

虚拟（增强）现实研究报告

（2025 年）

中国信息通信研究院

2026年1月

版权声明

本报告版权属于中国信息通信研究院，并受法律保护。转载、摘编或利用其它方式使用本报告文字或者观点的，应注明“来源：中国信息通信研究院”。违反上述声明者，本院将追究其相关法律责任。



前 言

在总体态势上，XR 产业已从初生期（XR1.0）、试错期（XR2.0）迈入发展“快车道”（XR3.0）。2024 年起，“AI+XR”双技术引擎成为核心特征，推动终端形态创新与融合应用落地。国内外政府积极布局虚拟现实产业政策，美韩持续推进沉浸式技术战略，我国以五年行动计划等产业政策引导融合应用与生态建设。

在技术趋势上，关键技术正快速演进以支撑 XR 体验升级。微显示方面，硅基 OLED 与 Micro-LED 分别成为 XR 终端的两条主要技术路线，其中面向高阶 AI+AR 眼镜的 Micro-LED 有望在未来几年迎来产能释放与规模应用的拐点。光学方面，MR 头显由菲涅尔转向 Pancake，AR 眼镜光波导技术路径未收敛，阵列与衍射方案竞逐。空间计算方面，XR 沿袭以手机为代表移动端计算架构，但这种范式转变对终端计算能力提出了诸多新要求，推动了多模态感知交互、视听多媒体、空间环境理解与构建、图形渲染多维能力的专项强化。原生音视频方面，3D 沉浸影音成为继超高清视频后重点发展方向，空间视频与体积视频分别成为两大代表性技术路径。内容生产方面，AIGC 赋能 XR 实景重建与虚拟资产生成从“能用”迈向“好用”。空间智能方面，XR 智能助手由语义因果迈向空间因果。

在产业趋势上，XR 终端整机演进呈现“加法”和“减法”两条路线。前者通过堆叠硬件性能来提升沉浸体验，后者聚焦形态轻量化与智能增强，其中以轻量化智能眼镜为代表的“减法”路线已成

为行业热点。“百镜大战”的竞争也正从硬件堆料转向“轻形态+AI能力+XR交互”的系统平衡。XR新器件驱动终端企业加速垂直整合，半导体、显示与光学产业间的协同创新正在从接口对接迈向共研共投的一体化范式。XR OS对硬件与应用的控制进一步前移，将空间计算等技术能力提升为系统能力，XR原生操作系统生态竞争日趋激烈。XR沉浸视听内容生态闭环初步成型，在内容采集端，XR沉浸影像采集进入专业化与大众化双轨演进。在制作与呈现端，XR沉浸影像正在收敛为“标准化编解码+完善工具链”的工业化流程。AR地图平台化发展架构初步成型，国际厂商的部署呈现“大平台+多应用”的发展格局。在工业、应急、教育、商贸等领域，应用从场景适配过渡到痛点破解，线下虚拟现实大空间文娱实践也从试点迈向规模复制。

在问题挑战与发展建议上，内容稀缺、价格偏高与形态偏重、显示与智能受限及眩晕等用户体验痛点影响普及。在商业模式和生态构建上，研发成本高、规模化与商业化路径不清、生态割裂，以及隐私安全与信息可信等治理难题待解。针对上述情况，本报告建议：一是聚焦技术攻关的靶点识别，应对技术创新的不确定性；二是聚焦产业链条上下游协同，构建新兴产业形成的内生机制；三是聚焦应用创新的多元探索，以规模落地破解商业闭环瓶颈；四是聚焦监管治理的预研预判，促进发展环境的健康有序。

目 录

一、总体态势.....	1
(一) 国内外政府积极布局虚拟现实产业政策.....	1
(二) 产业发展步入 XR 3.0 生态“快车道”.....	2
(三) 人工智能催生 XR 3.0 阶段终端新形态.....	4
(四) XR 领域迎来第三轮投融资热潮.....	5
二、关键技术趋势.....	7
(一) 微显示：硅基 OLED 与 Micro-LED 成为两类 XR 终端发展重点.....	7
(三) 空间计算：多维能力融合塑型 XR 新型计算范式.....	10
(四) 原生音视频：3D 沉浸影音成为继超高清视频后重点发展方向.....	13
(五) 生成式 AI：AIGC 赋能 XR 虚实内容生产.....	15
(六) 空间智能：XR 智能助手由语义因果迈向空间因果.....	17
三、产业发展趋势.....	18
(一) XR 终端演进呈现“加法”与“减法”双路线.....	18
(二) XR 新器件驱动终端企业加速垂直整合.....	21
(三) XR 原生操作系统生态竞争日趋激烈.....	22
(四) XR 沉浸视听内容生态闭环初步成型.....	25
(五) AR 地图平台化发展架构初步成型.....	26
(六) XR 大空间从试点走向规模化落地.....	29
(七) XR 赋能行业从“场景适配”迈向“痛点破解”.....	30
(八) XR 助推人工智能训练新范式.....	32
四、行业面临的问题.....	34
(一) 需求侧用户体验问题.....	34
(二) 供给侧产业生态问题.....	35
五、对行业发展的对策建议.....	37
(一) 聚焦技术攻关的靶点识别，应对技术创新的不确定性.....	37
(二) 聚焦产业链条上下游协同，构建新兴产业形成的内生机制.....	38
(三) 聚焦应用创新的多元探索，以 XR 规模落地破解商业闭环瓶颈.....	39
(四) 聚焦监管治理的预研预判，促进发展环境的健康有序.....	39

图目录

图 1 虚拟（增强）现实产业发展阶段.....	4
图 2 XR 终端双轨并行产品路线.....	20



一、总体态势

（一）国内外政府积极布局虚拟现实产业政策

国内外政府积极布局虚拟现实产业政策，但政策重心各有侧重。

虚拟（增强）现实（含混合现实、拓展现实，简称 XR）是新一代信息技术的重要前沿方向，是数字经济的重大前瞻领域，将深刻改变人类的生产生活方式，各国政府重视该领域的发展布局。在美国方面，2022 年 7 月美国国会通过了《2022 年芯片与科学法案》，将“先进的通信和沉浸式技术”列入十项重大技术领域之一。政府主要通过鼓励私营企业的研发投资来推动虚拟现实领域的发展。2024 年 5 月，美国白宫科技政策办公室（OSTP）发布《2024 关键技术和新兴技术（CETs）清单》，清单概述了可能为美国创新开辟新道路并加强美国国家安全的科技，其中，虚拟/增强现实作为重点领域之一。2025 年 3 月，美参众议员向国会提交《2025 年美国沉浸式技术领导力法案》，旨在授权商务部牵头成立了包含白宫科学技术政策办公室、国防部、能源部、教育部、劳工部、交通部、农业部等十余政府部门、企业、智库、学界在内的 XR 战略推进组，确保美国在该领域的技术竞争、国家安全、商业应用、社会发展方面的国际领导地位。在韩国方面，2020 年韩国科学与信息通信技术部（MSIT）等六部门发布《沉浸式经济发展战略》，旨在将韩国打造为全球五大虚拟现实（XR）领先国家，推动泛 XR 与数字经济发展的深度融合。2024 年 2 月国会表决通过了《虚拟融合产业振兴法》，同年 8 月正

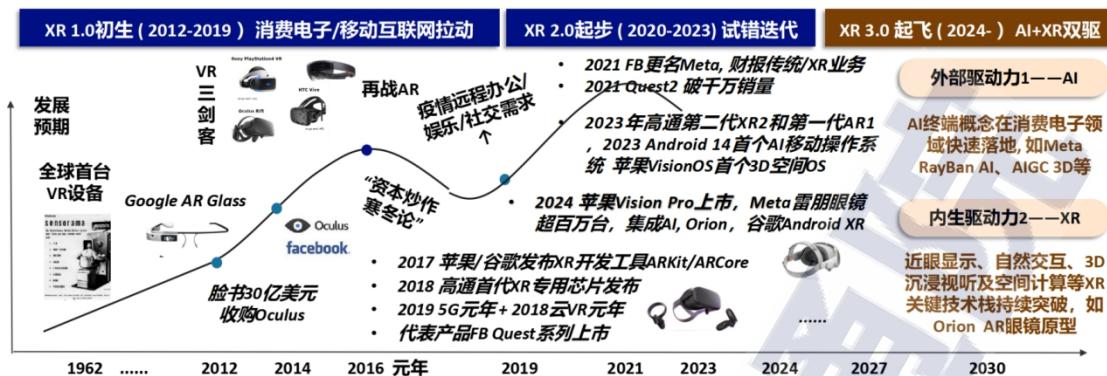
式生效实施，并由 MSIT 每三年制定一次产业振兴计划。该法律为 XR 与元宇宙产业提供了专门的法律框架和扶持措施，要求政府制定产业基本规划、建立产业统计与监管沙盒、培养专业人才等。法律及其施行令明确了政府可在信息提供、咨询、创业空间、市场渠道等方面对 XR 企业给予行政和财务支持，重点扶持中小企业发展。

我国重视虚拟现实产业发展。为深入贯彻《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》相关部署，提升我国虚拟现实产业核心技术创新能力，激发产业服务体系创新活力，加快虚拟现实与行业应用融合发展，构建完善虚拟现实产业创新发展生态，2022 年 12 月工信部、教育部、文旅部、国家广电总局与国家体育总局五部门联合发布《虚拟现实与行业应用融合发展行动计划（2022-2026 年）》（简称“《行动计划》”）。《行动计划》分别从推进关键技术融合创新、提升全产业链条供给能力、加速多行业多场景应用落地、加强产业公共服务平台建设、构建融合应用标准体系等方面提出了五大重点任务，设置了关键技术融合创新、全产业链条供给提升、多场景应用融合推广三大重点工程，明确了主要突破方向和重大事项。

（二）产业发展步入 XR 3.0 生态“快车道”

虚拟（增强）现实产业的发展可以大致划分为三个阶段，每一阶段都有不同的特征和里程碑。XR 1.0 阶段（约 2012–2019 年）是产业初生期，借助消费电子和移动互联网的东风，VR 设备从无到有

开始兴起。这一时期的 XR 产业主要依赖手机生态，例如早期的手机盒子式 VR，但整体体验存在明显短板。用户戴上头显常感眩晕，内容匮乏导致“技术兴奋、商业迷茫”，产业热度虽高但用户留存难。XR 2.0 阶段（2020–2023 年）是试错迭代期，行业在技术和应用上不断摸索改进。这一阶段出现了两个具有标志意义的事件：其一是消费级 VR 产品取得了突破性进展，Meta Quest 2 销量突破了千万大关，成为 XR 史上首个单品累计销量超过千万的设备，标志着消费市场初步被打开；其二是苹果公司的入局。2023 年苹果发布了高端 MR 设备 Vision Pro，宣示主流科技巨头对 XR 领域的重视，加速了产业成熟。总体而言，XR 2.0 时期行业在硬件迭代和应用探索上取得进展。进入 XR 3.0 阶段（2024 年起），XR 产业迎来了“AI+XR”双引擎驱动的新局面。人工智能技术与 XR 深度融合，推动 XR 产业发展进入“快车道”，AI+AR 眼镜、3D 空间照片、VR 大空间、碳化硅光波导等新兴细分领域频现。Meta 公司首次推出的 Orion 轻量化 AI+AR 眼镜原型，标志着 XR 3.0 阶段的技术特征正在加速落地。总体态势表明，经历了前期的孕育和探索之后，XR 技术成熟度和市场接受度显著提升。



来源：中国信息通信研究院

图 1 虚拟（增强）现实产业发展阶段

（三）人工智能催生 XR 3.0 阶段终端新形态

XR3.0 阶段的核心特征是空间计算与人工智能的深度介入，由此催生了混合现实(MR)头显和轻量化智能眼镜两类新型终端形态，并显著提升了用户体验。

MR 头显的兴起拓展了 VR 设备的功能边界。传统 VR 头显主要封闭式地为用户提供虚拟环境，而 MR 头显在此基础上加入了对现实环境的感知与重建能力，使设备能够实现虚实融合的体验。例如，Meta 公司于 2023 年发布的 Quest 3 和苹果公司推出的 Vision Pro，均属这一类高性能 MR 设备。它们配备了高分辨率摄像头和强大的计算芯片，可以将现实世界以数字形式实时呈现在用户眼前，再叠加虚拟影像，实现“透视式”的增强现实效果。MR 头显的出现表明虚拟现实终端正从纯粹的虚拟沉浸，向兼顾现实环境的方向演进，为用户带来更丰富的交互形式。

AI 赋能的智能眼镜迅速崛起，成为 XR 领域另一大创新亮点。

这类眼镜通常外形接近普通眼镜或太阳镜，但内置了麦克风、摄像头、语音助手等 AI 服务相关模块，主要侧重提供智能感知和信息服务，而不强调沉浸式可视化呈现。以 Meta 与 Ray-Ban 合作推出的智能眼镜为例，它具有拍照录像、语音交互等功能，能够通过内置的语音 AI 助手回答用户提问、提供导航或信息查询等服务。由于不包含高阶的近眼显示单元，这类眼镜重量更轻，续航时间更长。在市场表现上，无屏 AI 眼镜上市以来迅速得到消费者青睐，这反映出大众对于“随身智能助理”类穿戴设备有着显著需求。AI 智能眼镜的成功也在倒逼 AR 眼镜加快融合 AI，人们希望未来的增强现实眼镜不仅能显示数字内容，更能听懂看懂周围世界、成为个人日常生活助手。AI+AR 融合眼镜正是在这样的背景下应运而生。它将 XR 终端的近眼显示能力与无屏 AI 眼镜的智能交互能力合二为一，被认为是 XR 3.0 阶段终端形态的重点方向。从 2025 年下半年的业界实践来看，从国内外代表性 AI+AR 眼镜的产品发售与市场反馈较好。这些尝试预示着未来消费者可能戴上一副轻便的眼镜，就能同时看到增强的信息提示，并享受 AI 实时讲解和辅助决策。在 AI 的赋能下，XR 硬件形态创新活跃，设备智能化和轻量化趋势更加明显，为产业注入了新的活力。

（四）XR 领域迎来第三轮投融资热潮

2025 年 XR 领域启动了史上第三波投融资热潮。XR 产业加速发展的整体趋势也反映在资本市场的热情上。回顾前两轮 XR 投资

热潮，第一次出现在 2015-2016 年左右，当时 VR 元年概念兴起，创业投资活跃；第二次高潮则是在 2021 年前后，受“元宇宙”概念带动，大量资金涌入 XR 相关行业。而 2025 年的这第三波热潮，则是由“AI+XR”融合创新的前景所驱动。资本市场对该领域的关注度和投入规模均达到了新的高度。数据显示，截至 2025 年全球 XR 领域的融资总额已经超过了上一轮高峰（2021 年）的水平。不仅总体规模创新纪录，单笔投资的金额也显著提高，头部项目更受青睐，资本投向呈现出更集中明确的特点。

值得注意的是，本轮投资热潮中不同地区和领域的侧重点有所差异。从全球范围看，XR 产业链上下游——无论是终端硬件、底层技术，还是内容应用、平台服务——均受到资本青睐，投资分布相对均衡。国际科技巨头如 Meta、苹果等在硬件和内容两端都投入巨资布局 XR 生态，这对全球创业投资也起到了示范拉动作用。但在中国国内，XR 投融资呈现出“硬件尤热，内容平淡”的态势。目前国内资本更偏好于 XR 终端相关领域，尤其是空间计算、近眼显示等核心器件赛道。据统计，2025 年上半年国内 XR 融资项目中，光波导、微显示屏、终端整机等硬件领域重新回到了投资榜首。相比之下，面向 XR 的内容制作和应用服务类创业在国内获得的融资显著偏少，这与国外投融资结构形成对比。

二、关键技术趋势

（一）微显示：硅基 OLED 与 Micro-LED 成为两类 XR 终端发展重点

微显示技术是 XR 终端的关键技术之一，当前呈现出两条不同的发展路径。总的来看，VR/MR 与 AI+AR 两大 XR 设备对微显示技术的需求有所差异，前者深耕硅基 OLED 并不断改进其亮度、分辨率等技术指标，后者瞄准 Micro-LED 的长远优势与成本下探空间。对于 VR/MR 头显而言，硅基 OLED（又称 Micro-OLED）微显示屏能够在小尺寸上实现极高的像素密度和良好的图像质量，成为苹果 Vision Pro 等高阶 VR/MR 终端的主要选择。预计未来，硅基 OLED 微显示的技术发展方向可以概括为“更亮、更清、更稳、更经济”，即亮度提升、清晰度提高（像素密度和填充率更高）、稳定性可靠性增强，以及降低成本以利于规模量产。可以说，VR/MR 微显示技术正在朝着小尺寸下极致显示性能的方向不断突破。从技术选型看，VR/MR 头显微显示技术路线呈现硅基 OLED、Fast-LCD 高低搭配的市场格局。得益于硅基 OLED 在像素密度等显示效果上的优势，苹果等公司在高阶 MR 产品中普遍采用。传统 Fast-LCD 凭借成熟产能和成本优势，预计将在未来几年继续占据 MR 设备主要出货份额。

AI+AR 眼镜所追求的微显示技术路线则有所不同，业界聚焦 Micro-LED 发展。Micro-LED 被视为新一代近眼显示的理想选择，它在亮度、效率、色彩表现等诸多图像性能方面表现出众。对于可

日常佩戴的 AI+AR 眼镜而言，小型化的 Micro-LED 微显示模组有利于嵌入眼镜而不影响视野或增加过多重量。然而，Micro-LED 目前的量产难度显著高于硅基 OLED，无论是在外延生长还是在巨量转移方面，存在诸多技术挑战，其中一个关键难点在于全彩化。目前的解决方案多是使用三片单色 Micro-LED 通过光学合成来得到全彩图像，但这不仅使光学结构相对复杂，而且引入额外的驱动器件和硬件功耗。单片全彩成为当前业界攻关方向，相关代表性技术路径一是“三色 LED 垂直堆叠”，即在同一芯片上垂直集成红绿蓝发光层；二是“量子点色彩转换”，即用量子点材料将蓝光或紫外光转换成红绿光。此外，在 Micro-LED 制造中，实现高良率的巨量转移也是制约成本的关键环节，需要将成千上万颗微米级 LED 高效地植入驱动电路。受上述因素影响，Micro-LED 微显示的总体成本较高，各大显示厂商和科技巨头仍在积极投入研发。一旦单片全彩 Micro-LED 取得突破，AI+AR 眼镜的显示效果和外形设计都将迎来新一轮升级。从技术选型看，AI+AR 眼镜微显示技术路线呈现 Micro-LED、硅基 OLED 与 LCoS 高中低搭配的市场格局。硅基 OLED 遭遇双向挤压，面向高阶 AI+AR 眼镜的微显示技术 Micro-LED 将在未来几年进入产能释放和规模应用的拐点期。

（二）微纳光学：AR 眼镜光波导技术路径尚未收敛

VR/MR 头显已从传统的菲涅尔透镜进化到折叠光路的 Pancake 方案。过去 XR1.0 时期常采用菲涅尔透镜作为放大光学，其结构简

单但镜片厚重且视场受限，VR 用户佩戴时常能看到同心圆纹和光晕。进入 XR2.0/3.0 时代，主流 VR 厂商纷纷采用 Pancake 折叠光路，通过多层透镜和偏振光学设计将光线多次折返，从而大幅缩短光学模組的厚度，重量显著降低，同时像差等光学性能也得到改善，这一转变极大提升了头显的佩戴舒适度和便携性。

AI+AR 眼镜轻巧时尚的外形尤为重要，技术演进路线正从 BirdBath 等传统光学转向更薄更轻的光波导技术。当前业内主要探索两种波导实现路径：一是阵列光波导，代表厂商有 Lumus 等；二是衍射光波导，包括表面浮雕衍射波导和体全息波导两类，代表厂商如 WaveOptics、Digilens 等。上述光波导技术路线目前尚未完全收敛，各有优缺点。例如，阵列波导通过多层半反射镜实现影像传导，具有亮度高、视场适中等优点，但制造工艺复杂、成本高且存在条纹干扰的问题；衍射波导利用微纳光栅耦合出瞳，可以做到更大视场，结构也相对简单，但目前衍射效率偏低，常出现色散（彩虹纹）和杂散光，需要优化改进。

AR 光学技术发展方向可以概括为“看得亮、看得广、戴得舒适、看得舒适”。“看得亮”是指提高光效和出瞳亮度，使增强显示即便在户外强光下仍清晰可见；“看得广”意味着扩大水平和垂直视场角（FOV），理想状态下希望接近人眼自然视场；“戴得舒适”涉及减轻光学模块的重量厚度，使 AR 眼镜像普通眼镜一样轻便；“看得舒适”则指提高成像质量，减少鬼影、散斑、彩虹纹等视觉

伪影，以降低长时间佩戴观看的眼部疲劳。这些要求相互之间往往存在权衡取舍，例如增大视场可能会降低亮度、缩减体积可能引入像差，因此需要在设计上综合优化。此外，半绝缘型碳化硅等新材料和纳米压印、刻蚀等新工艺的应用正在逐步改善波导产品性能。随着光波导技术的逐步成熟和规模量产，AR 眼镜有望同时实现更明亮的成像、更宽广的视野以及更舒适的佩戴体验，为用户带来自然逼真的视觉效果。

（三）空间计算：多维能力融合塑型 XR 新型计算范式

XR 有望成为继 PC（个人计算）和手机（移动计算）之后的新兴通用计算平台。苹果公司在其 Vision Pro 中提出“空间计算”发展定位，强调消费电子在三维空间中进行计算交互的新范式。不同于传统 2D 终端局限于平面屏幕，空间计算设备通过将数字内容无缝叠加于真实世界，提供契合大众身处三维环境的交互体验。空间计算沿袭以手机为代表移动端计算架构，但这种范式转变对 XR 终端计算能力提出了诸多新要求，推动了多模态感知交互、视听多媒体、空间环境理解与构建、图形渲染四维能力的专项强化。

XR 多模态感知交互将更具自然感与智能性。当前 XR 终端集成了环境摄像头、深度传感器、惯性单元以及麦克风阵列等多模态传感器，实现对环境和用户的同步感知融合。同时，XR 设备强调以人为中心的自然交互，操作系统在底层整合手势、视线、语音、肌电、触觉反馈等多模态输入，为应用提供统一的接口调用诸多交互方式。

例如，Apple Vision Pro 配备了 12 个摄像头、红外传感器和 6 个麦克风，通过专用协处理器实时处理多路感知数据，在十余毫秒内完成从捕获到输出的闭环。此外，随着端侧 AI 能力日益增强，XR 设备可在本地运行生成式 AI 模型，显著提升了系统对用户意图的理解效率和响应速度，让 XR 人机界面更加贴近用户本能习惯。

XR 视听多媒体将更具空间感与个性化。在音频方面，XR 终端普遍采用区别于传统平面声效的空间音频技术，即通过头相关传输函数（HRTF）模拟人耳听觉线索，实现声音在三维空间中的定位感。目前高端 XR 厂商开始关注个性化 HRTF 以提升音频效果。首次佩戴时为用户建立听觉档案，结合头显内置的扬声器和动态头部追踪，实现真实舒适、千人千耳的沉浸声场。**在视频方面**，XR 设备正向着超高清和自由度更高的沉浸视频演进。一方面，全景视频的分辨率不断提升，从早期的 4K、6K 迈向 8K 乃至更高。另一方面，更高自由度的观影体验让用户转头观看、移动位置或是自由观察景深变化成为可能。为此，业界开始采用多视点视频编码方案，例如 HEVC 标准的多视点扩展 MV-HEVC 和 3D-HEVC 等，来高效压缩存储 XR 双目立体视频。

XR 空间环境理解将实现跨场景连续稳定的虚实对齐。让虚实融合成为可能的基础，是 XR 设备对空间环境的实时理解和重建能力。这首先体现在定位追踪上，由内向外的 6DoF 追踪已成为 XR 终端标配，业界在优化 SLAM 算法以降低算力消耗的同时，借助专用硬件

加速（如视觉处理单元、AI 芯片）确保在毫秒级完成环境更新，减少高速运动下的定位漂移。进一步的环境三维重建可生成房间的网格模型或点云，用于虚拟内容的物理落地和遮挡。近年来，随着高保真的全彩视频透视（VST）技术的启用，终端能够将真实环境摄像头画面叠加虚拟图形呈现，实现场景实时重建与渲染，VST 成为 VR 终端 MR 化的关键技术拼图。在持续定位方面，高端 XR 设备开始支持跨场景的空间锚定与记忆，业界积极探索云端共享锚点，实现多人协同的 AR 体验。此外，环境理解将从几何重建走向语义重建，XR 终端将识别房间中的墙面、桌椅等对象，并据此优化虚实交互用户体验。

XR 图形渲染将在“降本增效”的方向持续演进。XR 终端（以当前 MR 头显为主）在图形渲染上面临极高的负载挑战，近眼显示需要呈现高分辨率、广视场且高刷新率的立体画面，且对延迟容忍极低。为在有限功耗下实现流畅渲染，XR 领域发展出一系列软硬件协同的渲染优化技术，其中异步时间扭曲（ATW）和异步空间扭曲（ASW）已广泛用于 VR runtime，其机制是在应用未及时提交新帧时，由系统根据最新的头部姿态对上一帧图像进行畸变校正（ATW），或利用运动矢量预测生成插值帧（ASW），从而让显示端仍以既定高帧率输出画面。动态注视点渲染已逐步成为高端 MR 的标配，诸多系统级渲染优化围绕用户视线展开，旨在“高分辨率、高刷新率、低功耗”三者间取得平衡，当前 XR 芯片针对该技术进行了优化支

持，例如图形架构提供可变着色率（VRS）等能力，以便根据眼动数据动态调整像素着色密度。

（四）原生音视频：3D 沉浸影音成为继超高清视频后重点发展方向

XR 视听内容形态从平面超高清视频向三维沉浸影音升级，内容来源由手机、电视等传统终端的存量迁移向 XR 原生视听拓展。自上世纪 50 年代至今，音视频产业经过模拟信号到数字信号、标清到高清/超高清的两轮重大技术变革，带来了 4K/HDR 以及立体声/环绕声的大众化，当前这一从平面到沉浸的发展有望成为音视频技术的第三次革新，3D 和多自由度交互成为主要技术特征。在这一阶段，视听内容将不再局限于一个矩形屏幕，而是延展到用户周围的三维空间，声音也从传统立体声进化为真正空间音频，能够根据用户位置和朝向实时变化，使听觉匹配视觉的沉浸感。

业界将 XR 沉浸视听原生内容的发展归纳为近中期两大技术路径。一是以空间视频为代表，用户可以通过 VR 头显等设备自由转动视角来环顾拍摄的三维广阔场景。目前已经出现了 8K 甚至更高分辨率的 3DoF 空间视频内容，以及相应的 MV-HEVC 等视频编码标准来支持高效压缩和传输。空间视频技术相对成熟，成为近期虚拟现实原生视听内容的主导技术方向，越来越多的媒体平台和内容制作公司开始提供全景纪录片、体育赛事等 VR 空间视频。二是以体积视频为代表，作为进阶的 3D 沉浸视听内容形式，用户不仅可以环

顾，还可以自由移动位置来观察场景的变化。实现 6DoF 的体积视频需要对场景进行三维重建，可通过多机位摄影测量、深度摄像机阵列或者光场采集等技术获取，然后通过点云、体素或者神经辐射场（NeRF）等形式来渲染，生成的内容本质上是动态的 3D 模型。在实时传输方面，这类数据量庞大，仍处于研究和初步应用阶段。随着计算机视觉和图形渲染技术的进步，体积视频的制作门槛正逐步降低，在文博、教育、医疗培训等领域已出现试点应用。近期的技术进展表明了一些可行路径，例如通过 3D 高斯渲染（3D Gaussian Splatting, 3DGS）等方法，可将多张普通摄像照片快速重建出较为真实的 3D 场景。

3D 沉浸视听技术需要采集摄录、压缩编码、渲染播放等技术栈协同突破。实现 3D 沉浸视听内容，涉及一系列技术协同突破。内容采集方面，需要高分辨率 VR 全景相机、光场相机等新的摄录设备以及 AI 辅助的内容生成（如从普通视频推断景深信息，或者纯算法生成虚拟环境）。内容压缩与传输方面，大文件量要求制定高效的编码标准，如 MPEG 推出的沉浸式视频（MIV）标准，专门针对多视角体积视频进行压缩编码。播放渲染方面，XR 终端需要足够的计算能力来实时解码并渲染出用户视角对应的画面，这也推动了之前提到的芯片和算法改进。此外，沉浸内容往往还伴随交互性——用户可以与虚拟环境或角色交互，这又引入了对交互设计和实时反馈的技术要求。

（五）生成式 AI：AIGC 赋能 XR 虚实内容生产

AIGC 赋能 XR 实景重建与虚拟资产生成从“能用”迈向“好用”。虚拟（增强）现实内容的丰富与否，是决定其能否大规模普及的关键。长期以来，XR 虚实内容生产存在两类难点：一是 3D 真实场景的数字化成本高、周期长。二是 3D 虚拟内容的创作门槛高、专业性强，例如制作高质量的 VR 三维模型和环境，需要美术和程序团队投入大量时间。这些问题导致 XR 应用面临“巧妇难为无米之炊”的发展窘境，随着生成式人工智能（AIGC）的兴起，上述难题开始缓解，XR 内容生产正在出现范式转变，从过去勉强可以用（能用），跃升到更加高效易用且高质量（好用）的新阶段。

在 AR 内容生产方面，AIGC 大幅简化了实景三维重建流程，提高效率和精度。传统获取 AR 所需真实场景 3D 数据往往依赖专业激光扫描或人工建模，而生成式 AI 能够自动从二维图像推理出三维场景。例如，仅用手机拍摄多张照片，AI 算法便可生成逼真的该场景 3D 模型。神经辐射场（NeRF）和三维高斯散点等技术代表了这一方向的进展，它们利用深度学习从多视角图像重建出场景的体积模型。目前一些 AR 云服务公司已将 NeRF 用于构建城市级 AR 视觉定位所需的实景地图，将过去需要测绘团队数周完成的建模工作压缩到数小时内即可自动完成，且达到可用精度。这一能力对于城市 AR 导航、室内定位、工业数字孪生监控等大规模 AR 应用具有重要意义。

在 VR 内容方面，生成式 AI 同样成为虚拟资产创作的有力助手。

虚拟世界中需要大量 3D 模型、材质和场景设计，传统制作流程耗时耗力且要求专业技能。通过文生 3D 技术，开发者只需输入概念描述，AI 便能产出相应的三维模型雏形。利用扩散模型，AI 还能为粗略场景自动填充细节和纹理。例如 Meta 公司为自家 XR 社交平台应用 Horizon World 开发的 AssetGen 系列工具探索了此类应用，其 2.0 版本被认为是当前领先的 3D AIGC 系统之一。普通创作者无需精通建模软件也可快速生成所需的 VR 物体和环境，在专业内容制作中，团队也开始让 AI 生成初步场景布局，再由美术细化，这种人机协作提高了制作效率。

AIGC 的引入正在重塑 XR 内容制作流程，并极大丰富 XR 虚拟体验的供给。过去 3D 内容制作往往被拆分为多个阶段（2D 概念设计、3D 建模、导入引擎设定交互等），流程冗长复杂。而生成式 AI 有望将许多步骤压缩甚至合并，实现“一站式”内容生产。例如最新的研究已初步实现从自然语言直接生成可交互的 3D 场景雏形，被视作下一代智能内容引擎的雏形。当然，当前 AIGC 赋能内容生产也存在一些局限，如 AI 生成内容的一致性和可控性尚待提高，以及生成素材的版权归属问题尚无定论。然而总体而言，生成式 AI 正在为 XR 内容生态注入新活力，从社交、游戏到教育、训练，各类虚拟场景都有望因 AI 创作成本的降低而大量涌现。随着这一趋势发展，困扰 XR 产业的内容匮乏瓶颈有望逐步破解，技术与内容将形成齐头并进的良性循环。

（六）空间智能：XR 智能助手由语义因果迈向空间因果

大模型的能力正经历从图文理解到视频理解再到空间环境理解的跃迁，这一趋势成为 XR 领域空间智能技术发展的重要驱动力。早期的视觉 AI 模型擅长识别图像中的语义信息并与文本对应，随后出现的视频理解模型能够分析动态画面中的时序关系，为更深层的场景理解奠定基础。如今，业界进一步将人工智能扩展至三维空间的认知，谷歌、OpenAI、Meta 等公司都在探索融合 3D 环境感知与 LLM 的方法。例如，Google 提出了 3D-LLM 的概念，尝试让语言模型直接产生对三维场景的描述和问答，还有研究者开发了 PointLLM，将点云数据转化为 LLM 可处理的格式。另有项目如 LLaVA-3D 将图像对话模型拓展到三维场景的理解。这些研究证明了一个可行路径：通过增加专门的 3D 感知编码器，把视觉空间信息转换成类似语言的序列，再让 LLM 来处理推理。尽管 LLM 本身并不擅长处理几何信息，且 3D 感知结果往往数据庞大且噪声较多，但一些原型系统已能在部分场景下做到对物体和位置的提问应答。这种能力的拓展使得 XR 不仅是移动互联网的终端入口，更将成为大众视听共享、持续在线、理解环境的智能助手。

大模型认知范式正从纯粹的语义因果走向对空间和物理因果的掌握。AI 不再仅仅停留于语言层面的理解和推理，而是开始具备对物体在物理空间中的关系、变化与因果机制的推断能力。这种空间因果推理能力对于 XR 智能体尤为关键，当理解了真实世界中“为

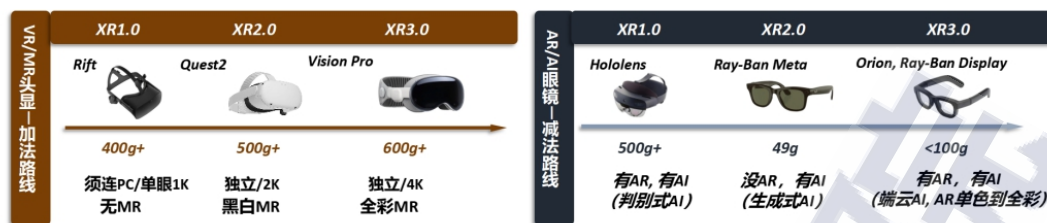
什么会发生”和“将会发生什么”，XR 终端才能对用户的所见所闻做出恰当的响应。以 Meta 公司推出的 SceneScript 模型为例，该技术成果被视为通往未来 AI+AR 眼镜的里程碑探索。它能够端到端地将物理空间布局解析为紧凑而完整的语言描述，被誉为连接物理世界和数字世界的桥梁。SceneScript 通过机器学习直接推断房间的三维结构，并以类似自然语言的形式表达场景元素（如墙壁、门窗、家具的位置及关系），从而为虚拟助手提供对现实环境的可解释记忆。该模型不仅显著压缩了环境表示所需的数据量，使场景记忆成为可能，还赋予大型语言模型理解物理空间的专门“词汇表”，这为个人助手进行空间推理和回答复杂的环境相关问题提供了必要基础。随着空间智能技术的推进，XR 智能体将具备“感知所处环境并长期记忆”的能力，为用户提供更加贴身和上下文相关信息服务，这标志着 XR 正迈向人机融合的新阶段，实现语义智能与空间智能深度融合。

三、产业发展趋势

（一）XR 终端演进呈现“加法”与“减法”双路线

XR 终端整机发展出现了两条演进路线。一是以 VR/MR 头显为代表的“加法路线”，XR 终端的功能配置日趋丰富，但也带来了重量和体积的增加。从早期 2016 年 Oculus Rift 头显（单眼 1K 分辨率）发展到 2020 年 Oculus Quest 2（单眼 2K、黑白 VST、手势识别等），再到 2023 年 Apple Vision Pro（单眼 4K、全彩 VST、眼动追踪、沉

浸视听等），每一代产品都在性能上“加码”。这一路线通过不断叠加新特性来提升沉浸感和交互性，如更高分辨率、更强空间计算、更自然地人机交互，但相应地也面临着重量体积上升和佩戴舒适性挑战。Meta、苹果、谷歌、三星等 ICT 巨头在这一路径上竞相投入，推动头显逐渐向专业化、高性能方向演进。当前，“加法路线”开启“便携形态”的新探索，即形态上由头盔头显向眼镜外观发展，一体式向分体式架构转向。二是以有无 AR 能力的 AI 眼镜为代表的“减法路线”，该路线下 XR 终端追求轻量小型化，以减配牺牲部分功能为代价换取终端形态接近日常穿戴，而后再逐步拓展能力。例如，2019 年 Microsoft HoloLens 2 终端重量逾斤，具备单眼 2K 全彩显示、眼动手势交互、FOV 升级等较为完整功能。与之相比的是 2023 年 Ray-Ban Meta AI 眼镜重量锐减至五十克左右，取消了 AR 显示，仅保留拍照摄像、音乐播放和语音助手等简化功能。随后业界推出的原型产品 Orion（2024）以及 AI+AR 眼镜 Meta Ray-Ban Display（2025）在维持轻便形态的同时重新加入全彩 AR 显示，新增肌电交互、光致变色等功能，并进一步强化 AI 端云协同。“减法”策略使得 XR 终端得以率先达到类似日常眼镜的佩戴形态，继而随着近眼显示、空间计算、自然交互与生成式 AI 等支撑技术的迭代演进而重塑产品定义。



来源：中国信息通信研究院

图 2 XR 终端双轨并行产品路线

以轻量化智能眼镜为代表的减法路径成为行业热点，“百镜大战”的比拼正从硬件堆料转向“轻形态+AI能力+XR交互”的系统平衡。行业普遍沿用“先减后增”的产品路径，最初通过取消近眼显示、压缩本地算力与传感器规模，优先确保佩戴外形与全天候使用，再以端云 AI 能力滚动升级为核心卖点。以业界标杆 Ray-Ban Meta 系为例，其在 2023 年底首发时仅提供语音控制、第一视角拍摄/录像等基础功能，尚未接入 AI。2024 年初引入“语音/照片问答”的 AI 能力；2024 年年底通过引入时序建模以理解动态事件，实现了 AI 眼镜从“看照片”到“看视频”的升级，相继启用了可持续记忆、轻量级智能体服务（如在眼镜端完成拍照并触发跨应用转发等）等智能服务。这一“AI 先行、显示后补”的思路，使得硬件门槛与佩戴压力显著降低，供应链可复用时尚眼镜与成熟声学件。当前，减法路线已进入“显示回填”的新阶段，即在不牺牲形态的前提下回补显示与输入通道，以使 AI 能力可被直观呈现与即时操作。从国内外先导产品实践看，内置 AR 显示与多模态 AI 合流的轻量化眼镜初步获得正向市场验证，该品类已从小众尝鲜走向规模化制造的拐

点，诸如实时字幕、无障碍可达、消息处理与导航等高频场景已显现“即时可用”的价值，这为后续在更大视场、更长续航与更强智能体调度上的代际跃迁提供了清晰抓手。

（二）XR 新器件驱动终端企业加速垂直整合

XR 新器件引发消费电子供应链结构变化。XR 终端引入了诸多区别于 PC、手机等传统消费电子的新型器件，包括微显示屏、新型光学、专用芯片以及感知交互组件等。在显示方面，传统智能手机多采用大尺寸玻璃基 LCD 或 AMOLED 屏幕，而 XR 设备采用的 Micro OLED 等微显示器件改为硅基晶圆作为基板，将驱动电路和像素阵列集成于一体，大幅缩小体积重量，提高像素密度与响应速度。光学方面，VR 设备使用的 Pancake 超短焦镜头体系，也不同于手机简单直射显示的方案，需要复杂光路折叠。AR 眼镜引入了光波导等新型光学器件，追求高透过率与大视场角。此外，XR 终端集成的摄像头阵列、深度传感器等环境感知模块和眼动追踪摄像头，也是传统手机所少有或性能升规的关键器件。一方面，这些 XR 新器件的引入重塑了终端整机的 BOM 清单与成本结构。同时，领先整机代工企业通过攻克 XR 产品装配工艺，实现了代工价值提升。另一方面，新器件还导致供应链版图生变，光波导、微显示、专用芯片等 XR 关键器件目前仅有少数上游厂商掌握，产业链尚不成熟，XR 终端厂商难以完全照搬手机采购模式——从现有成熟供应商处采购标准件。整机企业深度介入关键器件的开发与产能培育，以确保技术

性能和产品供应。

XR 新器件供应策略呈现垂直整合的产业态势。面对上述 XR 关键器件的新要求，近年来，各终端厂商纷纷调整零整协同的供应链策略，转向更高层次的垂直整合，通过收购、战略投资、联合研发自建产线等方式介入上游，以掌控核心器件的技术供给。例如苹果公司相继收购 Micro LED 微显示公司 LuxVue、眼动追踪技术公司 SensoMotoric Instruments、体全息光波导 Akonia Holographics，并专门设计了 XR 协处理器，以提升多模态感知交互的实时处理能力。Meta 收购了欧洲 Micro LED 微显示企业 InfiniLED、眼球追踪技术公司 The Eye Tribe、一体化 AR 光学公司 Luxexcel、肌电交互 CTRL-labs 等。通过这一系列“研、投、并”组合拳，各大 XR 终端企业正通过收购前沿组件公司和深度定制供应链，将所需“屏、芯、光、感”等关键环节纳入掌控体系，打造自身产品所需的核心零件库与差异化竞争力。此外，XR3.0 阶段，显示、光学与半导体间的协同创新正在从“接口对接”迈向“共研共投”的一体化范式。XR 终端开始通过协同设计和联合优化，实现微显示、光学镜片、芯片在架构设计、功耗亮度、光学耦合效率乃至外延工艺和模组形态上的双向约束与联合迭代。

（三）XR 原生操作系统生态竞争日趋激烈

XR 原生操作系统已从概念设计走向工程体系，多模态自然交互、空间计算原生支持、实时并行处理与 AI 深度融合成为核心技术特性。

在多模态交互上，与键鼠触屏等传统交互不同，XR 系统强调以人为中心的多模态交互。操作系统需整合视野内环境感知、手势追踪、眼动追踪、语音识别、肌电传感等多种交互方式，提供统一的系统级接口，让应用能够方便地调用这些自然交互手段；在空间计算原生支持上，操作系统直接管理三维场景、空间锚点、环境理解与空间音频，而非将其作为应用层库的可选项。这使应用在同一世界坐标中共享状态并保持持久化，形成稳定的空间体验基底；在实时并行处理上，XR OS 需要在低时延下并行处理图形渲染等任务，因而强调实时调度和多任务并行能力。例如，XR 系统在渲染管线中引入注视点渲染与时间扭曲/重投影等机制，以在有限算力下保障低延迟与高稳定性。这些能力被内建为系统服务而非单点 SDK，标志 XR OS 已在系统层面确立独立于传统移动 OS 的实时并行处理性能约束模型；在人工智能融合上，VisionOS2、AndroidXR 等新一代 XR 系统普遍内置 AI 系统级服务能力，用于环境理解、内容生成和智能助手等功能，并在端侧基础模型、微调组件、安全隐私等方面深化与 XR 终端的适配。

XR OS 对硬件与应用的控制面前移，空间计算等技术能力“上收”为系统能力，软件栈由“App 堆叠”走向“系统中枢”。从人工智能服务看，安卓将 Gemini Nano 部署为 AICore 系统服务，提供端侧生成式 AI 能力并统一调度专用硬件，以实现低时延、隐私保留与版本一致性。这一“系统级 AI 服务”思路体现了从应用侧集成到

OS 中枢托管的迁移路径。从图形与实时计算看，XR 运行时承担了应用提交帧后的关键环节，包括分层合成、畸变校正、时间扭曲与注视点渲染等，这些过去由应用分散处理的能力被纳入操作系统统一调度与质量控制，直接影响平台级的延迟、稳定性与能效指标。从空间计算看，系统级的场景理解、空间锚共享与空间音频等能力同样上收至运行时与系统服务，开发者通过稳定的系统 API 接入，而平台则以策略与权限管理保障安全与一致的体验边界。这种“能力上收”本质上把关键体验的确定性、性能与安全移交给 OS 统一治理，平台在体验质量与生态门槛上获得更强的主动权。

在 XR 原生操作系统领域，平台化阵营竞争正日益加剧。苹果的 visionOS、谷歌的 Android XR 和 Meta 的 Horizon OS 分别走出了不同的技术战略与生态路径。苹果与 Meta 阵营皆优先深耕 XR 原生操作系统的空间计算体验，而后逐步融入生成式 AI 能力，而谷歌阵营在 Android 14 中嵌入了 AICore 等底层 AI 能力，并携手高通、三星等于 2024 年共同推出 Android XR 平台，继承庞大的 Android 应用生态和开放特性。这一差异体现为苹果、Meta 以空间计算优先，谷歌以 AI 优先。此外，在生态开放度方面，苹果坚持垂直一体的封闭生态，强调端到端体验的可控性与一致性。Meta 凭借内容和社交优势占据市场并开始开放其操作系统授权给第三方 OEM 厂商使用。谷歌则依托 Android 开放生态联合硬件伙伴提供 XR 基础框架。总体上看，各大厂商围绕 XR OS 生态主导权的竞争日趋激烈，系统中枢、

标准接口、生态分发成为焦点抓手。其中，系统中枢统一掌控关键体验质量与安全边界，标准接口降低跨设备与跨形态的开发成本，生态分发与协作机制放大网络效应。

（四）XR 沉浸视听内容生态闭环初步成型

在内容采集端，XR 沉浸影像采集进入专业化与大众化双轨演进。在专业端，2018 年面向 VR 制作的多镜头相机已能录制 8K 3D 并支持 4K 级实时直播回传，2025 年面向影院制作推出的 URSA Cine Immersive 成为全球首台专为 Apple Immersive Video 打造的数字电影机，采用定制固定镜头与双 8K 传感器，最高支持单眼 8K，可以拍摄每秒 90 帧的 3D 沉浸视频，并且每帧都有 16 档动态范围，以确保细节和色彩准确性，其产业意义在于把影院级立体 180°沉浸内容的端到端 workflow 标准化，为专业内容供给建立可规模化的制作范式。在大众端，智能手机亦在探索基础的 3D 内容采集，如双目立体视频通过两颗后摄同步取景与系统级的时序/几何校准、编解码支持即可实现。虽然拍摄的 3D 效果有限，但其普及度有望通过 UGC 形式丰富基础沉浸内容来源。此外，基于手机拍摄视频进行多视角重建，即可实现分钟级 3D 场景生成，这类能力近来已在国内外主流移动操作系统应用商店快速兴起，显示出单机端三维重建在消费级的可用性。综上，业界现已形成“手机入门级空间视频、专业一体化 3D 相机与多机阵列相机”的采集梯度。入门侧快速扩充 UGC 供给，专业侧以一体化立体相机提升画质与制作效率，而面向更高视听自

由度的体积视频仍依赖多相机同步阵列与重建管线，近中期主要服务于演播室级制作。

在制作与呈现端，XR 沉浸影像正在收敛为“标准化编解码+完善工具链”的工业化流程。为传输 3D 立体及多视角的 XR 视频，国际上推出了多视角视频编码（MV-HEVC），作为现有 HEVC（H.265）标准的向后兼容扩展，被设计用以高效承载双目立体与多视角内容，其核心通过跨视图、跨层预测复用冗余信息，同时保持对单层 HEVC 解码模块的复用能力，便于在现有硬件与系统中集成与演进。目前，MV-HEVC 因技术成熟得到明显采用并具备产业化路径，高端 XR 终端如 Vision Pro 等已支持对 MV-HEVC 内容的硬件解码。同时，XR 内容制作工具也同步跟进适配，主流的视频剪辑软件如 DaVinci Resolve、Adobe Premiere 和 Final Cut Pro X 等相继支持 XR 沉浸式视频编辑功能。统一高效的编码标准和日趋完善的制作工具链，正在为规模生产和分发沉浸内容奠定产业基础。此外，各类播放平台和终端设备开始广泛支持沉浸视频形式。播放平台从平面视频库演进为可承载 3D 空间视频的社区形态，逐步形成上传、转码、分发、沉浸播放的发展链路。新一代 XR 设备将空间视频作为差异化卖点，围绕解码能力、像素密度、音画时延与舒适度优化展开竞争。

（五）AR 地图平台化发展架构初步成型

地图正从传统导航工具演进为运行应用与服务的空间入口。随着定位、感知和渲染能力提升，地图不再只是目的地路径与用户兴

趣点的静态呈现，而成为主动情境理解与内容分发，承载实时空间数据、叠加虚拟内容与服务的数字底座。在云 AR 架构下，智能眼镜及手机等移动终端对周围环境进行持续感知，将物理空间转化为可交互的数字空间，现实世界的地理坐标成为数字内容与服务的锚定节点，AR 地图平台角色更接近“空间操作系统”，以动态图层在正确的时间与地点分发信息与服务，弱化传统移动 APP 的孤立形态，有望成为线下数字服务的统一入口。上述变化在产业侧的共识已初步形成，并逐步从试验性应用过渡到平台化建设阶段。

业界构建 AR 地图的产业实践主要有多条路径。一是众包式视觉定位服务（VPS），依托手机与 AR 眼镜的规模化采集，利用视觉 SLAM 与重建算法生成三维实景与特征库，代表做法包括 LiveMaps 与 Niantic 早期的 Real World Platform，其优势在于时效性强、覆盖快；二是基于街景影像与测绘数据的 VPS，通过长期积累的街景图库提取稳定视觉特征并建立全球索引，设备在本地将摄像画面与点云匹配以获得高精度位姿，典型为 Google 的 Geospatial 能力。该路径发挥既有地图资产、便于标准化开放，但依赖影像更新与隐私治理；三是引入人工智能助手和空间语义标签体系，在地图之上实现对场景的智能理解与内容服务推送。其特点是在数字地图中不仅包含几何信息，还叠加对现实环境的语义标注和实时分析，从而使系统能够“看懂”场景并提供上下文相关的信息服务。

AR 地图包含从底层的定位基座到上层的智能交互的多层协同

体系。一是以统一坐标与高精地图为基底的真实地理位置层，它提供了现实世界的准确定位框架，是所有空间计算的参照系；二是将环境抽象为可理解的空间语义标签层，通过计算机视觉和传感器数据，平台为场景中的地点和物体附加语义信息；三是定位与锚定系统，这一层包括 VPS 和云锚点管理。定位系统利用上述地图和特征点云，实现设备在空间中的精准定位与姿态感知。锚点系统允许开发者在真实地理位置上放置虚拟对象，并通过云服务将这些锚点坐标共享给所有用户，让多个设备在同一物理地点看到一致对齐的 AR 内容，实现跨用户、跨设备的持久 AR 体验；四是直接面向用户的呈现层，包括各种叠加在现实空间的虚拟内容、交互界面和应用服务；五是位于最上层的是智能交互与代理模块，即 AI 驱动的空间助手。它结合上述语义地图和实时感知数据，扮演用户在物理空间中的智能导览和代理角色。该体系把“在哪里—是什么—能做什么”串联成闭环，支撑平台化的分发与运行。

国际厂商的部署呈现“大平台+多应用”的格局。Meta 以众包采集、XR 设备与 AI 一体化推进 LiveMaps，目标是把城市级地图升级为可部署第三方内容的“空间 OS”；Google 依托 Street View 与 ARCore/Geospatial API 向开发者提供全球锚点与定位接口，强调开放生态；Niantic 在 Spatial 中整合 VPS 与语义理解能力，面向行业输出平台服务；同时，Snap 与苹果等在本地 AR 与端侧定位上持续投入。为避免数据孤岛并提升互操作性，产业正通过标准接口与联

盟探索锚点格式、隐私与安全规则。总体看，AR 地图平台化方向明确但仍处于建设探索期，商业化规模取决于数据更新与质量、语义理解可靠性、标准成熟度及合规能力。

（六）XR 大空间从试点走向规模化落地

XR 大空间线下娱乐正从试点项目迈向规模化部署。全球市场此类体验的内容形态和运营模式已趋成熟，海外早期兴起的虚拟现实大空间多以高强度游戏竞技和社交互动为主题，国内则侧重文化 IP 和沉浸观影路径，这种差异源于受众偏好和市场环境差异。2024 年法国 Excurio 公司在推出的考古主题沉浸体验《消失的法老》，融合古埃及文化 IP 与尖端 XR 技术，让观众在数百平方米场馆中自由行走、探秘金字塔历史文化。这一爆款标杆成为全球首个线下 XR 大空间票房破亿元的项目，使 2024 年被视为虚拟现实大空间从小众走向大众文娱消费的元年，此后同类项目在全球迅速增多。业界实践证明，这种身临其境的沉浸体验兼具社交效应，引入大空间能够显著提升线下场所客流和坪效，有利于实现可持续的商业模式。

成熟运营模式与政策支持为虚拟现实大空间应用普及扩提供动力。目前业界已探索出多种成功的商业模式。例如，Sandbox VR 以独家内容结合加盟连锁在全球扩张，Zero Latency 输出 XR 技术方案与各地伙伴合作。这些实践表明，虚拟现实大空间可以依托标准化的优质内容或技术，实现跨区域的快速复制和规模运营。当前，我国沉浸式文娱市场快速发展，国内主要 XR 终端厂商和解决方案运

营商纷纷布局线下大空间应用。与此同时，政府主管部门积极引导，2025 年国家电影局发布了《关于促进虚拟现实电影有序发展的通知》，将虚拟现实电影纳入了电影管理和扶持范围，并颁发首张 VR 电影公映许可证（俗称“龙标”），强化对虚拟现实新业态的培育。总体来看，VR 大空间作为虚拟现实大众消费领域的先锋应用，率先跑通了用户付费盈利的商业闭环，实现了从试点示范向规模发展的跃升。未来随着优质内容持续丰富、用户体验不断优化，加之合理布局和政策扶持，VR 线下沉浸文娱有望像电影院、主题乐园等成为常态化消费形态，拓宽虚拟现实产业的大众市场空间。

（七）XR 赋能行业从“场景适配”迈向“痛点破解”

在工业生产领域，XR 技术正从辅助展示走向直接解决效率成本痛点。制造企业开始将 XR 用于核心业务培训和操作流程，而不仅限于展示。例如，宝马、PTC 等一些工业企业利用虚拟现实构建工业仿真培训系统，使新员工能够在虚拟环境中操作复杂设备，从而安全高效地传承技能。XR 还用于远程协同作业和维护巡检，一线技术人员佩戴 AR 设备即可实时获取后台专家的指导和设备的三维可视化信息，极大提高了维修效率。为支持这些应用，业界开发了低门槛工业 3D 内容生成工具，将 CAD 等工程数据自动转换为 XR 可用的三维模型。然而，XR 在工业领域规模落地仍有痛点亟待解决，如何将 XR 系统有机融入现有工业系统并产生成效，如打通 XR 设备与 MES、SCADA、WMS 等系统的数据链路，让 XR 真正融入工

业流程并发挥实质效益。这一整合难题横亘在商用推广前，促使行业重点从概念验证转向系统集成和标准建设，以破解推广瓶颈。

在安全应急领域，XR 使能“危极特恶”场景作业培训，显著提升应用安全性和可靠性。消防、矿山、核电等高风险行业长期存在培训演练成本高、危险性大的痛点。借助 XR 构建模拟仿真场景，学员可以在虚拟环境中经历火灾、矿难等极端事故情境，进行低风险操作演练。相较传统训练模式，XR 培训可以重复进行、数据可记录分析，培训效果更可量化。这些优势促使越来越多企业将其纳入常规培训体系。然而，目前不少应急培训 XR 系统仍以设备为中心设计，重硬件轻内容，导致学员体验和实效不足。未来需要转向以培训效果为中心，丰富虚拟场景的交互和评价机制，才能充分发挥 XR 在应急培训中的价值。

在教育领域，XR 正聚焦于提升教学质量和个性化培养。过去教育类 VR 应用多停留在虚拟展示和漫游层面，与教学大纲和考核需求脱节。现在一些教育科技团队将 XR 用于课堂教学和职业培训，有针对性地缓解教学痛点。例如在 K12 课堂，教师利用 XR 构建沉浸式教学环境，带领学生进入历史现场或微观世界，显著提升学习兴趣和理解。在职业教育中，虚拟仿真使学员可无限次练习复杂技能，不受实物设备和耗材限制，降低训练成本。一些团队研发 XR 教学内容自动生成和智能虚拟教员，以缓解师资不足并实现因材施教。目前 XR 教学内容与课程标准的融合仍不足，沉浸式教学规

范匮乏。随着落地案例增多，业界探索将 XR 内容纳入教学体系，推动 XR 在教育领域的规范化应用。

在商贸营销领域，XR 应用重心正从追求形式新颖转向注重投入产出比。如今商业运营方更关注 XR 能否带来实效，购物中心希望借 XR 提升客流、延长顾客停留并促进消费，一些企业推出覆盖众多商场的 AR 导购互动系统，提升顾客互动并带动消费。房地产行业借助 VR 看房和虚拟样板间提高看房效率，使客户更高效获取房产信息，提高成交率。这些案例表明，抓住商贸场景降本增收的需求痛点，XR 有望从既往的新奇噱头变成新型营销工具。但大规模推广仍需解决投入产出比（ROI）评估等难题，只有明确效果指标、建立可持续商业模式，才能打消企业对高投入回报慢的顾虑，推动 XR 在更多商业场景落地。

（八）XR 助推人工智能训练新范式

XR 对 AI 的催化效应将不断放大，正在促使人工智能训练迈向更通用、更高效的新阶段。XR 能够提供高度交互和仿真的三维环境，支持人工智能体在虚拟世界中学习感知和行动，从而加速空间智能的发展。虚拟机器人可以在高效逼真的 3D 仿真环境中接受训练，再将学得技能迁移到现实中。相比传统二维数据，XR 所提供的空间交互数据为 AI 算法注入了更加丰富的环境上下文和行为经验。一方面，海量沉浸式数据和仿真经验的积累有望催生机器人领域的“基础模型”，让 AI 具备跨任务迁移的通用技能储备。将人演示的直觉引

导与机器人运动规划算法相结合，可以显著降低智能体学习复杂技能的门槛，未来这种人机融合的数据迭代或将成为训练大模型的有效途径。另一方面，这一范式最终指向更智能的 AI 助手，它们能够适应人类生活场景的复杂性，实现真正的人机协作。例如 XR 仿真为 AI 提供了理解并融入人类环境的“练习场”，有望催生能在日常环境中与人协同工作的 AI 助手。当前，XR 赋能 AI 训练方式主要有 XR 沉浸遥控演示、多模态主视角数据采集等发展路径。

XR 沉浸遥控演示通过沉浸式遥操作示范为空间智能提供了高质量的模仿学习数据。操作员佩戴 VR 头显以第一视角远程控制机器人执行复杂任务并同步收集感知与动作示范。例如智能机器人 ALOHA 及其后续系统 AV-ALOHA，采用双手 VR 控制与主动视觉相结合，让人类演示者动态调整机器人摄像头视角以获得最佳任务视野。这一真人示教范式显著降低了训练数据获取的门槛，机器人无需复杂编程便可直接模仿人类完成精细操作，并在实际任务中取得较好表现。当前，越来越多研究采用 XR 沉浸式遥操作来收集机器人学习所需的数据，让示教者像操作自己身体一样控制机器人完成任务，从而获取精细的感知/动作记录。通过人类直接在机器人本体上的远程操控捕获专家演示，被视为收集模仿学习训练数据的重要途径之一。

XR 多模态主视角数据采集为空间智能训练提供了丰富的真实世界语料。在 XR 操作示范的基础上，引入多模态扩展正成为扩大

训练数据规模和多样性的关键手段，例如 Ego4D 项目在全球收集了数千小时日常生活第一视角视频，数据规模较以往提升一个数量级。这一趋势受益于新型可穿戴设备的普及，如佩戴 Meta Aria 智能眼镜等 XR 传感平台，用户在日常活动中可被动记录视频、音频、3D 手部轨迹和空间定位等多模态信息，为 AI 积累海量贴近人类视角的训练素材。近期高校院所研究者使用轻量化智能眼镜获取人类操作演示并与机器人遥操作数据对齐，训练统一的模仿学习策略，使额外增加一小时人类视角数据所带来的性能提升远超同等时长的机器人数据。通过这种第一人称“经验”反哺，AI 模型大幅提升了对新场景、新任务的泛化能力，并为具备人类视角理解的 AI 助手奠定了数据基础。

四、行业面临的问题

（一）需求侧用户体验问题

XR 产业虽然在快速发展，但行业仍面临诸多瓶颈与挑战。用户体验方面的“六道卡”问题尤为突出，即“用、贵、重、视、智、晕”六个维度的痛点，严重影响大众接受度和市场普及。其中，“用”是指内容供给不足削弱了 XR 对用户的吸引力。当前高品质 XR 内容和应用稀缺，内容生态尚处起步阶段。内容制作的常态化机制尚不健全，优质沉浸式内容的数量和更新频率尚不能满足用户需求。不少用户在初次体验新鲜感后，由于缺乏后续有吸引力的内容而失去继续使用的动力，设备陷入闲置；“贵”是指终端价格偏高提升

了大众消费门槛。当前部分高性能 MR 头显与 AI+AR 眼镜等终端售价仍然较高，少数设备价格接近或超过万元，影响了 XR 设备的规模化普及；“重”是指设备笨重影响佩戴舒适性。与手机等成熟电子产品相比，目前的 XR 设备在重量体积和外形美观上具备显著的优化空间。例如，苹果 Vision Pro MR 头显重量逾“斤”，Meta Ray-Ban Display AI+AR 眼镜重量仍数倍于普通近视眼镜。较高的设备重量降低了用户长时间使用的便利性和意愿；“视”是指显示效果上的局限。XR 设备的近眼显示效果距离人眼理想状态尚有差距，多数 MR 头显单眼分辨率仅 2K，像素纱窗效应依然明显。AI+AR 眼镜用户在入眼亮度、视场角、全彩显示与视力匹配等方面具备一定的预期落差；“智”是指智能交互能力欠缺。当前 XR 应用的交互方式和内容呈现仍不够“聪明”，难以根据用户与环境上下文做出智能交互，语音助理、手势识别等 AI 技术虽有应用但不够自然流畅，个性化与情景化的智能服务体验受限；“晕”是指多种因素诱发的生理不适。其成因复杂，包括晕动症、时延控制、辐辏调节冲突（VAC）、内容设计等因素。

（二）供给侧产业生态问题

研发成本较高，企业创新投入压力较大。XR 领域多种支撑技术交织复合，且产业链条较长，高昂的研发投入成为进入门槛。龙头企业如 Meta 近年来 XR 部门 Reality Labs 的年度投入高达百亿美元量级。相比之下，大多数中小企业无力承担如此庞大的研发投入，

导致 XR 重点技术研发后劲不足。在这种情况下，行业创新更多依赖于少数巨头推动，中小企业的创新动力和能力都受到限制，不利于产业的长期繁荣。如何降低创新成本、加强产学研协同以及共享基础研发成果，成为 XR 产业须面对和解决的问题。

应用规模化程度不足，商业变现路径尚在摸索。一方面，规模化、可落地的实践案例稀缺，导致示范效应有限且用户体验价值不足。多数厂商局限于技术视角，未能深度结合文化娱乐、工业制造、教育培训等垂直领域的业务特性，难以精准识别多元化的虚实融合刚需场景；另一方面，现有应用示范多停留于“展厅观摩式”层面，呈现“看似美好却难转化”的虚火状态。因缺乏与产业痛点匹配的可持续产出模式，商业化探索举步维艰，既难挖掘用户付费锚点，亦难建立可复制的盈利路径。

产业生态未贯通，协同效应尚未形成。XR 产业链条长且跨界融合特征明显，从底层器件、整机制造到内容制作、平台分发，环节众多且彼此依赖。当前 XR 生态存在各自为战、小而散的问题，硬件厂商、内容团队、平台运营商之间缺乏协同，产业集聚效应不明显。相比 Meta、苹果等构建起从设备到内容的完整生态，多数 XR 企业多聚焦于单点创新，缺少统一的发展蓝图和合作机制，生态飞轮效应尚未出现。

虚实融合互联网新机遇给社会治理带来了一系列新挑战。XR 技术对用户生物及行为数据的采集能力将上涨一个量级。攻击者可利

用不同的虚拟化身融入数字世界，在未经用户知情同意的情况下，进行隐私数据收集与交易。此外，虚实融合的沉浸式业态作为各种社会关系的超现实集合体，其价值理念、道德准则等复杂规则和现实世界间可能会出现明显分化甚至是异化对立，或试图将其中的规则应用于现实生活中，从而导致现实世界的道德伦理冲突。

五、对行业发展的对策建议

（一）聚焦技术攻关的靶点识别，应对技术创新的不确定性

聚焦重点领域开展技术攻关。支持我国骨干企业研发投入，开展共性技术研究与工程创新，聚力攻关 XR 关键技术，如专用空间计算芯片、微显示屏与光波导镜片等重点器件。大力发展 3D 沉浸内容采集、制作与编解码技术。探索基于 UGC、AIGC 的 3D 数字内容低成本规模化生产技术。推动人机交互向自然化、智能化方向发展，发展高性能、低功耗的手势识别、眼动追踪、环境感知与肌电传感等新型交互技术研发。研发面向眼镜形态的具备第一视角与持续在线特性的空间智能体。

深化产学研协同的合作机制。建立企业与高校、科研机构的常态化合作对接机制，推动产学研深度融合、协同创新。组建虚拟现实创新联合体，通过共建联合实验室、建设产业技术平台等方式，面向产业需求共同凝练科技问题，通过“揭榜挂帅”、“赛马争先”等项目组织机制推进产学研联合攻关。

（二）聚焦产业链条上下游协同，构建新兴产业形成的内生机制

打造标志性的产品体系。在新终端方面，发展 VR/MR 头显、轻量化智能眼镜等多形态头戴式智能终端，优化终端体积重量、续航散热、显示质量、自然交互、智能服务等指标。在新内容方面，加强我国在 3D 视听内容格式标准、拍摄采集、制作分发等多链条生态建设。在新平台方面，加强 XR 虚实融合业务平台建设，在云、网、边、端、用融为一体的创新体系下提升新型数字空间运营服务能力。搭建 XR 典型产品集成验证、产品中试与测试评价公共服务平台，培育一批新技术、新产品与新市场。

促进大中小企业融通发展。发挥领军企业在技术创新引领、产业合作协同等方面的重要作用，完善大中小企业间协同创新的融通渠道，通过供需对接、项目合作、联合研发等方式，推动更多创新企业融入龙头企业的创新链、产业链、供应链，带动中小企业协同发展，构建大中小企业链式融通创新格局，共同推动关键技术攻关突破和产业高质量发展。

优化产业发展支撑环境。强化产业融合导向，通过“强技术、搭平台、引育结合”加速要素互通，构建覆盖初创企业、专精特新企业至领军企业的梯次培育体系。鼓励专业化企业或机构统筹虚拟现实产业园区的建设与运营，推动产业功能与城市功能深度融合，实现技术研发、公共服务、市政配套的一体化布局。创新资金引导机制，统筹国家级基金与地方资源，探索“投补联动+投贷联动”等

财税支持模式。支持高校加强 XR 相关学科建设，推动校企联合培养。鼓励我国企事业单位参与国际标准化活动，积极贡献中国方案。

（三）聚焦应用创新的多元探索，以 XR 规模落地破解商业闭环瓶颈

深化虚拟现实对行业应用场景的融合赋能。从沉浸视听、智能交互、混合协作、工业仿真等数实融合焦点需求切入，在消费娱乐、工业生产、医疗教育等领域筛选可规模化变现的场景方案，“让能跑的先跑起来”，重点破解“展厅式孤岛、小众化雷同”的发展困境。具体实施中，通过打造集展示体验、孵化培训、供需对接等公共服务平台，组织沉浸式应用大赛、标杆案例征集等方式，驱动差异化需求匹配与商业潜力挖掘，最终培育出成规模、易推广、可持续盈利的实践成果，完成从“展示性应用”到“生产性应用”的发展跃迁。

（四）聚焦监管治理的预研预判，促进发展环境的健康有序

营造安全开放产业发展环境。结合 XR 沉浸式业态以人为中心、虚实融合、开放互通等发展特点，开展与之适配的新型政策监管治理研究。重点推动智能终端、虚拟化身、分发聚合平台等 XR 重点领域的用户隐私安全，探索研究用户知情同意的隐私交易机制，加速潜在创新所须的数据流动，公平补偿用户产生的生理、行为、环境等隐私信息。开展信息信任机制与社会规则研究，研究设计可验

证的信息信任机制，应对深度合成、虚假传播等信息失真问题。综合运用法律、伦理、技术等手段，由政府、行业、平台和公众多主体共同介入，以“共建共治共享”为目标推进行业发展。



中国信息通信研究院

地址：北京市海淀区花园北路 52 号

邮编：100191

电话：010-62302232

传真：010-62304980

网址：www.caict.ac.cn

