

Sora： 文化产业新机遇

■秦建新(独立策展人)

2月16日一早,OpenAI悄无声息地发布了第一个文本生成视频大模型——Sora,犹如一颗核弹迅速引爆全球各大媒体,朋友圈更是直接炸锅,其带来的冲击力,不亚于去年初爆火的ChatGPT。

去年笔者使用过一款Runway的AI应用来生成视频,当时的Runway已经是几款视频生成应用里的佼佼者,但与Sora的演示视频相比,简直不在一个等级,无论从视频生成时长(Runway只能生成几秒)还是动效和镜头运用,和Sora的60秒视频一比,就像弹弓和狙击枪的不同,代差不在一个数量级上。

Sora依托了两项核心技术的重大突破——Spacetime Patch(时空Patch)技术和Diffusion Transformer(扩散型Transformer神经网络)架构,OpenAI将Sora定位为一款世界模拟器的视频生成模型,为未来发展理解和模拟真实世界的模型奠定基础。官方宣称这次openAI训练的Sora所建立的是一种通用的视觉模型方法,而不再是细分的方式(之前的研究,图像与视频模型是不同的),因此既可生成视频又可生成图像,并且不止如此。当GPT学会语言的时候就学会了语言所传递的一切,同理,Sora也可以学会视觉所传递的一切。

笔者从2022年开始接触Midjourney(一款文本生成图片的AI应用),见证过文本生成图片,AI应用从一开始连手指几根都搞不清楚,甚至双眼视线都不能统一,短短一年时间就已经进化出无与伦比的成像质量。这次Sora在春节期间的炸裂亮相,给各界带来了耳目一新的震撼。

以往的AI文生图像和文生视频,局限在二维平面,Sora则实现了算法对现实世界的理解和对世界的模拟两层能力,跳出了二维平面范畴,展现了大模型的无限想象力。笔者认为,以Sora为代表的AI大模型不会带来行业的崩溃和消亡,恰恰相反,它将给短视频、影视工业,尤其是广告行业带来非常积极的正向推动。

众所周知,传统的广告行业TVC都是从创意脚本开始,继而是导演制片美术等执行拍摄以及后期剪辑工作,传统的流程既漫长又会在执行过程中出现很多不可控的变数。笔者2003年在BBDO担任创意工作时,曾经接手百事可乐的一部广告大片,当时品牌方百事可乐签约了著名球星姚明,客户希望由姚明与周杰伦两大巨星来共同拍摄一部广告片。当年我的创意是两大巨星间因百事可乐形成的超级感应:地球一端正在苦思冥想,进行创作的周杰伦,与地球另一端在球场刻苦练习的姚明形成心灵共振,合生出属于百事可乐的音乐节奏。就是这样一个得到各方认可的创意,在执行中,囿于难以呈现虚拟的灵感互动场景,只能省略了灵感能量在地球两端,透过各种介质传递的质感细节,简化成了姚明拍球,可乐罐震动,周杰伦得到灵感。最后,这支耗资超过600余万的片子充满了遗憾。记得我把成片转给一个业内朋友看,他第一句就问:是不是姚明在罐子里拍球?试想,如果有Sora这样的工具,很多广告人在创意阶段就可以生成Demo直接提案了,执行上也会更加精准,甚至可以直接用Sora来生成前文所述能量传递的特效镜头,拍摄成本势必也将大大降低。

其次利好影视、动画、游戏、视频创作等行业,例如现在比较风靡的短剧,非常适合用Sora来辅助制作。短剧,因为制作成本的限制,目前普遍质量不是很高,而Sora能够高质量的生成不少需要高成本拍摄或者人类比较难拍摄的场景画面,这样可以极大地提高短剧的质量水准,同时也有很大成本降低空间。

很多人可能想知道,那是不是丢一部小说给Sora就能生成一部电影了?这可能是夸张的一种说法,但可以预计,目前能生成60秒时长视频的Sora,在不久的将来,一旦Sora对外正式开放使用,会有更多第三方基于Sora的创作工具涌现出来,例如,将一个剧本分成上百个60秒,主线、辅线、暗线,不同灯光场景设置,主角配角造型服装、指定动作等多角度专业的细分镜头,加上对场景人物的特征记忆,相信在两三年内,完全有可能看到一部由AI生成的质量水准上乘的电影作品。

Sora的横空出世,令不少行业的从业人员惊呼自己快失业了,AI恐慌论也一时甚嚣尘上。笔者认为,AI终究无法取代人的想象力和创造力,特别是人的主观意愿。当AI的视频创作到了不需要拍摄、无需演员、灯光、美术等庞大的摄制组的状态后,审美能力、想象能力和导演思维却变得非常重要了。Sora无论进化到怎样完美的程度,使用工具的依然是人类本身,所谓“AI只会淘汰不懂使用它的人”。因此,任何有想法、有创意的人,都能以低门槛、低成本进入与视频相关的文化创意产业,这势必给该产业结构带来更大变革与蓝海机遇。

Sora的命名取自日语“空”,意为天空。OpenAI团队表示,选用该词即是因为其让人联想到无限的创作潜力。

Sora与文化艺术 创作者的身份主体性

■齐丽梅(中国国家画院)

最近一段时间,在忙碌的工作生活外,人们纷纷将目光投向人工智能巨头OpenAI在今年2月发布的几部短视频之上,诸如浪漫夜幕下东京街头散步的时尚女性、小小的咖啡杯中互相缠绕的两艘海盗船、几只雪地中玩耍着探出头的可爱幼犬、一列在微观的叶脉中缓缓行驶的小火车……尽管短视频早已进入我们的日常生活,不断影响着我们的知识获取与情绪表达的方式,但为何这几部短视频推出之后会再次引发热议?究其原因,主要是它们来自于OpenAI公司发布的首款智能视频生成模型Sora,以拟真态式与理解式的视频生产方式,推动了AI技术在现实物理世界中的实践运用,特别是在盛行的AI创作领域再次形成颠覆性的认知。

可以看到,Sora已经引发新型人工智能技术在影视、戏剧、美术、医疗、制药等领域的讨论热潮。尽管目前Sora的社会影响主要集中在视频生产领域,我们可以将其视为一种人工智能视频生成的工具,它扮演的角色更多的是知识的“模仿者”,但这种“模仿”显然较之以往的Meta、Runway、Pika等产品呈现出压倒之势,甚至完全能与人工精心拍摄创作出的短视频、影视剧等文化形式相比肩。这同时意味着Sora在某种程度上模糊了创作工具与创作者之间的身份,它能够主动理解指令文字中所包含的事物在现实世界中的存在方式,并赋之以合理的时空结构,从创作意识的自主性、身份的主体性等角度向文化界提出新的命题。质言之,Sora正带给文化创作者在身份层面的若干焦虑与纠结。

作为特殊的、社会性的艺术形态,文化创作通常包含多重环节。它最初源于创作者对已有文化知识的主动接受,至于接受何种类型的文化形态,以及对不同文化形态的接受程度如何,主要依赖创作者自身的知识水平和行为认知。但对于Sora来说,它在接受知识的过程中并不依靠自主意识,抑或其本身并不具备主体性,自然也就缺少对既有知识的主动筛选与阶段性目标的精细化理解。目前Sora的运作模式更多地依靠用户输入文字指令后,自动生成趋近于真实世界的视频动画,时间可达60秒。这虽然是以往产品无法达到的时间维度,但其中也存在着基于时间本体的弊病,以及在贯彻执行用户指令时的认识局限。

输入文字指令的方式类似于命题类的作文,Sora只能依据既有指令,在现有知识储备中进行画面的筛选拼接、行为主体的空间架构以及动作方式的演绎,一旦超出自身知识储备的范畴,必然会导致一些视频画面与技术运用层面的弊病。对于独立的文化创作者而言,他不仅接受既有的知识、数据、资料,而且在接受基础上形成个性化的思想、观念、意识,甚至能够较为自如地对某些超出知识范围之外的指令,与指令本身形成创造性的对话。当前已有评论者从“物理原理和因果关系模拟不足”“剧情和感情线问题”“战斗和操作问题”等方面,指出Sora所生成的短视频中存在的画面、操

作、情感等问题。这些问题从根本上还是由于Sora的工具属性所致,其本身并没有实现创作身份的转换,尤其当它在面对一些主体感与行为性较强的指令时,无法祛除固有的沉疴。

又如从画面呈现的角度来看,Sora生成的视频画面的识别度虽然较之以往的产品显现出优势,能通过十分逼真的场景创作出带有强烈现场感的图画,但这种生成只是在接受指令后进行的色彩组合与画面拼接,并非基于思想意识层面的主动创造。而现实的影视制作中通常会有专门的美术、设计专业人员,他们以团队的形式深入理解影视剧作所需要的画面形态,其中出现的每一帧画面都经过设计者的精心制作,在带给观者美感的同时引发着他们对故事本身的反思。

近年来,一些国产影片之所以能给观众带来巨大的震撼力,其根本的原因就在于影视创作者情感的巨大投入。以Sora为代表的人工智能技术虽然能带给传统艺术极大的冲击,但这种冲击更多地集中于技术效果与工作效率,尚未触及行为、思想、观念等认知层面。舞台美术、影视美术的设计者作为具有独立思想的个体,能在丰富的社会生活与不断变化的日常情绪之中精准地捕捉激发性的情感触点,这种意识自然是以Sora为代表的AI技术目前所无法拥有的。浙江大学人工智能研究院吴飞教授在谈及Sora的现实影响时指出:“Sora将重塑影视创作等行业的游戏规则,以文生视频模式颠覆电影创作过程,可能今后会有若干人通过AIGC相关技术来生成电影。”这就依然把Sora的颠覆性影响集中在游戏规则、影像技术等层面,而创作的身份主体依然是“若干人”。

从早些年的小简、小冰写诗,到后来ChatGPT在知识界的广泛运用,再到如今Sora带来的视觉轰动,每一次的技术革新都在影响乃至改变着我们的生活方式,使我们体验AI技术带来便捷与效率的同时,也产生了许多有关创作哲身份的焦虑与困惑。如2023年发生的好莱坞演员、编剧大罢工,以及最近影视、戏剧等行业发出的频频哀叹,都指向了人工智能技术取代人类工作的现实问题。但从根本上看,文化创作者只是受到AI带来的技术冲击,现在谈“取代”还为时过早。创作者的主体性以及自主意识是支撑与实施创作行为的关键要素,更是促进文化产业不断发展与突破的核心动力。总之,当前的Sora更多地作为文化的接受者,其创作者的身份属于表面的、浅层的身份,并不具备到真正的创作意识。我们甚至可以把它理解为指令的执行者,而非真正的文本意义上的创作者。也正是基于此,这种关乎创作者的身份纠结或许也可升华为一种对文化产业内部诸如创意内容、版权归属、原创故事的能力等若干关键问题的深入反思,某些评价还远不能成为定论,正如艺术与科技彼此的深广多维影响,还需在实践发展中慢慢体察与解读。