

附件 2

技术性文档项目建议书

计划编号：

项目名称 (中文)	通用沉浸式体验（AGI）技术研究报告		
项目名称 (英文)	Research Report on Advanced General Immersive (AGI) Experience Technology		
发布	☐ 内部	☒ 外部	
牵头单位	名称：咪咕文化科技有限公司		计划起止时间 2025 年 8 月-2025 年 12 月
	联系人：毕蕾		
	联系方式：18066077313		
共同申请单位	北京大学、上海交通大学、中国信息通信研究院、中央广播电视总台、国家广播电视总局广播电视规划院、中国电子技术标准化研究院、华为技术有限公司		
目的、意义	一、目的 （一）标准体系构建 针对沉浸式体验领域缺乏统一国际与行业标准的现状，构建通用沉浸式体验分级及评估框架。通过明确不同沉浸层级的技术特征、体验边界与评估维度，为沉浸式体验的技术研发、产品设计、体验测评建立规范化、可参照的行业准则，填补标准空白，引导行业有序发展。 （二）科研成果输出与影响力塑造 计划于 2025 年完成并发布沉浸式体验领域白皮书 / 研究报告，系统梳理通用沉浸式体验的科研技术路径。从基础理论、技术研发到应用实践，进行全链条剖析，为产业界提供技术研发指引，助力企业精准锚定发展方向、优化产品布局。同时，依托专业研究成果发布，强化发起机构（如北大视频与视觉国家工程中心）在沉浸式体验学术与产业领域的权威性，提升行业话语权，引领领域发展潮流。 （三）产业生态激活 以构建的分级体系与研究成果为基石，构建产业界与学术界协同创新的生态激励机制。通过明确技术分级目标与发展路径，激发企业开展技术创新、产品迭代；引导学术界聚焦关键科学问题（如超现实沉浸的感知理论、多模态交互的认知机制）开展前沿探索，形成“标准引领 - 研发推进 - 实践验证 - 理论反哺”的良性循环，驱动沉浸式体验产业持续升级。 二、意义 （一）行业治理维度 填补沉浸式体验领域标准空白，建立涵盖技术研发、产品评估、体验测评的统一规范。从技术研发端，规范不同层级沉浸式体验的技术指标与实现路径，避免企业因技术路线模糊导致的重复投入、资源浪费；在产品评估环节，提供多维度、可量化的评估标准，助力企业精准定位产品层级、优化产品性能；于行业秩序层面，构建有序竞争环境，推动资源向高价值创新环节聚集，促进沉浸式体验行业协同、高效、可持续发展。 （二）技术演进维度 清晰界定 L1 - L5 各层级沉浸式体验的核心技术特征与演进要求，为技术研发提供明确的阶		

	段目标与突破方向。通过聚焦关键技术，推动跨技术领域的融合创新。加速沉浸式体验技术从基础应用（L1-L2）向高阶形态（L4-L5）演进，持续提升体验的真实感、交互自然度与沉浸深度，引领数字交互技术的迭代升级。 （三）用户体验维度 构建多维度、分层级的沉浸式体验认知框架，帮助用户精准识别不同体验形态的技术特征与体验价值，实现体验内容与设备的理性选择。伴随技术标准驱动下的产业创新，用户可接触到更丰富的沉浸式应用场景，体验从单一感官刺激到全感官、虚实难辨的深度沉浸跃迁，重塑数字消费体验模式，拓展人类感知与交互的边界。 （四）产业生态维度 建立沉浸式体验产业上下游协同的“技术-产品-服务”关联标准。硬件制造企业依据分级标准研发适配设备；内容开发主体按照层级需求创作差异化内容；平台服务方基于标准搭建体验分发与交互平台。通过各环节的标准化协同，打破技术与产业壁垒，促进要素高效流通与配置，加速沉浸式体验商业化落地进程，培育新的经济增长极与产业发展生态。
范围和主要技术内容	三、范围 （一）体验层级覆盖 覆盖沉浸式体验从基础感官交互（初级沉浸：以3D电影、360°视频为代表，依赖视觉、听觉双通道，交互维度单一）到极致虚实融合（超现实沉浸：意识上传、完全虚拟世界构建，突破物理现实感知边界）的全谱系体验形态。纳入不同互动强度（低交互的传统影视内容消费、高交互的VR游戏竞技）、技术复杂度（低技术门槛的移动AR应用、高技术集成的全身跟踪VR系统）场景，对应L1-L5分级体系中各层级技术能力与体验特征（如L1级聚焦立体视觉基础能力，L5级探索多用户、多模态全真沉浸），实现对沉浸式体验发展进程的全周期覆盖。 （二）技术维度涉及 围绕体验强度（用户参与度、感官投入程度）、交互性（用户与虚拟环境的反馈机制、交互深度）、虚实融合度（现实与虚拟场景的边界模糊程度、融合模式）、技术复杂度（硬件设备集成度、技术实现难度）四大核心维度，构建沉浸式体验的多维度分类与界定模型。覆盖沉浸式体验的全技术链条：从内容生产端（多视点+深度信息采集、全景视频编解码）、呈现技术层（任意自由视点渲染、光场显示驱动）、交互手段侧（手势识别算法、触觉反馈硬件设计），到硬件支撑体系（VR头显空间定位、全身跟踪传感器组网），以及从基础产业化应用（L1）到前沿技术探索（L5）的全技术发展阶段，实现技术要素的全面涵盖与深度关联。 （三）标准化关联主体 关联国际标准化组织（ISO）中ISO/IEC JTC 1/SC 24 聚焦计算机图形、VR/AR技术标准以规范基础数据表示与图形渲染，ISO/IEC 23005 围绕感官信息处理来指导多感官交互系统设计；IEEE 标准协会推出IEEE P2048 系列标准，涵盖VR/AR术语体系定义、沉浸度评估模型建立及内容分类框架设计，其中P2048.1 夯实术语基础、P2048.2 完善分类分级逻辑；行业联盟方面，Khronos Group 主导的OpenXR 与World Wide Web Consortium（W3C）主导的WebXR 标准通过简化跨平台VR/AR应用开发流程以提升技术兼容性，Immersive Technology Alliance 推动沉浸式技术标准化普及来促进产业协同，Virtual Reality Industry Forum 制定VR内容制作与分发指南以保障内容质量与体验一致性；学术研究机构则开展沉浸度评估方法创新、用户体验认知理论及技术创新应用等前沿研究，为标准构建提供理论支撑与技术验证，各主体在标准制定、技术推广、理论研究等多环节形成协同机制。 四、主要技术内容

（一）分级体系构建

1. 层级划分逻辑

结合沉浸式体验的感官认知特性与技术实现路径，构建 L1 - L5 分级体系。L1 聚焦立体视觉基础能力，以多视点 + 深度信息处理为核心，实现基础沉浸式内容的产业化应用；L2 强化全视场呈现能力，依托全景视频技术，拓展用户视觉感知范围；L3 突破 6DoF 交互限制，通过任意自由视点 + 交互绘制显示技术，实现空间维度的交互拓展；L4 深化多模态交互，融合手势、触觉、光场显示、全景声、6DoF 等技术，构建多感官协同交互环境；L5 探索全真沉浸体验，瞄准多用户、多模态融合的极致虚实难辨场景，突破当前技术瓶颈。同时，从体验强度（低沉浸 - 高沉浸）、交互性（低交互 - 高交互）、虚实融合度（低融合 - 高融合）、技术复杂度（低复杂度 - 高复杂度）四个维度，细化各层级体验的技术特征与评估指标，形成可量化、可验证的分级框架。

2. 分级依据与验证

以技术实现难度系数（涵盖硬件集成、算法复杂度、系统功耗等）、感官参与度（视觉、听觉、触觉、本体觉等感官通道激活数量与强度）、交互深度（用户动作对虚拟环境的影响范围与反馈精度）、虚实融合水平（现实场景与虚拟场景的像素级融合度、物理定律一致性）为核心指标，结合实际应用案例（如科技冬奥项目中交互式自由视点技术对应 L3 级特征；Apple Vision Pro 设备的全景显示与空间交互能力关联 L2 - L4 级），建立分级指标的量化评估模型。通过实验室测试（如交互延迟测试）、用户体验测评等手段，验证分级体系的科学性与适用性，确保分级结果可衡量、可复现。

（二）关键支撑技术

1. 内容生产技术簇

（1）多视点+深度采集技术：采用多相机阵列与结构光/TOF 传感器融合方案，构建高精度三维场景重建系统，实现毫米级精度的空间信息采集，为沉浸式内容提供基础几何与深度数据支撑。

（2）全景视频编解码技术：基于球面投影与等距圆柱投影等映射算法，结合高效视频编码标准，研发低带宽、高分辨率的全景视频编解码方案，保障全视场内容的高质量传输与呈现。

（3）自由视点渲染与交互技术：运用光场渲染、神经辐射场等算法，实现用户自定义视角的实时渲染与交互控制，突破固定视角限制，增强沉浸式体验的交互自由度与场景真实感。

（4）AIGC 内容生成技术：①文本驱动三维场景生成：基于 Transformer 架构与扩散模型，通过自然语言解析模块将文本描述转化为语义特征向量，再利用三维生成网络生成高精度几何模型与材质纹理。该技术支持多模态信息融合，可根据用户输入的风格关键词快速生成对应场景，显著降低 3D 建模人力与时间成本。②图像/视频智能生成：采用生成对抗网络结合扩散模型，通过特征提取模块分析单目/多目图像、视频序列的空间与时间特征，再经生成器重构为动态虚拟场景。输入单张建筑照片即可生成其内部漫游场景，基于视频片段实现场景动态扩展，支持实时场景补全与内容延伸。③智能角色与动画生成：基于骨骼蒙皮与机器学习算法，将动作捕捉数据与 AI 生成模型结合。通过训练海量动作数据集，模型可自动生成符合物理规律的角色运动，如跑步、格斗等复杂动作；同时支持表情迁移技术，将真人表情实时映射至虚拟角色，提升角色交互的自然度与情感表现力。④场景智能优化与编辑：利用强化学习与自动编码器，对生成的虚拟场景进行实时优化。自动调整光照参数以适配场景氛围，通过语义分割技术识别场景元素并进行替换或重组；支持用户通过简单文本指令实现场景的快速编辑与迭代。⑤跨模态生成融合：基于跨模态通信架构，将 AIGC 生成的视觉内容与触觉、音频数据进行语义级融合。在生成虚拟森林场景时，同步生成风吹树叶的音效与触觉反馈数据，通过跨模态编码技术实现多维度内容的协同传输与

	渲染,增强沉浸式体验的整体性。⑥跨模态生成对抗网络:基于单阶段深度文本-图像融合框架,通过目标感知判别器与匹配感知梯度惩罚策略,强化语义与视觉特征耦合,实现高分辨率图像生成,支持跨模态通信系统内容生产。 2. 交互与感知反馈技术簇 (1) 手势交互技术:融合深度学习视觉算法与表面肌电信号采集,实现高精度、低延迟的手势动作识别与意图解析,支持自然交互指令的精准执行。 (2) 触觉反馈技术:基于压电陶瓷、形状记忆合金开发多维度力觉反馈装置,结合触觉渲染算法,模拟材质触感与物理交互力,构建逼真触觉交互环境。 (3) 光场显示与全景声技术:光场显示依托微透镜阵列与全息投影实现真三维空间成像;全景声基于波场合成与 HRTF 技术,营造三维听觉空间,强化视听融合体验。 (4) 6DoF 空间定位技术:集成激光雷达、IMU 与 VIO,实现亚毫米级六自由度运动跟踪,保障虚拟空间交互的流畅性与定位精度。 3. 硬件与环境支撑技术簇 (1) LBE 大空间交互技术:采用无线组网与分布式计算架构,结合 UWB、激光定位系统,构建超百平米多人交互空间,支持群体协同漫游与场景互动。 (2) 高端 VR 头显技术:配置高分辨率 Micro-OLED 显示模组与轻量化光学系统,结合眼球追踪与注视点渲染技术,优化视觉体验与佩戴舒适性。 4. 跨模态通信与传输技术簇 (1) 跨模态信息获取与表征技术:基于统一模型实现音-视-触多模态数据的普适化感知与混合编码,通过特征融合形成统一表征形式。 (2) 跨模态异构码流传输技术:采用边缘智能调度策略,动态调整多模态数据传输优先级与码率,保障复杂网络环境下的稳定传输。 (3) 跨模态信息重建技术:利用模态间语义融合机制,结合机器学习模型与用户反馈,实现远程场景下多模态信号的高效重建与还原。 (三) 标准化探索与应用 1. 现有标准适配分析 系统梳理 ISO、IEEE、Khronos Group 等国际组织发布的相关标准,剖析其在沉浸式体验分级、评估、开发中的适配场景。例如,ISO/IEC JTC 1/SC 24 的计算机图形标准,可指导沉浸式内容的几何数据表示与渲染流程规范;IEEE P2048 系列标准,为 VR/AR 术语统一、沉浸度量评估提供基础框架;OpenXR/WebXR 标准,助力跨平台沉浸式应用开发的技术兼容性实现。通过技术映射与需求匹配,明确现有标准的应用价值与适配边界,为构建沉浸式体验标准体系奠定基础。 2. 标准协同构建机制 建立“学术研究 - 产业实践 - 标准制定”的协同闭环。开展企业级技术验证与应用实践,提炼产业共性需求与技术痛点;联合学术机构(如高校人机交互实验室、计算机图形学研究中心),开展沉浸式体验基础理论,为标准构建提供理论支撑;依托 ISO、IEEE 等标准化组织,推动研究成果转化为国际 / 行业标准,从术语定义、分级框架、评估流程到技术应用规范,构建全链条、协同化的沉浸式体验标准体系,推动标准的全球范围内落地实施与行业应用
牵头单位	 (签字或盖公章)

年 月 日