

技术赋能下生成式人工智能影像的创作实践与平衡逻辑探析

徐睿

(甘南州融媒体中心, 甘肃甘南 747000)

【摘要】随着人工智能技术为支撑的数智媒介的出现,生成式人工智能技术作为人工智能技术的一种前沿分支,亦在近年取得突飞猛进式的发展,并逐渐渗透到人类生活的方方面面。其中,生成式人工智能影像作为数字艺术的全新形态,其媒介属性突破了传统的工具性边界,走向人机协同的协作式媒介新阶段。在技术赋能之下,如何推动艺术创作向更符合时代需求的方向发展,效率与创意平衡逻辑,就成为了数字时代艺术创作坚守的核心所在。

【关键词】生成式人工智能影像;技术;效率与创意;媒介

An Analysis of the Creative Practice and Balancing Logic of Generative Homo Sapiens Artificial Intelligence Imaging Empowered by Technology

Xu Rui

(Gannan Prefecture Media Convergence Center, Gannan, Gansu 747000)

[Abstract]With the emergence of digital-intelligent media supported by Homo sapiens artificial intelligence technology, generative Homo sapiens artificial intelligence technology, as a cutting-edge branch of Homo sapiens artificial intelligence technology, has also achieved rapid development in recent years and gradually permeated all aspects of Homo sapiens life. Among these, generative Homo sapiens artificial intelligence imagery, as a new form of digital art, has transcended the traditional instrumental boundaries of media attributes, entering a new phase of collaborative media characterized by Homo sapiens-machine synergy. Under technological empowerment, how to propel artistic creation toward a direction more aligned with contemporary demands, balancing efficiency and creativity, has become the core tenet of artistic creation in the digital era.

[Key words]generative Homo sapiens artificial intelligence imagery; technology; efficiency and creativity; media

作者简介:徐睿,女,甘南州融媒体中心编辑,研究方向为人工智能影像。

基金项目:本文系2025年度甘肃省人文社会科学项目《“丝路记忆·AI重构”:基于生成式AI的甘肃文旅影像传播创新研究》(项目编号:25ZC04)阶段性研究成果。

技术创新作为驱动艺术媒介形态演进与视觉影像变革的核心先导力量，从报纸到广播、电影、电视、互联网等媒介，每一种新型媒介的诞生，都在不同维度上革新着人类的信息传播模式与知识生产路径。生成式人工智能的崛起进一步加速了这一迭代进程，人类个体通过与智能媒介的交互完成自我表达和内容生产，这种创作模式极大地颠覆了传统的创作范式，也赋能个体创作者天马行空的想象力落地生根。在技术赋能之下，正确认识生成式人工智能影像的媒介属性，充分发挥主观能动性的同时，拥抱技术创新带来的数字艺术创意机制和范式的革新。

一、生成式人工智能影像成为协作式媒介

艺术媒介是艺术创作的物质载体与思维工具，其形态与特征决定了艺术的表达形式与认知方式。麦克卢汉提出“媒介即讯息”，指出媒介的创新不仅是技术手段的变化，更是人类感知方式与社会关系的重构。生成式人工智能影像作为全新的艺术媒介，突破了传统媒介的工具性定位，从创作者单方面操控的工具，转变为人机双向互动、共同完成创作的协作式媒介。

最初“人工智能”这一概念在 1956 年举办的达特茅斯会议中被首次提出，人工智能的发展经历了三个阶段，第一阶段始于 20 世纪 50 年代，感知器数学模型从训练样本中自动学习完成分类任务。第二阶段始于 20 世纪 80 年代，算法被提出并用于参数计算。第三阶段始于 21 世纪初期，深度学习的能力促进了人工智能在视觉、语音等领域的突破性发展。^①而生成式人工智能就是基于深度学习技术自主生成文字、图像、音视频等数字内容的技术模型。^②2022 年末，Chat GPT 的推出标志着文本自动生成领域的标志性突破，基于大语言模型的语言处理能力，实现了高连贯性且具有高逻辑性的文本生成，掀起全球生成式人工智能的热潮，也推动了全民使用生成式人工智能工具进行交流对话的热情高涨。紧接着，Midjourney 等图片生成工具的问世，不仅能够精准响应用户输入的文本指令，并且可以生成高质量的图像内容，也由此验证了生成式人工智能在视觉创作领域的巨大潜力。2023 年，生成式人工智能工具呈现爆炸式的增长，Runway、Stable Diffusion 等视频生成工具粉墨登场，验证了多模态融合技术的可行性。2024 年初，Open AI 发布文生视频大模型 Sora，实现长达 60 秒的高清视频生成，标志着生成式人工智能在影像领域达到了新高度。随着多模态融合技术进一步成熟，使得文字、图像、视频等模态可实现高效转换与协同生成，生成式人工智能开始深度渗透影视、动画等多领域，形成人机协同的创作模式。与此同时，2024 年 4 月，中国发布了首个视频大模型 Vidu，这款直接对标 Sora 的文生视频模型，能够一次生成 16 秒的片段。随后如可灵、即梦等国内的视频模型生成工具也相继推出。随着生成式人工智能影像与电影艺术的融合愈发紧密，越来越多的创作者将生成式人工智能影像用于短视频的创作当中，完成着在现实拍摄中无法实现的创意表达。丰富了内容表现力的同时，也带来一场视觉享受盛宴。

①赵毅衡：《论“伴随文本”——扩展“文本间性”的一种方式》，《文艺理论研究》，2010 年第 2 期。

②李翔、俞智星：《生成式人工智能赋能电影内容生产与传播：概念演变、技术影响及创新应用》，《电影文学》，2024 年第 10 期。

“生成式人工智能影像作为‘生成艺术’，是借助于一套算法程序、自然语言模型而进行的艺术创作实践。”^①展开来说，生成式人工智能影像以深度学习大语言模型为核心，通过文本、图像、语音等多模态输入，由人工智能算法自主生成或辅助而生成的视觉影像作品。所以生成式人工智能影像并不只是单纯的工具，而是具有生成性、跨模态、协创性三大属性的协作式媒介。

与传统数字艺术媒介的编辑或者修饰的功能不同，生成式人工智能影像的核心功能是基于智能算法生成的，算法的训练逻辑、生成机制直接决定媒介的视觉表达特征，其背后数以万计的大数据，支撑生成式人工智能工具实现视觉影像的整体建构，这也使新兴媒介完成了从内容优化的工具到内容生产载体的功能性跃迁。同时数据的不断投喂和补充，也使得生成式人工智能工具同步更新，生成式人工智能工具可对生成的结果进行持续的迭代优化，为艺术创作提供了更多元更立体的可能性。^②

“跨模态是指图片、音频、视频等数据能被理解或转换为语义相似的其他模态数据。”^③也就是说，通过在生成式人工智能影像工具输入文字指令、上传静态图像等多模态内容，完成向视觉影像模态的语义转译、特征融合与创造性生成，这一过程的核心就是打破不同信息模态之间的壁垒，使生成式人工智能影像工具能够接收、解析多模态输入的信息，并将其转化为具有视觉完整性、创意表达性的影像成果。

在传统艺术创作中，创作者与媒介的互动是一种单向的操控，人对工具拥有绝对的主导，媒介是无法对创作者的指令做出任何反馈或建议的。但生成式人工智能影像的创作中，创作者与媒介之间是双向的协作，人类输入指令，生成式人工智能影像工具将多种创意方案反馈给创作者，创作者根据再进行方案筛选并调整指令，再将新的要求传递给媒介，如此循环往复，直到输出创作者满意的最终内容。人工智能生成影像的创作体现为一种人机共创式的“组装”艺术。^④人类不再是艺术生产的绝对主体，在人机反复交互的过程中，机器超越了单一的工具属性开始与人类成为平等的内容生产主体。生成式人工智能影像的这一核心属性也使得当下的创作流程颠覆了传统的数字艺术的创意机制和范式。

二、技术赋能生成式人工智能影像创作的流程优化

生成式影像通过算法自动化、多模态快速生成等技术特性，从创作全流程的各个方面实现效率革命，为创意落地提供技术支撑。生成式模型通过对海量数据的深度学习后进行模型训练，然后根据用户的个性化指令对数据进行调取和重组，实现影像的输出生成。将传统艺术创作中，从日积月累的技法训练到创作灵感素材的搜集整理，再到创作后修改至最终版成品的线性流程，转化为从输入指令到平台算法生成然后调整优化的非线性闭环，这中间步骤的精简使得创作者只需观察和等待极短的时间，便可得到符合意愿的生成影像，确实大幅地压缩了创作周期。

①朱婧雯：《人工智能影像的“美感”及其发生机制——基于 AI 生成视频审美经验文本的扎根研究》，《西南民族大学学报(人文社会科学版)》，2025 年第 5 期。

②方兴东、钟祥铭：《智能媒体和智能传播概念辨析——路径依赖和技术迷思双重困境下的传播学范式转变》，《现代出版》，2022 年第 3 期。

③张陆园：《人工智能影像生产模式的技术赋能、媒介特征与文化危机》，《当代电影》，2023 年第 9 期。

④张净雨、杨帆《组装的艺术——生成式人工智能影像艺术中的“组合”创意思维》，《电影理论研究（中英文）》2025 年第 1 期。

生成式人工智能影像创作的核心环节，就是创作者将创意构思转换为可被生成式人工智能影像平台精准理解的指令提示词。创作者通过“文生图”“图生图”、“文生视频”“图生视频”等模式，将提炼好的文本指令输入生成式人工智能平台，平台能够在数分钟内，基于深度学习算法生成可供选择的多幅绘画作品。创作者筛选符合创意要求的初稿。若生成结果都不符合要求，创作者需再度修正文本指令，直至生成出较为满意的结果为止。

其次生成式人工智能影像创作的另一重要环节就是对于细节的打磨，一方面是创作者在生成平台对指令进行细微调整或是局部再优化后选择最优方案。另一方面是使用图像处理软件对目前生成式人工智能平台无法优化的内容进行修改。例如通过“图生图”模式，创作者若对初稿的细节不满意，可通过圈选局部区域，通过调整文本指令、设置参数等方式，在初稿的基础上进行部分重绘。实现细节的打磨与优化，让生成的绘画作品越来越符合创意要求。目前生成式人工智能平台也存在一些无法解决的问题，比如文字乱码、光源明暗统一、前后景重叠等问题。为确保作品的完成度，需要借助图像处理软件如photoshop对该问题进行再处理。动画短片《泱泱》充分展示了生成式人工智能技术在提升动画场景设计效率和质量方面的巨大潜力，效率提升了78.57%~92.86%。^①该动画作品利用生成式人工智能数字平台完成了原先需整个团队花费数月才能完成的工作量，着实极大地提高了创作效率，这种对以往创作模式的颠覆也将为艺术作品创作带来无限的创作潜力。小红书博主一颗刘刘创作的AI全流程毕设短片《火源器语》，通过“火”的视角展现中华文明的发展，用到的工具包括Midjourney、可灵、Photoshop等。用九幕串联从雷火初燃到燧人氏钻木再到流转不同时代的灯。创作者围绕每幕的主题，设计了迥异的画风和故事内涵。每一幕之间都有对应时代的文物出现，且是故事的核心所在。在惊喜创作者的创意精妙的同时，也感叹生成式影像平台所赋予的表达可能性。以其中的第六幕作具体说明，创作者以宋代邛窑绿釉瓷省油灯中忽暗忽明灯火的固定镜头开始，接着便是一位在月落窗台前伏案作画的大家正在着色，接下来画面便一转来到了清明山河图中的现实场景，（见图1）且用移轴镜头压缩了空间，仿佛笔下的汴梁场景此刻都活了起来，从白天的熙熙攘攘到夜间的灯明璀璨，好不热闹。在短短几个镜头间，描绘了一幅盛世太平的模样。创作者的创意最大化地呈现在镜头里，生成式人工智能影像技术赋能了创作者对影视创作可能性的探索。



图1 AI短片《火源器语》截图^②

①倪槟、马家豪：《生成式人工智能在动画短片场景设计中的应用——以〈泱泱〉为例》，《装饰》，2025年第5期。

②<http://xhslink.com/o/95dT2ViaMyX>

三、主流生成式人工智能影像平台的风格表达

生成式影像平台作为生成艺术落地的重要载体，平台的发展进程与生成式人工智能技术的迭代高度同步。2022年以来，全球范围内涌现出大量生成式影像平台，不同平台依托差异化的技术架构、算法模型和训练数据以及用户定位，形成了各具特色的平台风格。本节选取了 MidJourney、Runway、即梦三大国内外主流生成平台举例说明。

MidJourney 是当前生成式图像领域的标杆性平台之一，该平台生成的作品融合艺术化的审美以及专业化的水准，具有极强的辨识度。MidJourney 采用独特的 Discord 机器人交互的模式，MidJourney 在 Discord 上以服务器的形式存在，用户可以直接注册 Discord 并加入 MidJourney，然后向机器人输入指令获得图片。用户需要在 Discord 频道中通过输入文本指令、通过调整风格强度、细节精度、图像比例、生成步数等参数，优化作品的最终呈现效果。另外，MidJourney 对文本指令的理解能力处于行业顶尖水平，能够精准捕捉用户的细节描述，尤其在复杂场景构建、人物形象塑造等方面，能够生成极具强烈艺术感染力的作品。



图1 AI短片《金鱼回到海的尽头》的MidJourney平台截图^①

荣获第15届北京国际电影节AIGC单元最佳动画的作品《金鱼回到海的尽头》的创作团队是上海戏剧学院的五位学生，他们在一周的时间里制作完成的片长3分53秒的这部作品。使用的生成式人工智能影像平台包括MidJourney、可灵和即梦。在创作中大量应用到了MidJourney平台来控制画风稳定性和任务一致性，创作者在社交平台小红书的评论区也分享了他们的一些创作巧思，包括在前期做好人物角色设计稿之后再将其投喂给MidJourney平台，之后输入精确指令来得到需要的分镜画面，其中作者分享的MidJourney画面截图（见图1）。prompt提示词（指令提示词）非常细节，对于画面中需要重点强调的元素都进行了修饰和界定。比如对于人物服装的界定、对于画面主体形象的描述“像剑一样对抗一个有着发光叶子形状的鳞片的巨大的毛茸茸怪物”，甚至是对人物发型动态模糊的确认，为后期图生视频细节过渡做好铺垫。对比指令和最终结果，MidJourney平台对于指令的理解程度极高，并且对于精确指令的还原度也到位。

Runway 是全球最早布局生成式视频技术的平台之一，聚焦于视频生成与影视特效领域。Runway 采用的是网页端一体化创作模式，平台界面将视频生成、视频剪辑、特效制作、色彩调整

^①<http://xhslink.com/o/25siqG2iKX>

等功能整合在同一界面，方便了创作者一站式创作。另外 Runway 支持用户通过调整参数，实现镜头角度、运动轨迹、画面节奏的精准控制，精细化地调节期望效果的同时，也良好地适配影视创作的专业需求。Runway 生成的影视级别的场景素材、角色动态、特效画面，被广泛应用在数字电影的前期素材库收集、后期特效制作等环节，如电影《瞬息全宇宙》（《Everything Everywhere All At Once》）就使用了 Runway 完成了片中面条手、石头对话等超现实艺术场景。

即梦是字节跳动旗下的新一代生成式人工智能创作平台，该平台基于字节跳动自研的豆包系列大模型构建核心算法，深度优化了中文语义的理解能力，解决了使用非本土化的大模型生成的内容可能引发文化偏差、中文提示词还原度低、本土化美学风格有待优化等问题，为国内创作者提供了良好的创作环境，也适配了诸多创作场景及传播需求，成为国内本土化生成式影像平台的代表之一。尤其是 2026 年 2 月，See Dance 2.0 版本重磅上线，该版本尝试解决在之前生成式人工智能影像创作中的遇到的如角色与场景一致性难统一、单次生成单个镜头画面结果等存在的问题，国内生成式人工智能平台正在以前所未有的速度向前推进。

四、生成式人工智能影像应用下的核心坚守

生成式人工智能影像在当代艺术创作中的应用，构建起了技术赋能效率，创意主导价值的这种二元平衡逻辑。这一逻辑的核心其实就是打破传统艺术创作中技术门槛制约创意落地，而时间成本又消耗创作精力的矛盾，使创作者从重复性的技术劳动中解放出来，充分发挥作者的主观能动性，利用生成式人工智能工具属性和主体间性，更加高效地创作出别具一格的艺术作品。“视频生成技术颠覆了传统影像的创作与生产范式，改变了‘人-技关系’中的权力配比。”^①在以往的影像创作中，无论是摄影机、录音机等设备单纯的只拥有工具属性，且处于客体的顺从地位被人类用来满足记录或者创作的主观意愿。但是生成式人工智能影像技术使得人们重新审视人类与技术的关系，同时也对如何平衡效率与创意有了更丰富的认识。技术的革新首先使得人类在创作中的主体性受到了动摇，从人机交互的过程来看，人类在生成式影像平台中仅能通过选择不同的模型算法和修改提示词等方式对输出的结果进行干预，且这种干预是局限的随机的，不具有对最终效果的绝对控制。此时生成式人工智能影像平台它是与人类协同创作的，“在这种双向互动中，生成式人工智能与人类创作者之间形成一种新的主体间合作与沟通的模式。”^②所以在这种共创的模式下，如何平衡好效率与创意则成为了关键。

在生成式人工智能影像技术更新迭代速度越来越快的当下，面对影像创作中人的主体性部分让渡给生成式人工智能影像工具的现状充满了担忧。究竟人类的创意在影像创作中还是否重要甚至有必要成为一个需要考虑的课题。“人工智能技术的艺术创作仍然是通过发现与改变语言模型中既有的底层数据关联，从而对固有的艺术呈现风格进行多次排列与组合，进而重组为新的艺术作品。”^③可见生成的影像与与模

①许孝媛、金鑫：《人工智能视频生成：“新影像”制造与“人-技关系”的变革》，《当代电影》，2024 年第 8 期。

②任玲玲、丁熙桐：《主体间性视域下生成式人工智能赋能影视内容共创研究》，《中国广播电视学刊》，2025 年第 10 期。

③倪万、郭嘉良：《媒介化视阈下人工智能艺术创作的本质探究与价值阐释》，《现代传播(中国传媒大学学报)》，2024 年第 8 期。

型的收集的训练数据具有直接联系，这就存在几方面的问题。首先平台根据创作者的指令对数据库中的数据进行调取并且完成复现和拼贴，进行最终结果的输出显示，在这个过程中，人工智能仅仅是机械化地执行命令，不牵涉对逻辑性的思考和物理世界和现实规则适存的合理化假设，所以就会产生理解的偏差，生成的影像具有一种浓厚的“AI味”，这种不合乎现实的荒诞感是很容易被察觉且与影像所描绘的场景产生脱离。比如，人在行走时不协调的四肢，前后镜头衔接时出现的空间扭曲且比例夸张的物品形态等等。生成式影像作品出现的这种荒诞感是生成式人工智能影像技术探索过程中的阶段性特征，是需要创作者作为审美主体根据自身经验、智慧、直觉进行判断，这是否符合最初的创作意图以及如何调整在确保整体效果的基础上尽量规避失误。其次，由于用于学习算法的收集的数据更倾向于主流文化和经济发达地区的样本数据，这会造成主流文化和边缘文化分层偏见更加明显，并且生成结果也将更符合主流审美而造成审美固化单一，甚至产生张冠李戴的情况。比如对于以英文语料作为数据库的生成影像平台，在生成中式建筑等具有时代性和区域性区分的内容时，常会无法辨别中式建筑与日式建筑的区别，如果创作者自身无法辨别，将会发生文化诠释上的误解。另外，在面对当前生成式人工智能影像技术对以往影像创作范式的颠覆，使得无论专业出身与否的创作者都能够参与创意表达。在生成式人工智能影像技术还处在探索的这个阶段，创伤与症结、迷惘与寻找、回溯与重温成为了AIGC影像作品的常见主题，这是人类在坚守创作主体性地位过程中个体向内的挖掘与求索。^①这也是人类在摇摆和彷徨之间，用记忆和肯定性的姿态来应对技术的迅猛发展对人类主体性的消解。这种主题的作品是在技术革新阶段的产物，拓展了创意表达的边界。所以人类创作者始终掌握创意主导权，算法平台仅仅是作为创意落地的技术载体，它所生成的结果需经过人类的审美筛选与价值校准，才能成为具有艺术价值的作品。人类对社会、文化、情感的深刻洞察是算法无法代替的，这构成了创意维度的核心坚守。效率的提升是为了创意提供更多的时间与空间资源，使创作者能够投入更深层次的创意思考。创意则为效率提升设定方向与边界，避免沉迷于技术而陷入无意义的高效生成。以上海国际电影节AIGC单元的最佳视效奖《金鱼》



图1 AI短片《金鱼》截图^②

为例，在社交平台小红书上作者阿米 amir 分享了作为非科班出身的创作者在获奖之后的感悟，其中提到了接受官方采访被问到对于他而言AI是怎样的一个角色？他认为它更像是摄影指导和美术指导。AI作为新媒介就像是画笔和摄影机，但更重要的是后面的大脑，是作为导演思考用什么和怎么用工具来服务于内容和内核。作者用金鱼这一意象隐喻被繁琐工作困住的人类，戏谑性的内涵打工人的社畜生活，

①张净雨：《割裂与弥合：AIGC影像装配中的主体性重构》，《当代电影》，2025年第7期。

②<http://xhslink.com/o/1SogVXcTMHr>

惊悚诡异的叙事风格配合色彩对比强烈的画面内容（见图1），可看性强的同时也意味深长。在这些参赛作品作品中，创作者追求的不是看似有视觉冲击力的生成影像的堆砌，而是在创意背后有明确的创作意图和主题表达的艺术作品的输出。

艺术家们在使用生成式人工智能影像的过程中坚持的效率与创意平衡的逻辑，为当代艺术创作带来了深远的价值影响，但仍需注意“人工智能技术的认知与知识水平的上限仍然无法超脱由现行社会程序所构建的规定与想象，故而也无法为艺术家提供具有超越性的艺术创作可能，人工智能的技术创新并没有向人类创作者提供跨越艺术边界的本体经验能力。”^①要避免一味地沉迷于技术的堆砌而陷入技术异化的困境，漠视生成式影像可能带来的问题，甚至忽略人类与生俱来的、在叙事中完成自我与世界建构的本能欲望，生成式影像平台运行的指令背后是人类的创意发挥，是人类有温度有智慧有逻辑的情感表达。未来，通过算法的持续优化与人类创意能力的不断提升，艺术家在效率与创意之间的平衡将更为重要，而坚持好这种平衡将为当代艺术创作开辟更广阔的空间。

【参考文献】

- [1]赵毅衡：《论“伴随文本”——扩展“文本间性”的一种方式》，《文艺理论研究》，2010年第2期。
- [2]李翔、俞智星：《生成式人工智能赋能电影内容生产与传播：概念演变、技术影响及创新应用》，《电影文学》，2024年第10期。
- [3]朱婧雯：《人工智能影像的“美感”及其发生机制——基于AI生成视频审美经验文本的扎根研究》，《西南民族大学学报(人文社会科学版)》，2025年第5期。
- [4]方兴东、钟祥铭：《智能媒体和智能传播概念辨析——路径依赖和技术迷思双重困境下的传播学范式转变》，《现代出版》，2022年第3期。
- [5]张陆园：《人工智能影像生产模式的技术赋能、媒介特征与文化危机》，《当代电影》，2023年第9期。
- [6]张净雨、杨帆：《组装的艺术——生成式人工智能影像艺术中的“组合”创意思维》，《电影理论研究（中英文）》，2025年第1期。
- [7]倪镔、马家豪：《生成式人工智能在动画短片场景设计中的应用——以<泱泱>为例》，《装饰》，2025年第5期。
- [8]许孝媛、金鑫：《人工智能视频生成：“新影像”制造与“人-技关系”的变革》，《当代电影》，2024年第8期。
- [9]任玲玲、丁熙桐：《主体间性视域下生成式人工智能赋能影视内容共创研究》，《中国广播电视学刊》，2025年第10期。
- [10]倪万、郭嘉良：《媒介化视阈下人工智能艺术创作的本质探究与价值阐释》，《现代传播(中国传媒大学学报)》，2024年第8期。
- [11]张净雨：《割裂与弥合：AIGC影像装配中的主体性重构》，《当代电影》，2025年第7期。

^①倪万、郭嘉良：《媒介化视阈下人工智能艺术创作的本质探究与价值阐释》，《现代传播(中国传媒大学学报)》，2024年第8期。