

来源：量子比特

大模型模式正在新一波AIGC中再次得到验证。

从AI绘画的出圈到现在ChatGPT的流行，面向大众的爆款产品界面背后，无一不是大模型技术的突破和创新。

当这个技术路径的值为“努力创造奇迹”变得越来越突出，行业内外的都可以不禁好奇：继

AI绘画和ChatGPT版本搜索之后，下一个大模型的爆发点会出现在哪里？

遵循技术规律推测，可以预见的是视频领域的技术革命指日可待：

从技术角度来看，在大语言模型迭代进化的同时，Google、Meta等大公司在视频自动生成领域进行了更深层次的探索。从商业角度来看，中信建投在报告中指出，AIGC在文字、音频、视频、游戏等行业有巨大的增长空间。

量子比特智库也预测，五年后AI生成的视频将被广泛应用。

△图源：量子位智库

AIGC下一站：视频生成

Let’s；让我们先看看相关领域的技术进展。

Google:Phenaki和ImagenVideo

正当ChatGPT在刷屏的时候。GoogleAI生成的一个视频突然爆了，一时间引发了很多讨论。

背后的AI模型叫做Phenaki。这种文本到视频的模式只需要提供一个提示，就可以在几分钟内生成长达两分钟的视频。。

不仅是长期超早期的Vincent视频模型，Phenaki生成的视频也相当有故事。

比如给它一个场景描述：一只逼真的泰迪熊在潜水；然后慢慢浮出水面；去海滩；

这是放大图。泰迪熊走过海滩上的篝火。

你可以得到这样一个视频剪辑：

除此之外谷歌也推出了基于扩散模型的Imagen视频。其特点是分辨率高，可以理解不同的艺术风格和3D结构。

Meta : Make-A-Video

根据“用嘴制作视频”，Meta也做了一些安排。

Meta’s文生视频模型称为Make-A-Video，也是文本图像生成模型的升级版，主要由三部分组成：

文本图像生成模型P，时空卷积层和注意层用于提高帧率，两个超模型

用于提高图像质量。制作视频不仅给出了一句“马饮水”，但是可以生成纪录片图片：

这个AI模型还具有将静态图像转换为视频，基于前后两张图片生成视频，以及基于原始视频生成新视频的能力。

百度 : VidPress

在国内，百度也已经将文心大模型的能力应用于智能视频合成平台VidPress。

VidPress可以实现文本到视频的自动转换，即文本脚本、视频内容搜索、素材处理和音视频对齐。以及五个编辑步骤自动化。

语义分析、素材相关性评分等等涉及的环节都是基于文心大模型的训练。

从技术的发展可以看出继语言模型和图像模型之后，多模态模型已经成为一种新的趋势。视频就是其中一个有代表性的应用领域。

从商业化的角度来看，文化、娱乐、教育、传媒等诸多领域。对基于人工智能的视觉内容有着强烈的需求。

根据中信建投's计算未来各类可以AI制作的内容比例，当视频成为信息的主要表达载体，无论是在游戏、短视频、直播、电影等领域，AI视频内容生成将成为AIGC的主要关注点。

并且在2022年在"轰炸"在DALLE、Imagen、StableDiffusion等众多高质量的大模型中，一个新的市场规律得到了验证：

当生成的图像质量提升到专业水平时，原有的AI生成内容的商业化瓶颈，比如变现难度，将被打破。

哪些公司是潜力股？

机会自然属于有准备的人。

比如图像视频的核心团队，早已离开谷歌创业。。据VC爆料，首轮估值1亿美元。

除了上面提到的拥有革命性技术储备的科技巨头，还有两类公司值得关注。

第一，是有数据基础和应用场景的公司。

在这方面，行业内的大公司一般更有优势。。比如国外的网飞和迪斯尼。以网飞为例。公开资料显示，2012年，网飞有数十亿的会员评论，每天可以添加上百万的视频播放信息，包括观众'的观看时间，播放设备等等。

事实上，基于如此庞大的独家数据，网飞试图在产品中用AIGC取代标准内容制作。例如，电影的缩略图是网飞使用AI算法从电影中提取符合用户的图片生成的'的观看习惯和需求。

就在今年2月初。网飞还发布了一部AIGC动画短片《犬与少年》。其中动画场景的绘制全部由AI完成。

同样国内短视频行业的两大巨头TikTok和AautoQuicker的动向也值得关注。

目前，字节跳动的视频编辑工具和AautoQuicker的云编辑工具已经上线，具有图文转视频的功能。用户只需要输入几个关键词或者一段文字。，人工智能可以自动收集素材并编辑视频剪辑。

其次是在垂直细分赛道上有技术储备的公司。

比如国内智能视觉技术企业影谱科技，就是以人工智能视觉技术产业化为主要目标成立的。早在2018年就发布了基于生成式AI技术的AGC智能图像生成引擎，这也是国内较早提出的生成式AI的技术框架。

影谱技术的AGC使用MCVS(来自视频系统的电影书籍运动捕捉)来提取、理解、关联和预测现有视频的关键帧。，将视频内容分为像素和亚像素维度的结构化数据，自动完成标注，形成对视频中各种内容的自动理解和标注。接下来，利用计算机图像技术如三维虚拟重建。Mape(电影书自动制作引擎)生成一个全新的视频，集成了人工智能多模态语义理解。并利用深度学习实现自动视频处理和同步视频生成，创新了视频图像的制作模式。

这使得AGC可以在极短的时间内生成个性化的视频内容，还可以对拍摄的视频进行重构，比如自动锚定关键帧。根据帧内容，生成不在原图中、没有矛盾感的内容，进而智能生成一段AI视觉内容。

公开数据显示，影科科技AI生成引擎生成一个60s视频的总成本比传统方式降低了79.8%。最高生产率可提高100倍以上；检索一个60s视频中相似帧图像或特定图像的总成本比人工降低99.73%，错误率降低10倍以上。目前主要应用于政府服务、企业、科教、泛娱乐、传媒、文化旅游等领域。

例如在视频采集制作阶段，可实现主体识别、跟踪拍摄、画质修复、自动剪辑、视频自动生成等功能；在分发阶段实现智能审核和个性化推荐；在用户体验方面，融入了数码孪生技术。实现数字内容、数字空间、数字人的高效生产和视觉交互的功能。

目前，AGC已经在很多行业得到了广泛的应用。随着数字化与产业融合的深入，未来其作用会越来越明显。

P.S.上面提到的网飞AIGC的短片，还有萧冰公司的参与。

ChatGPT的大火正在加速AIGC的成熟

不得不说，尽管chatGPT的火爆大模特时代的AIGC获得了前所未有的关注，但AIGC行业只能说是“一个新生儿”。

根据Gartner发布的2022年新兴技术成熟度曲线，生成式AI仍处于“技术萌芽阶段”目前。

估计还有5-10年才生产成熟。

但是随着大模型等技术基础作用的日益凸显，也越来越受到重视。不可否认，这种新生的增长趋势正呈现加速状态。

而且已经得到了文生图和ChatGPT的验证。AIGC在其他垂直领域的可拓展空间已经可以预见。

量子比特智库预测，2030年，AIGC市场规模有望突破万亿人民币。

在这个过程中，如何抢占先机将成为中国工业界迫切需要关注的一个命题。

至少，像盈普这样的垂直赛道种子选手已经得到了国有资金的关注：他们有全国社保基金、央企、国有银行、地方国资的背景。并且D轮融资13.6亿元。创下中国人工智能视觉制作领域最高融资纪录。

这次我们能保持领先世界吗？

—结束—