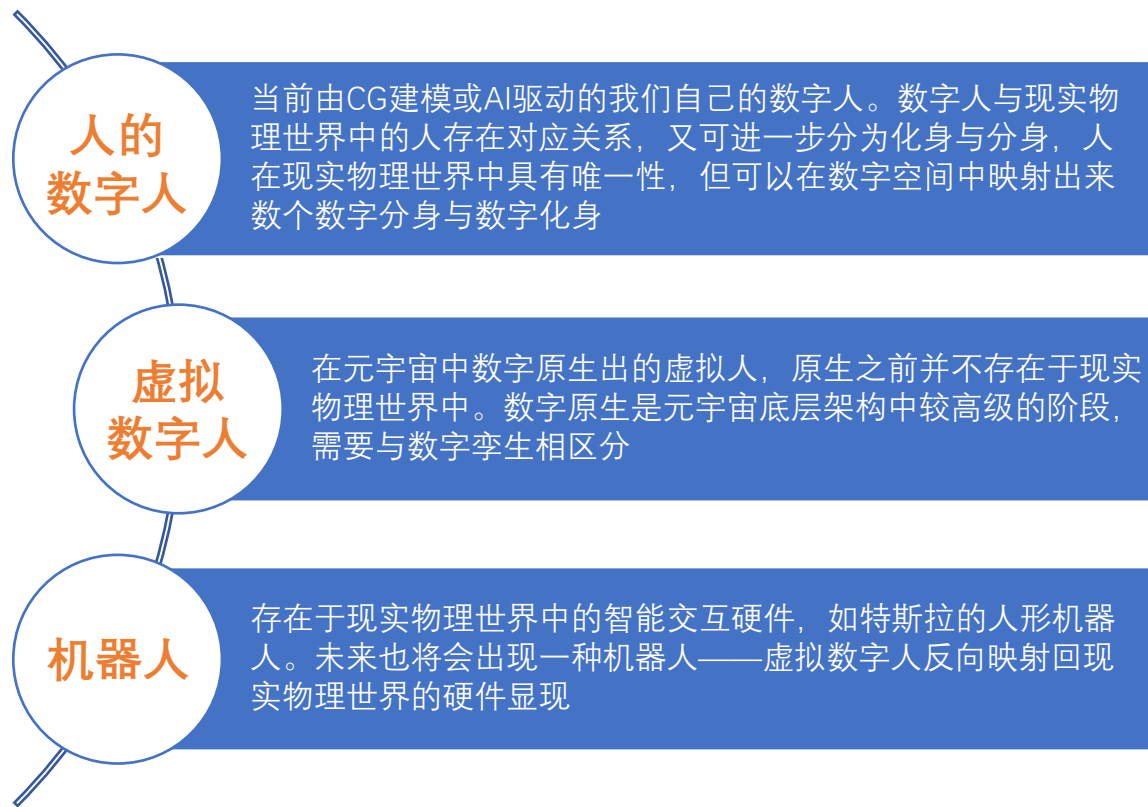
从这个角度，“人”在元宇宙时代，软硬件均需要升级，因为混合平台的实现，技术路径之一，即以人为硬件载体。元宇宙对“人”的挑战，是“人”的软件需要适配升维至3D的虚拟现实（元宇宙）；智能的真正实现，亦需要“人”硬件的升级（甚至是重塑）。

从用户的角度，互联网时代的PC、智能手机是人的一个器官，努力去拉近人与网络（虚拟世界）的距离，虽然在努力拉近，但最大的限度只能去靠近、逼近，元宇宙则是将用户直接拉进虚拟世界中，故虚拟人的职能有三：

- ✓ 第一，是人的数字人，肩负着将人拉进元宇宙的使命；
- ✓ 第二，是在现实物理世界中不存在实体的虚拟数字人；
- ✓ 第三，是机器人，机器人的存在对于智能的真正实现（通用人工智能平台的构建）非常重要，未来现实物理世界中很多的角色、职能、功能将由机器人来协助甚至代替人来承接；浅层次上，机器人的存在将在经济学意义上降低商品与服务的价格，尤其是服务，解决诸多日益沉重的社会问题。

人 VS 人的数字人、虚拟数字人、机器人

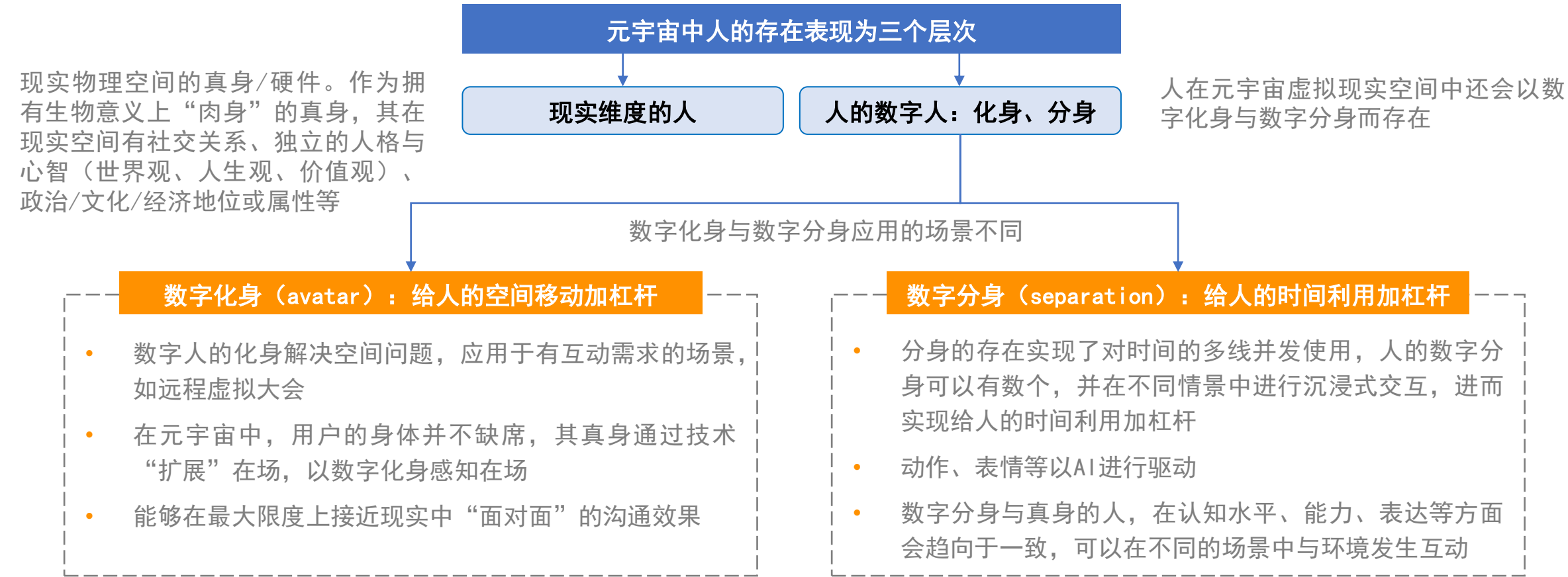
- 从交互对象的角度，未来元宇宙囊括现实物理世界后，人的交互对象增加了三类，分别为人的数字人、虚拟数字人、机器人。其中，AIGC即AI创作，这里的创作者可以是人的数字人、虚拟数字人、机器人——他们本质上都是AI，其创作均可以称为AIGC。



互联网	元宇宙
2D	3D
人	→ - 人 - 人的数字人 - 虚拟数字人 - 机器人
↓	↓
交互种类	
人与人交互，1种	C_4^2 / A_4^2 6或12种

人的数字人：分身与化身

- 元宇宙是虚实融合的世界，那人如何进入元宇宙参与各类活动？人进入元宇宙的数字空间需要有一个数字身份，数字人是构建数字身份的基础。



数字分身与数字化身的存在，让人在时空环境中均能加杠杆。工作效率、劳动生产率因而得以成倍扩大、增长，这就是元宇宙可能带来的增量价值。分身与化身不是目的，杠杆作用于人的创造力、放大这种创造力的效用与效益才是目的。

机器人也将会是虚拟数字人在现实世界的显现

- 元宇宙的后半场，现实物理世界中的人、物、空间这三者均会被重塑，未来将会出现基于情感需求投射的分布式垂类硬件，机器人是分布式垂类硬件之一：
 - » 一是，人在元宇宙中将会被开发出新需求，尤其是非物质需求，未来元宇宙中的 AI 将会发挥越来越重要的作用，2022年AIGC发展迅速且获得市场广泛的关注，AIGC即人工智能自动生成内容，被认为是继专业生产内容（PGC）、用户生产内容（UGC）之后的新型内容创作方式，人的潜在需求会越来越的被 AI 的供给所满足，另外 AI 将进一步，显现在现实物理世界中，以新硬件的形式承接人在未来现实物理世界中的更多需求；
 - » 二是，虚拟数字人有望反向映射回现实物理世界，即显现为机器人/新硬件而存在，其在与人共享的现实物理空间中的需求与体验同样需要被满足，也就是说，分布式垂类硬件的产生，以“现实物理世界”为空间，以 AI 为生成与驱动的机制，是“人”与“物”受“机器人”与“新硬件”供给与重塑的结果。
- 在我们看来，机器人或新硬件仅是外在的表现形式，内核仍然是服务于人交互的 AI，包括各类主体之间基于情感或利益需求的交互。
- 元宇宙较上一轮移动互联网，增加了 AI 生成与驱动的机制，移动互联网时代，交互的内容 / 对象基本上都是由真实的人（软件工程师、创作者等）设计与渲染出来的，但在元宇宙时代，AI 成为元宇宙世界里的一大新增生产要素，将会大量存在于供给、需求的各个环节，数字人、虚拟人等就是 AI 的诸多应用之一。
- 案例：北京冬奥会上的各类机器人、特斯拉人形机器人“擎天柱”、人形仿生机器人CyberOne……

人与人的交互 VS 新增6类交互

- 我们认为人与自然的交互、人与物的交互，是人对外交互的基本层面，人与人的交互则是人对外交互的更高层面，故研究序列中，人与人的交互在绝大多数情况下，是我们关注的重点。
- 通过梳理人与人交互的历史脉络，我们发现在历史进程中，任何一个方向的升维都有可能推动人与人的交互发生翻天覆地的变化，呈现出愈加复杂的特性。整体而言，人与人交互脉络最为核心的主线，是技术的发展推动人与人交互的空间范围不断扩大，交互的反馈变得迅速（交互时间缩短），在交互空间折叠、交互节奏持续加杠杆过程中，人与人交互的媒介、渠道持续迭代，交互行为的范式也变得更加丰富与复杂。
 - » （一）传统农业时代：主要依赖面对面交流，重要信息靠书信传递
 - » （二）电报/电话时代：即时、在线沟通交流逐渐成为普通老百姓的常态
 - » （三）PC互联网时代：真正实现了全球用户的连接，由通讯走向娱乐等多元的交互行为
 - » （四）移动互联网时代：随时随地线上交互，基于位置衍生出一系列全新的交互行为

人类交互历史梳理

传统农业时代

电话/电报时代

PC互联网

移动互联网

元宇宙

空间范围

- 邻居亲友面对面交流
- 特定阶层可远游（大航海）

- 邻居亲友面对面交流
- 远距离电信/电话沟通

- 与世界各地的人交互（存在上网场景限制）

- 与世界各地的人交互（随时随地）

反馈周期

- 线下实时交互
- 书信往返约半个月

- 写信5-10天
- 电话/电报可即时在线交流（费用高，适用于部分人）

- 实时在线交互（文字、图片、视频）

- 实时在线交互（图文、视频、直播）

交互行为



贸易



农业生产



- 1844年
莫尔斯发出人类历史上第一份电报



- 1993年
中国无线寻呼业务向社会开放



- 1877年
第一条电话线路连接了波士顿和萨莫维尔



- 1994年
广东开通第一个省级数字移动通讯网

社交

- 1999年
BBS、QQ上线

- 2005年
博客上线

- 2009年
微博上线

娱乐

- 1998年
网文《第一次的亲密接触》在BBS连载

- 2005-2006年
视频网站PPS, 土豆、优酷网相继上线

- 2007年
《传奇》最高在线人数突破60万

电商

- 2003年
淘宝网上线
- 2009年
首次举办双十一购物节

社交

- 2011年
微信上线

- 2010年
快手上线

- 2016年
抖音上线

娱乐

- 2010年
爱奇艺上线

- 2015-2016年
直播大战

电商

- 2020年
手淘改版 推出逛逛

- 2021年
抖音兴趣电商

其他

- 外卖

- 打车

- 移动支付

元宇宙

- 透过上述从传统农业时代到移动互联网时代四个阶段的历史梳理，我们总结人与人的交互沿着四条脉络发生了翻天覆地的变化：
 - » （1）从交互**空间**范围来看，从亲人邻里的面对面沟通，到远距离交流，再到互联网时代能够与世界范围内的人自由交互，交互的范围不断扩大；
 - » （2）从交互**时间**来看，人与人进行远距离时，从最开始的书信、电报，到电话、微信时代，交流的反馈周期越来越短，信息的交互越来越同步；
 - » （3）从交互**媒介**来看，从阅读文字、图片、视频到直播，承载交互信息的媒介实现了从一维到三维的升级，现形式更加立体，用户在单位时间内能够看到的信息量变得更加丰富，也调动了更多的感官进行交互，包括视觉、听觉等，使我们能够更好的感受对方的情绪、表情等。
 - » （4）从交互**内容**来看，从最早的获取资讯，发展到线上娱乐、购物，甚至衍生出许多原来并不存在的行为，比如线上打车、外卖等，涵盖了生活的方方面面，行为的复杂性有了全方面的提升，而且交互的需求也从满足物质性需求的目的发展至服务于精神性需求的目的。
- 这四条脉络里人与人交互的空间、时间、感官的变化是表层的原因，而交互内容的升级及丰富度的提升是前三个层面的升级所导致的结果，每一个维度的升级都会带来现实物理世界的生活，其巨大的变化甚至是变革。

交互空间的变革

面对面沟通→远距离交流→移动互联网时代的随时随地

交互时间的变革

书信、电报→电话、微信
反馈周期越短，信息的交互越同步

交互媒介的变革

书籍→电视/PC/智能手机→VR/AR
一维→二维→三维

交互内容的变革

图文→音频、视频→直播、购物
内容越来越丰富

元宇宙时代新增6类交互

- 元宇宙时代将沿着前序的脉络，进一步重构人所处的空间与时间，人与人交互的场景的复杂度将全方位被放大。
 - » 首先，现实与虚拟的3D融合，人更被裹挟入虚拟时空里，一方面是空间上会出现与现实世界孪生的虚拟世界，另一方面基于用户的创造将衍生出现实世界不存在的空间，而且在不同虚拟空间中用户可以自由移动。
 - » 其次，在时间层面，虚拟与现实的边界被消融，人随时随地甚至无感进入虚拟空间，在线时长将不受限制，甚至可以24小时沉浸式进入线上。这种空间的自由流动与时间的极限沉浸叠加，有望创造出极为繁荣的新场景，从而释放出巨大的产业价值。
- 仅从人与人的交互，我们就可预见未来将出现的不可思议的变化，更何况在元宇宙这一计算平台上，交互主体已经不限于人与人之间。元宇宙中的人、人的数字人、虚拟数字人、机器人这四类用户，能衍生出6类交互方式：



人与人的数字人的交互

人的数字人与虚拟数字人的交互

人与虚拟数字人的交互

人的数字人与机器人的交互

人与机器人的交互

虚拟数字人与机器人的交互

元宇宙时代新增6类交互

人与人的数字人的交互

在二维平面的互联网/移动互联网中，人的身份象征体现为一个数字头像或数字人物，人与数字头像或数字人物之间的交互感很弱，源于二维屏幕不足以重建现实。

而在虚拟实现或元宇宙中，我们人的身份象征则体现为人的数字人，是非常具象的、立体的形象，因而此时人与人的数字人之间的交互感，则体现得较为强烈且真实。如初次体验VR头显设备时，注册登陆设备账号时，需要构建一个或数个3D的虚拟形象，来适配于不同的应用场景，以西装的形象参加线上会议，以休闲的形象参与线上社交活动，以运动的形象参与线上健身等。

人与虚拟数字人的交互

虚拟数字人特指在虚拟世界中数字原生出的虚拟人，与现实世界的人不存在对应的关系。

虚拟数字人将会是元宇宙中的重要组成元素，原因在于元宇宙中的交互界面不可能再依赖于文字，而必然要回归人与“人”之间的交流，虚拟数字人便承担了连接元宇宙虚实共生世界的主体作用。但目前，人与虚拟数字人之间的交互仍然存在壁垒与差异，因为虚拟数字人的本质是人工智能，人与虚拟数字人之间的交互非常依赖于人工智能的发展进程，以及配套的算力算法，人与虚拟数字人之间的交互未来会趋向自然与顺畅。

人与机器人的交互

与虚拟数字人的本质相同，机器人的内核也是人工智能。人与机器人之间的交互并不是新鲜事，目前我们可以在商场、餐厅、酒店等地方看到诸多的移动机器人，这些机器人大多都是起到客服或导航的作用。当前绝大多数的机器人还无法做到与真人一样进行自然的交互，但方向是明确的，即拟人的设计、更加的智能化，甚至是情感表达与交流。随着深度学习技术的发展，2022年智能机器人发展提速，如特斯拉人形机器人“擎天柱”，代表了人工智能技术落地的新方向。未来，人与机器人之间的交流将逐步从指令交流升级为无特定具体指令的正常交流，这种交互的进化过程中，将伴随着系列化的难题与众多挑战。

- 除了以上三对交互关系外，还存在人的数字人与虚拟数字人的交互、人的数字人与机器人的交互、虚拟数字人与机器人的交互。因而，元宇宙实现了社交主体的拓展，交互对象的范围也扩展至虚拟人与机器人这一非生物类群体，达到万物互联，这意味着人在社交生态中的身份、地位与功能都将发生变化。最重要的是社交主体的拓展必然带来社交主体之间关系的重构，社交生态更为丰富，社交关系也更为复杂。元宇宙时代，在各种非生物类群体成为人的交互对象的背景下，这些以往的“物”也就不再只是交往的客体，而具有了全新的主体身份。这对于社交生态而言，具备颠覆性意义，一切均将被重构。

虚拟人首先是人进入元宇宙的身份

- 总结人的身份特征的变化，我们可以梳理出较为清晰的主线，即从书信时代的笔名、PC互联网的QQ头像、QQ秀，到移动互联网时代的微信头像、脸萌、ZEPETO，代表人们进入虚拟世界的身份角色形象逐步向更立体的方向升级，从一维文字到二维的图片，目前逐步在向三维形象升级。

人进入虚拟世界时身份演进的变化



虚拟人首先是人进入元宇宙的身份

- 随着元宇宙计算平台的发轫，更多的互联网平台在打造自己的虚拟空间，比如Horizon Worlds、Decentraland、百度希壤等，每个用户进入虚拟空间就会拥有至少一个可个性化装扮的虚拟化身(avatar)，用以在虚拟空间中行动体验，或其它用户交流互动。这些化身可以理解为虚拟人的初级形态，但是这也像我们展示了未来人进入元宇宙世界的一种可能的交互形态。
- 如多人在线VR社交游戏VRchat的大火，游戏借助外置可穿戴设备，支持音频唇形同步、眼球眨眼追踪、运动范围捕捉等多方面的状态采集。
- 元宇宙这一计算平台的持续发展，预计虚拟人在表现形式上将继续呈现出巨大的升级，其关键的方向包括以下几个方面：
 - » 第一，虚拟化身形象将进一步由2D、2.5D向3D方向的升级；
 - » 第二，不同主体间的交互将得到全面的实时反馈。

元宇宙未来
经典内容的
四个发展阶段
(以游戏为例)



数据量级倍增，对底层技术要求越高

虚拟人其次为用户新一轮匿名化社交的开始

- 从PC互联网到移动互联网的发展，虚拟世界中的用户的对应身份由**匿名化走向实名化**，表现为微信的熟人社交属性、基于现实职场的在线办公等线上化交互场景都在将真实世界的用户身份与虚拟世界进行对应，且线上的交易行为需进行实名认证。实名化趋势在过去十余年，受益于移动互联网的快速普及，到目前为止几乎完全实现了所有人的身份在现实物理世界到虚拟世界的一一对应。
- 我们认为虚拟人的价值不仅仅是作为人进入虚拟世界时所需要的身份，延续从PC互联网时代的QQ到移动互联网时代的脸萌、ZEPETO所展现出的表现形式升级，其更深的价值在于：一方面，虚拟人作为人的化身进一步突破现实物理世界时空的限制使人更沉浸于虚拟世界，从而**增加在线时长**；另一方面，虚拟人作为人的分身，通过撬动时间杠杆的方式帮助用户通过创作**增加可支配收入**。
- 其中，我们认为最重要的，则是继PC互联网时代的匿名化社交、移动互联网时代的实名制社交后，**元宇宙是新一轮匿名化社交的开始**，叠加**2D向3D的升维**、**更多感官体验的增加**，匿名化社交预计能释放出更多的社交活力，去新一轮的抢占用户“时长”与“可支配收入”。

元宇宙带来的机遇



用户在线时长将进一步提升，借助分身去创作进而增加可支配收入

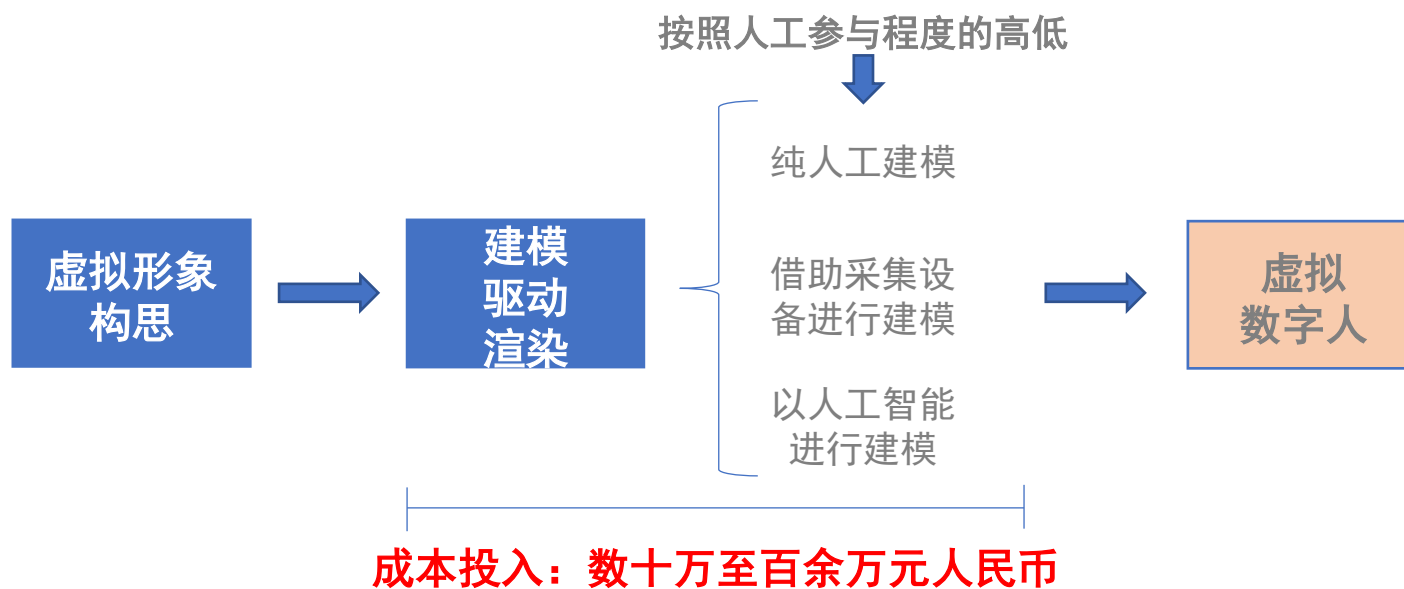
交互场景更多元且复杂，将衍生出更具吸引力的业态

新一轮匿名化社交的开始

虚拟人再次是AI的定向产业化：AI生成与驱动

- 虚拟人的本质是AI，这里的AI与“AI作画”的AI未来是同样的通用平台。故，虚拟人可以看作AI的定向产业化，即定向于元宇宙中人的身份映射这一方向的产业化过程，背后包括AI生产虚拟人形象、AI驱动虚拟人表情、动作等。
- 上游虚拟人的生成涉及计算机图形学、图形渲染、动作捕捉、深度学习与语音合成、人工智能等诸多技术，根据制作过程可归纳为建模、驱动、渲染三大关键环节。其中虚拟人的建模生成是核心步骤之一，也是较为困难的一环，奠定了未来使用的基石。
- 目前虚拟数字人的建模生成主要有三种，按照人工参与程度的高低，依次为**纯人工建模**、**借助采集设备进行建模**、**以人工智能进行建模**。同时涉及到相关的软硬件，包括建模软件、驱动软件、渲染引擎、拍摄采集设备、光学器件、显示设备等。因此，**虚拟人的上游投入集中在专业人才与相关的软硬件成本**。

图：虚拟人的生成主要分为建模、驱动、渲染三大环节



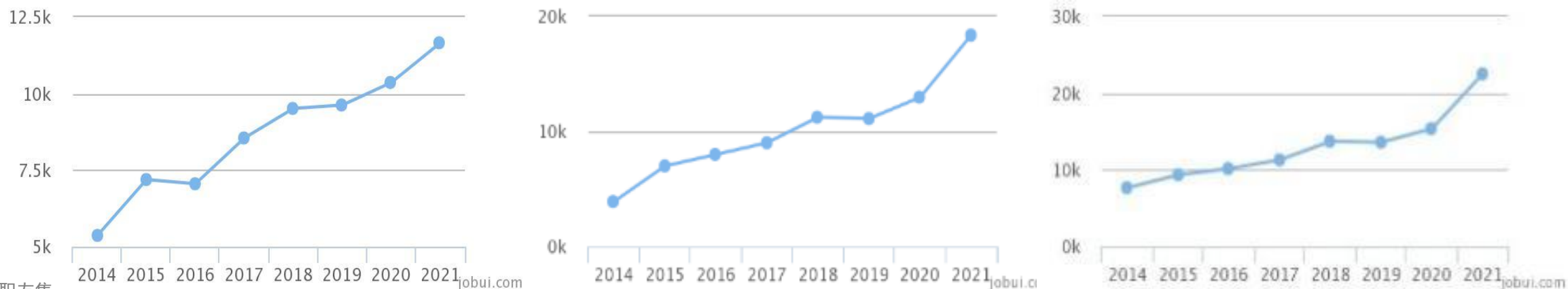
纯人工建模方式——成本高、产能低

- 纯人工建模分为三个步骤：

- » 1) 运用电脑软件画设计稿，描绘建构出人物的三视图；
- » 2) 根据设计图纸确定人物三维的图形，运用MAYA、3DMax、Zbrush等三维建模软件建立基本的三维模型；
- » 3) 除了建模师，虚拟人的运营需要团队的协作，虚拟人制作的其他环节还涉及到贴图师、绑定师、动画师、道具师、三维场景设计师等。

- 纯人工建模方式人工制作周期较长，且成本非常高。目前纯人工建模方式仍在广泛使用，市场上一些超写实的虚拟人均可以基于这种建模方式做出来，但人工制作周期较长，且成本非常高，源于该种建模方式的成本主要跟人力相关，一是取决于制作团队的配置大小，二是取决于建模的精度要求，虚拟人越逼真、越精细，所需建模时间与成本越高。
- 虚拟数字人相关制作人员平均薪资涨幅明显。2021年平均薪资涨幅明显，随着元宇宙热度的升温而一路走高，虚拟人相关制作人员的平均薪资水平在2021年涨幅明显，随着元宇宙热度的升温而一路走高，建模师、绑定师、动画师、道具师、3D场景设计师的月平均薪资水平分别为10-15k、10-30k、10-20k、8-15k、15-30k。

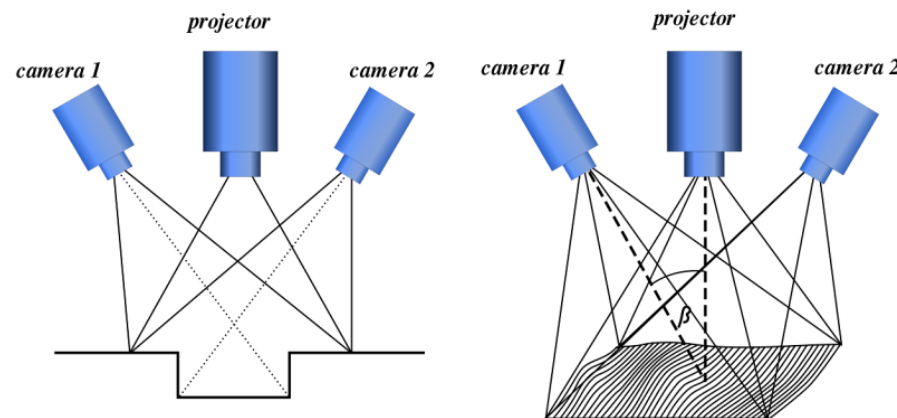
图：2021年建模师、绑定师、3D场景设计师的平均薪资水平出现明显的上涨



借助采集设备进行建模——成本适中、应用广泛

- 市场上大多数虚拟数字人是成本相对低的卡通画风形象，各类游戏中的游戏角色是早期的虚拟人，可以被称为**次世代游戏级别**。次世代游戏级别的虚拟人的渲染在电脑端/手机端就可以完成，其与超写实虚拟人在细节上有一定差距，即牺牲了一定的精细度，保证了画面的实时渲染。
- **次世代游戏级别虚拟人的建模大多是借助于外部扫描设备采集模型数据，数据的输入方式大致分为两种：**
 - » 1) **相机阵列扫描**：使用上百台相机进行360度的环绕拍摄，构建三维模型，基于模型进行数据的处理，然后转化成可以编辑的格式，再去进行相关的贴图、绑定等操作。
 - » 2) **结构光扫描**：利用三维扫描仪对人体、外部环境等进行扫描，建立点云数据（point cloud data），扫描精度越高，得到的点云数据越多，制作出的模型就越精细。
- **借助采集设备进行建模的成本主要为设备与人员成本，在几十万至一两百万元不等，可以应用于游戏、电影（不需要超写实、高精度）等领域。**

图：相机阵列扫描与结构光扫描的原理



利用AI进行建模——成本低、技术有待提高

- **人工智能建模的优势在于能够大幅降低成本。**通过人工智能技术进行建模，依据的是算法与机器学习，能将人力成本与时间成本都将降到最低。人工智能可以在很短的时间内快速生成虚拟数字人，建模的过程无需人工，后期修缮环节需要人工参与，进而将建模成本控制在极低的范围。
- 以近期大火的AIGC（人工智能生成内容）为例，相比于人进行绘画的创作，以Stable diffusion、Disco Diffusion为代表的AI作画工具只需要数秒就可以依据文字描述生成图片内容，一定程度上可以辅助人进行创作，节省大量的时间与成本。
“AI作画”是人工智能生成内容、技术落地的典型应用案例之一，展示了AI技术的突破进展，虽然还较为局限于二维的图片，随着技术的迭代，AIGC也会在三维建模方面发挥越来越大的作用。更重要的是，**AIGC会引导相关行业完成生产力与生产关系的重构。**
- **人工智能建模的难点在于关键技术的突破。**人工智能生成与驱动的虚拟数字人所呈现的效果受到语音识别、自然语言处理、语音合成、语音驱动面部动画等技术的共同影响。目前人工智能建模还远不能完全取代纯人工建模与借助采集设备进行建模的结果。人工智能建模结果就像Siri等虚拟助理对话一样，存在一些差距，但仍在不断改进之中。
- 我们认为，未来虚拟数字人的发展方向在于其能够应用于广阔的C端领域，因此：
 - I. **未来想要实现批量化生产虚拟数字人，重要前提是大幅降低制作门槛与成本；**
 - II. **人工智能技术是推动虚拟数字人规模化应用的重要基础，以人工智能为代表的技术将决定未来虚拟数字人的高度。**

交互型虚拟人的驱动，AI的重要性突显

- 虚拟人的驱动同样也涉及到技术路径的选择。虚拟人可以分为交互型、非交互型两种，其中交互型虚拟人则需要靠驱动技术来驱动动作、表情等，以达到虚拟人和外界环境进行互动与反馈。交互型虚拟人的驱动进一步又包括传统驱动方法与智能驱动方法。
 - » 传统驱动方法分成光学动作捕捉、惯性动作捕捉、Track 设备+IK 算法的动作捕捉等方法。现阶段，光学式和惯性式动作捕捉占据主导地位。传统驱动方法一般需要”真人+动捕设备”来进行驱动，这个后台的真人又称为“中之人”；
 - » 智能驱动方法是指通过AI技术，例如CV、ASR、TTS等来对虚拟人进行驱动的方法。该方式造价成本低，可以无限拓展，在未来有很大的想象空间。不过现阶段AI技术有限，一般需要结合合适的场景，通过较多垂直领域的训练才能达到商业可用的效果。
- AI驱动的虚拟人所呈现的效果受到语音识别（ASR）、自然语言处理（NLP）、语音合成（TTS）、语音驱动面部动画（ADFA）等技术的共同影响。

AI驱动的虚拟人技术	领域内相关公司
语音识别	领域内的公司如科大讯飞、百度、腾讯、阿里均有布局
语义理解	国内外在语义理解领域的进步较慢，语义理解这一交互环节的难度比语音识别高了数倍，相对做的好的公司如谷歌、IBM沃森机器人
语音合成	诸多大公司的客服机器人已能做到语音回复，但还不够真实、自然。若要做到每个人的声音个性化，并且能够快速生成自己的声音，仍有较大的工作量
人的形象驱动	目前市场上的巨头公司，如搜狗、科大讯飞、百度、腾讯等推出虚拟人的难度并不大，但难度在于，1）不同于卡通形象，在高逼真、高拟人的要求下，用AI驱动高仿真虚拟人的表情或动作仍有较大提升空间；2）快速生成虚拟人的同时，如何能够降低成本以及提高精确度

2022年虚拟人市场持续升温

- 2021年，国内虚拟人市场热潮开启，经过近两年的发展，新的虚拟人不断涌现，新的入局方不断加入，依靠自身的资源禀赋切入，或基于AI能力、运营能力等，探索行业应用发展。
- 与此同时，虚拟人相关投融资市场一直维持着高景气。根据雷报统计数据，截至2022年9月，今年国内虚拟人领域共计发生22起相关投融资事件（2021年全年为25起），其中世优科技与魔法科技均于2022年完成了过亿元的融资，其中魔法科技两轮融资共计1.3亿美元。

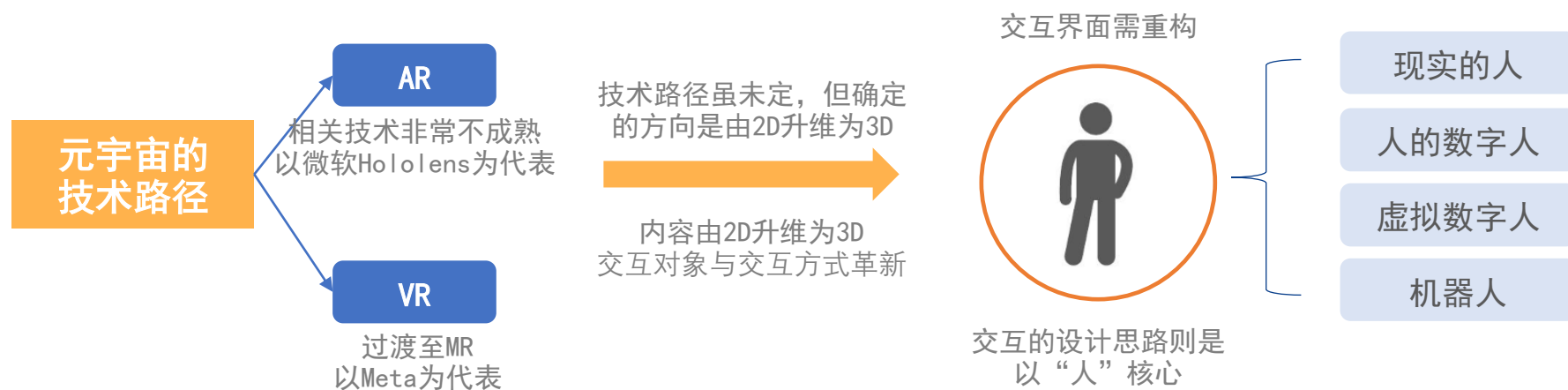
表：2022年新推出的虚拟人项目（部分大厂）

时间	名称	开发项目
2022. 9. 26	高德地图	宣布推出虚拟人导航“小高老师”，在现有的导航语音基础上融入了智能预测、AI学习、情绪感知力
2022. 9. 1	快手	在世界人工智能大会上推出“快手虚拟演播助手(Kuaishou Virtual Studio(KVS) ”
2022. 7. 21	百度	百度计划推数字人直播平台。2022百度世界大会正式开启，AI数字人希加加重磅带来百度智能云曦灵数字人直播平台，聚焦电商直播、品牌营销、互动娱乐等领域，实现超写实数字人24小时纯AI直播
2022. 6. 15	华为	在合作伙伴暨开发者大会推出数字内容生产线MetaStudio，以及全新版的数字人“云笙”，满足了各行业对数字内容生产、协同、融合以及应用的广泛需求
2022. 4. 27	快手	上线虚拟偶像歌手“神奇少女张凤琴”
2022. 2. 4	央视	推出专业服务类虚拟数字人AI手语主播
2022. 1. 22	京东	首个智能数字人客服芊言正式上岗
2022. 1. 1	蓝色光标	首个自有知识产权的虚拟人“苏小妹”正式推出并在北京卫视春晚亮相

资料来源：各公司官网和公众号

虚拟人将是时空构造中基础设施建设的“圆心”

- 元宇宙的技术路径，严格意义上是尚未定型的（目前仅基于视觉的技术探索），但目标是实现虚拟现实，即模糊掉虚拟与现实之间的边界。目前主流的技术路径为，一是直接以AR的技术路径去实现，但现阶段AR相关技术均非常不成熟；二是先VR再迭代至MR，以MR作为过渡。布局元宇宙较为激进的巨头Meta选择了以VR为主的技术路径，同时探索AR的发展；Apple预计将于2023年推出首款MR头显。
- 从2D到3D，交互界面需要重构，一方面是交互范式需要重构，另一方面是交互体验更贴近于现实物理世界的交互。交互的设计思路则是以“人”核心，最大限度的去实现人与周围环境的自然交互。这里的“人”不止是真实的人，还包括虚拟人（人的数字人、虚拟数字人）、机器人。



参照现实物理世界的交互方式，虚拟人很可能是元宇宙时空构造过程中，基础设施建设的“圆心”，虚拟数字人的作用在于最广泛承接2D（图文）升维成3D（视频）的交互界面；换个角度，任何入局方，无论是平台生态的构建方、B端企业的入局方案，以虚拟人为切入口或虚拟人的交互脉络为建设的主脉络，将被普遍采用的方式。

NFT：数字世界的魔法棒

NFT的魔法棒：从数字内容到数字资产

- » NFT是数字世界的“注册制”
- » NFT作为一种机制，金融属性最为突出
- » NFT是元宇宙世界运行的精髓之所在
- » NFT作用于“未来”的魅力更大

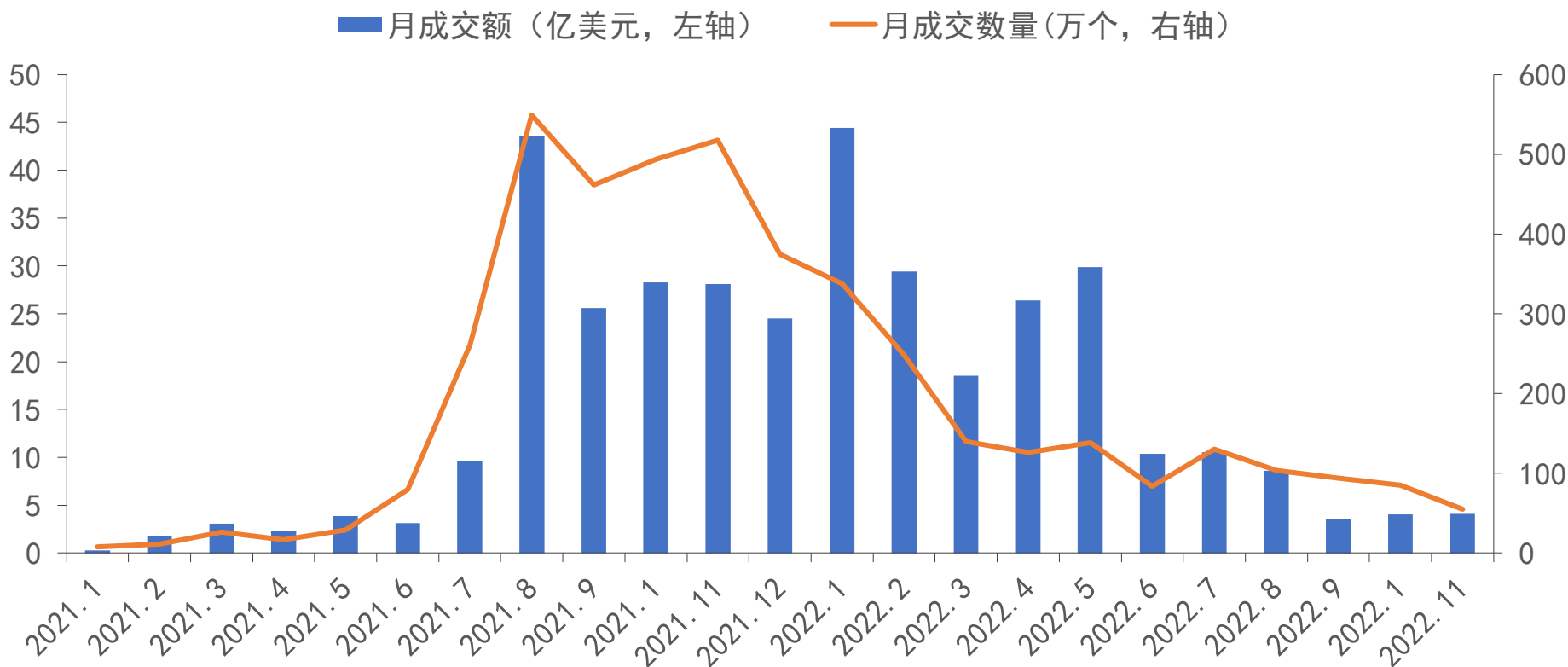
NFT的魅力在于孵化与孕育未来的各类“明星”

- » NFT的命门一定是“向前看”，去孵化与孕育新“明星”
- » 旧有的IP+NFT模式，效用仅能发挥出四分之一
- » NFT的交易链条越长，价值重估空间越大
- » NFT机制下，警惕顶级IP的过度泡沫化

2021-2022年NFT市场行情复盘

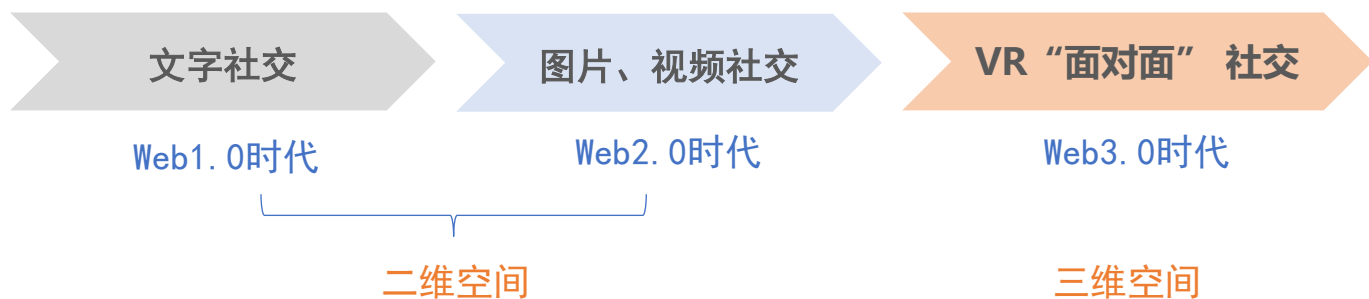
- 根据Nonfungible数据显示，2022年以来，NFT交易量持续下滑，尤其是自3月以来，NFT交易量较去年最高点时（2021年9-10月）缩水了80%。
- 海外相关上市公司的股价走势也相应走低，如Coinbase、Riot Blockchain、MARATHON DIGITAL等。

图：2021年1月-2022年11月NFT月度成交额与成交量



NFT是元宇宙世界运行的精髓之所在

- 当下国内仍处于深度数字化的过程中。数字经济时代，**新型生产要素——数据**一跃成为这一时代下的核心生产要素之一。社会经济随着生产要素的变革，快速推进从生产要素到生产力再到生产关系的全面、系统变革。
- 但数字化程度越高，安全的挑战就越大。在数字化时代，区块链提供了一个很好的信任机制。社会的数字化程度越高，对于区块链的依赖性越高、应用范围越广。未来元宇宙中数据资产将暴增。元宇宙一个高度立体且复杂的数字化时空，相应的数字资产也前所未有的丰富，若没有新技术、新机制的加持，数字资产的安全性、流动性都会面临系统风险。
- 如何来衡量各个创作者对于数字世界的贡献程度？目前来看，NFT的去中心化属性及金融属性，天然就与元宇宙的开放性相契合，区块链技术背书的NFT可以为元宇宙中的数字资产实现高效确权、交易。



- 元宇宙是新的计算平台，是三维数字时空
- Web3.0是新规则，一套去中心化的生产资料分配规则
- NFT是新要素，是在元宇宙场景中根据Web3.0规则分配的要素，具体可以表现为商品、数据、权益等

NFT是元宇宙世界运行的精髓之所在

- 案例：从文化产业的角度出发，NFT的出现正在为文化IP带来两大新变革。
 - » 一是，NFT开拓了IP授权新模式，提升了IP合作效率。
 - » 二是，NFT打破巨头垄断利润的格局，实现收益的共创共享。
- 这种以发行NFT来进行IP孵化与商业化的模式正在被越来越多的公司所采用，以共创+版权自主运营+社群运营的方式进行IP的运作，这正在为文化产业带来巨大变革。

传统互联网的运行模式下

传统IP授权

链条长、效率低、不透明

中心化的平台或机构获取大部分收益

传统IP创作和运营的模式下，IP的版权与大部分的收益往往归属于头部的中心化机构，强者恒强



元宇宙的运营模式下

以NFT的方式进行IP授权

大幅简化IP授权的手续和链条，可以极大提升IP合作的效率

NFT机制助力实现收益的共创共享

NFT带来了新的IP孵化与商业化的思路，群策群力，共创共享，使得个人创作者、中小机构或平台有了更多被“看到”的机会



NFT作用于“未来”的魅力更大

- 区块链相关技术的应用前景与各行业的数字化程度相关。对于数字化程度还不够高的行业，需要先对行业进行数字化改造，然后进行区块链的改造。因此，从区块链相关技术的应用来看，第一层是行业的数字化改造，第二层是区块链的改造，第三层是资产化的改造，最后才是金融化的改造，即资产的证券化，此时的NFT才能充分发挥其作为通证的价值。

区块链技术的应用层级

第一层 · 行业的数字化改造

第三层 · 资产化的改造

第二层 · 区块链的改造

第四层 · 金融化的改造，即资产的证券化

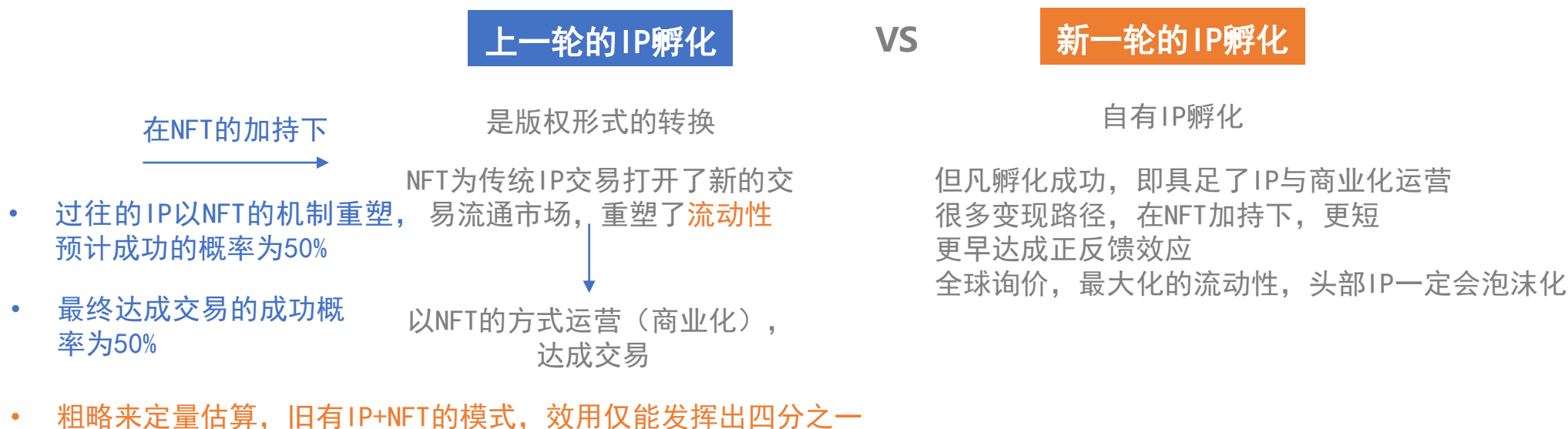
- 2021年，NFT的出圈使得人们过多地关注NFT的货币金融特性，而忽视其立足于实体经济、为实体经济服务的价值。当下的NFT更多是聚焦于现有的一些标的产品，如IP与NFT结合，形成既有IP的NFT化，以“协同方”的身份取得一批批NFT，进而构建各自的商业模式。这种方式的可取之处，在于赋予过往知名作品以切合当下的流动性，通过增加流动性创造未来的更多可能性。
- 当下NFT的潜力还未充分开发出来，相对于作用于已存在的IP，我们判断其作用于“未来IP”的魅力更大，即先构建业态，基于业态催生用户创造新的作品（有望成为IP），在创造的同时标配以NFT的机制，一方面，NFT在激发用户的创造性，另一方面，将NFT作为一种机制，可以重塑生产关系，进而刺激出更多商业模式与创新。

NFT的命门一定是“向前看”，去孵化与孕育新“明星”

- 海外NFT市场在模式创新上更具潜力、发展更加多元化，进而能焕发出比较大的活力，有非常活跃的NFT交易市场；而国内NFT市场的文化属性比较重，在版权保护、传统文化宣传领域进行发力，创新性不足。
 - » 海外NFT应用生态繁荣，有如OpenSea、Rarible为代表的综合化NFT平台，也有如SuperRare、Nifty Gateway为代表的专注精品化艺术品赛道的NFT平台。发展至今，NFT产品品类从最传统的NFT赛道，发展到游戏，再到二次元、元宇宙等应用场景中，并与新的元素结合，创新发展模式；
 - » 现阶段国内数字藏品大多聚焦于文创、IP领域，生命力尚未真正展现。
- 在元宇宙中，创造力是唯一的“增值性资产”。NFT的命门一定是“向前看”，去孵化与孕育新“明星”，而不是聚焦当下已有的IP进行NFT化，而是着眼于未来进行创作或基于已有IP去做一些创新后重塑这一IP的未来生命力。
 - » 确切的作品是用户未来行走元宇宙的通行证；
 - » NFT作为一种机制，绝非直奔向作品，而是孵化出作品的土壤，共识、机制、技术、生态，极具活跃度与创新性是结果；
 - » 作用于过往的IP，多数情况下是“投机主义”；过往的IP需要通过创新性的手法重塑“金身”。

旧有的IP+NFT模式，效用仅能发挥出四分之一

- 2015年起，IP泛娱乐化的新生态产业链的概念兴起；2021年，NFT迎来发展热潮，IP的运营逻辑发生变化。
- NFT化的IP有新旧之分：
 - » 新IP指的是在新技术条件下（如元宇宙）诞生出来的新内容，具备非常大的活跃度与价值；
 - » 旧IP是指在过去时代背景下所诞生出来的内容，借鉴新IP的打造方式进行NFT化，以期获得更多流动性。
- 新旧IP的NFT化的效率大不相同：
 - » 传统IP授权模式的授权流程长、标准化程度低。因此，旧IP的NFT化具备一定的门槛，这不仅需要跨行业，还涉及到标准授权、IP价值评估、商业价值计算、内容开发等一系列环节；
 - » 未来的新数字内容，具足IP与NFT运营（商业化），在诞生的初期就创新了IP授权模式，链条更短、变现效率更高。



NFT的交易链条越长，价值重估空间越大

- NFT推动内容资产价值重估的三重逻辑，分别是确权、数字资产价值得以充分被估算、扩大数字资产的范畴。随着NFT加持的数字资产越来越多，流动性也会快速分化。
- NFT的灵敏的价格反应机制是如何实现的？一是强势思维，即强者通吃，两极分化更严重；二是IP化运营，即一个NFT作品要想实现流动性的放大，需要IP影响力的加持；三是确切的作品，即确切的作品反哺IP影响力，能撑开更高的变现价值。
- （一）确切的作品交易链条越长，价值空间越广阔：以单一某IP为例，元宇宙时代确切的作品，并非一定是实物，其授权等业务的交易链条越长，价值空间越广阔。同样是2022年冬奥会火爆的冰墩墩与谷爱凌，他们自身即是“确切的作品”，展现出交易链条长短对价值空间的推高效应——冰墩墩这IP难以顺畅拓展广告代言与电商直播，而谷爱凌则顺畅的拓展了交易链条。
- （二）参与主体多、价值链条长、流动性好，则生态供应链的价值生产越大：以乐高产品为例，乐高的参与主体非常多，乐高的“确切的作品”在持续推新，在教育、电影、游戏等多个领域均有涉猎，且流动性极好。反观移动互联网时代的中国动漫产业，产业链很长，但产业链与价值链极为不匹配，难以商业化的诸多环节，流动性较差，而这是NFT可以明显起作用的地方；而交易链条越长，意味着有商业化可能性的环节越多，单一环节的商业化路径清晰且顺畅，最终的商业空间则不是单纯的“加法”，而是“乘数杠杆效应”
- （三）流动性折价、被低估，即是外界对“确切的作品”的认知不充分：NFT赋予“确切的作品”的流动性，类似于股市中单个股票的“股性”。任何“确切的作品”，不仅是创作者的创作，创作者更要花精力在作品的“显现”上，显现即被关注，有关关注度、关注度高，创作与显现，是元宇宙时代创作者经济，运行逻辑的奥义之所在。

NFT机制下，顶级IP一定会过度泡沫化

- NFT作为一种机制，虽然平权（针对于所有用户而言），但并非“平均”（对不同的作品或IP而言），NFT的机制会让“资产”估值两极分化更严重，顶级的IP一定会过度泡沫化；
- 商业化作为一种正常的市场运行行为，本来就遵循“二八原则”，加之NFT对流动性的推高，“二八原则”会走向更极致；从另一角度，过度泡沫化的顶级IP为所有的IP商业化打开原有的“天花板”，对整个市场的商业化而言，具有非常突出的示范效应；
- 以明星为例，衡量明星的商业价值有三个维度：粉丝数、路人缘（潜在粉丝数）、持续时间，前两者是根据现有数据让商家在衡量明星商业价值时的数据支持，后者则是建立在前两者之上的“弹性部分”，是预测数据。弹性部分”反而是发挥乘数效应的最重要的部分；
- 移动互联网带来的流量惯性，元宇宙时代的初期亦会承接，甚至马太效应会更明显——流量时代粉丝数量决定代言价值。年薪、奖金与代言、综艺活动、个人品牌、社交媒体价值，NFT加持下各项收入都有放大空间，甚至基于正反馈效应，各项收入都有“高估”的可能，故顶级IP一定是在马太效应加剧下，过度泡沫化的。

虚拟人与NFT亦会重塑企业的三张报表

虚拟人与NFT，据安信观察，很多入局方将数字人录入为费用，但它本质上将成为企业的“资产”，费用在利润表中，资产在资产负债表中；很多用户将NFT当成一种消费方式，体现为个人的支出，但它的价值呈现，是服务于个人创作的商业化变现，即创收，收入与支出分别对应现金流入与流出，是截然相反的性质。

1. 虚拟人作为未来的“明星”，相较于当下的明星演员，落座于资产负债表的“资产”而非“商誉”中；
2. 虚拟人的制作会增加企业的“成本”，但大幅降低企业的“销售费用”；
3. NFT的价值重估作用，将明显放大资产负债表中“资产”的量级；
4. NFT将带来丰沛的经营性现金流。

虚拟人的分类：身份型虚拟人与服务型虚拟人

身份型虚拟人与服务型虚拟人有着不同的商业模式，各有优势：

- **从变现能力的角度看：**身份型虚拟人的变现能力更强，通过品牌推广或代言、参演节目、直播打赏、发布音乐专辑、售卖周边等方式，依托粉丝经济进行变现，其特点是能带来较高的收入增量；而服务型虚拟人，其主要经济效益在于降低人力成本，虽然创造的经济效益较小，但服务型虚拟人的内核是人工智能，是元宇宙的基础构成要素，长期来看有非常可观的B端市场空间；
- **从数字资产的角度看：**国外是将虚拟人当作数字资产在进行运营，但是国内很多的服务型虚拟人的应用是将虚拟人当作一项成本进行投入，更多是对现有行业的赋能，在短期内实现降低成本的作用。

	服务型虚拟数字人	身份型虚拟数字人
定位	功能型，主要面向B端提供服务 ● 替代真人，提供简单功能服务 ● 多模态AI助手，提供日常陪伴、关怀等服务	身份性，用于娱乐和社交 ● 虚构的虚拟形象，生产虚拟内容 ● 真人的虚拟形象，用于社交、娱乐、游戏
应用	● 功能类服务：虚拟客服、虚拟导购、虚拟讲解员 ● 情感类服务：虚拟陪玩、虚拟陪聊、虚拟恋人 ● 直播带货：AI虚拟带货主播	● 虚拟IP运营：虚拟偶像、虚拟主播等 ● 人的虚拟化身：解决空间问题 ● 人的虚拟分身：给时间加杠杆
价值	● 对公司来说，作为成本项而存在 ● 降低已有的服务业成本，为存量市场降本增效 ● 内核是人工智能，是元宇宙的基础构成要素	● 对公司来说，作为数字资产项而存在 ● 在增量市场创造新价值增长点 ● 元宇宙中虚拟化身/分身的C端社交需求有待释放

身份型虚拟人的商业模式分析

- 收入端，虚拟偶像与真人明星类似，计入相关收入项：

- » 对于真人明星而言，经纪公司主要通过安排签约艺人参与商业活动，例如代言、商业宣传活动及其他商业活动；提供娱乐内容服务，例如出演电影、剧集及综艺节目，为包括企业客户、媒体平台、内容制作商及广告传媒公司在内的客户提供服务来产生收入；
- » 对于虚拟偶像而言，其打造的核心也在于IP的运营、人格化的培养，通过发布音乐作品、参加节目、产出内容来增加曝光度，积累一定的粉丝数量后再通过演唱会、广告代言、周边衍生、直播等路径进行变现，获取相应的收入。

- 成本端，虚拟偶像与真人偶像存在较大的不同：

- » 对于经纪公司来说，真人偶像或艺人的主要成本为支付给艺人的分成结算款。根据艺人管理合同，经扣除相关开支，公司会按艺人合同所签订的百分比与签约艺人分成收入，由于艺人影响力以及活动内容不同，该百分比因合同条款而异。根据头豹研究院的数据，虚拟偶像成本主要集中于前期的制作方面，占比可达60%，后期的规划运营与商务成本为40%；
- » 相较于真人偶像，虚拟偶像在前期需要投入一定的制作成本，后续可以持续使用，且其规模化或将带来边际成本持续下降。同时，无需考虑虚拟偶像的时间、精力等因素，虚拟偶像可高频次出席品牌活动、参与影视剧制作，公司所得收入不需要再与虚拟偶像进行分成，影视片酬与代言成本等均可能大幅降低，后续整体运营成本或将远低于真人偶像。

身份型虚拟人可以沉淀为公司的数字资产

- 虚拟人将革新所谓“轻资产”公司的资产负债表，大幅优化利润表。具体在一些财务指标上，虚拟人可以沉淀为公司的数字资产，落座于资产负债表的“资产”而非“商誉”中。“轻资产”公司对“轻资产”的收购或经营，是资产负债表中的商誉或应付账款，虚拟人则真正成为了公司的“重资产”；虚拟人要么降成本要么降费用，提升净利率或毛利率。
- 虚拟人更重要的是重塑商业模式。虚拟人既用于作品内容，也用于变现内容，将重塑内容产业的业务模式与商业模式。相关公司的业务模式、盈利模式将实质性改变。

	在公司中的定位	成本	收入与利润率
真人偶像	作为企业的员工而存在，且需签订劳动合同（5-10年不等，存在走红后流失风险），企业需要给艺人支付劳动报酬	成本主要包括前期培养费用与活动收入的分成，成名艺人的话语权变大，预计会要求提高自身的收入分成比例	所带来的边际效用不会有大幅增长的可能，若艺人要求分成比例的提升，毛利率则会下降
虚拟偶像	可以沉淀为数字资产，作为无形资产而存在	成本主要为虚拟人的前期制作与后续持续的驱动费用，相关营运成本提升幅度不大（收入无需与虚拟偶像分成）	随着知名度与影响力的提升，会带来公司收入的大幅增长，成本较为固定，进而带来边际成本的大幅下降，毛利率提升

虚拟人的制作会增加企业的“成本”

- **身份型虚拟人与服务型虚拟人两者的制作技术基础及应用场景均不同。**根据精细度的不同（卡通or超写实，2D or 3D等），虚拟人的制作成本介于几万元至数百万元不等。超写实虚拟人的制作成本根据质量的高低而有较大的差异。
- **制作成本方面：**
 - » 高质量虚拟人的制作过程较为复杂，由于对精度的要求高，需要采集光感的面部数据，成本超过百万；而正常质量虚拟人的制作成本在60-120万之间。其次，在驱动环节，高质量虚拟人最重要的部分是光捕，优质的光捕设备价格为200-300万，差一点的光捕设备价格为100万，光捕的成本取决于光学相机的数量与质量；
 - » 卡通形象虚拟人的制作成本要相对低很多。相较于超写实虚拟人的制作，卡通形象虚拟人无需在面部细节方面投入过多。此外，卡通形象虚拟人的角色细节较少，真实感较低，所以渲染的工作量也会较低。一般来说，卡通形象虚拟人的制作成本大多在10万元以内。
- **制作周期方面：**
 - » 高质量的超写实虚拟人的制作周期3个月，低质量的超写实虚拟人的制作周期为1-2个月；
 - » 高质量的卡通形象虚拟人的制作周期正常在1-2个月，低质量的形象的制作周期为1-2周。

身份型虚拟人可以大幅降低企业的“销售费用”

- 虽然虚拟人的制作会增加企业的“成本”，但可以大幅降低企业的“销售费用”。不同于身份型虚拟人，服务型虚拟人的外在形象并不是核心要素。在经济效益上，两者最大的区别在于，身份型虚拟人可以为企扩大营收水平，服务型虚拟人可以降低企业运营成本。
- 目前服务型虚拟人瞄准的应用场景为虚拟带货主播、虚拟客服、虚拟导购等，其主要经济效益不在于直接创造增量收入，而是在于降低人力成本。以AI虚拟主播为例进行分析，其作为服务型虚拟人而存在，此类虚拟人主播可以定位为“工具人”，在人设构建初期即可上岗产生效益。
 - » **虚拟主播的成本低**：对于品牌方而言，邀请中腰部以上的网红达人进行直播带货，所需付的坑位费、佣金等费用并不低。品牌方大多有自己的直播间，需要经常直播，根据前程无忧招聘信息，普通带货主播的月薪资水平约1万-2万元。目前已经有一些工具化的平台出现，如京东云数字人主播平台、百度智能云曦灵数字人直播平台等；
 - » **且虚拟主播可以24小时在线**：真人主播的时间、精力有限、流动性也较强。对于中腰部品牌方而言，一年的虚拟主播服务费在几千元到几万元不等，可以作为日常运营，补充真人未播的时段，如凌晨档，实现降本增效。



百度智能云曦灵数字人直播平台，聚焦电商直播、品牌营销、互动娱乐等领域，实现超写实数字人24小时纯AI直播

NFT将明显放大资产负债表中“资产”的量级

NFT最大的潜力就是重塑了资产的流通性，并带来了价值重估。相比于传统交易，NFT提高资产流动性。

- **NFT使得传统艺术市场有了互联网化能力，释放出交易频次与市场规模。**比如对于对于书画、文物、瓷器等实体藏品而言，再怎么稀缺，没有流动性也就没有价值实现，NFT将流动性带入交易稀少或不属于传统投资组合（例如艺术品、房地产、IP）的资产类别。
- **NFT重塑IP的价值与商业模式。**在资本市场，账面价值与公允价值有时会存在较大的差异。以影视IP为例，NFT与IP的结合为影视公司在商业模式的创新上提供了一个新的思路。在实务中，正常情况下，影视公司或视频网站对影视剧版权价值根据无形资产版权摊销政策予以结转，主要为两种摊销方法，加速摊销法与直线摊销法。不管采用何种摊销方法，一般来说影视剧的价值主要在第一年或前几年体现，即会贡献较多的收入来源，之后价值逐年递减，摊销结束后，其并不体现在公司的资产负债表中。但IP本身并不会消失，在一定的条件下，仍能够贡献收入，比如在NFT的加持下，一些老IP资产能够被盘活，被盘活的老IP已无成本，未来将成为名副其实的“现金牛”。

在商业模式上，NFT对影视公司而言，最直接的效益是能够盘活现有的IP资产，老IP焕发新生，0成本再次变现。这对于以虚拟资产为主的公司而言，NFT将会明显放大资产负债表中“资产”的量级。

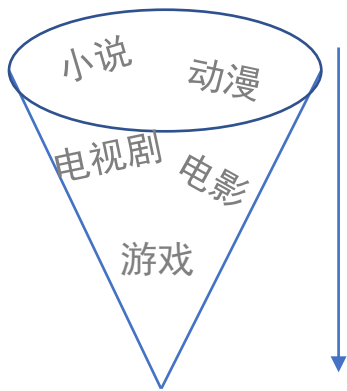
NFT将带来丰沛的经营性现金流

- 从区块链相关技术的应用来看，第一层是行业的数字化改造，第二层是区块链的改造，第三层是资产化的改造，最后才是金融化的改造，即资产的证券化。从区块链应用的层次来看，目前大家过多地关注了NFT的货币金融属性，对于NFT的认知也大多停留在数字藏品的范畴内，而忽视其立足于实体经济、为实体经济服务的价值。
- 从赋能实体经济的角度，NFT通过将非流动性资产转化为可交易证券，证券化有助于将现金流引导至本应成为借款人的创作者，改变其创作机制。
 - » 以具体案例来看：对于个体来说，每个人都有大量冗余的数据，但通过区块链与NFT，这些数据能够转化为个人数字资产。数字经济时代，新的数据要素市场产生了新的生产资料，生产资料依赖区块链技术解决数据可用不可见的问题，再通过NFT的价格发现机制（流通与交易），资产才更容易出售并转化为流动资产，例如现金。
- **NFT背后的技术释放了数字经济中一直未被开发的部分，区块链与Web3.0带来新的协议，也带来新的数据作为生产资料的价值分配。价值分配将为中小企业带来丰沛的经营性现金流，支持其生存发展，很平权，也应该能起到一定的共同富裕的引领作用。**

虚拟人与NFT的交叉是新一轮的IP孵化与商业化

- 过往一轮（2010-2019）的IP孵化与商业化，主体单一、链条过长，导致变现效率低，进而又拖累整个IP产业链的良性发展，IP产业作为孵化成功IP的土壤，整体上很“贫瘠”。
- 新一轮的IP孵化与商业化，不同于以往的原因在于：
 - » 1) IP的主体创作者种类大幅增加，即由单一类的创作者，增加至四类创作者；
 - » 2) 商业化环节缩短，当IP的商业化有了NFT的加持，链条明显缩短，能给更多的IP带来成功的机会。虽然门槛有提高但变现效率明显提升，变现效率提高是一种正反馈机制，滋养整个IP产业链，形成“肥沃”的孵化沃土。

过往的IP产业链，
是极为经典的“漏斗机制”



IP产业链条长，复杂的制作环节、发行环节

只有集齐了天时地利人和的大IP才有望成功，尤其是近十年的新IP中，成功的大IP凤毛麟角

新一轮的IP孵化与商业化



IP孵化：专业度、辨识度、局部影响力

- 从NFT加持IP的商业化的视角，这一轮的IP孵化需要有非常特别的专业度、辨识度、局部影响力。
- 新IP的孵化，**门槛在提高**（需要用更高形态的内容，如视频成为标配，甚至需要更高阶的内容形式，且要更符合算法的需求）、**竞争更激烈**（用户种类增加、算法驱动的竞争会更残酷）。
 - » 在迭代形式方面，如图文创作者向视频领域迁移转型正在成为普遍现象，但元宇宙这一虚拟现实的网络世界，未来的内容形态，一方面是3D沉浸式的、另一方面会增加互动性甚至是用户的主观能动性，未来创作者必然要从2D视频这一内容形式迁移至元宇宙的新内容形式；
 - » 在契合算法方面，一个具有创造能力的生成艺术家，他的天赋是创造出世界需要的精彩作品，虽然分工上并不一定符合商业化，但创作上仍然需要照顾“商业化”合作伙伴的潜在偏好，如抖音推荐上热门核心算法。
- 未来超级天王、巨星般的新IP，相较互联网、移动互联网为时代更难出现，因为随着观众细分领域的划分，很难出现超一线的作品了，明星与作品之间互相赋能，才能产生出琴瑟和鸣的效果。但会出现基于专业度、辨识度、局部影响力的各类IP。同时，新IP昙花一现的概率会大幅提高。

新一轮的IP将彻底区别于过往的“好莱坞IP”

过往一轮的IP孵化与商业化，是好莱坞模式在中国的移动互联网时代的落地；新一轮的IP孵化与商业化将彻底有别于好莱坞模式。

面向新一轮IP孵化，好莱坞IP的运营方式将存在很大的挑战：

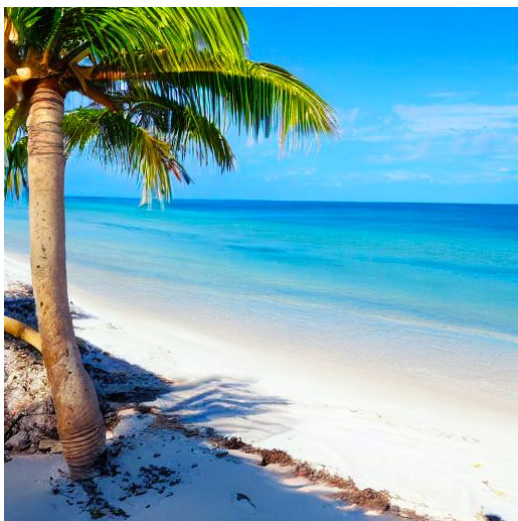
- 1. 现实与虚拟融合，时空边界被打破，导致传统的普世价值正在遭受巨大的挑战；
- 2. 内容形式及渠道也将发生变化，好莱坞IP也需要跟进快速变化的环境对内容做出合适的调整；
- 3. 好莱坞IP运营模式的成功得益于其内核是普适的情感，具有普遍的观众认可度与共情力；而元宇宙世界，具有专业度、辨识度、局部影响力的IP将为主流，专业内容具备认知门槛，不同背景的受众间不具有普适性，跨受众的转化存在难度。

过往的“好莱坞IP”	VS	新一轮的IP
发展于移动互联网时代		发展于元宇宙时代，基于新的规则
打造全球大群体都喜欢的IP		打造基于特定群体认同的分布式IP
以全球为市场，3分内容、7分周边衍生		看似将真正实现以全球为市场，实则依托小圈子进行商业化变现，7分内容、3分周边衍生
机制上是高举高打、强者恒强——以电影这一最高等级的内容形态打造新IP（制作成本巨高、失败概率不低）、全球范围内的IP知名度来做庞大的周边衍生变现、全球范围内各种维度的商业化变现反过来又强化IP的知名度。		机制上是锚定“商业化”来释放创造力——NFT的机制赋予每位创作者均等的“呈现/显现”机会、更大概率形成链条较短的正反馈、小步快跑的方式来激励创作者，从而在社会面上充分释放创造力
好莱坞的模式，更易打造垄断的商业模式，故全球的顶级IP阵营，基本是美国的六大来垄断		元宇宙模式，更易打造充分竞争的商业模式，形成层出不穷的新IP
好莱坞的盈利模式，少数成功IP来兜底多数不成功的IP，且单一成功IP的70%盈利靠周边衍生		元宇宙的盈利模式，则是新IP成功概率都拉高但较难出全球大爆款，单一成功IP的盈利，预计主要靠内容本身，而非周边衍生

新一轮IP商业化的门槛会整体提高

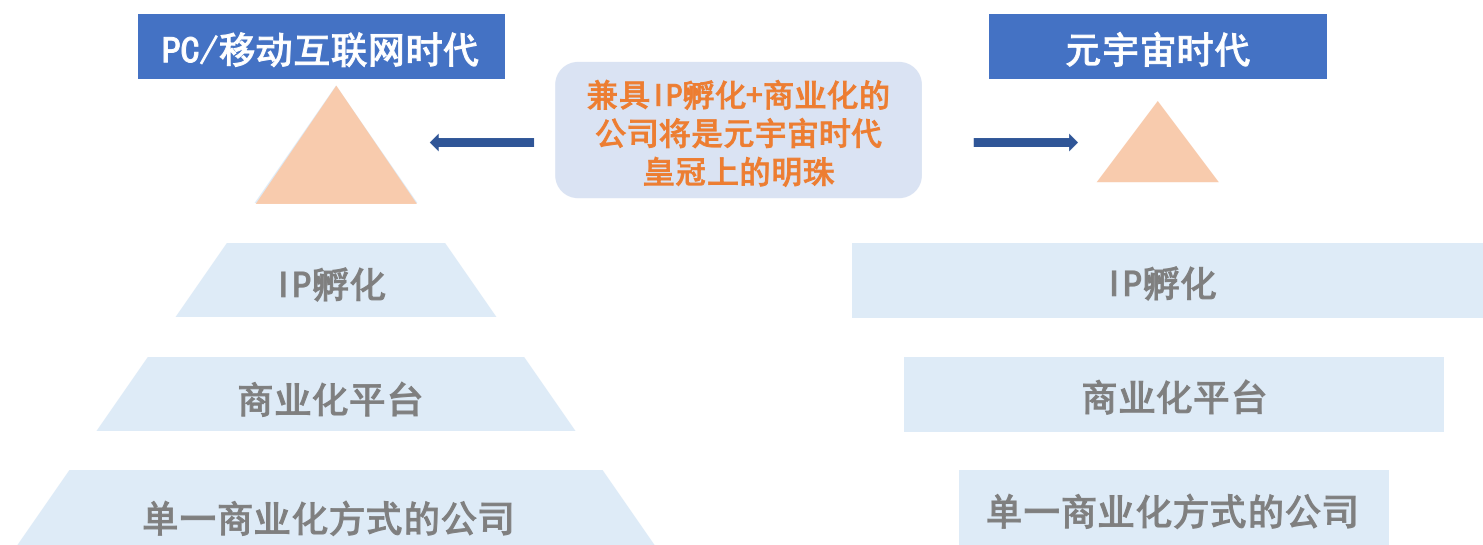
- 移动互联网时代经典的商业化方式，包括但不限于广告、游戏、电商、直播等方式。不同于过去的IP创作，在新的环境下，即在元宇宙时代，对于创作者而言，算法机制下的IP孵化门槛会整体提高，体现在效率、审美、创新，甚至是学习成本等方面，对未来创作者的综合实力要求会整体提高（比如文字创作者转化为短视频创作者，门槛不低）。
- AIGC出现确实大幅降低了人们的创作门槛，使得元宇宙时代“人人都能成为创作者”成为可能。但AIGC技术与工具带来创作便捷性与内容井喷的同时，也带来了内容创意的竞争加剧；在内容领域，长远来看，底层技术与模型并不是壁垒，内容本身的创意是壁垒。
- 因此，NFT加持下的商业化门槛，看似“平权”，但整体门槛提高，且昙花一现的IP将成为常态（来不及商业化或商业化不成功），故，反过来对IP的要求是必须持续出“确切的作品”，即人为延长“花期”。

图：Stable Diffusion生成的图片案例



新贵崛起：商业化平台类公司

- 本轮IP孵化与商业化的新逻辑中，有四类公司：
 - ① 兼具IP孵化+商业化
 - ② 只有IP孵化
 - ③ 商业化的平台型公司
 - ④ 只有某一项IP商业化能力
- 寻找元宇宙时代的“六大投资方向”，兼具IP孵化+商业化的它们将是元宇宙时代皇冠上的明珠——真正的王者、价值最高。
- 内容公司的估值永远且明显高于变现类（某一项商业化能力）公司——为何IP公司的价值高于商业化的公司？元宇宙时代可能会更加呈现出“内容为王”的凌厉，纯粹的商业化变现类公司，可能在估值上会进一步下降。



商业化平台类公司是最大的新机遇之所在

- 为何商业化平台类公司，将是元宇宙时代最大的新机遇之所在？

- » 首先，移动互联网时代，UGC或PUGC极大丰富，故流量后期的分发环节，竞争极为激烈，算法的机制下，“初面”定生死。元宇宙作为新一代计算平台，一开始会承载着打破这一局面的希冀；
- » 其次，历经持续多年的经营与积累，以平台为核心的分发体系，已建立起相对稳定的进入壁垒，对于内容创作者而言，当下的分发环节，各平台充分竞争已成型的纵深沟壑，已成为隐形的进入门槛。元宇宙时代，**打穿不同分发平台的商业化公司，有望应运而生**（有产业价值）；
- » 再次，商业化平台的商业模式，在元宇宙时代有望明显优化。商业化平台在移动互联网时代，需要耗费大量的精力去维系内容创作者的聚集与稳定，且盈利较为困难，这一状况有望在元宇宙时代彻底改变——虚拟人足够稳定；
- » 最后，相较于算法为灵魂的移动互联网平台，元宇宙时代的商业化平台，会叠加更多技术要素的加持，如AI、虚拟现实等。更多技术要素的积累将成为商业化平台的竞争门槛，长时间的技术积累带来抗风险能力明显提高。

- **被希冀、有产业价值、商业模式更稳健、更具抗风险能力的商业化平台，将成为元宇宙时代IP产业链上，最大的新机遇之所在，但它们的命门在于——打穿不同分发平台。**

◀ PART 02 ▶

对 Web 3 的警惕原因

对比元宇宙与 Web 3

对 Web 3 可以建立一种国际视野

● Web 3 在国内是2021年12月中旬开始，运作出来的概念

1. Web 3 的概念来自于2014年以太坊的联合创始人Gavin Wood，在他看来，Web 3 被描述为一个“基于区块链的去中心化生态系统”；区块链是Web 3 技术的基石，区块链技术通过构建一个去中心化的共识网络，使得没有一名网络参与者能够控制网络上的数据和协议；
2. 基于Web 3，人们能够不用一个中心化的平台来制定和监督游戏规则，一群参与者联合在一起就能构建一个社区(DAO)；人们之间的支付和清结算不需要银行这样的金融机构来提供“信用”，点对点的支付在Web 3 上能够直接完成；人们自身的身份和数据不用存放在一个中心化的平台，人们能够真正“拥有”并处置自己的数据；人们可以创建“非同质化代币”(NFT)，使得自己创作的歌曲、数字图片也能像以往的雕塑、邮票一样被收藏、流转、拍卖；
3. Web 3 的内核，是“治理”，这一点在海外看似“行得通”，在国内？
4. “去中心化”只是“反抗有理”，对“去中心化”的真实需求？

Web 3的概念是一个“去中心化的生态系统”，这个“去中心”既是计算上的“去中心”，也是存储上的“去中心”；

去中心化自治组织(DAO)：去中心自治化组织中没有一个中心人物（例如总经理，幽默的是多数入局方都想去做实质上的总经理）来对组织进行管理，一切组织的运行规则都写在智能合约中，由智能合约自动运行。

● Web 3 在国内是2021年12月中旬开始，运作出来的概念

1. “单向”的Web 1：1991年到2004年，特征是从互联网上获取信息，信息和知识由商业机构生产和组织；
2. “双向”的Web 2：从2004年开始，进入Web 2 时代，知识和信息更多地由用户进行生产，商业机构搭建平台进行信息的组织和分发，并由此诞生了诸如Facebook、Twitter、阿里、腾讯这样的超大型的平台公司；
3. 2022年火爆的“王心凌”、“小学教材”事件，本质上是Web 2 正在被反噬：过往天王巨星的作品，都是经过唱片公司（专业机构）严格筛选，网红爆款则是大众投票；老一代的教材编写、插画是顶级作家、画家；
4. 拥护者多数为币圈玩家，本质上是金融属性；延伸出来的NFT，在国外正在进行可能真正有意义的创新；
5. 马斯克也表示“Web 3 更像是一个营销活动，并不能变成现实”；概念的背后可能在模仿“扎克伯格”，中国版的“扎克伯格”的大旗是“Web 3”，对抗海外“扎克伯格”的大旗“元宇宙”；“扎克伯格”若能成功，则是“新（虚拟）航海时代”的哥伦布。

● 对Web 3 向来警惕的核心原因

1. 元宇宙作为下一代计算平台，与互联网、移动互联网的本质性区别，是真正开始作用于人的躯体；历来所有的技术都是作为工具、最高级别是作为人的“器官”，作用于人的时空观——折叠空间（也包括反向折叠）、给时间加杠杆（也包括去杠杆）；
2. 元宇宙的虚拟现实，是消除或自然化“真实与虚拟的边际”，涉及到两个问题：1）技术上怎么做到的？2）为什么用户会分不清？
3. 用户“分不清”，涉及到西方哲学，大致的结论是：用户分不清，是用户作为“人”，其意识与五感官直接交互；
4. 元宇宙的“真命硬件”不可能只作用于“五感官”，更重要的是触发人的意识与五感官的交互；
5. 元宇宙可以基于虚拟现实技术构造出生态，Web 3.0更偏底层/共识，其“七寸”在于金融属性的全球流动性。

元宇宙正由技术演化成生态

● 硬件入口的命题：基于人为何会“分不清”，去真正做到虚拟与现实的边界模糊

1. 框架：硬件入口&操作系统、底层架构、后端基建、AI、内容与场景、协同方；
2. 体系：元宇宙上半场（建设过程）、元宇宙下半场（反向叠加加入现实物理世界）；
3. 这样的框架与体系，将最难的“怎么实现”虚拟现实的问题留在“硬件入口”部分；这一部分可能与“脑机接口”未来会有直接关系；
4. 虚拟现实与人机协同/共生，是并列的两条技术路径，两者共同外推而演化成元宇宙生态；元宇宙只是智能的“内在部分”，与混合平台一起去实现真正的智能。



六大板块的产业轮动之2023

继2021年的硬件入口销量破1000万台、2022年底层架构的升级、2023年核心是MR这一硬件入口与操作系统（交互）的正式上场；

此外，2023年AI与后端基建有望进入景气上行期。

● 六大板块的产业轮动顺序

- » 首先，硬件与内容先行，硬件作为第一入口，硬件之上需要配套的内容相互促进发展，内容则以VR游戏、链游等元宇宙初级内容形态为主
- » 其次，底层架构要开始发挥作用，新内容/场景的制作、生产、运行、交互，依赖底层架构的大力升级（游戏引擎/工具集成平台等）
- » 再次，随着底层架构的升级带动数据处理的量级大幅提升，后端基建与人工智能才能真正发挥大的功效
- » 数据洪流下，即物理世界充分数字化后，人工智能的作用将越来越大，人工智能不仅依赖于底层架构与数字基建的完善，也非常依赖于内容与场景丰富的程度，此时AI将替代或辅助人去发挥建设性的作用，成为元宇宙中的核心生产要素
- » 最后落脚到内容与场景，相较于其他板块，内容与场景的变数最大，元宇宙将会催生出远超我们当下所预期的新内容、新场景、新业态，重塑内容产业的规模与竞争格局
- » 过程中有大量繁荣整个生态的技术、服务方，协同于每一轮轮动



元宇宙的跟踪研究，可以建立框架与体系

	前端硬件	后端基建	底层架构	人工智能	内容与场景	协同方
Meta	√		√	√	√	
字节跳动	√		√		√	
Unity		√	√		√	
Roblox		√	√		√	
Epic Games			√		√	
Nvidia	√	√	√			
高通	√		√		√	
Sony	√				√	
网易					√	√
腾讯		√	√		√	√
Amazon		√			√	
阿里巴巴		√	√		√	
Microsoft	√	√	√		√	
Google	√	√		√		
百度		√		√	√	√
Apple	√		√	√	√	
华为	√	√	√			
小米	√				√	
HTC	√				√	
Valve	√				√	
以太坊			√		√	

1. [Meta：最激进的元宇宙巨头](#)
2. [字节跳动：角力腾讯、效仿Meta，硬件+内容All in的打法](#)
3. [Unity：不只是游戏引擎，更是综合型创作工具](#)
4. [Roblox：简化版本的元宇宙世界雏形](#)
5. [Epic Games：游戏引擎+内容平台](#)
6. [Nvidia：深耕GPU技术，构建元宇宙技术底座](#)
7. [高通：元宇宙世界之“芯”](#)
8. [Sony：PS VR独树一帜，海量优质内容有望迎来价值重估](#)
9. [网易：RPG能力与虚拟人技术尤为突出](#)
10. [腾讯：元宇宙拼图较为完善，多赛道投资布局](#)
11. [Amazon：引领云计算行业发展，聚焦底层技术开发](#)
12. [阿里巴巴：云计算领域先发优势，展望电商多场景应用](#)
13. [Microsoft：最先提出并专注于企业元宇宙](#)
14. [Google：AI+云+安卓，元宇宙底层架构关键入局方](#)
15. [百度：决胜AI时代，积极布局元宇宙方向](#)
16. [Apple：最具潜力将VR/AR/VR设备推向主流](#)
17. [华为：河图，底层架构的“中国之光”](#)
18. [小米：VR硬件布局仍在推进，重点布局云游戏分发](#)
19. [HTC：打造五大VR产品线，构建硬件—内容良性闭环](#)
20. [Valve：硬件、内容与平台实力协同](#)
21. [以太坊：元宇宙世界经济基础](#)

Web 3 的跟踪，可以建立一种国际视角

● 2021年——：美元基金+美元基金看得懂的Web 3

- 逆全球化、退出通道（中概股退市潮）、Web 3 在中国的落地生根；
- 新增众多互联网、移动互联网新贵→国内本土的风投机构。

» 2019年——：中美贸易战开启的国家队

- 部委、省市、高校等的投资基金、种子基金；
- 上市公司、互联网/移动互联网巨头的产业基金；
- 新增众多互联网、移动互联网新贵→国内本土的风投机构。

Web 3 的跟踪，可以建立一种国际视角

● 1997-2017：美元基金+美元基金看得懂的互联网+走通上市

1. 1997年四通利方650万美金的融资，此后收购了华渊视讯（新浪网）；
2. 1998年搜狐融资220万美元；同年网易成立；
3. 新浪曹国伟、盛大张勇，探索出VIE架构；
4. 1999年中华网上市，盛大（上海）融资300万美金、阿里（杭州）融资500万美元、腾讯融资220万美元；易趣成立（最早的电商）；IDG投出15例互联网；ChinaRen（3000W么美金卖给搜狐）、易车网、当当、卓越网（7500万美金卖给亚马逊）先后成立；
5. 2002年开启Web 2.0：千橡互动、空中网；
6. 2005年百度上市、阿里收购雅虎（中国）；去哪儿、校内网、土豆、优酷成立；徐新（今日资本）、阎焱（赛富）、GGV纪源、熊晓鸽等IDG七长老、红杉、高瓴；
7. 2005-2008年A股大牛市，但互联网比较平淡；百度、腾讯、阿里上市；
8. 2006年DCM、经纬成立。

● 1997-2017：美元基金+美元基金看得懂的互联网+走通上市

- 2010年移动互联网元年，百亿、十亿独角兽陆续涌现：美团（2010）、小米（2010）、快手（2011）、字节跳动（2012）、滴滴（2012）、饿了么（2009）、陌陌（2011）、知乎（2011）、YY（2012）、小红书（2013），红杉等从淘金子→捡金子；
- 2010年开始，PC互联网新贵，开始成立自己的风投机构；
- 2011年俄罗斯Yuri 15亿美金投资京东，开启了投资靠抢、“上车关门”的模式；
- 2011年真格基金、2014年梅花天使、2015年更多互联网合并案；
- 2013年百度19亿美金收购91无线，“不能错过”；
- 2010年开始、一直到2020年的共享单车、在线教育，风投资本被反噬；
- 创业公司经过十年的“被教育”，亦被反噬。

● Web 3 最重要的应用——NFT

1. 国内外NFT发展极为不同

首先底层技术架构不同海外的NFT平台大多架构在去中心化的公链之上，而国内的主流NFT平台基本依托于联盟链运营，并非完全去中心化；

其次，NFT与数字藏品的定位及商业模式差异大，体现在NFT的价值定位、发行方式、定价方式、交易方式等方面，海外较为开放的创作者环境以及完全放开的二级市场，使得NFT产品的供给非常丰富；

再次，海外NFT发展路径更具创新性与多元化——海外的公链NFT与国内的联盟链数字藏品的禀赋不同，因此两者的发展路径相异，前者在模式创新上更具活力、发展更加多元化，后者在版权保护、传统文化宣传上正在发挥积极作用。

2. NFT在海外在真正引领出各个环节的创新

NFT一定会向着“功能性”发展，如映射现实资产的通证，映射虚拟道具的资产，映射链上身份的密码卡等；从收藏这一属性向工具、身份标识、实体权益凭证等更多维度发展。

PART 03

AI与后端基建将于23年下半年景气上行

AIGC：定位、产业价值及节奏上的参考

AI能力突破主要源于AI大模型化

AIGC将成为元宇宙时代的内容供给范式

2023年AIGC对产业链的三大提振

21年预判AI与后端基建将于23年下半年景气上行

- 我们21年年底预判，根据六大板块的轮动顺序，2022年核心在于底层架构的升级，以支撑23年年中开始的AI与后端基建的景气上行；
- 目前看，23年将以AIGC为主要业态，牵引出AI与后端基建的业务量上行，即景气度上行；
- 六大板块中，硬件与内容、AI与后端基建、底层架构与协同方，可以在研究上看作三组CP；
- AI与后端基建，尤其是AI，当下的现实物理世界并不一定需要它们，但元宇宙作为重构的时空，天生需要它们；
- AIGC是供给形态（可以作为内容、应用、场景、模式等具体形态），供给决定需求，新的供给将带来供需的正反馈，进而拉动AI与后端基建；
- AIGC只是AI与后端基建的应用方向之一，从后验的角度，应该较大/主要的一个应用；
- 2022年底，我们对AI与后端基建将于23年年中开始景气上行的判断不变。

AIGC：定位、产业价值及节奏上的参考

定位

- 从研究框架上，AIGC的**技术**支持是“AI”，但生成的**AIGC**属于“内容与场景”，背后需要“后端基建（**渲染**）”“底层架构（**引擎**）”的支持
- AI参与内容生产有**两类**：1) 凭借AI的高效率，**替代人**进行内容的生产；2) AI作为工具**辅助人**，或**人辅助AI**进行内容生产

产业价值

- 近期产业内（动漫/动画圈/艺术圈）对AI作画的巨大争议，且二级市场**对硬件产业链**的关注，转移至Web 3/AIGC
- 产业价值：元宇宙会释放用户的创造力/同时竞争会加剧，一条起跑线上的**创作者**（人的竞争对手）**增加了AI（虚拟数字人、机器人）**；借助于AI的协作，人作为用户的**创造能力及创作范围**将大幅提升

逻辑节奏

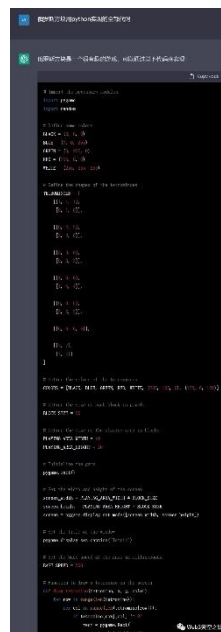
- 投资逻辑上，AIGC与数字人更有关联（本质上都是AI），与Web 3并无新增的关联；投资节奏上，2023年我们认为是硬件的大年，**2024-2025**是应用/场景/模式/内容的景气上行期，届时AIGC将大放异彩
- 门槛高，2023年或有局部小爆款应用

AI 绘画、ChatGPT 火热

- 2022年底AIGC、ChatGPT展现出更强的智能：
- ✓ 2022年8月美国新兴数字艺术家竞赛中，参赛者Jason Allen使用Midjourney完成的AI生成绘画作品《太空歌剧院》获得了“数字艺术/数字修饰照片”类别一等奖，引起科技界对AIGC的热议；
- ✓ 2022年11月30日OpenAI推出聊天机器人ChatGPT，表现出流畅的对话能力，上线五天全球注册用户数超过100万；
- ✓ 相关公司获得融资：Stability.AI获得1.01亿美金；Japer.AI获得1.25亿美元A轮融资等。



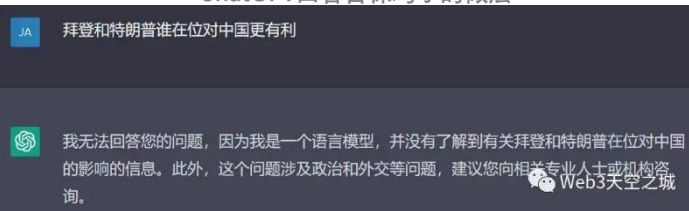
《太空歌剧院》



ChatGPT用python实现俄罗斯方块代码



ChatGPT回答宫保鸡丁的做法



ChatGPT能屏蔽敏感问题

AI 绘画、ChatGPT 实质为 AI 生成式内容（AIGC）

- AI 绘画、ChatGPT 聊天机器人均为 AI 生成式内容（AIGC）——其实质是“C”，但需要“AI”去“G”。
- AIGC 涉及到元宇宙六大版图中的四个——本质是**内容与场景**，需要**AI**与**后端基建**共同去生成，此外有大量的应用会被**协同方**给蹭出来。



内容与场景

协同方

人工智能

后端基建

内容与场景

协同方

AI 模型

算法

数据

算力

70余年AI进化史——从ANI正逐步走向AGI

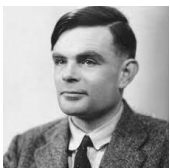
- 从1956年达特茅斯会议提出人工智能至今70余年，AI正在从弱人工智能（ANI）走向强人工智能（AGI）：

第一阶段 (1943-2006年)

特点

这是奠定人工智能发展的理论知识积累的时期，期间也曾产生过一些阶段性成果，但总体而言受限于数据规模与算力，人工智能更多还是停留在研究及理论阶段。

典型事件



- ✓ 1950年，艾伦·图灵提出“图灵测试”，给出判断计算机是否具有智能的方法



- ✓ 1956年夏天，美国达特茅斯学院举行了历史上第一次人工智能研讨会，被认为是人工智能诞生的标志。



- ✓ IBM于1985年开始研发深蓝，专门针对国际象棋的专家系统。1996年，击败了国际象棋大师加里·卡斯帕罗夫（Garry Kasparov）。

第二阶段 (2006-2016年)

算法、算力与数据规模都较前期有了质的飞跃，三者效用叠加成果显著，AI在计算机视觉、语音识别等多个领域的应用突破技术红线，具备商业化应用能力。

算法

- 2006年Geoffrey Hinton等人提出深度学习算法
- 后又迭代为DNN、CNN深度学习算法
- 形成较成熟的算法框架如Pytorch、Tensorflow、

数据

- 互联网的发展为人工智能模型训练提供了庞大的数据积累

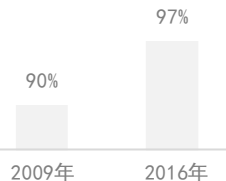
算力

- 芯片发展遵循摩尔定律，每18个月性能提升一倍
- 2012年全球最快的计算机红杉的计算速度为1.6亿亿次，较19个月前提升6倍

2014年人脸识别率提升至98.52%



2016年语音识别率提升至97%



第三阶段 (2016年至今)

以Alpha Go打败李世石为里程碑事件，AI赛道掀起投资热潮，在政策扶持、资本推动的背景下，AI被应用到越来越多的场景中，如安防、交通、自动驾驶、医疗等。

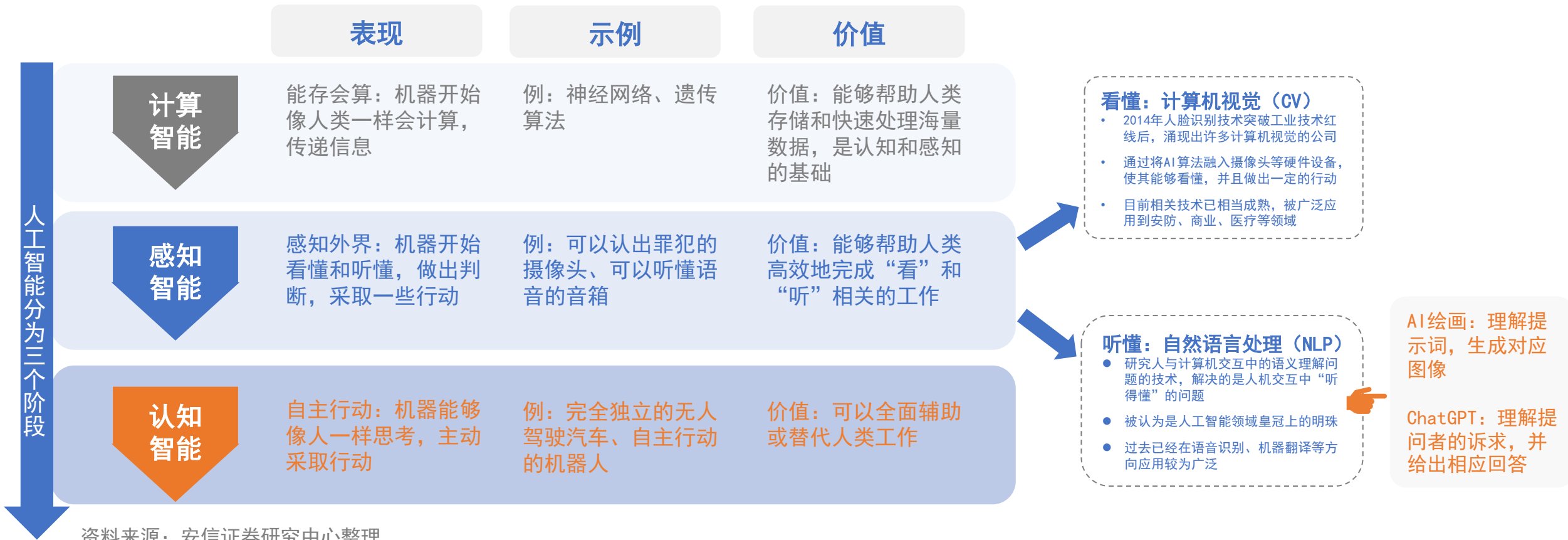


- ✓ 2016年3月由谷歌（Google）旗下DeepMind公司开发的AlphaGo（基于深度学习）打败了人类围棋世界冠军李世石



AI GC表明机器不仅能“看懂”，也将能“听懂”

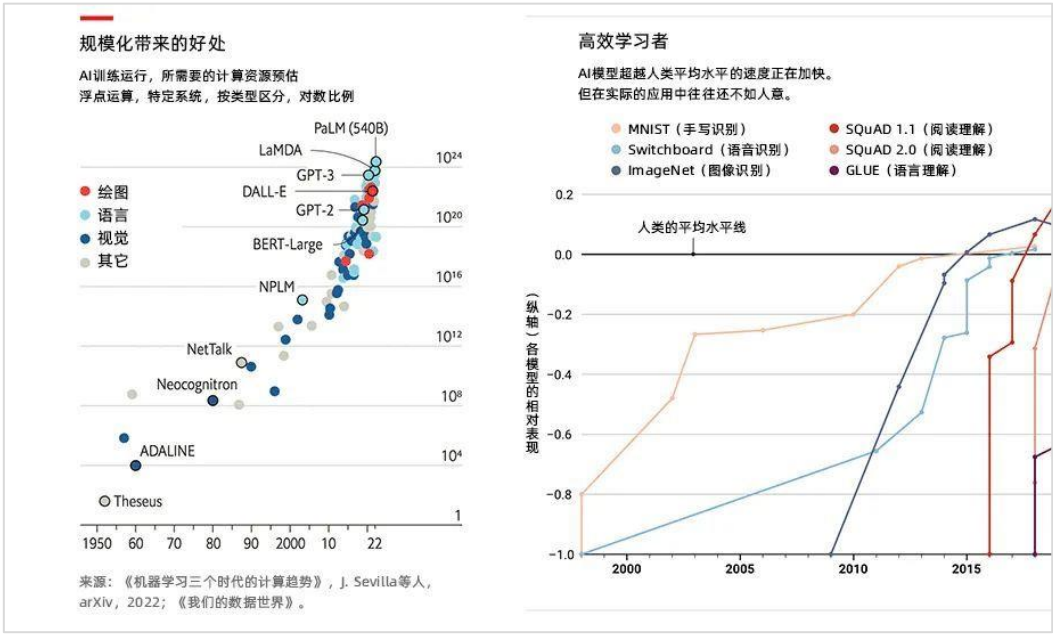
- 人工智能发展分为计算智能、感知智能、认知智能三个层次，目前正处于提升感知智能能力的阶段；
- 在感知智能阶段，其核心技术包括计算机视觉（使机器能够“看懂”）、自然语言处理（使机器能够“听懂”）等；
- 上一轮AI发展热潮源于计算机视觉技术突破使机器具备“看得懂”能力，而AI绘画、ChatGPT的表现反映出自然语言处理技术正在突破，机器或将具备“听得懂”的能力。



AI能力突破主要源于AI大模型化

- 研究发现神经网络越底层数据越抽象，越上层数据越与具体任务有关，因此可以通过预训练泛化大模型、针对不同任务微调实现效率提升；2017年提出Transformer模型能够实现自我监督学习，无需标注数据，使训练数据规模可以大幅提升。基于此，以OpenAI推出GPT系列模型为标志，AI大模型逐步成为共识，互联网巨头纷纷走上自研AI大模型的路径。
- AI大模型表现突出，且其能力正加速迭代中，目前已经在手写识别、语音识别、图像识别等领域超过人类平均水平。

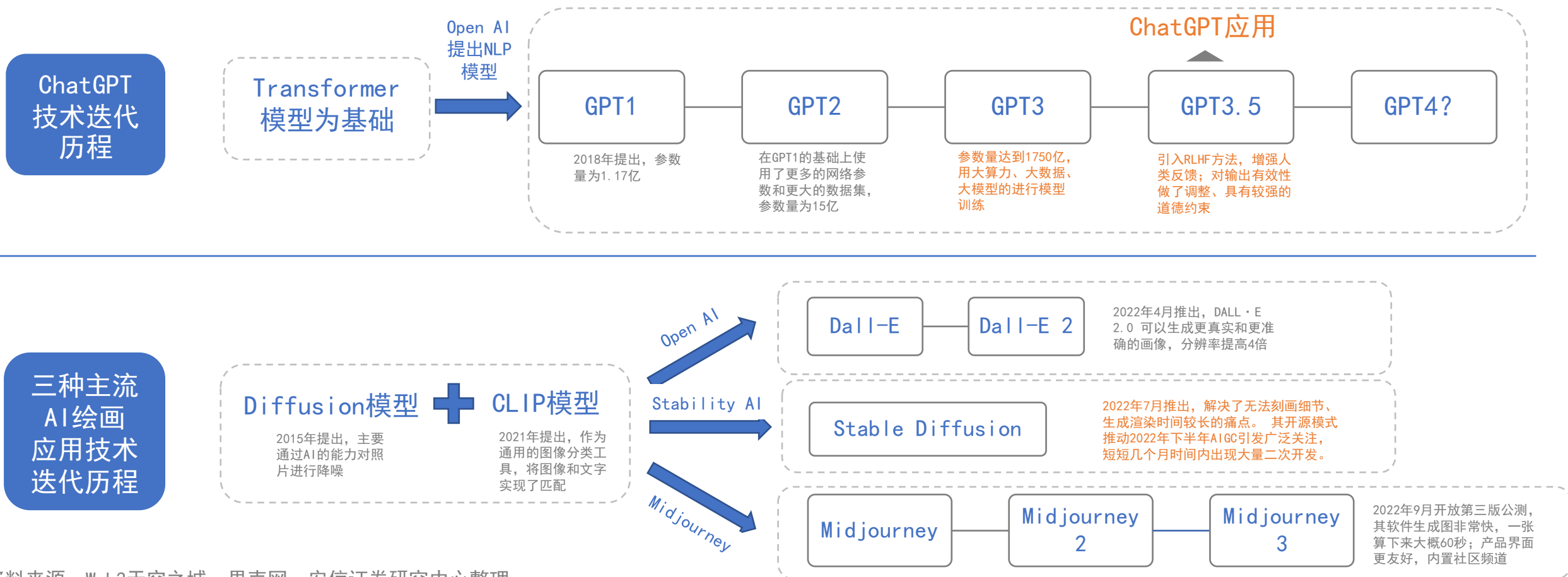
公司	模型	发布时间	参数	预训练数据量	模型类型
OpenAI	GPT	2018年6月	1.17 亿	约5GB	自然语言模型
	GPT-2	2019年2月	15 亿	40GB	自然语言模型
	GPT-3	2020年5月	1750 亿	45TB	自然语言模型
谷歌	BERT	2018年10月	-	-	-
	Switch Transformer	2021年3月	1.6万亿	-	-
华为	鹏程·盘古	2021年5月	2000亿	1.1TB	中文预训练语言模型
智源	悟道2.0	2021年6月	1.75万	4.9TB	双语多模态预训练模型
浪潮	浪潮源1.0	2021年9月	2457亿	5TB	通用NLP预训练模型
微软&英伟达	Megatron-Turing	2021年10月	5300亿	-	自然语言生成
阿里云	达摩院M6	2021年11月	10万亿	-	跨自然语言、图像的多模态AI模型
百度	鹏城-百度·文心	2021年12月	2600亿	-	NLP大模型



得益于大模型的支撑，
AI在多个领域获得突破

ChatGPT、AI 绘画以AI大模型迭代升级为基础

- ChatGPT、AI绘画是大模型迭代升级的产物，AI大模型让AI生成和理解能力有了显著提升，推动内容形式从单一走向多元，内容质量也较以往大幅优化；
- 目前模型的迭代升级呈现加速状态，如2022年先后推出的Dall·E2、Stable Diffusion、Midjourney 3间隔时间较短，生成效果及交互性均有显著提升。



或可期待AIGC生成3D内容？

- AIGC的本质是AI 赋能内容生产，AI 绘画、ChatGPT获得高度关注的原因在于AI 生产的内容在文字、代码、图像等领域都具有较高的可用性，而且体现出一定的创造性。
- 从写诗机器人到AI 绘画模型，目前AI 已经能够实现文字到图片、图片到图片的转化，而且Google发布模型已具备文字到视频的生成能力。
- 随着模型能力的提升，推断AI 或将逐步具备生成3D、数字孪生等内容的能力。

	2020年之前	2020 年	2022 年	约 2025 年?	约2030 年?	约2050 年?
文本领域	诈骗垃圾信息识别 翻译 基础问答回应	基础文案撰写 初稿	更长的文本 二稿	垂直领域的文案 撰写，实现可精 调（科学论文等）	终稿，水平高于 人类平均值	终稿，水平高于 专业写手
代码领域	单行代码补足	多行代码生成	更长的代码 更准确的表达	支持更多语种 领域更垂直	根据文本生成初 版应用程序	根据文本生成终版 应用程序，比全职 开发者水平更高
图像领域			艺术 图标 摄影	模仿（产品设计， 建筑等）	终稿（产品设计， 建筑等）	终稿，水平高于 专职艺术家、设 计师和摄影师
视频/ 3D/ 游戏领域				视频和3D文件 的基础版/初稿	二稿	AI版Roblox 可依个人梦想定 制的游戏与电影

大型模型实现程度 ● 初级尝试 ● 接近成熟 ● 成熟应用

资料来源：红杉汇

text2text



eg: AI 写诗

text2img



eg: Dall·E根据提示词生成图片

img2img



eg: 将动画中的人物形象用AI生成更具细节的人物

text2video



eg: Google发布的AI模型 Phenaki 生成30s视频只需要22s

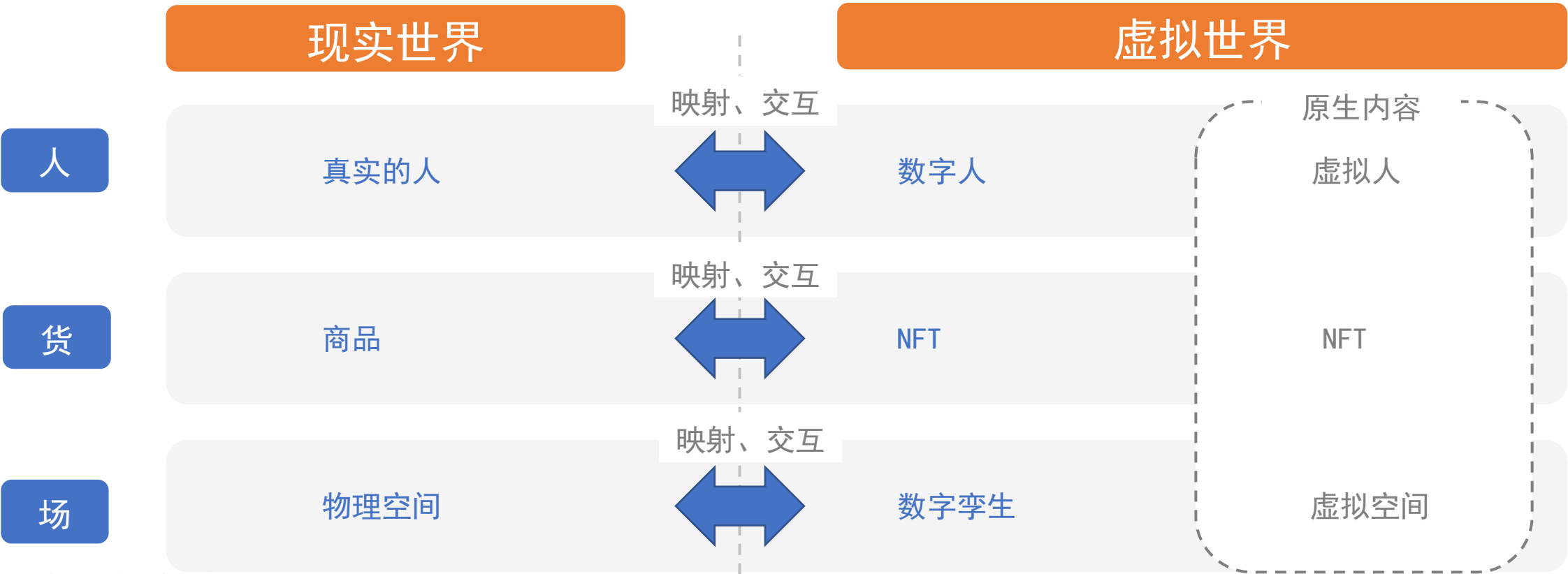
3D? 数字孪生?



元宇宙时代内容需求将指数级增长

随着计算平台的迁移，元宇宙时代对内容的需求将迅猛提升：

- 现实物理世界的人、货、场均需虚拟世界有一一对应，同时将会有大量的虚拟世界原生内容，如虚拟人、NFT、虚拟空间等，内容呈现形式变得多样；
- 元宇宙时代交互主体增加，一方面虚拟人的内核由各种内容所支撑，其次需要输出内容与其他主体交互，因此对交互内容的需求增加。



AIGC将成为元宇宙时代的内容供给范式

- 随着几代计算平台的持续升级，用户对内容的需求迅猛提升，供给端的范式也会持续升级，继互联网时代之前的PGC、互联网时代的UGC，AIGC是元宇宙时代的内容供给范式——有条不紊的提供数量级飙升且内容品质有保障的内容。
- 目前AIGC处于仍助手阶段，如通过AI绘画方式帮助创作者生成创意等，随着技术能力的提升，预计AIGC将作为内容供给方自主生成创意内容，从而大幅提高生产效率，并降低内容生产成本。



AIGC：以内容/场景或垂类硬件广泛存在

● AIGC的终局——元宇宙世界中广泛的内容、应用、场景，以及未来现实物理世界的分布式垂类硬件：

- ✓ 元宇宙：虚拟3D交互世界是由大量的数字内容构成，主要为AI生成式内容，体现为虚拟人、3D场景等广泛的内容；
- ✓ 物理世界：AI生成内容为内核，依场景需求借助一定的硬件形态呈现出来，即为垂类硬件。

内容、应用、场景



养老陪伴



迎宾



生产协作



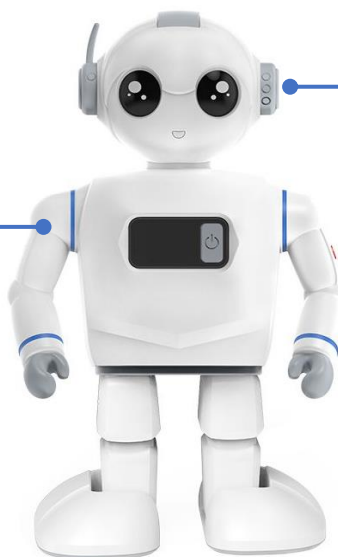
客服

.....

分布式垂类硬件

外在：体现为不同形态的垂类硬件

- 具备不同形态以应对不同场景的需求
- 具备灵活性、同时满足特殊的材料要求



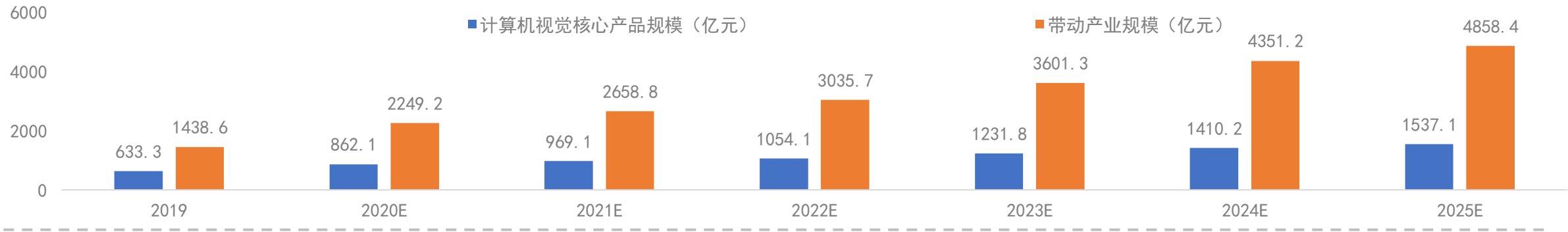
内核：AI 模型+数据+算力赋能

- 能够看懂、听懂外界的环境
- 具备对话能力（取代搜索）
- 能够做出交互行为（AI绘画、编写代码、发出指令等）

AIGC或预示下一个千亿产业机遇

AIGC获得关注也反映出NLP技术在商业化应用层面取得突破，参考上一轮计算机视觉技术突破所带来的产业机遇，我们认为新一轮市场机遇已在酝酿中，有望诞生下一个新的千亿市场：

- ✓ 根据艾瑞咨询，预计2025年计算机视觉核心产业规模将达到1537亿，带动相关产业规模达到4858.4亿；
- ✓ 行业中涌现出一批新兴的AI创业企业（如Alpha Go背后的DeepMind），同时巨头业务重心向AI方向转移（如百度等）。



典型AI创业公司（计算机视觉方向）



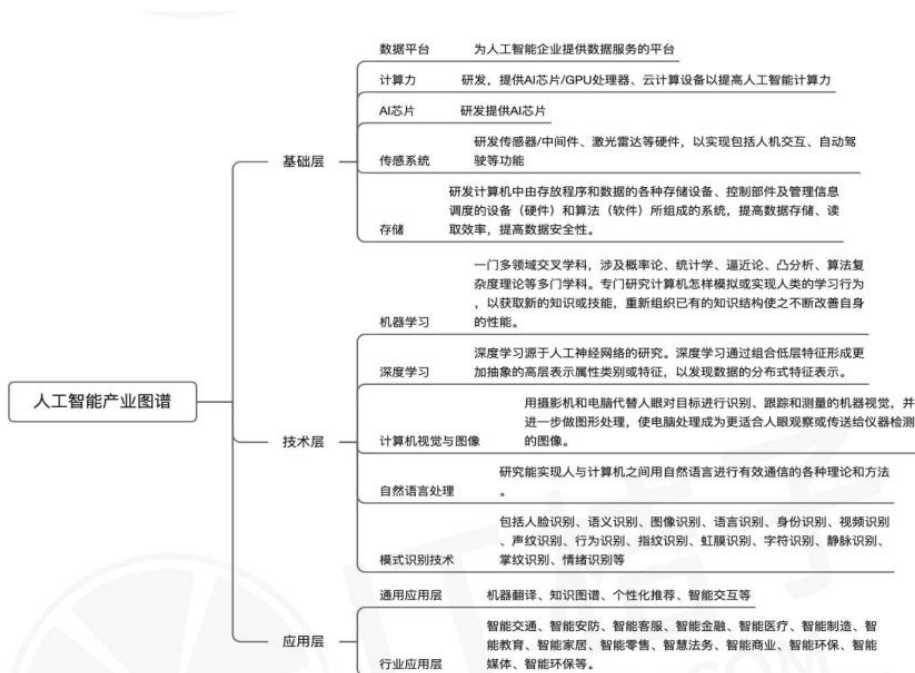
	公司	时间	主要内容
海外	Google	2017年10月	在Google O/I开发者大会上提出，公司未来发展战略将从移动为先（Mobile First）调整为“人工智能为先”（AI First）
	Facebook	2016年4月	扎克伯格在F8年度开发者大会上发布Facebook未来十年规划，将人工智能、VR/AR、连接作为公司未来5-10年战略重点
	微软	2017年5月	人工智能取代“移动为先、云为先”战略成为微软新愿景；
	亚马逊	2017年4月	亚马逊首席执行官杰夫·贝佐斯在股东信中表示将人工智能纳入公司长期发展规划
	Apple	2017年10月	苹果高管Heff Williams表示人工智能将会成为今后苹果产品的重要基石，包括iPhone和Apple Watch产品线
国内	百度	2016年7月	李彦宏在百度云战略发布会上发布了百度云“人工智能+大数据+云计算”三位一体的发展战略；2016年百度大脑平台开放，飞桨上线；
	腾讯	2017年11月	在腾讯全球合作伙伴大会上公布了“AI in All”战略，包括“基础研究——场景共建——AI开放”三层AI架构，落地社交、内容、游戏、医疗、零售等八大场景。

资料来源：虎嗅、经济观察报、环球网、Wind等

预判2023年AIGC对产业链的三大提振

● 我们预判此轮AIGC的火爆将主要对产业链带来以下三方面的提振：

- ✓ 算力及基建：AIGC应用需求增加及其对精度要求提升，将推动其背后的超大规模AI模型对算力的需求呈现急速增长，带动相关算力基础设施建设；
- ✓ 算法及模型：GPT系列模型突出表现使基础模型的商业价值被认知，Mass或将成为更适合的商业化模式，能使更多公司专注底层核心技术研发；
- ✓ 内容与场景：此前AI主要应用于To B/G的场景，而AIGC的本质是内容，预计这一轮AI在To C场景的应用将被充分挖掘，带动AI技术进一步向C端消费者普及。



- 预计2023年
提升至70%

Figure 1: Proportion of Basic Computing Power, Intelligent Computing Power, and Supercomputing Power from 2015 to 2021. The chart shows that the proportion of intelligent computing power has increased significantly, reaching 50% in 2021.

Year	Basic Computing Power (%)	Intelligent Computing Power (%)	Supercomputing Power (%)
2015	98	1	1
2016	95	3	2
2017	92	5	3
2018	85	12	3
2019	70	28	2
2020	55	45	0
2021	45	50	5

算力及基建：2023年或仍为智算中心建设“大年”

- 截止2022年2月，全国建成并投入运营的智算中心为9个，规划及建设中的智能计算中心为18个。考虑2022年疫情反复影响项目施工进度，预计2023年仍为智能算力中心建设“大年”。

国内建成并投入运营的智算中心项目

	智算中心名称	目标规模	建设进度	时间	签约/中标企业	所在城市
1	武汉人工智能计算中心	200P	投运	2021. 05	华为等	武汉
2	合肥先进计算中心	双精度计算峰值 12PFLOPS 整数计算峰值 256PFLOPS	投运	2021. 06	曙光等	合肥
3	南京智能计算中心	800P	投运	2021. 07	浪潮、寒武纪等	南京
4	中国电信京津冀大数据智能算力中心	-	投运	2021. 08	中国电信等	天津
5	未来人工智能计算中心	300P	投运	2021. 09	华为等	西安
6	浙江（长三角）新一代全功能智能超算中心	-	投运	2021. 09	曙光等	嘉兴
7	中原人工智能计算中心	100P	投运	2021. 10	华为等	许昌
8	哈尔滨人工智能先进计算中心	100P	投运	2021. 12	-	哈尔滨
9	商汤科技人工智能计算中心	3740P	投运	2022. 01	商汤科技	上海

国内规划或正在建设的智算中心项目

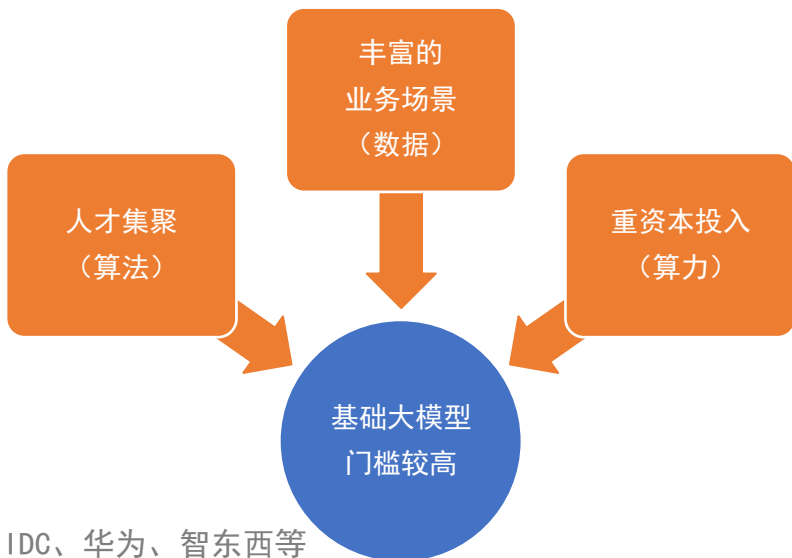
	智算中心名称	目标规模	建设进度	时间	签约/中标企业	所在城市
1	北京数据经济算力中心	300P	发布相关规划	2021. 04	-	北京
2	腾讯智慧产业长三角（合肥）智算中心	-	开工建设	2021. 07	腾讯等	合肥
3	庆阳智算中心	-	开工建设	2021. 08	-	庆阳
4	海南人工智能计算中心	-	签约	2021. 09	华为等	海南
5	沈阳人工智能计算中心	总规模300P 一期100P	开工建设	2021. 09	华为等	沈阳
6	大连人工智能计算中心	300P	开工建设	2021. 09	华为等	大连
7	成都智算中心	-	项目建设中标公告	2021. 09	华为等	成都
8	徐州国科智算中心	-	项目揭牌	2021. 09	中科院虚拟经济与数字科学研究中心等	徐州
9	克拉玛依浪潮智算中心	-	揭牌	2021. 09	浪潮等	克拉玛依
10	无锡智能计算中心	-	签约	2021. 10	浪潮等	无锡
11	昆山智能计算中心	500P	中标公告已发布	2021. 11	寒武纪等	昆山
12	杭州人工智能计算中心	首期40P	中标公告已发布	2021. 12	天宽科技等	杭州
13	深圳市人工智能融合赋能中心	-	开工建设	2021全年	-	深圳
14	中国-东盟算力中心	-	规划建设	2022. 01	-	南宁
15	广州人工智能公共算力中心	-	中标公告已发布	2022. 02	广电运通等	广州
16	长沙人工智能公共算力中心	1000P	开工建设	2022. 02	-	长沙
17	青岛人工智能计算中心	100P	主要招标工作完成	2022. 02	华为、杰正泽湖等	青岛
18	河北人工智能计算中心	100P	揭牌	2022. 02	华为、神舟新桥等	廊坊

模型层：不同层级模型的准入门槛不同

- AI模型采用预训练-微调的思路，对应产业应用中逐步形成基础大模型+行业大模型思路，目前百度、华为等均采用这种布局方式。
- 我们认为基础模型与行业模型对新公司的门槛不同：
 - ✓ 基础大模型：门槛较高，中美均以互联网巨头或其投资的公司为主，也包含部分政府支持的科研机构等，主要是押注AI作为新一代经济的基础设施；
 - ✓ 行业大模型：相对基础大模型而言，对创业公司更友好，其可专注垂直领域对其进行深度赋能。



行业大模型：创业公司更灵活，可专注垂直领域进行深度赋能
+
基础大模型：巨头为主



互联网巨头为主要玩家



模型层：Mass模式有望推动AI 公司价值重估

部分AI公司的商业逻辑或将重塑，进而推动其商业价值重估：

- AIGC特别是文生图的应用使模型作为服务（Modle as a Service, Maas）商业模式获得关注，以OpenAI为例，如果调用其DALL·E模型生成图片，收费标准为0.016-0.020美元/图；如调用其语言模型，收费标准为0.0004-0.02美元/tokens；
- 过往AI行业主流的商业模式表现为AI是一种能力而非产品，因此AI公司不得不投入巨大成本提升研发实力构筑壁垒但无法凭其获利，又需要深入场景提出满意的产品（产品力），高研发投入与过长的变现链条形成商业化困境；
- Maas作为新兴商业模式能够为模型本身定价，使专注于模型研发的AI公司能直接获得收入，有望极大改善其商业化问题，同时GPT-3作为案例为市场展示了其极大的商业潜力，或将推动该类公司价值重估。

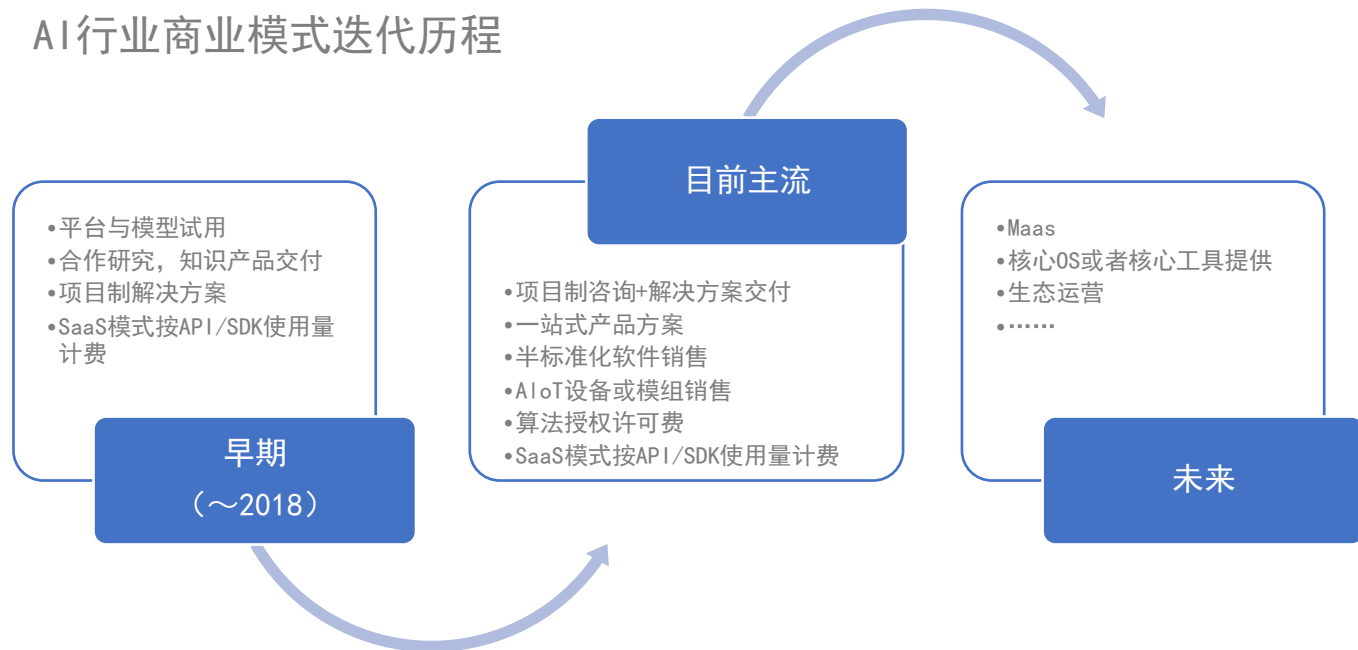
OpenAI的图像模型收费模式

RESOLUTION	PRICE
1024×1024	\$0.020 / Image
512×512	\$0.018 / Image
256×256	\$0.016 / Image

OpenAI语言模型收费模式

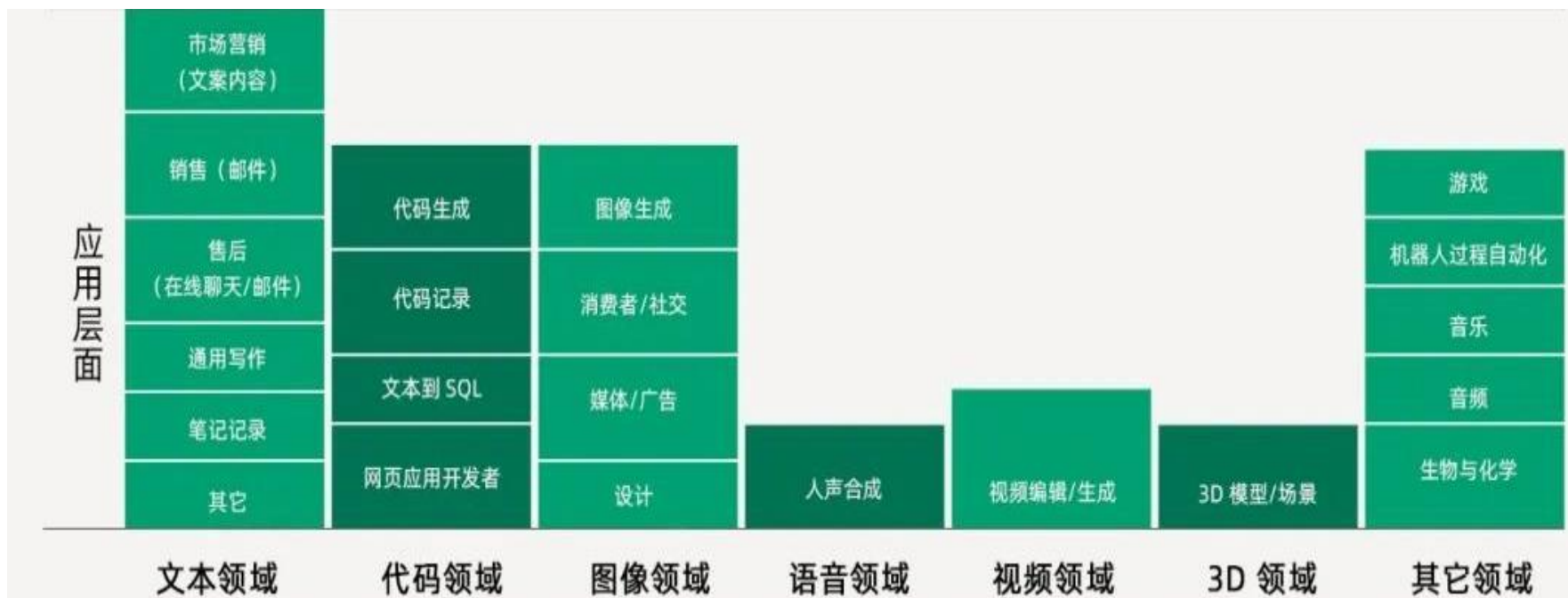
Ada Fastest	Babbage	Curie	Davinci Most powerful
\$0.0004 / 1K tokens	\$0.0005 / 1K tokens	\$0.0020 / 1K tokens	\$0.0200 / 1K tokens

AI行业商业模式迭代历程



场景：AIGC激活C端场景，有望迎来百花齐放

- AIGC应用处于发展初期，技术突破及认知普及有望助推2023年AI应用场景景气上行：
 - ✓ AIGC已经具备文字、图片甚至视频内容的生成能力，在生成创意、内容创作等方面对创作效率的提升非常显著，可以应用于游戏、艺术创意、广告营销等场景；
 - ✓ 在专业领域，AIGC可以用于构建三维模型、工业设计等，能够替代人工承担繁琐的工作，降低人工成本；
 - ✓ 目前AI绘画多按照图片数付费，使用成本相对较低，将会撬动一部分C端用户，进而扩大用户覆盖面。
- 据不完全统计，海外可能已经至少有200多家创业公司专注于生成式AI的方向，同时在国内已经有小几十家新创公司投身于这一热潮当中。



场景：海外巨头底座扎实，新公司产品力突出

- 海外AIGC领域的玩家包括Google、OpenAI等头部的AI实验室，也包括新成立的独立实验室如Jasper、Midjourney、Stability AI；
- 巨头具有更扎实的技术底座，而新公司在产品化方面表现更加突出，比如MidJournery单张图片生成仅需60s，且采用社区模式，用户交互体验更顺畅。

	产品名称	主要功能	收费方式	隶属公司	公司介绍
	Jasper. AI	利用 OpenAI的技术为博客文章、社交媒体帖子及网页等平台生成文字内容；	订阅制付费：2W字，29美元/月；5w字，59美元/元	Jasper	2021年1月由 Dave Rogenmoser、Chris Hull 和 John Phillip Morgan 在美国得克萨斯州奥斯汀创立。目前有102名员工。获得了1.25亿美元的融资，估值达到15 亿美元
	Disco Diffusion	发布于 Google Colab 平台的一款利用人工智能深度学习进行数字艺术创作的工具	现阶段免费，开源模型	Google	业务范围涵盖互联网广告、互联网搜索、云计算等领域的跨国科技大公司，Colab是Google Research 团队开发的一款产品。在 Colab中，任何人都可以通过浏览器编写和执行任意Python代码。它尤其适合机器学习、数据分析和教育目的。从技术上来说，Colab是一种托管式Jupyter笔记本服务。
Midjourney	Midjourney	文生图应用，社区模式将应用加载在聊天服务器上	免费生成25张图后，10美元/月生成200张图，或30美元/月无限生成	Midjourney	是独立研究实验室，自筹资金运作，目前有11名全职员工，专注于AI、基础设施等。目前公司已实现盈利
	DALL · E2	文生图应用	免费生成25张图后，每月免费60张图，15美元生成460张	OpenAI	OpenAI成立于2015年底的AI实验室，总部位于美国旧金山，创始人 为马斯克以及Sam Altman，其创立动机是出于对强人工智能潜在风险的担忧。2019年接受微软10亿美元投资，双方合作Azure 云端平台服务开发人工智能技术。
	Stable Diffusion	文生图模型/应用	免费生成200张图后，2英镑/200张图。完全开源模式	Stability AI	由Emad Mostaque创立于2020年，致力于构建开源AI项目，包括文生图模型、语言、音频、生物学方面的模型开发。2022. 10月获得1.01亿美元的种子轮投资，投后估值10亿美金。

资料来源：第一财经，各公司官方网站

场景：国内仍处于跟随阶段

- 百度推出文心·一格技术较为领先，特别是针对中文语义理解能力更强，与海外模型及应用相比在中国风作品上表现更突出；
- 整体来看，当前时点国内仍处于跟随海外的阶段，主要是在已有的应用场景中借鉴AIGC的技术以提高业务效率，如盗梦师、即时AI等，部分公司如TIAMAT也于近期获得数百万美元投资。

	产品名称	主要功能	收费方式	隶属公司	公司介绍
	文心一格	AI作画工具	免费生成100张图后，9.9元50张图；15.9元100张图；49.9元400张图	百度	百度基于知识增强的文心跨模态大模型语义理解技术为代表，2022年推出了基于文心大模型的 AI 艺术和创意辅助平台——文心·一格
	异次元的我	用户上传照片一键生成专属异次元形象	现阶段免费	腾讯	以“绝悟”为代表，腾讯AI Lab基于自己的多模态学习及生成能力在游戏领域进行了全流程的布局。
	TIAMAT	生成式AI平台	公测阶段，可免费生成140张图，之后充值方式尚未明确	上海退格数字科技有限公司	成立于 2021 年，是国内第一批入局 AI 生成领域的技术团队，为个人用户和企业客户提供创意生产、广告、数字藏品等场景形式下的服务。2022年10月获得数百万美元投资。
	6pen art	文生图产品	最多免费生成100张图后，5元10张图；30元100张图，100元400张图，500元2500张图	北京毛线球科技有限公司	2022年4月成立，致力于打造为一个有趣且有价值的AI生成艺术平台。
	盗梦师	AI绘画小程序	免费生成5张图后，5.5元25张图，24.9元125张图；或会员制：99元/月 660张图，299元/月2160张图	西湖心辰（杭州）科技有限公司	由蓝振忠团队创办的人工智能辅助精神健康服务的科技公司。致力研发一款实现人工智能软件与心理障碍人群高效互动交流、长期陪伴，并能够有效缓解心理焦虑、抑郁等症状的专业心理咨询机器人。
	即时AI	支持图像联想和中文语义描述的人工智能设计工具	现阶段免费	北京雪云锐创科技有限公司	成立于2016年9月，先后上线「即时原型」、「即时设计」平台，已累计服务数百万用户。已完成4轮数亿人民币融资，目前团队已接近百人规模。

资料来源：第一财经，各公司官方网站

＜PART 04＞

场景的本质：时空与节奏

八大场景模型



场景：2024-2025年的“八面来风”

元宇宙是重构时空，场景是如何切割新时空：

场景与内容、应用、模式，是新计算平台上，未来将跑出来的各类创新。其中：

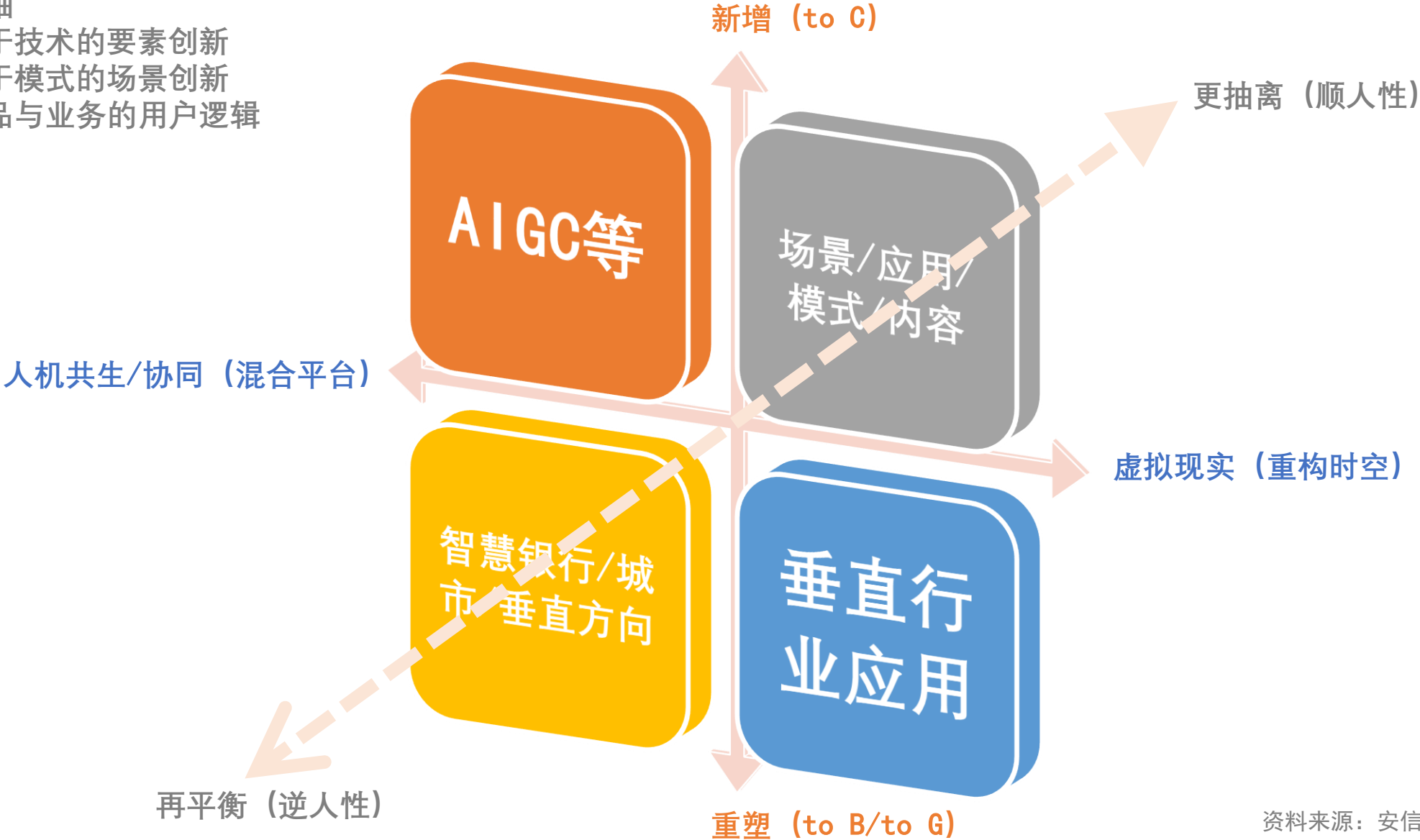
1. 内容最具象；
2. 其次是各类应用；
3. 模式相对抽象；
4. 场景则是宏观层面的切割方法，理论上场景能覆盖所有的内容、应用、模式。

此外：

1. 内容是相对于文字（在线阅读）、声音（音频）、图片（动漫）、视频（长视频/短视频/直播）而言的下一代新内容，目前看会兼具沉浸式、互动式等特征；
2. 应用指服务于用户的各类独立产品，如PC时代的新浪、网易，智能手机时代的各类APP；
3. 模式则侧重于运营逻辑，即独特的经营、盈利模式等；
4. 场景则是能抢占用户时长、可支配收入的各类新时空。

场景：2024-2025年的“八面来风”

三个坐标轴
横向：基于技术的要素创新
纵向：基于模式的场景创新
斜向：产品与业务的用户逻辑



场景：2024-2025年的“八面来风”

场景所“置身”的三维度：

对场景这个“面”的客观、真实理解，须沿着“技术+模式→产品或服务”这条线来分析；

技术维度的要素创新，包括虚拟现实、人机共生/协同两个方向，其中虚拟现实本质上重构时空，人机共生/协同本质上是混合平台；

模式维度的场景创新，可以简单分为to C或to B/G两个方向，to C与to B/G在产品与服务落地时均能发挥新增与重塑的作用，不同产品与服务在新增与重塑的作用上的权重不同；

基于技术与模式的任何产品与服务，底层逻辑均是作用于用户的“人性”，“顺人性”与“逆人性”是此消彼长的两大方向。顺人性更多的是情感需求，逆人性更多是利益需求；

任何供给方的供给，均能落入八大场景模型之一：

1. 供给方根据自身技术禀赋选择一种业务模式再设计具体的情节，交付于用户后，用户在其中投放自身时长与可支配收入；
2. PC与智能手机时代的场景，主要是基于用户想看世界、了解世界，求知欲与对看世界的渴望异同迸发。元宇宙时代的场景，则侧重于用户希望自己被看到，表达自己的渴望；
3. 场景一定会极具区域化特色的，中国在场景这一层面极具竞争力；
4. 趋势上，时间征服空间。

技术要素与顺逆人性界定出八大场景模型



虚拟现实



人机共生/协同



to C



to B/G

技术



模式

新产品
新服务

顺人性

- 用户时长
- 用户可支配收入

逆人性

- 用户时长
- 用户可支配收入



- 虚拟现实 + to C → 顺人性：各类娱乐类、社交类应用、内容、产品等
- 虚拟现实 + to B/G → 顺人性：各类办公类、业务类应用、产品、服务等
- 人机共生/协同 + to C → 顺人性：如家庭端的服务型机器人
- 人机共生/协同 + to B/G → 顺人性：如应用于医疗领域的脑机接口
- 虚拟现实 + to C → 逆人性：如学生的网课
- 虚拟现实 + to B/G → 逆人性：如考勤/规范类的应用、产品、服务
- 人机共生/协同 + to C → 逆人性：如机器人管理员类的应用、产品
- 人机共生/协同 + to B/G → 逆人性：如管理类的应用、产品、服务

北京大学汇丰商学院商业模式研究中心 安信证券元宇宙研究院 元宇宙三十人论坛

魏炜、焦娟、王晶晶、冯静静、王利慧、易欢欢

谢绝未经许可的转载！若有转载、交流等需求请联系：xyjzy2021@163.com

希望我们的研究能为您带来些许的信息增量与认知启发

风险提示

- 硬件层面的进展低于预期
- 应用领域的“劣币驱逐良币”
- 诸多进展带来的伦理争议
- AI失控