

人机共创——AIGC时代影视业模式创新

■ 李本乾 陶婷婷

人类传播技术在经历传统媒体、互联网之后，又进入到AI这一崭新时期。特别是美国当地时间2024年2月15日，OpenAI推出首款文本生成视频（text-to-video）的Sora模型，更是令人惊叹不已！那么，如何抓住人工智能技术革命的机遇，大力发展影视新质生产力？这一课题亟待我们深入思考和大胆探索。

文生视频Sora发布前，ChatGPT已经在报业掀起了一波巨浪。美联社、路透社、彭博社、《华盛顿邮报》《纽约时报》等媒体已经将ChatGPT大模型应用于内容采集、数据分析、内容制作、个性化报道和受众互动运营等媒体场景中。全球首个由人工智能生成新闻报道的平台NewsGPT，声称其不受广告主、个人观点的影响，可以每周7天、每天24小时为观众提供准确可靠的新闻。谷歌人工智能工具Genesis备受瞩目，据谷歌发言人表示，其“目标是让记者利用那些提升工作效率的新兴技术”。OpenAI开发的ChatGPT-4内容审核，突破了过去被认为机器无法进入的禁区。由此可见，未来在媒体“策、采、编、审、发、营、评”各个环节，AI赋能已是大势所趋。

与此同时，美国正在爆发一场“数据起义”。好莱坞、艺术家、作家、社交媒体公司和新闻机构作为反抗者，都把矛头指向ChatGPT和Stable Diffusion等生成式人工智能工具，它们被指在未经许可或未提供补偿的前提下，非法利用内容创作者的作品训练大型语言模型。由此引发的传媒数据保护和数据安全问题，应引起传媒业界的高度重视。

与ChatGPT相比，文生视频Sora的技术更加复杂。简单了解其关键技术创新，有利于我们更好运用Sora赋能影视业。

一、Sora 的新技术及新功能

1.Sora的新技术。文生视频Sora主要采用如下技术：首先，Sora使用了视觉补丁（patches），类似之前的文本标记（tokens），可通俗地理解为使用图像块将影片编码，以便后续训练生成模型。其次，Sora将视频的普遍特征最大限度地高度概括到低维度空间，即潜在空间（latent space），以节省算力。再次，视频生成缩放（Transformer）。它利用时空补丁（spacetime patches）、Transformer进行训练，实现对不同分辨率、长度和长宽比的影片生成。最后，将DALL·E3中的字幕技术应用于视频，能够准确地按照用户提示生成高质量的视频。

2.Sora的新功能。首先，Sora涵盖了文本输入、图像生成、视频合成、优化与调整等多个步骤，可以实现文生视频、图生视频、视频生视频，以及图像生成动画等产品的生成。其次，它具有在不同视频要素、场景，以及在不同主题和场景构图的视频之间创建无缝过渡的强大功能；还能实现将静态的图像动态化、向前或向后扩展视频、视频到视频编辑、视频无缝连接等。再次，Sora的技术创新，最大突破在于“世界模拟器”（world simulator）。相较于此前 Runway, Pika Labs 等多家人工智能公司推出的文生视频大模型，Sora不仅仅是技术的简单升级，而是能够生成动态摄像机移动的视频，模拟以简单方式与世界互动（Interacting with the world）和模拟数字世界（Simulating digital worlds）等。

二、Sora 对影视行业的新挑战

尽管Sora的目标是世界模拟器，但目前还难以准确地模拟“玻璃破碎”等物理过程。显然，要实际投入更加复杂的、能够反映人类精神世界的影视专业制作，其技术尚待进一步的发展。尽管如此，我们有必要对其带来的挑战，做一前瞻性的预测分析。

1.颠覆传统的影视市场结构。随着Sora文生视频模型的大规模推广、普及与应用，视频制作的门槛将逐步降低，精英制作将逐步演变为大众制作。Sora将首先直接冲击自媒体、短视频、直播及社交媒体视频；其次将冲击自制短剧、微剧、新媒体广告；最后一波将冲击专业媒体、电视媒体、新闻广播、综艺直播。尽管专业媒体并未首当其冲，但前两波产生的海量文本、视频，势必分散用户注意力，将导致传统主流媒体市场下滑。Sora等生成式AI模型通过直接和间接两种方式，将导致主流媒体市场占有率进一步下降，主流媒体或将再次遭受重大冲击。

2. **颠覆传统的影视生产模式。** ChatGPT、Sora输入文字、图片或者视频，能在极短时间内实现文生文本、文生视频。Sora具有超强的语言理解能力，从文本、图片到视频的生成能力，复杂场景、多角色、多镜头的生成能力，以及3D和4D动画视频生成能力。算法、模型介入影视内容生产，其一是算法、模型将逐步取代人工，冲击我国现有的宣传文化队伍；其二是将逐步颠覆过去人类单一的影视生产模式。

3. **颠覆传统影视价值链体系。** Sora不仅可文生视频，创造长达一分钟的视频，且具备多镜头的一致性。Sora让“扔进一部小说、出来一部大片”的梦想，变得如此接近现实。倘若如此，传统影视生产价值链中，去头掐尾后，中间环节或都存在被替代的风险。显然，Sora将对传统影视产业价值链体系造成强烈冲击。

三、创新影视“人机共创”模式

1. **整合视听行业数据，创建影视数据平台。** 影视制作的数字化和智能化，已成为一个不可逆转的发展趋势。整合视听行业数据，创建影视数据平台，优化新型影视生态，才能以更高的效率去创造更大的价值。广电行业数据资源质量高且规模庞大，并且有视频标准规范、易于标注等优势，为此，整合行业数据、创建影视数据平台，是广电行业智能化发展的基础。在此领域，上海广播电视台着手进行的语料库建设、多模态数据对齐、大视听垂类模型研发、自主开发空镜素材智能推荐系统等尝试，都是值得充分肯定的。

为迎接人工智能技术革命带来的机遇与挑战，广电行业或应积极建立广电语料数据联盟，制定全行业数据标准，创建影视数据平台，推进广电行业智能化水平。

2. **开发智能软件工具，赋能影视内容生产。** Sora只需简单的自然语言提示，即可生成高清视频；其次，Sora技术简化、省略了传统动画特效复杂的数学模型、建模软件、VR技术，将使特效制作更加简单高效。上海广播电视台推出Scube智媒魔方，“开箱即用”，高效赋能两会报道。阿基米德应用音频指纹识别，生成全国广播热歌榜和综合排名榜单；应用自然语言处理（NLP），提取节目的语义和结构信息；研发音频类算法，实现音频分类、声纹识别和说话人识别；研发虚拟主持人、有声书生产、AI电台。第一财经建构融媒体智能生产平台，实现财经融媒体内容生产辅助和智能图库建设，AI可辅助写稿和视频内容生产。上海广播电视台（SMG）的以上努力取得了良好的效

果。尽管受到硬件和算法的制约，这些模型还有待完善，但他们在此领域积极探索的方向，给未来的传媒实践提供了经验。

3.创建“人机共生”生态，建构“人机共创”模式。继UGC和AIGC之后，随着文生视频Sora的到来，人类将不折不扣地迎来H-AI-GC（Human-AI Generated Content），即人类与AI共同创造内容的时代。为此，汲取中国文化“天人合一”思想，有助于创建人机共生的影视新生态，实现人机合一、人机合德、人机和谐、人机共生，共同推动AI技术与影视内容的深度融合，共创影视行业高质量发展。以此视角来考量上海广播电视台旗下东方明珠新媒体的实践，它提出的广电+文旅行业赛道生态建设，是着眼于未来“人机共生”生态的一个初步尝试，虽仍是以概念为主，但方向感的确立，有利于未来传媒行业朝向“人机共创”模式的进化。与此同时，将人的作用引入AI系统中，建构以人为中心、人类智能与人工智能混合的、人机共创的影视内容生产模式，有利于发展影视新质生产力，实现行业的高质量发展。

（作者简介：李本乾，上海交通大学智能传播研究院院长、特聘教授；
陶婷婷，上海交通大学媒体与传播学院助理研究员）