

快手可灵:扩展视频行业想象边界

□本报记者 李雪昆

伴随Sora的面世,“挑战者”也正前赴后继地涌现,其中不乏国产大模型的身影。最近,在国内社交平台以及海外推特等平台,大量网友在“求”一个中国视频生成应用的内测机会,这个让国内外网友纷纷关注的大模型就是快手发布的自研视频大模型——可灵。截至7月初,已有累计超过70万人排队申请内测。可灵大模型的发布,不仅代表着快手在AI领域的又一重大进展,也预示着短视频行业即将迎来一次全新的变革。

《中国新闻出版广电报》记者体验后发现,可灵正以其强大的生成能力和便捷的创作模式,为用户提供一个全新的视频生成解决方案。用户只需输入简单的文本指令,即可让可灵生成高清、高质量、长达3分钟的视频,其分辨率达1080P,帧率达到30fps,为用户的创意赋予了无限可能。其不仅支持自由调整的视频宽高比,还能够生成符合真实物理特性的运动场景。无论是模拟自然界的物理规律,还是将复杂的概念或想象转化为生动的画面,可灵均表现出了不错的能力。

创意表达打开新天地

技术的创新和应用始终是推动行业发展的核心动力。可灵的诞生,不仅实现了技术上的突破,更为用户打开了创意表达的新天地。在技术创新的推动下,可灵以其独特的优势,为用户提供了全新的视频生成解决方案,使得短视频制作不再受限于传统的制作流程和技术壁垒。

长达3分钟的视频是如何生成的?简单而言,用户使用可灵大模型可以通过文生视频和图生视频两种模式,生成约5秒钟视频,此后可使用视频续写功能,让每次续写都能将视频延续约5秒,最终可实现长达3分钟的视频。同时,每一段续写还能够融入用户的创意和想法,实现场景转换和过渡,用户可对每一段延长的视频添加不同提示词,生成连续长视频。

对此,天风证券认为,在输出视频的自由度方面,可灵3分钟的内容长度远超Sora的60秒长度。在对文本提示词的体验中记者发现,可灵具备了将复



7月6日,在2024世界人工智能大会期间,可灵应用再次受到关注。

资料图片

杂概念和想象可视化的能力。通过文本提示词,用户可以引导可灵生成特定主题或情绪的视频,这使得创意不再受限于制作者的技术水平,而是可以个性化地在视频中展现创意。

例如,当使用一张小女孩的照片生成视频时,模型默认生成的效果可能是微风中她发丝轻轻飘动,周遭花朵摇曳,水面闪烁着细腻的波纹。而如果用户具体输入“女孩转身背对镜头”,视频便会智能添加女孩优雅转身的动作,并细腻捕捉到转身瞬间的微笑和眨眼,展现了自然而生动的细节处理能力。

再比如用巨龙图片生成视频,输入“正在捕食的巨龙,飞扬的沙子,奇幻风格的电影”,视频便呈现出巨龙昂首展翅掀起漫天黄沙,嘴角张开仿佛在咆哮的生动效果与奇幻场景。通过这种文本与图像的紧密联动,可灵可以按需求定制视觉叙事提升到一个新高度。

这也意味着,用户能够更方便地创作出内容连贯、情节丰富的长视频,这无疑扩展了视频创作的边界。

美国《麻省理工科技评论》也刊文谈道,AI或将颠覆短视频的内容创作场景,将在短期内使创作者受益。从长

远来看,以快手为代表的短视频平台有可能接管视频制作,直接为用户生成定制的内容,从而减少平台对明星创作者的依赖。

实用效果源于技术支撑

在一段测试倒牛奶的视频中,力学方面的重力规律、液面上升都较符合现实,甚至倒液体时泡沫一直在最上层的特性也被考虑其中。由此可见,可灵大模型在模拟真实世界时,所生成的视频是较符合物理规律的。

据快手大模型团队介绍,可灵整体框架采用了类Sora的DiT结构,用Transformer代替了传统扩散模型中的U-Net,这一决定大幅提升了模型的处理能力和生成能力,同时也增强了模型的扩展性。在该团队看来,一个优秀的视频生成模型,需要考虑四大核心要素——模型设计、数据保障、计算效率和能力扩展。这其中,在技术创新方面,可灵采用了3D VAE机制和DiT模型来建模复杂的时空运动。这种先进的技术架构使得可灵能够更真实地模拟物理世界中的运动 and 变化,无论是物体的自然运动,还是复杂的

场景交互,都能够生成符合现实世界规律的视频。

得益于自研模型架构及Scaling Law激发出的强大建模能力,可灵大模型构建起了一个无限逼近现实的想象空间,无论是真实世界的光影反射、重力影响下的流体运动,还是与物理世界的交互,可灵大模型都能够生成符合物理规律的视频。比如,只需上传一张牧羊犬追球的静止照片,模型就能生成一段生动的视频:牧羊犬活灵活现地奔跑追逐网球,其耳朵轻轻晃动,毛发随风飘扬,而网球则在空中弹跳,视频整体自然且流畅。

在AI领域,数据是模型训练不可或缺的基础。快手可灵大模型之所以能够在视频生成方面取得不错的反响,得益于其在数据建设上的深入工作。记者从快手研发团队了解到,他们构建了高效的大规模自动化数据解决方案,其中较为完备的标签体系,可以精细化筛选训练数据,或对训练数据的分布进行调整。该体系从视频基础质量、美学、自然度等多个维度对视频数据质量进行刻画,并针对每一个维度设计多种定制化的标签特征,由此助力了视频内容的高质量呈现。

就目前而言,市场对可灵的评价普遍积极,他们认为,可灵是真正经过验证且可以直接使用的中国版Sora。同时,在商业前景方面,市场分析机构对可灵同样持乐观态度。摩根士丹利和天风证券等机构分析认为,可灵的成功发布将有望使快手的用户规模和活跃度进一步提升。而随着可灵技术的不断完善和功能的扩展,其在短视频、电商、广告等多个领域也将带来巨大的应用潜力。

业界普遍认为,生成式AI在视频创作和世界模型的大踏步进步将实现对视频/3D/游戏等下游应用场景的渗透。对此,天风证券预测,在短视频、创作工具、游戏等下游领域,可灵、Sora等AI原生产品有望融入工作流,增强用户体验,降低用户使用壁垒,进一步降低创作成本,并极大拓展创作者能力边界。因此可灵大模型的推出,不仅为用户带来了全新的视频创作体验,也将为短视频行业带来技术革新机遇。

综艺速递

“新声一代”崭露头角

□本报记者 杨雯

婴儿呱呱坠地时的嘹亮啼哭,戏剧表演中情感动人的精彩对白,令人感到陶醉的放声歌唱……各种声音在现实之中如影随形,向世人展现着声音存在的重要意义与强大魅力。与此同时,一群锐意进取、极富个性的“声音工作者”,也在属于自己的领域崭露头角,绽放出耀眼的光彩与色彩。

7月5日,新生代声音人物纪录片《新声一代》在腾讯视频、爱奇艺、优酷同步上线。该纪录片共八期,每位嘉宾两期,分为正片和采访篇,主要聚焦不同声音领域年轻而富有实力的新生代青年,分别以歌手、音乐剧演员高杨、配音演员吴韬,男高音歌手马佳、配音演员谷江山为主人公,深入他们的工作和生活之中,捕捉他们台前幕后的精彩瞬间,将他们的成长故事和心路历程娓娓道来。

该纪录片不仅传递出“新声一代”新锐、个性化的人生态度,以及他们对艺术与生活的无限热爱,同时也展现出中国声音领域的青年力量,彰显了新时代中国的文化自信与时代风貌。

聚焦多元声音领域

说起这部纪录片的缘起,则要追溯到2022年年初。由于有引进海外纪录片的业务,为了扩大纪录片受众群体,该片出品人、总制片人、总策划彭侃尝试去请一些跨界名人来给纪录片配音。巧的是,当时的片子正好请到了高杨来进行配音,结果收到了很好的效果,这便引起了彭侃的关注。

“现在有很多音乐剧演员、广播剧演员,他们都有忠实的受众,这些受众过去可能不是纪录片的受众,但有可能通过跨界合作的方式,吸引他们来观看纪录片。”彭侃表示,如今,在不同声音领域内,一系列深受年轻人关注和喜爱的垂类圈层文化正在蓬勃兴起。希望这部纪录片搭建起声音垂类的圈层文化与大众之间的桥梁,让更多人了解音乐剧、广播剧等领域,也看到嘉宾们所展示的中国青年一代的文化自信。

在嘉宾选择上,彭侃介绍,有3个重要标准:其一,嘉宾要有一定的人气。其二,嘉宾正在经历人生中比较重要的事件,比如他职业生涯中的机会或者转变等,只有这样片子才会有“重头戏”,有可以去表现的东西。“《新声一代》是一个轻量级的项目,它不是那种长线的跟拍,因此,我们需要借助嘉宾正在经历的一些大事件,使之成为纪录片内容的主要部分。”其三,嘉宾要能够代表其所在的圈层文化,并且能够有自己独立自由的表达和灵魂。

深入贯穿台前幕后

从开始筹备到最终上线,纪录片《新声一代》耗时超两年时间,拍摄贯穿台前幕后,囊括嘉宾们在职场与生活中的方方面面,他们背后的付出和汗水,他们未曾表现出的顾虑与犹疑,他们充满反差感的不同侧面,都在片中得到了真实的记录和表达。

该片导演、剪辑指导谢福坤透露,现场拍摄时会尽量不打灯,也不用大的挑杆去收音,避免对着脸拍而使用长焦远距离拍摄,更不会刻意调动人物摆布,尽可能减少对拍摄对象的干预,以确保更自然的人物状态。

从工作角度出发,创作团队深入他们工作和生活的各个环节,包括音乐剧演出、个人音乐会、专场见面会、专场巡演、线下见面会、录音棚配音、歌曲录制等多个场景的展现,传达出他们对自己“声音事业”的投入、坚守与热爱;从生活领域出发,《新声一代》将镜头对准4位新生代青年放松休息的业余时间,钓鱼、打篮球、打台球、做发型……深度探寻他们闲暇时的个人爱好与习性,生动呈现他们在镜头前不曾展露过的鲜活瞬间。

“我没有想到音乐剧、配音竟然都是挺‘卷’的行业,他们每一个人都比我想象中要更努力。”在做完这部纪录片后,谢福坤感叹道,“他们虽然外表看着光鲜亮丽,但其实平时也是要熬夜的,也是要加班的。”

传递“90后”新锐态度

据彭侃介绍,片名中的“新声”首先指该纪录片聚焦声音领域,致力于表现音乐剧、广播剧等中国近年来受到青年群体追捧的新兴声音文化;其次,片中嘉宾都是“90后”,与前辈相比都是各自专业领域的“新生代”;再者,“新声”也寓意着这部纪录片致力于传递青年声音,凸显嘉宾们独特、自我的表达。

除了全面跟拍记录之外,《新声一代》还针对4位拍摄对象进行了每人近两个小时的深度采访。该片没有使用旁白进行故事延展,而是选择用人物的采访自述串联始终,让他们如同跟老友聊天一般,以更加松弛自如的状态讲述自己的成长故事、心路历程、思想观念、对现在和未来的想法等,让人们看到“新声一代”最真实纯粹的样貌。

“我始终觉得纪录片有一个功能,就是为当代的人立传。”谢福坤说道,通过影像记录这些优秀的新生代青年的蜕变与成长,假以时日,可能他们中的谁也会成为未来的时代记忆。“未来100年、200年后的人,怎么看待现在的音乐剧演员、配音演员,我觉得纪录片是一个很重要的影像资料。”

短剧集《三星堆:未来启示录》

探索AI影视剧制作新流程

□本报记者 杨雯

7月8日,AIGC科幻短剧集《三星堆:未来启示录》在抖音正式播出。作为全国首部AIGC生成的连续性叙事剧集,《三星堆:未来启示录》由博纳影业组织构思开发,联合抖音和即梦AI共同打造。从事电影工业制作的专业团队,在创作过程中应用了十几种最先进的人工智能视频技术,全面展示了AIGC技术在影视创作领域的广泛应用。

博纳影业深耕影视行业多年,具有成熟的电影制作与发行经验,今年成立的AI工作室——博纳影业AIGMS制作中心,致力于影视内容生成等AI领域的技术应用实践,此次推出的《三星堆:未来启示录》正是该制作中心的全新作品。这不仅是一次文化与科技的有益结合,更是人工智能与影视艺术的双向奔赴。

应对快速迭代周期

AIGMS,即AI自动生成影视剧。与目前行业热门的AIGC不同之处在于,AIGMS多了“影视剧”这一个维度的内容,具有连续性叙事的特征。因此,相比上半年中央广播电视总台推出的AI全流程微短剧《中国神话》等作品,《三星堆:未来启示录》讲述了一个更完整的、连续性的故事,拥有一个整体的大世界观。按照规划,《三星堆:未来启示录》立项了两季,每一季12集,一共24集。按照每集4分钟计算,整体正好是一部电影的长度。

用AI来自动生成这样一部短剧乃至长片,首先要做的是做什么?答案是:研究软件。

在项目启动之初,博纳影业AIGMS制作中心负责人曲吉小江以为工作流程很快就能梳理清楚,可是真正操作起来,“每天睁开眼睛后,第一项工作就是研究全球今天又迭代了哪一款软件,或者是开发了哪一款大模型。”由于各款软件“内卷”很厉害,每天都在迭代升级,加上中心自身也要在影视

领域这一垂直品类上开发各种应用软件,以解决实际应用中的一些难点和痛点,因此,“一边研究、一边开发”成为创作团队的工作常态。

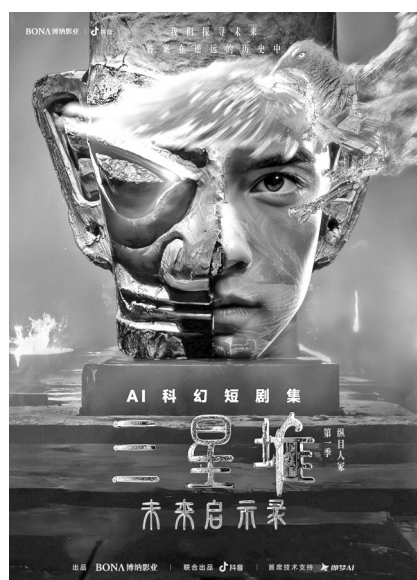
“内容本身其实是表层的东西,里层的东西是我们要把现在所有能够用到的人工智能研究透彻,并让它能够适用于我们的创作。如果在应用过程中出现了盲区,那我们就需要去做研发工作,将盲区填充。”对创作团队而言,《三星堆:未来启示录》的完成度,也反映着AI工具与电影工业化的结合度。

《三星堆:未来启示录》制作完成后,创作团队怀着些许迫切的心情,想让观众看看目前人工智能视频软件所能达到的极限,更关键的是,“人工智能技术的发展速度非常快,出乎了我们的意料。我们原来以为半年会是一个迭代周期,现在看来,基本两三个月就是一个周期了。”创作团队表示,等到之后第二季出来的时候又将会是另一个技术水准,解析度、清晰度、人物细腻度、情感表达度都将再上一个台阶。

不是+AI而是AI+

AI在影视剧中的应用已经不是新鲜事,在即将于7月12日上线的博纳影业新片《传说》中,制片方运用了最新的AI技术,重现27岁的成龙形象来饰演征战沙场的汉代将军,并完成精彩的动作场面。这正是典型的+AI——在电影中运用了一些AI工具,使我们实现了一些以前要用很高成本才能实现的效果。

而博纳影业AIGMS制作中心则与之不同,它是AI+,即以人工智能作为新质生产力,AI在这里是合作伙伴而非简单的工具。“AI+和+AI是两个概念,它们在生产流程上是完全不同的。”曲吉小江说,“当我们把AI当作合作伙伴时,它就有了灵性,虽然它还缺乏一些创造力,但即便这样,我们仍



是一个团队,真正实现了人机共创。”

《三星堆:未来启示录》探索了从“AIGC纯生成式短剧”到“AI+实拍长剧集”再到“AI+工业化电影”的三步走模式,通过AIGMS制作中心独创的AI影视剧制作流程,控制人物的动作、场景变换和情感表达,打破AI生成影视作品的静态叙述限制,实现了角色和场景的一致性,形成了影视IP开发的“N+2”闭环模式,为观众带来独特、创新的视听体验,为影视创作者提供了更多的创作自由和创新可能。

在如今的AIGMS制作中心,粗粗一看,与传统的电影公司并无太大不同,造型部门、美术部门、合成部门、剪辑部门仍然存在并不可或缺;而细看过去,似乎又有一些不一样,比如拍摄部门不复存在,比如有些传统职位改换了新名称——编剧在这里被称为“硅基语言高阶数学集编译师”,他们需要熟悉AI技术知识,对剧中可能运用AI技术完成置景合成或情节合成之处做到心中有数、预先设置……其实,不论是哪

一个岗位,其背后的逻辑是一样的:这个岗位上站着的人,都是在传统的影视领域中最快掌握AI工具的人。正如创作团队所说,“人工智能一个很大的特点是,它门槛很高,但又不是那么高。这其实就是技术平权,它给予了愿意拥抱这项技术的年轻人很多而且没有上限的机会。”

人机共创不能没有“人”

在《三星堆:未来启示录》的创作过程中,在对AI技术一层一层的“解码”后,曲吉小江反而越来越感受到人的重要性,这是在研发之前他完全没有预料到的。“其实大家现在都有一个误解偏差,以为AI生成式就真的完全是AI自发生成的。其实并非如此,我们是将人工智能视频技术运用在剧本创作、服装设计、图像到视频转换、视频编辑和媒体内容增强等多个环节中,结合电影工业制作的专业团队,来生产一系列的影视作品。”

具体到《三星堆:未来启示录》,其中一些最根本的地方仍然需要人来进行把控。比如三星堆是华夏文明重要的文化IP,瞎编肯定是不行的,它的考古内容、审美价值、思想内涵是什么?拍摄时需要顾忌什么?在国际上要怎么传播我们的中华文明?“在这些问题上,AI完全没有办法代替人的工作。”曲吉小江表示,“所谓人机共创,如果没有‘人’,那么‘机’的能量就是零,但是如果‘人’的作用发挥好了,‘机’的能量就会呈现几何级数的增长。”

任何电影精品作品,其实都有一种哲学的高度蕴含其中,虽然AI有时候能够给予人类启发和灵感,但是电影作品的最终力量,还是来自人类内心的。可以说,“截止到我们对话的这个时间点,AI还没有达到它的奇点,也许达到了奇点,它就无所不能。但目前来看,AI是不如人的。”创作团队表示。