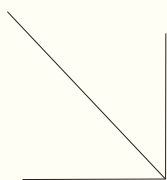


基于 AIGC 的数字人生成与应用研究

文 ◆ 中南财经政法大学 余树乔



应用演进趋势，为数字人技术的标准化、系统化发展提供理论依据与实践参照。

1 AIGC 驱动下的数字人核心技术

1.1 多模态生成技术

AIGC 技术的本质是大模型驱动的内容生成，涉及语言、图像、视频、音频等多模态协同。当前，Stable Diffusion、Make-A-Video、AnimateDiff 等模型可实现高质量的图像与视频合成，而语音生成模型（如 VALL-E、Bark、XTTS）则支持多情绪、多语种、多音色的语音表达^[2]。多模态生成系统正在推动内容从静态走向动态、从单一走向融合，为数字人提供了语义理解与表达能力的双重支持。

1.2 语音驱动与动作建模

语音驱动下的表情和肢体动作同步是实现数字人“拟人化”的关键。当前技术通过语音信号到音素的映射，再结合唇形同步技术（如 Viseme Mapping）、骨骼动画系统与动作迁移模型（如 Audio2Motion、StyleGestures），实现自然语言驱动的实时面部与身体表达^[3]。此外，还可通过深度学习网络对语义内容进行情感分析，使数字人在语调、表情、动作表现上实现情境感知。

1.3 视觉建模与神经渲染

高保真度的数字人需要依赖精细的视觉建模技术，如 3D 建模、纹理贴图与面部表情捕捉。神经渲染技术（Neural Rendering）通过引入神经网络对图形渲染进行优化，可在保证真实性的同时提高渲染效率。同时，结合表情迁移与头部重定向（Head Reenactment）技术，可实现任意人物形象在不同语境下的自然表达。

1.4 知识驱动与语义理解

为实现更自然的人机对话，AIGC 数字人须具备复杂语义理解与知识生成能力。大型语言模型通过预训练与微调机制，具备强大的语境适应能力。此外，通过 RAG（Retrieval-Augmented Generation）机制，数字人可结合外部教学资源、数据库与实时信息，实现更高可信度的内容输出。

1.5 实时渲染与边缘部署

为了适应多场景应用，数字人必须具备实时响应能力。Unity、Unreal

引言

数字人（Digital Human）作为人机融合的产物，借助 AIGC（Artificial Intelligence Generated Content）技术在形象生成、语音驱动、动作捕捉、情感交互等方面取得了显著突破。自 ChatGPT、Sora 等生成模型面世并不断发展迭代以来，数字人逐步从“高成本、低智能”的特效模型，演变为“低门槛、高可用”的虚拟个体，被广泛应用于新闻播报、虚拟客服、品牌代言以及教育陪伴等领域^[1]。本文以 AIGC 为技术基础，从底层能力到行业落地进行系统研究，力图揭示数字人从“生成”到“赋能”的路径逻辑，以期探索其在多元行业场景中的

【作者简介】余树乔（1996—），男，湖北黄冈人，硕士研究生，助理工程师，研究方向：教育信息化。

Engine 等游戏引擎结合 NVIDIA Omniverse、WebRTC 与 5G 网络，实现了数字人图像与语音的低延迟渲染与传输。此外，轻量化部署方案不断涌现，使数字人真正走进日常应用环境，通过边缘计算与分布式架构，用户可在资源有限的终端实现流畅交互^[4]。

2 数字人场景应用分析

2.1 数字媒体与智能传播

在数字媒体领域，AIGC 驱动的数字人正在重塑内容生产与传播逻辑。数字主播、虚拟主持人、AI 记者等应用形态广泛出现在新闻与媒体平台中。例如，新华社的“新小微”、韩国 MBC 的 AI 主播，能够不间断自动播报新闻、生成多语言内容，对特定用户群的个性化推送实现粉丝粘性与传播影响力的同步提升。技术上，这类应用依托文本生成（NLG）、语音合成（TTS）与表情驱动（Facial Animation）融合，实现在不同场景中的高效输出。

2.2 智慧政务与服务辅助

在政务系统中，数字人逐渐承担起服务引导、政策讲解、办事辅导等多重角色。例如，政务大厅上线的“数字导办员”，可提供人脸识别登录、语音交互引导、材料填写辅助等功能，提升政务效率与体验满意度。技术上，这类数字人常集成语义理解（NLU）、人脸识别、自然语音对话系统（NLDS）以及业务知识库，具备语境适应与多轮对话能力。在政策宣贯场景，虚拟代言人可承担复杂内容的可视化解释任务，通过动画演示、即时问答等方式帮助民众更好地理解法律法规与政务信息。

2.3 智慧教育与沉浸式教学

教育是当前 AIGC 数字人应用中最具变革潜力的领域之一。在智慧教育环境中，数字人不仅可承担虚拟教师角色，还可作为学习伙伴、课程助手、评测反馈系统的核心组成。例如，基于课程内容生成的虚拟讲解员可通过多模态输入（语音 + 图像 + 文字）实现知识点讲解、难题解析与实验演示，尤其适用于理科类、工程类课程的抽象内容展示。在个性化教学方面，数字人可结合学生的学习数据生成差异化学习路径，提供有针对性的问答与知识补强。同时，数字人还可实现考后错题分析、心理健康引导、跨文化交流等功能，进一步拓展教育的服务边界^[5]。此外，在特殊教育与边远地区教育场景中，数字人提供了优质资源的远程复制能力，能够解决师资不足、教学资源不均的问题。因此，AIGC 数字人所打造的沉浸式、互动式、可迁移式的教学环境能提升教育公平性与普及率。

2.4 数字人直播“带货”与商业创新

在商业领域，数字人正成为营销推广的重要工具，尤其在直播电商场景中展现出独特优势。数字人直播带货结合了 AI 智能生成内容与虚拟形象驱动能力，突破传统主播的人力成本与时间限制，实现“全天候不下播”的营销模式。技术上，这类应用主要依赖语音合成（TTS）、动作捕捉驱动（Motion Capture）、知识图谱构建与实时语义识别，实现与用户的语义互动、表情同步与个性化引导。相比真人主播，数字人可根据用户画像与购买行为数据，动态调整话术、推荐逻辑与展示节奏，提

升转化率与用户粘性。同时，数字人形象可高度定制，支持品牌化 IP 构建与多风格适配，从而形成持续化、差异化的品牌传播路径。更重要的是，数字人具备可复制、可扩展的优势，适用于跨区域、跨平台的多场景部署，尤其适合中小品牌在初期拓展市场、提升曝光方面的应用需求。

3 面临的挑战与发展趋势

随着 AIGC 与数字人技术的快速发展，其在各行业的应用潜力日益显现，但仍面临多方面挑战。本研究通过综述文献资料与实际应用情境，将数字人技术存在的挑战与发展趋势分为三大类，通过分析这些技术和应用瓶颈，以期推动数字人从概念验证走向广泛落地。

3.1 拟真度与身份识别方面

数字人在拟真度与身份识别方面较前些年已经有了大幅进步，但是仍然存在诸多发展瓶颈与问题。首先，计算机运算得到的高精度视觉还原的挑战。真实人脸拥有丰富的细微肌肉运动、皮肤光影变化和动态皱纹，这些都需要通过高分辨率 3D 扫描、动态纹理映射和神经渲染等技术来模拟。神经渲染（Neural Rendering）和神经辐射场（NeRF）技术正逐步解决视角切换和光照一致性问题，但计算资源需求高，实时渲染仍是瓶颈^[6]。其次，语音自然性和情绪表达。目前语音合成虽然已经实现多音色、多语调的生成，但如何做到符合人物身份、上下文以及情绪的高度拟合依然困难。多情绪语音生成和自适应说话风格模型（如基于情感编码器的 TTS）是未来发展重点。再次，语音的情绪识别与反馈机制

尚不完善，限制了数字人与用户的深度情感交流。最后，身份伪造与伦理风险在数字人的应用中也值得重点关注。深度伪造技术（Deepfake）带来的身份欺诈风险日益严峻，数字人的身份验证、版权保护和隐私安全亟须建立技术和法律双重保障机制。利用区块链技术进行身份溯源、多模态水印以及数字指纹技术是当前研究热点，能有效提升数字人内容的可信度和可追踪性。

3.2 多轮交互与情感理解

不论是简单的人工智能对话助手，还是高阶的数字人模型，目前大多数数字人只能处理单轮或短对话，缺乏对长时间、多轮交流的上下文记忆和语义理解能力。大型语言模型（LLM）虽然具有强大的语言生成能力，但如何结合对话状态管理、知识库检索（RAG 技术）和用户画像实现个性化、持续性的智能交互仍是技术难点。此外，为了提升用户体验，数字人需结合用户行为数据、偏好和情感状态进行动态调整。基于多模态感知（语音、面部表情、肢体动作）的情感计

算（Affective Computing）和行为预测模型正逐步被引入，但如何在保护隐私的前提下高效收集和利用数据仍是瓶颈，因此融合多源传感器数据、强化学习驱动的情感表达优化以及模拟人类社会行为的伦理框架设计是数字人应用的发展趋势。

3.3 泛场景部署与系统集成

在数字人应用和实践过程中，多数数字人应用依赖定制化开发，导致系统难以复用和扩展，且平台化与模块化设计不足，限制了多场景快速部署和功能迭代。未来需构建基于微服务架构的模块化平台，实现语音、视觉、动作、知识库等能力的松耦合集成，支持跨行业、跨场景的灵活组合。此外，实时性与算力平衡也是需要重点考虑的问题，数字人应用对实时响应性能要求极高，融合边缘计算、云端分布式协同和硬件加速技术（如 GPU、TPU）是未来趋势。数字人依赖语音、视觉、动作等多模态数据协同工作，保证各模态间的语义一致性和时间同步至关重要。多模态融合算法、时序同步机制和数据标准化方案仍处于发展阶段，亟须提升模型的协同感知与实时适应能力。

人工智能的飞速发展使得大众对安全隐私保护的关注度越来越高，数字人涉及大量用户数据和交互记录，因此必须严格遵守数据安全规范。而如何设计安全可信的交互协议、数据加密方案以及用户隐私保护机制，是数字人平台必须面对的长期课题。

结语

随着生成式人工智能的快速发展，数字人作为其重要应用形态之一，正在各行业场景中广泛落地。本文从 AIGC 核心技术出发，梳理了数字人生成的关键环节与技术演进路径。通过该技术的主流应用场景，本文分析了 AIGC 数字人在其中的作用与局限，最后提出 AIGC 数字人未来在情感计算、身份拟真与自主交互等方面的发展方向。未来，数字人将在教育、传媒、政务等领域持续深入融合，成为人类认知与交互方式的重要补充。面对技术挑战与伦理问题，亟须从算法优化、制度设计、场景创新多维协同，推动 AIGC 数字人健康、可持续发展。

引用

[1] 郭全中.虚拟数字人发展的现状、关键与未来[J].新闻与写作,2022(7):56-64.

[2] 赵星,乔利利,叶鹰.元宇宙研究与应用综述[J].信息资源管理学报,2022,12(4):12-23+45.

[3] 程思琪,喻国明,杨嘉仪,等.虚拟数字人:一种体验性媒介——试析虚拟数字人的连接机制与媒介属性[J].新闻界,2022(7):12-23.

[4] 钟世镇,李华,林宗楷,等.数字化虚拟人背景和意义[J].中国基础科学,2002(6):14-18.

[5] 李楠.基于媒介场景理论的陪伴型虚拟数字人研究[D].广州:广州美术学院,2022.

[6] 谢新水.虚拟数字人的进化历程及成长困境——以“双重宇宙”为场域的分析[J].南京社会科学,2022(6):77-87+95.

