

文章编号: 1673-1646(2025)04-0057-10

颠覆与重塑: AIGC对微短剧行业的变革

郑晓君, 刘若瑾, 李毓瑄

(暨南大学新闻与传播学院, 广东广州 510632)



摘要: 人工智能生成内容(Artificial Intelligence Generated Content, AIGC)技术的快速发展正重构传统内容生产逻辑,微短剧行业成为其创新实践的前沿领域。本文聚焦AIGC技术与微短剧创作的深度融合,多模态模型技术在剧本生成、视觉设计、审美表达等全流程环节的应用范式,揭示其对创作思维革新与视听语言重构的核心价值。通过实证研究,多维剖析AIGC赋能下微短剧行业在效率跃迁、成本优化、创作民主化及用户定制化维度的突破性优势,批判性提出技术现阶段存在的生成内容同质化严重、伦理边界模糊、数据隐私风险及情感共鸣阈值不足等关键挑战。研究指出,伴随算法迭代与跨模态交互技术的演进,AIGC将驱动微短剧行业形成“智能生产-个性分发-生态协同”的新型产业架构,重塑内容消费体验与创作者生态,为数字文化产业的范式转型提供实践路径。

关键词: 人工智能生成内容; 微短剧; 创新变革; 多模态模型

中图分类号: G206 **文献标识码:** A **doi:** 10.62756/xbsk.1673-1646.2025080

引用格式: 郑晓君, 刘若瑾, 李毓瑄. 颠覆与重塑: AIGC对微短剧行业的变革[J]. 中北大学学报(社会科学版), 2025, 41(4): 57-66.

Disruption and Reinvention: AIGC's Revolutionary Approach to the Micro-Drama Industry

ZHENG Xiaojun, LIU Ruojin, LI Yuxuan

(School of Journalism and Communication, Jinan University, Guangzhou 510632, China)

Abstract: The rapid development of artificial intelligence generated content (AIGC) technology is reconfiguring the logic of traditional content production, and the micro-drama industry has become the forefront of its innovative practice. The paper focuses on the in-depth integration of AIGC technology and micro-drama creation, analyses the application paradigm of multimodal model technology in the whole process of script generation, visual design, aesthetic expression, etc., to reveal its core value to the innovation of creative thinking and the reconstruction of audiovisual language. Through empirical research, the paper analyses the breakthrough advantages of AIGC-enabled micro-drama industry in terms of efficiency leap, cost optimization, democratisation of creation and user customization, and critically puts forward the key challenges of the technology at this stage, such as the serious homogenization of generated content, the blurring of ethical boundaries, the risk of data privacy, and the insufficient threshold of emotional resonance. The study prospectively points out that along with the evolution of algorithm iteration and cross-modal interaction technology, AIGC will drive the micro-drama industry to form a new type of industrial structure of 'intelligent production - personality distribution-ecological synergy', and reshape the experience of content consumption and the ecology of the creators, providing practical paths to the transformation of the paradigm of the digital culture industry.

收稿日期: 2025-03-08

基金项目: 2024年广东省高等教育教改项目: 人工智能赋能新媒体艺术人才培养的创新与实践

作者简介: 郑晓君(1983-), 女, 副教授, 博士, 硕士生导师, 从事专业: 新媒体艺术。E-mail: hotideas@jnu.edu.cn.

Key words: artificial intelligence generated content(AIGC); micro-dramas; innovation revolution; multimodal model

人工智能生成内容(Artificial Intelligence Generated Content, AIGC)由AI制作完成,采用生成式人工智能(Generative Artificial Intelligence, 简称GAI)技术,依靠复杂的生成模型、大量计算资源生成与人类指令相匹配的内容。目前该技术处于快速发展阶段,冲击并应用于艺术、广告、教育、影视制作等诸多领域^[1-2]。

2024年2月16日,Open AI发布的首款文生视频模型Sora在全球范围内引起了广泛关注,Open AI声称Sora本质上是一个世界模拟器,代表了世界模型的发展方向^[3]。作为多模态大模型的集大成者,Sora可以根据提供的文字,生成长达一分钟的高连贯性、高保真度、高真实性视频,“中国数实融合50人论坛”智库专家洪勇在采访中表示“Sora大模型如果应用于短剧行业,就会对该行业的传统工作流程带来不小的冲击”^[4]。AIGC技术为微短剧产业带来了前所未有的创新可能性,重塑了内容创作者与观众之间的互动生态,预示着微短剧行业将迎来一个以技术创新为驱动,高效协同与多元化叙事并重的新时代。

1 AIGC微短剧创作的全流程自动化革命

AIGC是指生成式人工智能生成的内容,而生成式人工智能是基于多模态模型构建的,其中的“模态”作为感知或表达自然现象的一种方式,为人类提供了多样化的信息获取与交流方式。

生成式人工智能模型核心在于模型对数据的生成与处理。该模型一般可分为单模态模型和多模态模型,单模态模型专门处理和分析单一类型的数据信息,模型接收的输入指令与它生成的输出内容在模态上是一致的,模型从特定类型的数据中提取特征。而多模态模型能够处理和分析来自不同模态的数据,接收跨模态的指令,捕捉不同模态之间的关联和互补性,生成不同模态的结果^[5]。AIGC结合多模态数据,在不同的模态数据之间形成互动和转换,由此进行跨模态生成工作。

1.1 单模态生成模型

单模态生成模型只接收人类单一类型数据(文

字、图像、音频)的输入。其中包括“生成语言模型”,如生成式预训练转换器(Generative Pre-trained Transformer, GPT),可基于自回归解码器进行语言的生成,它的自注意力机制可以使模型捕捉到单词和单词之间的依赖关系^[6],通过上下文预测训练^[7],提高模型对自然语言的理解能力,增强生成文本的连贯性和可读性。

除了“生成语言模型”,单模态生成模型还包括“生成视觉模型”和“生成音频模型”,生成对抗网络(Generative Adversarial Network, GAN)作为一种新的机器学习模型获得较广范围的应用与改进,其中生成器部分从随机数据出发创建新的合成图像,鉴别器部分则用来判断区分真实图像与生成图像^[11]。以GAN为基础,结合卷积神经网络^①工作的深度卷积生成对抗网络(Deep Convolutional Generative Adversarial Network, DCGAN)模型展示了图像生成领域高分辨率图像超分辨率重建的潜力。Ledig等人研究使用低分辨率图像作为DCGAN的条件输入,生成器网络采用深度卷积结构,逐步生成高分辨率图像;鉴别器通过多层卷积、池化操作和非线性激活函数提取图像特征,判断输入是否为真实图像^[8]。基于GAN模型的音频波形生成对抗网络(Wave Generative Adversarial Network, WaveGAN)生成器接受从高斯分布或均匀分布中采样的随机噪声向量作为输入,从底层的高频特征逐步构建到高层的音频整体结构;鉴别器则致力于提高其区分真实音频与生成音频的能力,生成器和鉴别器两个网络迭代训练,使生成器能够生成与训练数据分布相似的高质量音频^[9]。

1.2 跨模态生成模型

跨模态生成模型作用于多模态模型基础上,是多个模态之间更为复杂的变换。李白杨等学者指出“跨模态生成的本质是文本、视觉、听觉乃至脑电等不同模态的知识融合”^[10]。跨模态生成模型可从多模态模型中提取出能够跨越不同模态的通用或共通表征,即“跨模态表征”,这些表征具备在不同模态间保持语义一致性和可迁移性的特点,使模型在一种模态下学习的知识或技能能够有效地应用于另一种模态^[11]。根据近些年计算机视觉与模式识别领域的顶级会议

^① 卷积神经网络(Convolutional Neural Network, CNN):一种深度学习模型或类似于人工神经网络的多层感知器,常用来分析视觉图像。

(Conference on Computer Vision and Pattern Recognition, CVPR)、国际计算机视觉大会(International Conference on Computer Vision, ICCV)、欧洲计算机视觉大会(European Conference on Computer Vision, ECCV)、神经信息处理系统大会(Conference on Neural Information Processing Systems, NeurIPS)、国际机器学习大会(International Conference on Machine Learning, ICML)和国际表征学习会议(International Conference on Learning Representations, ICLR)这六大顶尖学术会议的发布数据,从2022年开始,跨模态生成的论文数量呈现出上升的态势^[12],这意味着使用AIGC跨模态生成的作品数量持续上升,学术界的关注度也随之增强。这些研究成果在顶级学术平台上的出现,表明该领域的研究活动已经进入了一个活跃的阶段。

跨模态生成过程中,多种模态之间发生了相互作用,比如文本到图像的生成,2022年Ding Ming提出CogView 2^②模型,该模型引入一种新的名为跨模态通用语言模型(Cross-Modal General Language Model, CogLM^③)的预训练方法,利用预训练的CogLM生成一批低分辨率图像并进行筛选,去除质量较差的样本,再将生成的低分辨率图像通过CogLM微调得到的直接超分辨率,而后利用迭代超分辨率模块对高分辨率图像进行细化^[13]。此外,文本到视频的跨模式生成模型也在持续探索中。Imagen Video是一个基于视频扩散模型级联^④的文本生成视频系统,它继承并扩展了图像扩散模型的框架,将文本到图像的扩散过程推广至时间域。该模型的基础模型负责生成初步的低分辨率视频,后续的超分辨率模型逐步提升视频的空间分辨率和帧率,直至达到高清标准^[14]。

除了跨模式生成图像、视频等基本操作,3D形态也可跨模式生成。如Set-the-Scene模型^⑤可以实现文本生成三维场景的转换,该模型是一种基于神经辐射场^⑥的三维场景生成方法,利用全局-局部训

练框架与对象代理技术,实现了对文本提示驱动的3D场景合成的高度可控性及编辑能力^[15],用户可以对场景中每个对象的放置、外观乃至几何细节进行精确调整,保持输入文本与场景整体风格的协调一致。Zhang Longwen等学者2023年提出的DreamFace三维人脸生成模型仅需文本提示即可定制具有特定形状、纹理和动态效果的3D人脸,实现与文本描述相一致的多样且连贯的人物外观生成,该模型生成的人脸包含细腻面部特征、逼真皮肤质感及微表情变化的高保真特征,极大地促进了生成面部资产的广泛应用^[16]。

1.3 AIGC微短剧跨模态全流程创作

AIGC生成式微短剧是由多种模态模型在指令下协同工作生成的。换句话说,在AIGC微短剧的创作过程中,除了使用单一模态模型生成单一深度内容外,还涉及多模态数据的跨模态转换生成。例如《影视制作》杂志复盘了2024年3月央视推出的《AI我中华》AIGC两会全国文旅宣传片的创作过程,其中包括“文生剧本”“文生图”“图生视频”等多个跨模态的创作方式。《AI我中华》创作团队前期通过AI分析整合了中国34个省份的特色文化数据,从大量的自然历史照片中进行筛选,运用文心一言、ChatGPT等大语言模型确定视频风格并完成脚本撰写;创作中期使用Midjourney^⑦、Dalle3^⑧、ImageFX^⑨等模型制作单帧图像,利用Stable Diffusion图生图模型、PS(Photo Shop)工具进行细节优化,而后用runway、SVD等图生视频模型生成简短视频、Deforum生成图片到图片的转场效果视频,最后使用Topaz video放大视频画面;创作后期使用AE进行视频的剪辑合成并添加配音和配乐。

除了《AI我中华》外,另一个具有代表性的中国首个AIGC微短剧《中国神话》也实现了美术、分镜、视频、配音、配乐全流程的AI化,涉及“文生剧本”“文生图”“文生视频”“文生音频”“图生视频”

② CogView2: 一种基于大规模语言模型的文本图像跨模态生成模型,可根据文本描述生成高质量图像。

③ CogLM: 是一种跨模态通用语言模型,旨在通过结合不同类型的数据(如文本、图像、视频等)来进行多模态学习。它能够同时处理并理解来自多个模态的数据,进行跨模态的推理和生成。

④ 级联结构: 一种将一个机器学习问题分解为一系列子问题的设计模式,可以解决涉及异常情况或不同类别的数据问题

⑤ Set-the-Scene 模型: 通过理解图像或文本的背景信息,帮助生成更准确的输出。

⑥ 神经辐射场(Neural Radiance Fields, NeRF): 一种面向三维隐式空间建模的深度学习模型,可利用深度学习技术从多个角度的图像中提取对象的几何形状和纹理信息,生成一个连续的三维辐射场,从而呈现高度逼真的三维模型。

⑦ MidJourney: 是一个基于人工智能的图像生成工具,能够根据用户提供的文字描述生成高质量的艺术风格图像。它擅长生成具有独特视觉效果和创意的艺术作品。

⑧ DALL·E 3: 是OpenAI开发的图像生成模型,基于文本描述生成图像。它在理解复杂文字描述和生成更高质量、更符合用户需求的图像方面有显著改进,尤其擅长生成细致的细节和准确的图像内容。

⑨ ImageFX: 是一个图像编辑和生成平台,提供多种图像处理功能,如风格转换、图像增强和特效生成等。它允许用户通过调整和添加特效来修改和生成视觉内容。

五个跨模态创作,这彻底改变了传统影视行业在完成一部微短剧制作过程中所面临的一系列复杂而庞大的任务,包括构思创意、人物和故事的塑造、脚本编写、视频拍摄、后期剪辑和宣传推广等。这种颠覆了传统行业的全新创作方式反映了“消费社会理论”^⑩框架下生产与消费逻辑的融合,这不仅是一种技术革新,更是消费逻辑向文化生产领域的深度渗透。AIGC利用算法模型,自动完成剧本创作、角色设计、场景构建乃至后期制作,通过消除传统内容创作中的某些人为限制(创作时间、成本和资源限制),使得内容生产变得更加高效和可扩展。这种生产模式的变革,反映了消费社会对于效率和可获取性的追求,同时也预示着文化消费模式的转变,观众可以更快地接触到多样化的内容,在有着真实的人物、生动逼真的场景、沉浸式的互动交流、音乐特效的绚烂的微短剧里感受前所未有的感官体验。

2 AIGC微短剧的内容创新维度

AIGC微短剧是从传统的微短剧创作中汲取灵感,创造出更具创意的作品,这种巧妙地在原创和模仿之间游走的创作手法,展现了“消费社会理论”中符号创新与复制的双重特性。一方面,AI生成技术能够创造出前所未有的剧情设定和视觉效果,丰富了文化表达的多样性。另一方面,这种创新是基于对过往成功模式的学习与模仿,本质上是对既有符号系统的再生产。鲍德里亚的“拟像”概念^⑪恰到好处地描述了这一现象:AIGC微短剧是现实的模仿品,它们是高度编码和体系化的产物,满足了消费社会对新鲜感和差异性的渴望。

2.1 融合创新:AIGC与人类智慧的碰撞火花

与人类大脑相比,AIGC在学习上展现出了明显的优越性。依赖于先进的算法技术,AI有能力在较短的时间内对大量样本数据进行消化和分析,通过学习大量数据(绘画、音乐、文本等),自动识别并建立起作品中元素(颜色、线条、音符、词汇等)之间的复杂关系与模式。生成新作品时,AI算法能够依据学习到的关联规则,创造出新的元素组合和联结方式,

形成创意作品,这个过程体现了“创意金字塔”^⑫中“不断发掘新的关系与联结”的核心理念。Anantrasirichai和Bull曾指出:“样本数据的质量、数量以及标签准确性对人工智能性能具有显著影响。在创意行业中,机器学习驱动的人工智能算法依赖大量的训练数据,尤其是图像和视频数据,以便准确地学习和提取语义特征。”^[1]例如,在内容创作、信息分析、内容增强、信息提取和数据压缩等领域,高质量且多样化的数据集能够有效地提升AI模型的表现力和泛化能力,这很大程度上避免了潜在错误,加快了工作效率。AIGC创作是基于其数据库的信息储备,而训练数据库则是优化机器学习流程性能的关键工作。微短剧在制作前期,AIGC工具会获取大量的影视数据样本,然后进行不同故事、场景、人物的分析,经过人类的干预不断训练,从而为创作新剧本提供参考。AIGC技术系统拥有生产独特创意成果的能力和优势,这对传统微短剧行业的简单的人工创作内容产生了冲击。

但仅仅使用AIGC技术创作的作品会在一定程度上缺乏艺术家或是创作者的主观性和创造力,由于支撑AIGC的开源软件基本工作原理是既定的,无法根据创作者的主观思想进行改变,缺少了创作者的思想和灵魂,AIGC的创作就成了机械地重复堆叠,显然,完全依靠AIGC是无法完全满足创作者的需求的。在《中国神话》的创作初期,使用AIGC制作的画面更倾向于欧美的审美风格,这与创作者想象中具有东方特色的神话角色和场景有显著的不同。对此,艺术创作者需要进行上百次的提示词打磨和调整,引导AI准确地理解中华传统文化的词汇和创作者的意图,从而最大程度地提升角色的风格形象和情感表达。此外,微短剧创作者和艺术家还可尝试设计定制化AI生成模版,提供包含艺术家过往作品和风格元素的数据集,对AIGC模型进行再训练,使其更精准地捕捉艺术家的创作习惯和偏好,更好地将艺术家和创作者的创意思想融入到生成结果中。

2.2 视觉精进:微短剧创作多样化风格

人工智能对于不同风格的数据学习创新能力非常强,它可以识别未被标记的数据,学习其常见结构

^⑩ 消费社会理论:法国哲学家让·鲍德里亚在《消费社会》中提出,揭示了现代社会中消费已不再局限于满足基本生活需求,而是转变为一种符号价值的追求和身份构建的手段。

^⑪ 拟像:法国哲学家让·鲍德里亚提出,拟像不再是对某个领域、某种指涉对象或某种实体的模拟。它无需原物或者实体,而是通过模型来生产真实,这种真实被鲍德里亚称为“超真实”。

^⑫ 创意金字塔:在赖声川撰写的《赖声川的创意学》中提出,强调创意的核心在于事物之间的联结,即不断发掘新的关系与联结。

和视觉上一致的风格集群，捕捉有用的视觉关联，然后对重复性强的部分进行知识的组织和处理，生成新的内容^[17]。随着各种各样风格数据读取和分析，人工智能可以生成多样化视觉风格，具有丰富的视觉色彩和艺术价值。比如图像风格迁移，即某个图像的风格属性转移到另一个图像上，详见图 1。该技术可用于艺术创作、内容增强、后期制作流程、信息提取与增强、数据压缩等场景。在艺术创作中，AI可以根据用

户需求快速生成多种风格的艺术作品，同时，对于实时视频流也可以实时完成风格迁移，极大提高了创作效率和可能性。这也进一步验证了深度学习技术在图像处理领域中对于特征表达、统计匹配和灵活变换的有效性^[18]。这种 AI 的高效生产模式确保了内容的快速迭代，适应了消费社会中不断变化的市场需求，实现了即时消费的可能。



图 1 风格转移举例

机器善于从视觉数据库中学习，而这些视觉数据库都代表着人类的视觉经验。生成式 AI 在模仿既有视觉风格和结构的基础上，通过算法的变异、随机性或对抗训练等方式，产生不同于已有作品的新颖形态。尽管 AI 可能在初期更擅长模仿已知风格，但随着技术进步，如 CAN[®] 等算法能够在模仿基础上产生显著的原创性风格，展现出“从有生有”和“从无生有”的双重创造性风格，这与“创意金字塔”中追求不断发现和创造新联结的精神相吻合。可见，AIGC 可以实时自主理解艺术风格和设计，凭借其强大的学习吸收和创造能力，生成多样化风格的内容，从而达到更加精细丰富的视觉效果。这种基于 AIGC 创造新风格的创作方法也同样适用于微短剧行业，技术可以根据某些已存在的风格，创造出类似的场景、人物和视频风格。这种方法大大减少了创作者在选择和绘制不同风格的人物场景时的工作量，其内容的冲击力甚至超过了手工绘制的视觉效果，AIGC 为微短剧行业提供更生动、多样化的创作平台。

2.3 审美迎合与创意实现：AI 拟人化艺术的精妙构建

视觉艺术家 Hannu Töyrylä 曾在开发创作艺术

作品的 AI 工具时表示“技术创作内容与人类的审美偏好之间存在着一场拉锯战，技术创作要与人类的审美迎合，才会被人类所接受”^[19]。一般而言，审美趋势会受到客体属性的影响，人类对 AI 创作的情感和接受度也会因 AI 外观拟人化程度而改变，AI 创作的真人化风格作品更易激发用户的正向情绪^[20]。从技术角度看，AI 模型的拟真能力越强，作品在视觉和听觉层面的真人化程度越高，理论上更容易获得较高的感知接受度。因此，人类对二维风格接受度较低，对真人风格接受度较高。《中国神话》中神农氏尝百草、女娲补天、嫦娥奔月等情景均由 AI 自动情景生成，其形象与真人写实风格更为匹配，形象拟人化，极大程度上反映和迎合了当代观众的审美趋势和偏好，提升了用户接受度。

除了拟人化的审美趋向，AI 生成微短剧角色的情感表达也尤为重要，有微表情变化、语音语调变化等动态表达的角色会带给用户更强的视觉和听觉体验。如 Crazytalk 技术[®]可以追踪角色面部动画和表情，这项技术的核心在于利用先进的面部识别算法捕捉和模拟真实人脸的表情变化，然后通过口形同步技术匹配音频或文本输入，使角色的嘴唇动作与语音内容精确对应。此外，CrazyTalk 还允许用户通过简单的

^⑬ CAN (Conditional Adversarial Networks) 是一种生成对抗网络 (GAN) 的变体，通过引入条件信息 (如标签、文本或图像特征) 来指导生成过程，使得生成的结果可以根据这些条件进行控制。它主要用于根据给定的条件生成特定的图像或进行图像转换。

^⑭ CrazyTalk: 一款基于人工智能和计算机视觉技术研发的软件，主要用于将静态照片转化为能够进行实时面部动画和语音交互的 3D 或 2D 动态角色。

操作调整角色表情、头部运动以及身体动作,从而创建出如同真人般的对话效果。另有FOMM^⑤运动模型接受一个静止的目标图像(如一个人脸的照片)和一个包含动作信息的参考视频(如另一个人讲话时的视频)。模型通过学习,能够把参考视频中的动作特征提取出来,并应用到目标图像上,这样原本静止的目标图像就仿佛动了起来,比如做出讲话、微笑或者其他复杂的面部表情变化。这种技术在电影特效、游戏角色动画制作和社交应用等方面有着广泛的应用潜力。可见,人工智能的创作内容不仅要满足拟人化的审美标准,还需确保生成的角色能够有效地传达情感,使角色更具亲和力。因此,AIGC微短剧只有满足现代观众的审美需求时,才能与观众在感官体验和互动交流中实现审美共振。

3 AIGC微短剧的优势

3.1 效率提升:AI全流程制作的时代表红利

人工智能技术在提升创作效率方面具有显著的价值,众多学者关注人工智能技术应用于视频制作领域的最新进展。AI通过自动文本生成、智能场景生成、AI运动控制技术、AI语音生成和深度伪造等手段,几乎渗透到视频生产的每一个环节,如文本内容生成、拍摄材料收集与组织、非线性编辑、动画插入、视频特效、视觉包装、字幕添加等。传统的微短剧行业中,团队在创作初期需要投入大量的时间和精力进行资料的筛选、调查和分析,以便从中找到一个微小的创意灵感。有时为了创作出一个令人满意的剧情脚本,甚至需要进行二次的返工和复调查,但在AIGC技术的帮助与驱动下,系统可以自动分析不同模态的资料,将初始文本材料转换为不同模态形式,不仅可以生成具有丰富文本内容的故事,还可以直接从文本生成视频。

在传统微视频制作的中后期阶段,表演拍摄和合成剪辑是一项相当大的工程任务。然而,AI技术能够将原先需要进行后期处理的工作整合到生产执行阶段,从而使得视频制作能够在执行阶段基本完成。这不仅简化了传统微短剧的制作流程,还避免了后期制作的需求,详见图2^[5]。中央广播电视台制作的文生视频动画《千秋诗颂》主创成员陈格平在接受采访时透露,与传统的制作周期1个月制作1集视频不同,AI制作6集的过程仅需2个月

的时间^[21]。经验表明,使用AIGC进行视频创作可以优化工作流程,从根本上重塑视频生产过程,提升了生产效率,缩短了制作周期。

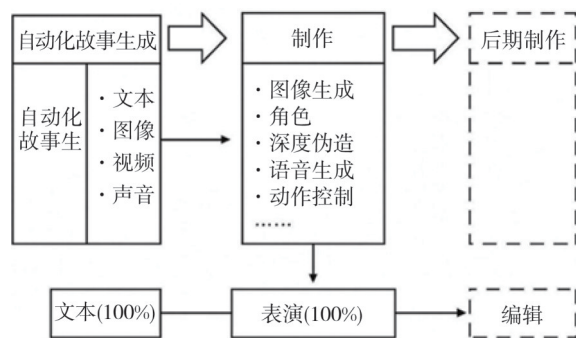


图2 AI视频制作流程

3.2 经济优势:前所未有的高效成本管控

2023年初,一部传统的微短剧的制作成本大约在20万~30万元之间,但2024年传统微短剧制作成本已经显著增加,达到了原来的两到三倍^[22]。AI可以利用大模型帮助艺术家快速创造视觉背景和环境,缩短手工绘制的时间,减少试错成本。AI在视频分析、制作与流媒体传输中可以简化编码过程,减少计算资源的消耗,同时也降低硬件设备投入和能源消耗的成本;此外,AI编辑工具具有够快速编辑视频的能力,从而减少人工干预需求,降低人力资源成本;AI能够模拟复杂场景和特效,可以替代部分昂贵的实景拍摄和后期特效制作,节省真实拍摄场景的搭建、外景转场的时间和成本^[23]。从整体上看,人工智能在视频制作过程中,通过全方位的自动化、智能化流程优化,精准内容管理与高效分发,实现了显著的成本节约,展示了高性价比的技术优势,这符合文化产业对快速迭代和成本控制的内在要求,与消费社会中对经济性和实用性的追求相契合。

此外,AIGC技术在降低经济成本的同时,也为小众文化的传播创造了条件。过去受限于高昂制作成本的小众艺术或地域性文化内容,现在也可通过AI技术以更低的成本进行制作和推广,这不仅丰富了消费市场的多样性,也让消费者能在海量同质化内容中找到与自己文化认同相契合的产品,反映了“消费社会理论”提倡的文化的多样性和包容性。从消费社会理论的角度来看,AIGC技术不仅体现了经济的理性和实用主义,而且也是推动

^⑤ FOMM: 全称为First Order Motion Model,是一种用于图像动画化(Image Animation)的人工智能技术,它是基于生成对抗网络(GANs)的一种变体实现。

人文关怀和文化多样性的重要力量。

3.3 用户生成：视频创作的民主化浪潮

Web 1.0 时代是专业生成内容时代，用户单项使用互联网，不分享提供自己的内容，互联网中的信息主要由相关领域的专业人士创造。随着时代发展，Web 2.0 时代，用户生成内容逐渐取代专业生成内容，用户纷纷在社交媒体上生成内容信息分享、评论。而现在 Web 2.0 到 Web 3.0 的过渡时期，AIGC 成为权衡专业生成内容和用户生成内容的工具，利用 AIGC，任何人都可以创建属于自己的作品。用户在利用 AIGC 创作微短剧时只需输入文本，就能得到与之匹配的智能图像和视频，人们无论是想恢复过去的影像，还是想创造未来的场景，都可以在 AI 生成技术的帮助下完成。这种自创自用的方式使得内容创作者和消费者之间的界限变得模糊。现阶段 AI 生成技术快速进步，特别是深度学习、神经网络等底层算法已获得优化提升，使计算机能够更高效地理解和生成复杂的人类情感表达，包括面部表情、语音语调等。这大大降低了内容生成的复杂度和技术难度，非专业人员也能便捷地利用 AI 生成工具创作高质量的个性化内容。伴随着各类 AI 生成服务平台和工具的商业化趋势，大多 AI 生成平台提供了用户友好的界面和丰富的预设选项，用户无需具备深厚的技术背景即可轻松使用 AI 生成技术。这种用户友好的设计不仅降低了用户参与 AI 生成过程的学习成本和操作门槛，而且有效激发出用户的探索性创作欲望，每个人都有机会成为自己故事的讲述者和传播者。AI 创作广泛渗透至大众领域，引发了艺术与科技交融的公众参与热潮，同时也突出了消费社会中个体创造力的价值，以及人们对于平等参与文化消费的渴望。

3.4 定制化创作：打造专属的用户标签

普通用户借助 AI 生成工具参与到内容创作中，生成包括社交媒体帖子、短视频、艺术作品、营销材料等多种形式的个性化内容，这反映了用户对 AI 生成技术的主动应用与个性化需求。由于 AIGC 微短剧不受传统剧集制作的限制，因此 AI 能够根据不同受众的各种兴趣喜好生成各种题材和各种风格的视频内容，从而满足受众群体多样化的需求，增强用户的观感体验感。人工智能生成定制化视频需经过大规模数据集的训练，如第三代通用和预训练转换器(General Pre-trained Transformer 3, GPT3)生成模型，需涵盖各类视频元素、视觉风

格、叙事结构等内容的训练，确保其具备生成符合用户要求视频素材的能力。此外 AI 生成对抗网络技术可以生成高质量、逼真的图像或音频内容，创造独特个人特色的视觉风格或声音效果；AI 可以利用计算机视觉技术分析图像、视频内容，生成个性化的视觉元素，定制个性化的角色设计、场景渲染、动画制作^[24]。AIGC 具备基于用户精确需求定制微短剧视频内容的能力，借助大数据算法分析给用户标记兴趣需求属性标签，以确保产出高度契合目标受众的偏好和期待，提升用户在微短剧制作过程中的参与度，打破传统微短剧行业视频是由专业人士制作的模式。AIGC 通过算法理解个体消费者的细微差别，创造出能够满足特定群体需求的微短剧，增强消费者与内容创作之间的情感联结和个性化体验。AI 个性化创作模式下，内容不仅仅是消费品，更是消费者表达自我和内在欲望、构建身份的符号工具，体现了消费社会中对个性化和定制化的追求。

4 AIGC 技术的争议与挑战

4.1 质量与控制：AI 创作的技术瓶颈

目前，AI 技术生成的图像真实度还存在很大的局限性。通过对生成图像和真实照片的组合评估，有学者发现客观上，现有最先进的图像质量模型和特征并不能直接应用于预测哪些 AI 生成器被使用或区分生成图像与真实图像。为了进一步分析生成图像的吸引力，他们设计了一项在线主观测试，该测试涉及对图像吸引力、真实感以及图像与文本提示匹配程度的评分。结果显示，目前可用的客观模型和特征与图像吸引力和真实感的相关性较低。因此如何提升图像视频的真实感和吸引力是未来工作中需要重点考虑的问题^[25]。同样，由于目前 AIGC 的技术算法还不够成熟，在微短剧制作中无法彻底解决模拟真人对话口型对不上的问题，很多时候 AIGC 创造的人物脸部也存在不够自然的情况，如《中国神话》一经推出，就有网友发表“感觉人物呆滞”“没有灵魂和感情”“人的画面很尴尬”“嘴不自然”等评论，这些都需要技术算法的不断优化与学习，以此提高生成内容的质量与效果，但算法优化需要不断地尝试训练和调整参数，是一个复杂长期的过程。

除了生成内容缺失真实感，AI 技术的自身局限性也让其生成内容缺乏独特的创意，同质化问题

严重,大多作品呈现相似性,使观众感觉似曾相识甚至千篇一律;画质低、画面连贯性不强等不足也影响了观众的观赏体验。AI机器不是人,没有人的思想和感情,生成的内容缺乏文化内涵,这就可能导致观众在古色古香的场景中看到主角拿手机的不合时宜的画面,一定程度上影响了文化的传播。尽管未来AIGC在微短剧行业的应用会越来越多,但内容的价值体现在人身上,而非机器身上,以人为本,技术为辅才能生成高质量的作品。

4.2 伦理与法律:版权与创作者身份的新问询

由于AIGC强大的记忆能力,大模型在处理数据信息时可能会直接复制集中训练的图像,或会出现未经他人授权而使用他人数据和素材的情况,从而引发侵犯版权或版权纠纷等问题,这也加剧了制作AIGC微短剧的法律风险。此外,AIGC制作微短剧通常是在预先存在的数据和创意基础上将元素混合生成内容,这就导致AIGC是否为独创内容的界限非常模糊,而AIGC模型会在创作中生成大量的内容,因此此时每件作品是否侵犯版权也变得困难,这引发了人类对谁拥有AIGC生成内容的知识产权的担忧,到底谁才是内容的创作者引发讨论。

生成式AI并非机器独立自主地“创作”艺术作品,而是通过与人类的交互过程共同形成新作品。在这种情况下,理解并界定AI在创作过程中扮演的角色及其贡献,以及相应的版权归属和法律责任至关重要。我们可以把这种人机混合形成的作品理解成是人类作者和技术共同创作下完成的,而AIGC创作也只是一个新趋势,对于谁是创作、版权归谁本质上都没有改变。鼓励AIGC平台和用户使用已经获得授权的素材库,或者使用公共领域和知识共享许可的素材,也可以有效降低版权风险。AI生成的伦理问题引发人们思考,但归根结底,AIGC只不过是人类用于创作的工具,我们不应夸大人工智能的独创能力,应将注意力转移到如何利用人的思想创造更具创新性的东西。微短剧行业的长久发展需以人的思想创意为核心,而技术的加持只不过让更多更有意思的人的想法成为了现实。

4.3 隐私与安全:AIGC技术的双刃剑

AIGC借助大模型生成创意内容,同时也通过大模型处理用户提供的数据信息,通常情况下,AIGC模型要求用户提供多模态的数据输入(图像、

文件等),而一些涉及隐私、机密等信息会在大模型一次又一次的训练中所记住,例如人工智能聊天机器人(Chat Generative Pre-trained Transformer, ChatGPT)可以记住与用户的对话及用户输入的所有信息,其过于透明性的技术加大了隐私数据被泄露的风险。而用户制作AIGC微短剧想要获得高质量的生成内容,就不可避免的需要将大量多样的样本信息输入进去,在跨模态技术的交互下,机器根据数据预测用户的偏好并绘制他们的个人资料,这增加了机器滥用用户数据的可能性和风险,而如何缓解这类隐私风险问题,在效用与隐私之间实现权衡仍有待人类发掘研究。

此外,AIGC的快速发展也使得制造和传播虚假信息变得越来越容易,其“深度伪造”技术可以生成高度逼真的微短剧视频内容,对人们进行错误或虚假信息的欺骗,甚至会出现身份盗窃、网络欺凌、人工智能诈骗、知识产权盗窃等严重安全问题。在互联网发达时代,AIGC生成的虚假微短剧一旦传播,就可能对公众舆论产生影响,操控虚拟社交群体,影响社会的安全和谐。因此如何利用好AIGC这把双刃剑重塑传统微短剧行业是人们需要谨慎思考的问题。随着AI在多模态领域内实现了飞跃式发展,相关产业及行业更需要在变革中保护用户的隐私与信息安全,才能为AIGC技术的应用与发展提供更加坚实可靠的平台。

4.4 AI数字人:用户的恐怖谷效应

AI生成式作品模拟真人化程度逐步提高,但如果作品过于真实,可能会引起观众的反感和排斥,这种现象被称为“恐怖谷效应”,该效应描述的是当非人类实体(如机器人或动画人物)与人类相似度极高时,人类对其产生的感情反应可能会经历一个先升后降的过程,到达类似程度的峰值后迅速跌落,然后再随着相似度进一步提高而逐渐恢复好感的现象^[20]。这是因为当相似度达到一定程度但还未完全达到无法区分的程度时,细微的差异会被放大,使得人们更容易察觉到非人类特征,从而导致情感上的排斥。例如,京东创始人刘强东的AI数字人在直播活动中展示了人工智能在模拟真实人物方面的能力,但同时也暴露了AI技术面临的“恐怖谷效应”挑战。对于AI刘强东这样的数字人而言,尽管采用了先进的AI驱动的大姿态动作捕捉和合成技术,但在实际应用中可能会出现表情、动作、语言交流等方面的不自然或瑕疵,触发观众

的恐怖谷效应。即使技术进步能够实现高度逼真的视觉效果,数字人依然可能存在难以跨越的情感联结和深度理解方面的鸿沟。因此,在创造 AI 数字人时,开发者需要特别注意调整其表现细节,确保既不过于机械引起疏离感,也不至于仿真以至于触及恐怖谷效应的敏感区域。当实体越来越接近真人,直至达到难以区分的地步时,人们的好感度理论上会再次升高,这是因为高度拟真的实体能够更好地激发人类的共情反应和社交互动本能。然而,实际的技术挑战在于要跨越恐怖谷并达到完全无感知差异的状态是非常困难的。即便在现今技术下,人工智能和机器人技术已经在很大程度上缩小了与真实人类之间的差距,但仍存在一些微妙的细节和动态表现上的不足,使得即使在高相似度的情况下仍可能出现令人不适的感觉。

因此,“谷”的出口并不明确,而是随着技术和艺术表现力的进步而逐渐显现。科学家、艺术家和技术开发者不断尝试优化算法、材料和表现手法,力求创造出能够无缝融入人类社会并触发正面情感反应的高保真度人造实体。同时,公众对人工智能和仿生技术的认知和接纳程度也在随着时间和环境的变化而演变。在未来一段时间内,AI 最有可能发挥的作用将是增强人类的创造力,而非取代人类在创新过程中的核心地位。

5 AIGC 微短剧的未来展望

AIGC 作为一个持续演变的竞赛领域,在微短剧产业的崛起和竞争日益激烈的背景下,技术的支持变得尤为关键,使得 AIGC 在国际市场上迅速取得了主导地位。美国的谷歌、微软,英国 Stability AI、Open AI 等科技公司纷纷在 AIGC 领域抢占布局,利用 AI 文本生成视频,为实现 AIGC 大模型高质量发展倾注了大量的资源;中国市场包括爱奇艺、优酷、腾讯视频在内的部分大型企业和平台已经表示,他们将加强技术内容的制作,提高微短剧创作的效率。虽然中国在生成式大模型的应用上取得了飞速的进展,但与英美等国相比,其在生态建设和商业领域仍存在巨大的成长潜力。

AIGC 微短剧正以前所未有的态势颠覆传统微短剧行业,开启一个创作全流程自动化的新纪元。这场革命的核心驱动力源自单模态与跨模态生成模型的创新应用,以及在内容创新、效率提升、经济优势、用户参与等方面的全面变革。AIGC 跨模态生成微短

剧模式彻底改变了传统微短剧依赖人工逐层创作、多环节协作的繁琐流程,微短剧制作经历了从传统的复杂手工制作到简单的技术辅助,再到 AI 全流程自动化制作的转变,这极大地提升了内容产出的效率与一致性,降低了制作微短剧的成本。同时,AIGC 微短剧通过 AI 与人类智慧的深度融合,实现剧本创意的高效生成与精准定制,打破传统创作瓶颈,拓宽题材边界,强化个性化体验,并借助智能制作技术革新内容呈现形式,在内容创新维度上颠覆了微短剧行业的传统运作模式,为微短剧行业的未来发展指明了更为智能和多样化的方向。未来 AI 将赋能微短剧行业构建智能化、个性化、高效化的全新生态,推动行业从内容智能生产到个性化分发、消费生态协同的全链条升级。我们期待看到一个由 AIGC 引领的微短剧黄金时代,它将深度融合科技与艺术,重塑内容产业格局,为广大观众带来无尽的创意惊喜与沉浸式消费体验。

参考文献

- [1] ANANTRASIRICHAI N, BULL D. Artificial intelligence in the creative industries: A review[J]. *Artificial Intelligence Review*, 2022, 55(1): 589-656.
- [2] HUANG Y F, LV S J, TSENG K K, et al. Recent advances in artificial intelligence for video production system[J]. *Enterprise Information System*, 2023, 17(11): 4618-4636.
- [3] WANG F Y, MIAO Q, LI L, et al. When does Sora show: The beginning of TAO to imaginative intelligence and scenarios engineering[J]. *IEEE/CAA Journal of Automatica Sinica*, 2024, 11(4): 809-815.
- [4] 李静. 并购短剧公司微盟借力 AI 大模型入局短剧赛道[N]. *中国经营报*, 2024-03-18(D04).
- [5] ARSLAN A K. Exploring challenges in applying foundation and generative models in AI[J]. *Journal of Research in Engineering and Computer Sciences*, 2023, 1(2): 01-09.
- [6] LIANG C, DU H, SUN Y, et al. Generative AI-driven semantic communication networks: Architecture, technologies and applications[EB/OL]. 2023-12-30[2025-04-11]. <https://arxiv.org/abs/2401.00124>.
- [7] CAO Y, LI S, LIU Y, et al. A comprehensive survey of AI-generated content (AIGC): A history of generative AI from GAN to ChatGPT[EB/OL]. 2023-03-07[2025-04-11]. <https://arxiv.org/abs/2303.04226>.
- [8] LEDIG C, THEIS L, HUSZÁF F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//IEEE Computer Society. *Proceedings of the*

- IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), Honolulu, USA, 2017. Los Alamitos, CA: IEEE, 2017.
- [9] DONAHUE C, MCAULEY J, PUCKETTE M. Adversarial audio synthesis [EB/OL]. 2018-02-12 [2025-04-11]. <https://arxiv.org/abs/1802.04208>.
- [10] 李白杨, 白云, 詹希旒, 等. 人工智能生成内容(AIGC)的技术特征与形态演进[J]. 图书情报知识, 2023, 40(1): 66-74.
- [11] VAN TULDER G, DE BRUIJNE M. Learning cross-modality representations from multi-modal images[J]. IEEE Transactions on Medical Imaging, 2018, 38(2): 638-648.
- [12] FOO L G, RAHMANI H, LIU J. AI-generated content (AIGC) for various data modalities: A survey[EB/OL]. 2023-08-27 [2025-04-11]. <https://arxiv.org/abs/2308.14177>.
- [13] DING M, ZHENG W, HONG W, et al. CogView2: Faster and better text-to-image generation via hierarchical transformers[J]. Advances in Neural Information Processing Systems, 2022, 35: 16890 - 16902.
- [14] HO J, CHAN W, SAHARIA C, et al. Imagen video: High definition video generation with diffusion models [EB/OL]. 2022-10-04 [2025-04-11]. <https://arxiv.org/abs/2210.02303>.
- [15] COHEN-BAR D, RICHARDSON E, METZER G, et al. Set-the-scene: Global-local training for generating controllable NeRF scenes [C]//IEEE/CVF. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV 2023), Paris, France, 2023. Los Alamitos, CA: IEEE, 2023.
- [16] ZHANG L, QIU Q, LIN H, et al. DreamFace: Progressive generation of animatable 3D faces under text guidance [EB/OL]. 2023-04-06 [2025-04-11]. <https://arxiv.org/abs/2304.03117>.
- [17] GRIEBEL M, FLATH C, FRIESIKE S. Augmented creativity: Leveraging artificial intelligence for idea generation in the creative sphere[EB/OL]. 2020-06-10 [2025-04-11]. https://aisel.aisnet.org/ecis2020_rip/77.
- [18] HUANG X, BELONGIE S. Arbitrary style transfer in real-time with adaptive instance normalization[C]//IEEE. Proceedings of the IEEE International Conference on Computer Vision (ICCV 2017), Venice, Italy, 2017. Los Alamitos, CA: IEEE, 2017.
- [19] CHOI S K, DIPAOLA S, TÖYRYLÄ H. Artistic style meets artificial intelligence[EB/OL]. 2021-04-02 [2025-04-11]. <https://doi.org/10.2352/J.Percept.Imaging.2021.4.2.020501>.
- [20] 宋章通, 朱永琼, 方浩. AI化身形象风格的审美偏好: 情境因素的调节作用[J]. 包装工程, 2024, 45(10): 332-338.
- [21] 李丹. 首部中国原创文生视频AI系列动画《千秋诗颂》播出: 一场中国诗词与人工智能的“双向奔赴”[J]. 影视制作, 2024, 30(3): 24-28.
- [22] 陈汉辞. AI造剧, 你会追吗?[N]. 第一财经日报, 2024-04-02(A01).
- [23] JAYANTHILADEVI A, RAJ A G, NARMADHA R, et al. AI in video analysis, production and streaming delivery [C]//Journal of Physics: Conference Series. IOP Publishing, 2020, 1712(1): 1-6.
- [24] BURLACU C. The impact of AI-powered content generation on customer experience [EB/OL]. 2023-07-01 [2025-04-11]. <https://essay.utwente.nl/95515/>.
- [25] GÖRING S, RAO R R R, MERTEN R, RAAKE A. Analysis of appeal for realistic AI-generated photos[J]. Ieee Access, 2023, 11: 38999-39012.
- [26] MORI M. The uncanny valley[C]//MacDorman K F, Minato T. Proceedings of the CogSci-2005 Workshop: Toward Social Mechanisms of Android Science. Tokyo: Energy, 1970.