证券研究报告

海外行业报告 | 行业专题研究

2023年10月15日

生成式AI下一颗最闪亮的星：视频+引擎

作者：

分析师 孔蓉 SAC执业证书编号：S1110521020002

分析师 李泽宇 SAC执业证书编号：S1110520110002

天风证券

[综合金融服务专家]

摘要

1）内容创作工具的重要性如何？我们认为关键在于拉动远比工具本身更大的市场，类似于短视频时代的前身GIF创作工具，如剪映这种创作工具和抖音这种内容生态，Maya这种创作工具和动画的内容生态，视频与建模工具将进一步大幅拉动生成式AI的需求。

2）能力或出现明显分化。我们认为当前Diffusion开源模型下各公司生成图片的能力尚未明显出现分化，但建模和视频更重要的在于和传统工具与工作流的结合 。

3）接下来半年关注什么？我们认为从能力来看，图片生成的可控性快速提高或将出现，矢量图、结构、视频、3D模型生成也将提速。尤其关注Unity与Apple的合作，在AI生成内容/建模/App适配上将会如何塑造空间计算的内容与应用的标准生态。

风险提示：生成式AI发展不及预期；算力成本及硬件发展不及预期；相关应用产品上线后效果不及预期。

目录

1、生成式AI在视频/3D/游戏等领域的渗透加速

2、生成式AI下游应用场景展望

3、风险提示

1.1 生成算法模型不断突破创新，下游应用场景不断拓展

基础的生成算法模型不断突破创新，基础能力日新月异，从图像向视频和3D扩展，更广泛地应用于下游应用场景

生成对抗网络（GAN）是早期最著名的生成模型之一，尽管在图像生成上产生了卓越的效果，但其训练常常受到梯度消失和模式崩溃等问题的影响。与GAN相比，**扩散模型（Diffusion Model）**只需要训练“生成器”，不需要训练别的网络（判别器、后验分布等），训练时仅需模仿一个简单的前向过程对应的逆过程，实现简练过程的简化。**扩散模型相对GAN来说具有更灵活的模型架构和更精确的对数似然计算，生成图像质量明显优于GAN，已经成为目前最先进的图像生成模型。**

此前扩散模型主要适用于生成2D图像，**23年Runway的最新研究成果将扩散模型扩展到视频领域**，在未加字幕的视频和配对的文本-图像数据的大规模数据集上训练出视频扩散模型。

NeRF（神经辐射场）的出现为3D场景生成带来了新的可能性，进一步拓宽生成算法领域下游的应用场景。NeRF（Neural Radiance Field）是一种基于神经网络的 3D 重建技术，不同于传统的三维重建方法把场景表示为点云、网格、体素等显式的表达，NeRF将场景建模成一个连续的 5D 辐射场隐式存储在神经网络中，输入多角度的 2D 图像，通过训练得到神经辐射场模型，根据模型渲染出任意视角下的清晰照片。

图：生成式算法模型对比

	GANs	Diffusion	NeRF
原理	生成对抗网络（GAN）是一种深度学习模型，通过同时训练两个相互对抗的网络（一个生成器网络和一个判别器网络）来生成与真实数据分布相似的新数据	扩散模型是一种概率生成模型，通过添加噪声逐步解构数据，然后学习逆转扩散过程来生成样本。	基于2D图像输入，将3D场景展现为一组可以学习且连续的神经辐射场，不直接生成3D模型，而是由输入视角+位置来生成密度+色彩信息，从而生成新视角的模型
优势	能够生成较高质量的样本，一般只需要一次通过网络就可以生成一个样本，比较快速	只需训练生成器而无需训练判别器，能够生成细节清晰的数据样本，质量明显优于GANs模型	更准确还原3D场景中细节和颜色，比网格和其他几何表征更容易优化
缺陷	GAN的训练过程需要生成器和额外的判别器，且稳定性较差	扩散模型的训练缓慢且计算量密集，且需要大量的数据进行有效训练	训练复杂，且无法对生成的场景进行直接编辑
应用场景	GANs常用于图像生成、图像超分辨率、风格迁移等任务	可用于生成建筑方案，游戏人物、场景设计	游戏，电影和虚拟现实： 可用于创建高度逼真的虚拟世界 建筑和城市设计：可用于创建比真的建筑模型并实现可视化效果

资料来源：NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis（B Mildenhall等），Generative AI meets 3D: A Survey on Text-to-3D in AIGC Era（C Li 等），量子位公众号，品览pinlan公众号，腾讯云开发者社区，datagentech等，天风证券研究所

1.2 Runway：生成式AI内容平台，Gen-1可实现用文字和图像从现有视频中生成新视频

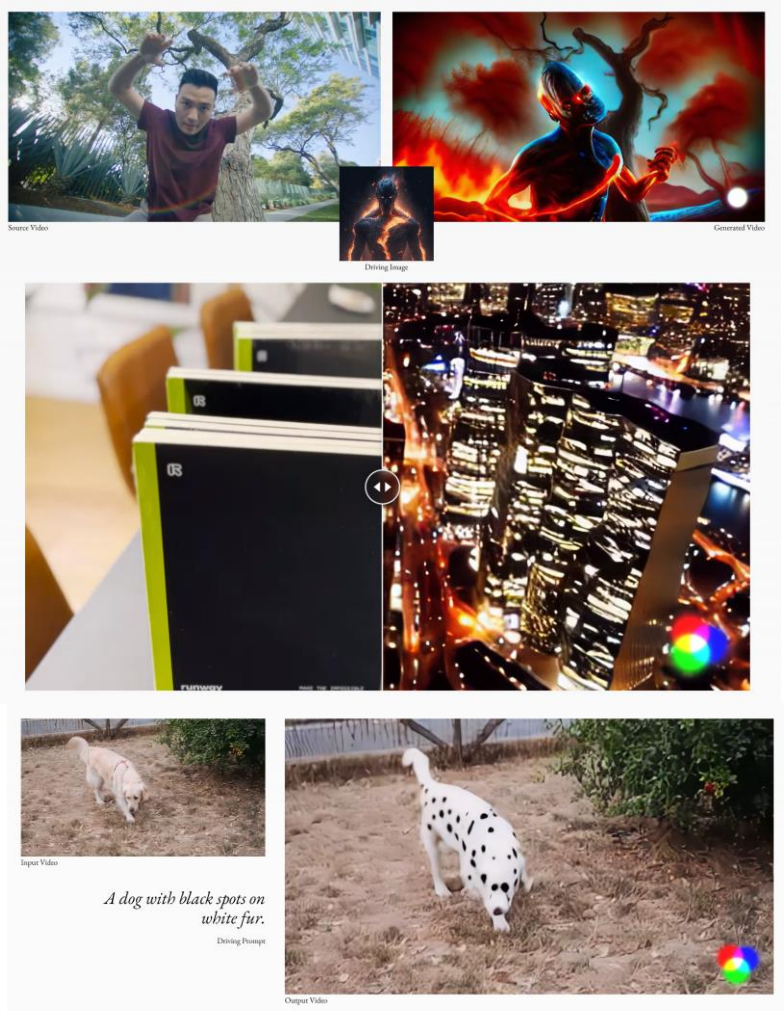
Runway是一家生成式AI内容平台，致力于让所有人都能进行内容创作。Runway创立于2018年，总部位于纽约，提供图片、视频领域的生成式AI服务。**Runway得到众多资本青睐，获得谷歌领投的D轮融资。**创立以来，Runway已获得Felicis、Coatue、Amplify、Lux、Compound等顶级投资机构投资。23年6月，Runway获得由谷歌领投的1亿美元的D轮融资，这笔融资交易包括三年内7500万美元的谷歌云积分和其他服务，估值达到15亿美元。

主要产品为Gen-1和Gen-2，Gen-1可实现用文字和图像从现有视频中生成新视频，Gen-2可实现文字生成视频功能。
Gen-1：不需要灯光、相机和动捕，通过将图像或文本提示的结构和风格应用于源视频的结构，逼真且一致地合成新视频，且具有表现力、电影感和一致性。

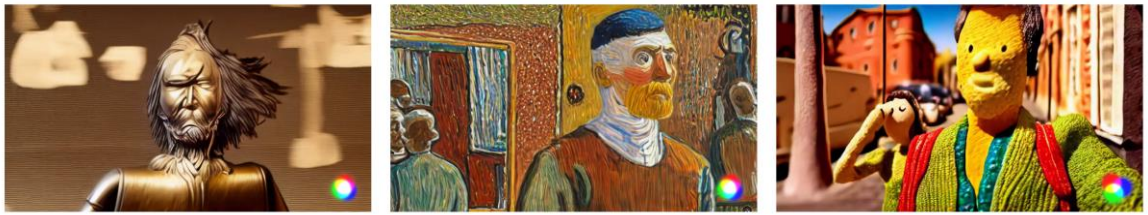
Gen-1提供给用户五种不同的视频制作模式：1）Stylization：将任何图像或提示的风格转移到视频的每一帧；2）Storyboard：将模型变成完全风格化和动画的渲染。3）Mask：隔离视频中的主题并使用简单的文本提示对其进行修改；4）Render：通过应用输入图像或提示，将无纹理渲染变成逼真的输出；5）Customization：通过自定义模型以获得更高保真度的结果，释放 Gen-1 的全部功能。

Gen-1的性能优势：基于用户研究，GEN-1 的结果优于现有的图像到图像和视频到视频的转换方法，比Stable Diffusion 1.5 提升 73.83%，比 Text2Live 提升 88.24%。

图：Gen-1三种模式演示：Stylization（上）、Storyboard（中）、Mask（下）



图：Gen-1和Gen-2生成的作品



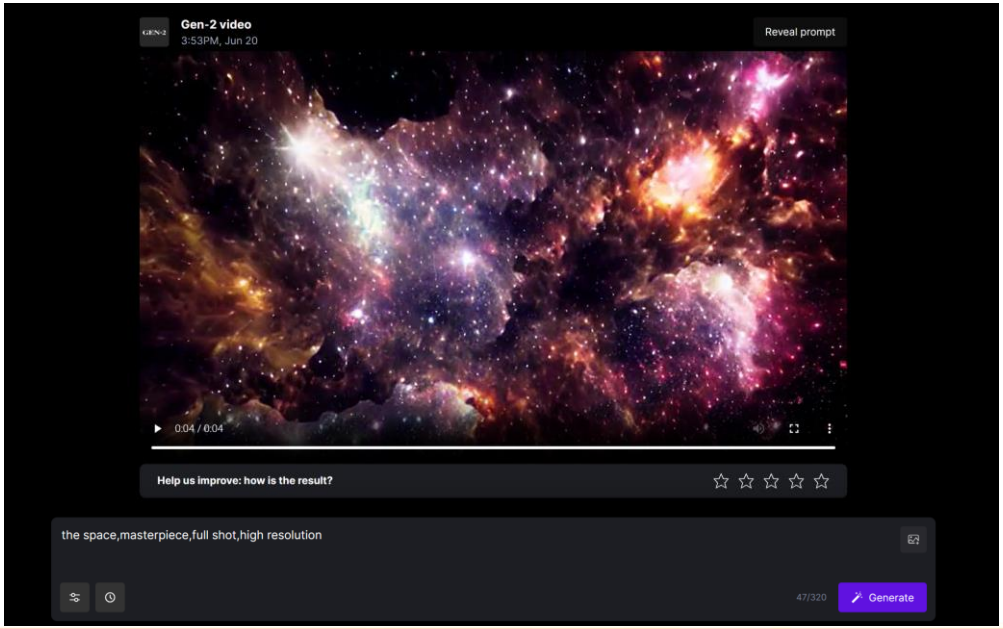
1.2 Runway：生成式AI内容平台，最新产品Gen-2可实现文字生成视频功能

Gen-2是一个多模态的人工智能系统，可以用文字、图像或视频片段生成新颖的视频。Gen-2在Gen-1的基础上迭代，保留通过将图像或文本提示的结构和风格应用于源视频的结构合成新视频的功能，**新增了只用文字便可生成视频的的功能。**

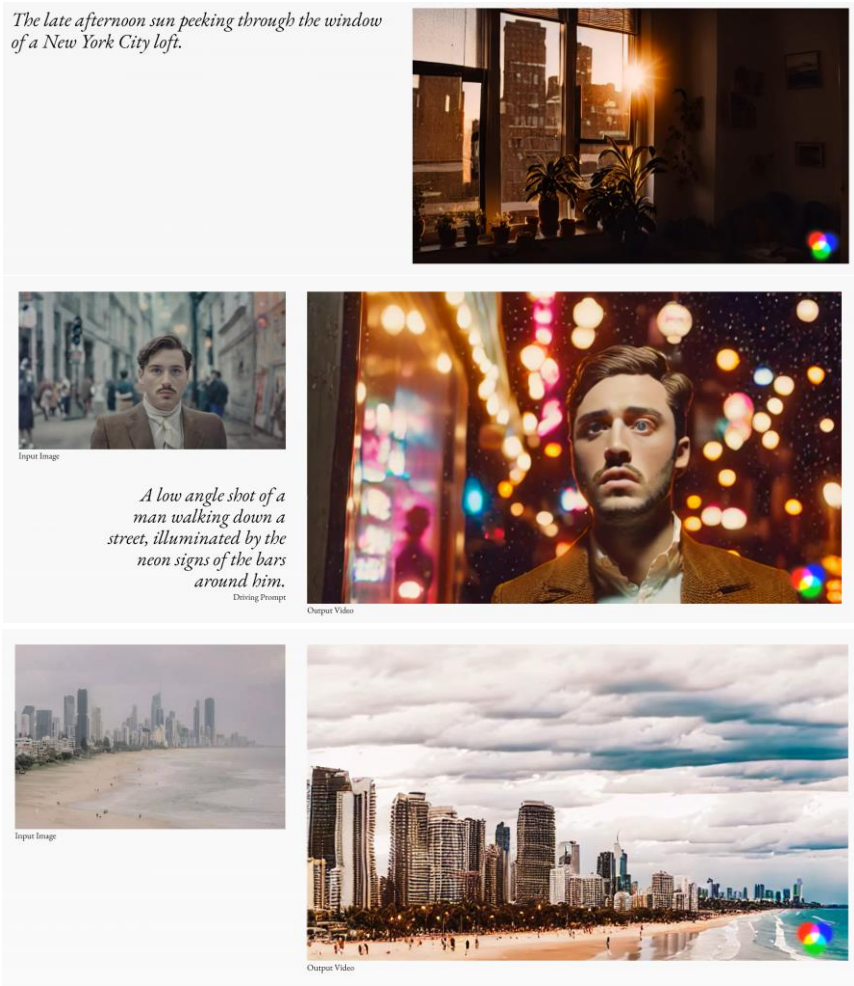
Gen-2在Gen-1的五种视频制作模式上新增了三项新的模式：1）Text to Video：**仅通过文本提示合成任何风格的视频；**2）Text + Image to Video：**使用图像和文本提示生成视频；**3）Image to Video：**用一张图片生成视频（变体模式）。**

Gen-2已于2023年6月上线，用户可以通过网页端和移动端App免费体验文字生成视频的功能。收费模式上，Runway采用订阅模式，分为Standard和Pro两个版本：Standard 15美元/月；Pro 35美元/月。订阅版本提供更高的credits（制作视频消耗credits）、更长的视频长度、更高的分辨率等。

图：Gen-2 创作界面：直接通过文字生成约4秒的视频



图：Gen-2新增的三种模式演示：Text to Video（上）、Text + Image to Video（中）、Image to Video（下）



资料来源：Runway官网，天风证券研究所

1.2 Runway技术路径剖析：Gen-1基于扩散模型的视频生成技术

Runway的Gen-1是一种基于扩散模型（Diffusion Model）的视频生成技术。用户可以通过文字和图像来生成新的视频，同时保留现有的视频结构和内容。
扩散模型是一种概率生成模型，通过添加噪声逐步解构数据，然后学习逆转扩散过程来生成样本。

- 去噪扩散概率模型利用两个马尔科夫链：一个前向的链将数据扰动为噪声，一个后向的链将噪声还原为数据。前者通常为手动设计，旨在将数据分布转换为一个简单的先验分布（例如，标准高斯分布）
- 而后者的马尔科夫链通过学习由深度神经网络参数化的转换核来逆转前者。新的数据点随后通过首先从先验分布中抽样一个随机向量，然后通过向后马尔科夫链进行祖先抽样来生成。

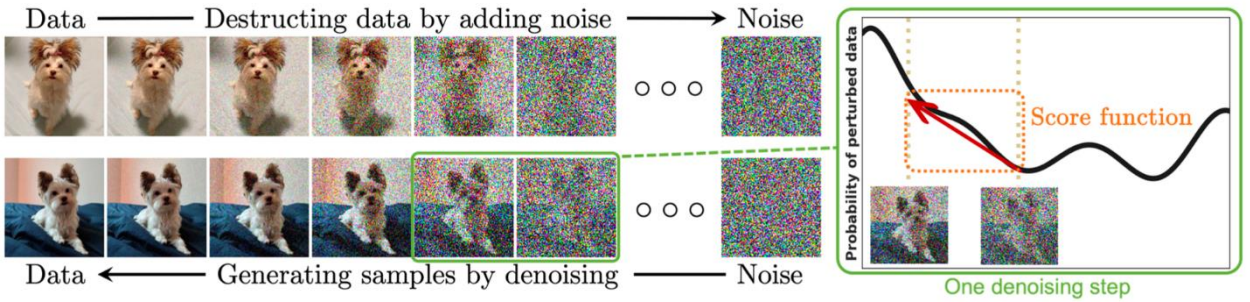
扩散模型的技术优势：

- 可以生成高质量、细节清晰的数据样本
- 使用成熟的最大似然估计进行训练
- 既可以用于生成图像，也可以用于生成音频

扩散模型的技术缺陷：

- 扩散模型的训练缓慢且计算量密集，且需要大量的数据进行有效训练
- 这些模型需要进行多次迭代才能生成高质量的样本，因为生成过程（从噪声到数据）模拟了前向扩散过程（从数据到噪声）的逆过程，这可能需要数千步

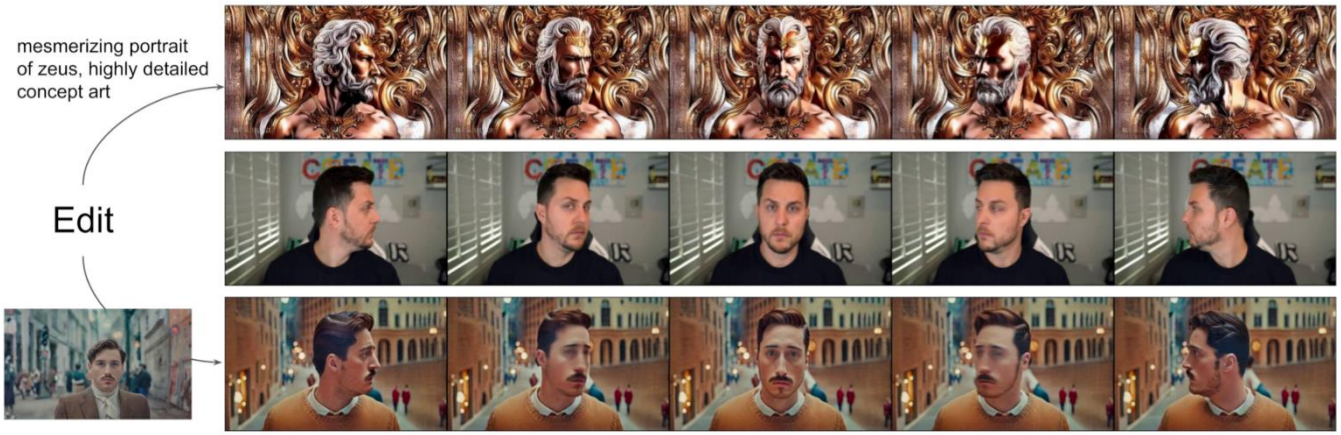
图：扩散模型（Diffusion Model）通过添加噪声对数据进行平滑扰动，然后反转这一过程来生成新数据。



1.2 Runway技术路径剖析：Gen-1-基于扩散模型的视频生成技术

Gen-1提出了一个可控的结构和内容感知的视频扩散模型，将时间层引入预训练的图像模型，将扩散模型扩展到视频生成，在大规模的未标注视频和配对的文本-图像数据上进行训练。
Gen-1用单眼深度估计来表示结构，用预训练神经网络预测的嵌入来表示内容。编辑完全在推理时进行，无需额外的每个视频训练或预处理。
Gen-1实现了对时间、内容和结构一致性的完全控制，首次展示了在图像和视频数据上联合训练可以在推理时控制时间一致性。对于结构一致性，Gen-1在生成效果中的不同细节级别上进行训练，实现高度自定义化推理过程。

图：Gen-1在保持输入视频（中间）的结构的同时，通过文本（上）或图像（下）描述的内容来指导视频（上和下）的合成。



Gen1视频生成技术的实现路径

- 首先，类似于图像合成模型，Gen-1训练模型使得推断出的视频的内容，如外观或风格，匹配用户提供的图像或文本提示（见上图）。
- 其次，由扩散过程主导，Gen-1将结构信息模糊化，以便模型选择以多大程度贴合给定的结构
- 最后，Gen-1通过自定义的推断方法，调整推理过程，以便控制生成片段时间的一致性。

资料来源：Structure and Content-Guided Video Synthesis with Diffusion Models（P Esser等），Diffusion Models: A Comprehensive Survey of Methods and Applications（L Yang）等，天风证券研究所

1.3 Luma AI：3D内容解决方案平台，基于NeRF上线文字转3D和视频转3D功能

Luma AI是一家3D内容解决方案平台。Luma AI创立于2021年，总部位于加州。公司创始人在3D视觉、机器学习、实时图形学上有所建树：CEO & Cofounder Amit Jain曾于苹果任职，在3D计算机视觉、摄像头、ML、系统工程和深度技术产品方面有者丰富经验； CTO & Cofounder Alex Yu致力于推动神经渲染和实时图形领域的发展，研究成果包括Plenoxels、PlenOctrees和pixelNeRF。

Luma AI深耕3D领域，也发布多项3D生成产品。22年10月开放网页版Luma；22年12月推出文生3D模型功能；23年1月iOS版App开始支持NeRF Reshoot；23年2月推出网页版全体积NeRF渲染器；23年3年iOS版App支持AR预览，同月推出视频转3D API。23年4月发布Luma Unreal Engine alpha，帮助开发者在Unreal 5中进行完全体积化的渲染，无需对几何结构或材质进行修补。

主要产品：

Luma App：目前只推出iOS客户端，可以通过iPhone上传视频，基于NeRF生成3D场景。Luma App支持导入视频，以及引导模式和自由模式三种：导入模式，和Web模式功能类似，对设备和视频理论上要求最低；引导模式，需要360度拍摄，App将具体提示框提醒拍摄视角、拍摄位置；自由模式，支持非360度（部分视角）拍摄，App不会给出明确提示框，需要尽可能拍摄多个角度。

网页端：目前集成了三大主流功能：网页版Luma、文字转3D模型、视频转3D API。**网页版Luma：**上传照片、视频来进行三维重建，网页版可以上传更大的文件，目前视频和图片（ZIP压缩包）体积最大限制5GB；**文字转3D模型：**输入文字描述生成对应的3D模型。**视频转3D API：**效果基本和网页版一致。收费模式为按次收费，转换一个视频费用为1美元，转换时间在30分钟左右。

图：Luma AI产品发布时间线

图：Luma AI 视频转3D功能演示

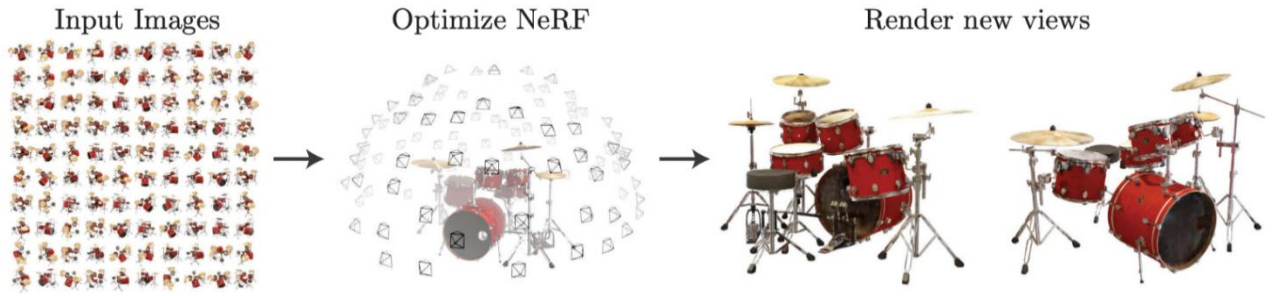
图：Luma AI 文字转3D模型产品演示

资料来源：Luma AI官网，青亭网公众号，XRtoday，天风证券研究所

1.3 Luma AI技术路径剖析：NeRF-基于神经网络的隐式3D场景展现法

NeRF，即神经辐射场（Neural Radiance Field），是一种基于神经网络的隐式3D场景展现法。基于输入的2D图像，NeRF能够生成和渲染逼真的3D场景。NeRF可以从任何新视角生成2D图像，而无需生成完整的传统3D模型

图：NeRF通过周围半球上随机捕获的100个架子鼓的输入视图，呈现了架子鼓的两个新视图。

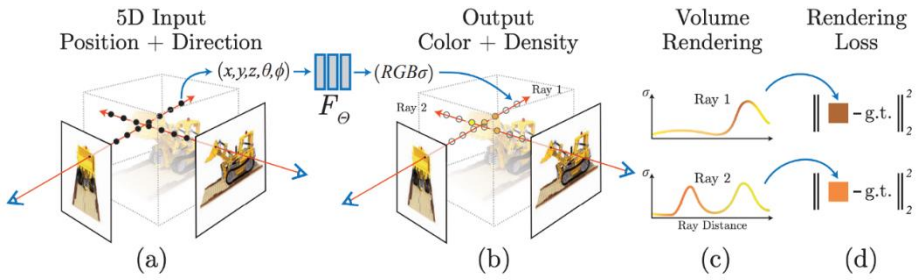


技术原理：

- 通过场景投射相机光线以生成一组采样的3D点
- 使用这些点及其对应的2D观察方向作为神经网络的输入，以输出一组颜色和密度值
- 使用体渲染技术将这些颜色和密度累积成任意角度的2D图像

该过程自然可微，可以采用梯度下降的方式，即最小化每个观察到的图像和表示的相应视图之间的误差来优化这个模型。在多视图中最小化这个误差即可以预测出一个连贯的场景模型，将体积密度和准确的颜色分配给涵盖真实底层场景内容的位置。

图：NeRF场景表示和可微分渲染程序的概述



（a）通过沿着相机光线采样坐标（位置和视觉方向）来合成图像；
（b）将这些位置输入到一个MLP中以产生颜色和密度；
（c）使用体渲染技术将这些值合成为一个图像；
（d）通过最小化合成图像和真实观察图像之间的残差来进行优化。

1.3 Luma AI技术路径剖析：NeRF-基于神经网络的隐式3D场景展现法

与传统的摄影测量技术需要从各个角度获取重叠图像以建模相比，NeRF只需要少量照片便可构建3D场景。它能够通过训练小型神经网络来重建场景，并通过预测3D空间中任何位置的任何方向的光线颜色来自行填补空白。除此之外，NeRF技术具有一系列其他优势：

- **高质量建模：**NeRF具有高质量和逼真的三维模型生成的优势，在任何角度和距离都能呈现出真实的物体表面和纹理细节。
- **不需要输入数据的提前处理或标记：**NeRF可以从任何数量的输入图像中生成三维模型，且不需要对输入进行特定的处理或标记。
- **可以在低功率设备上运行：**神经网络经过训练后可以在低功率设备上运行：但高质量的神经场可以在手机甚至网络浏览器上进行渲染。对比来看，多边形光线追踪（Polygon Ray tracing）以高帧率渲染高分辨率和真实的场景，需要昂贵的图形卡。

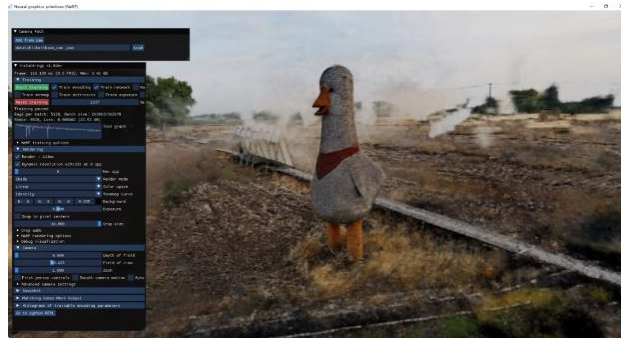
NeRF的技术缺点：

- **生成时间较长：**NeRF需要大量的计算资源和时间进行训练；
- **无法应用在复杂、大规模场景：**NeRF难以处理大规模的场景和复杂的照明条件；
- **无法直接用于3D资产生成：**NeRF不能生成结构化的数据，难以直接应用于3D资产；

Luma AI与Epic合作将NeRF引入了Unreal Engine实现实时渲染，可应用于游戏和电影应用场景

- Unreal Engine是一款强大的游戏开发引擎，被广泛用于视频游戏制作, 电影和电视业、建筑可视化、实时渲染等领域。
- Luma Unreal Engine插件使用户可以将这些Luma Field文件导入Unreal Engine 5，该插件可以将这些Luma Field文件导入到虚幻引擎中作为环境使用，自动生成一组Blueprints来照亮并显示这些捕获物。
- 用户可以从捕获的场景中选择照明或从虚幻引擎场景中选择动态照明。其他蓝图会自动裁剪捕捉到的中心物体，并生成一个天空环境。
- 目前，插件是作为两个虚幻引擎项目样本的一部分提供的，一个是为游戏和实时项目定制的，另一个是为电影和电影项目定制的。

图：NeRF场景重建后导出至Unreal Engine再创作



资料来源：Generative AI meets 3D: A Survey on Text-to-3D in AIGC Era（C Li等），大智营城公众号，cgchannel，天风证券研究所

1.4 Unity：制作和运营交互式实时 3D (RT3D) 内容平台，结合AI大模型赋能游戏业务

Unity是一家**全球领先的制作和运营交互式实时 3D (RT3D) 内容的平台**，也是**全球最大的游戏引擎公司**。收购ironSource之后，其主营业务包括**与开发相关的引擎类产品Create**和**与广告营销相关的产品Grow。**

Unity 成立于 2004 年，起初为 Over the Edge Entertainment 并进行游戏开发工作，2005 年公司在游戏开发基础上转型工具， 并于 2005 年发布 Unity1.0 版本。20余载，Unity 先后登陆并支持苹果IOS平台、OS平台、Windows平台等，伴随着iPhone以及整个移动互联网的发展，Unity迎来用户数量的快速增长。同时，经过长期的迭代升级以及并购，公司逐步建立起游戏以及其他领域的业务，形成当前公司的主要业务架构，实现全平台全产业链覆盖的高兼容特性。

2023年，公司发布AI产品：Unity Muse、Unity Sentis，宣布结合AI大模型赋能游戏业务。

主要产品：

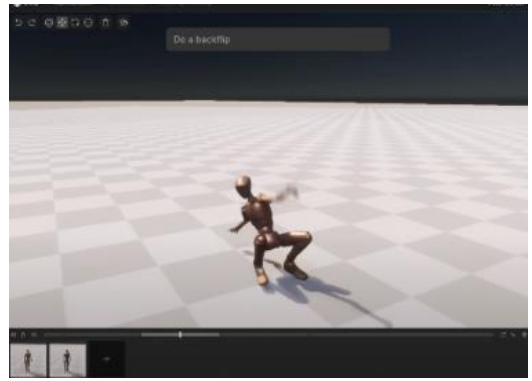
Unity Muse：提供AI驱动协助的扩展平台，它加速了像视频游戏和数字孪生这样的实时3D应用和体验的生成。在Muse上，用户能够通过自然语言在Unity编辑器中开发游戏，打破技术壁垒。

Unity Sentis：嵌入神经网络，解锁全新实时体验。在技术层面， Unity Sentisl连接神经网络与Unity Runtime，因此，AI模型能够在Unity运行的任何设备上运行。Sentis是第一个也是唯一一个将AI模型嵌入到实时3D引擎中的跨平台解决方案。Sentis在用户的设备而非云端运行程序，因此其复杂性、延迟和成本都大大降低。

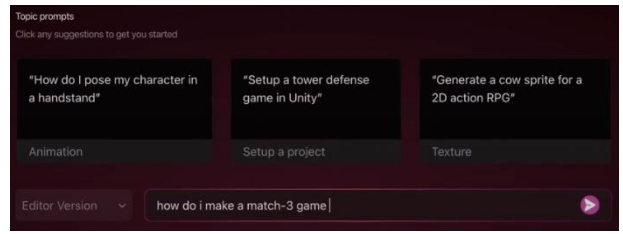
Muse Chat：基于AI，用户可以搜索跨越Unity文档、培训资源和支持内容，以获取来自Unity的准确且最新的信息。Muse Chat能够帮助开发者快速获取相关信息，包括工作代码样本，以加速开发进程和提供解决方案。

Unity Asset Store：Unity**与第三方展开了紧密合作**，在Unity Asset Store中为用户提供高质量且兼容的第三方AI解决方案，使用户能够无缝使用各类工具。其供应商包括Leonardo AI，Replica， Atlas等，涵盖图像生成，人工智能对话，以及投资领域等。

图：Unity Muse模型执行自然语言“后空翻”指令



图：Muse Chat对话界面



图：Unity与第三方合作情况

资料来源：Unity官方公众号，ironSource官网，Unity 官网，Unity Blog，Unity Bilibili账户，界面新闻，人民日报客户端，天风证券研究所

1.4 Unity：制作和运营交互式实时 3D (RT3D) 内容平台，与苹果合作开发Apple Vision Pro

2023年6月，Apple 发布了旗下首款**MR（混合现实）眼镜Apple Vision Pro**。这是一台革命性的空间计算设备，突破了传统显示屏的限制，为用户带来全新的 3D 交互体验。Vision Pro 搭载了**全球首创的空间操作系统 visionOS**，该系统通过用户与数字内容互动的模式，让数字内容如同存在于真实世界。

Unity 作为本次大会上公布的 Apple visionOS 原生合作方之一，为全新发布的 Apple Vision Pro 提供了被开发者熟知且性能强大的创作工具，用于创建沉浸式游戏和应用，这意味着除了 **AR Foundation 和 XR Interaction Toolkit** 等广为人知的功能外，开发者还能在自己的应用中加入**诸如 Pass-through（穿透）和 Dynamic Foveated Rendering**（动态注视点渲染）等功能。

通过 Unity 新的 **PolySpatial 技术与 visionOS 之间深度集成**，应用程序可以在共享空间（Shared Space）中与其他应用程序一起运行。并且通过将 Unity 的创作工具与 RealityKit 的托管应用渲染相结合，开发者将能轻松使用 Unity 在 Apple Vision Pro 上创作内容。

图：Unity XR 交互系统

图：Apple Vision Pro产品介绍图

“Apple Vision Pro 重新定义了计算平台的可能性。开发者可使用他们熟悉的强大框架着手构建 visionOS app，并利用 Reality Composer Pro 等新的创新工具和技术进一步推进开发工作，为用户设计全新体验。”Apple 全球开发者关系副总裁 Susan Prescott 表示，“空间计算技术利用用户身边的空间为开发者解锁了全新机遇，支持他们构想全新方式帮助用户联络彼此、提升效率、享受新型娱乐。我们迫不及待地想见证开发者社区的奇思妙想。”

资料来源：Apple官网，Unity官网，天风证券研究所

1.5 Open AI：3D生成技术Point-E与Shap-E的更新迭代

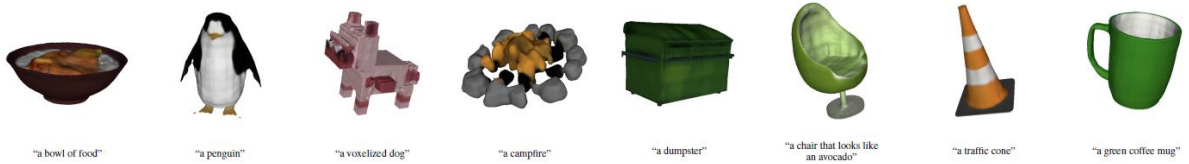
Point-E是一个3D模型生成器，可以在几分钟内生成3D图像。Point-E是一个机器学习系统，可以通过文本输入制作3D物体，由OpenAI于2022年12月发布到开源社区。Point-E本身包括两个模型:GLIDE模型和image-to-3D模型。前者类似于DALL-E或Stable Diffusion等系统，可以从文本描述生成图像。第二个模型由OpenAI使用图像和相关的3D物体进行训练，学习从图像中生成相应的点云。

图：使用给定的文本提示由Point-E生成的点云

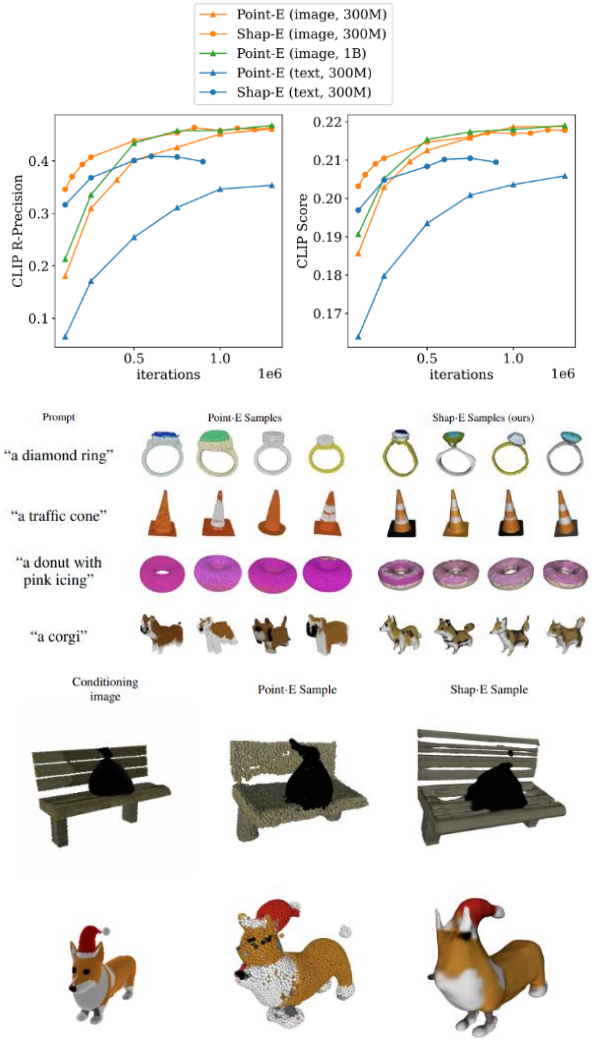
2023年5月，OpenAI再次发布了一款升级模型Shap-E。Shap-E的训练分两个阶段:首先，训练一个编码器，生成隐式表征，然后，在编码器的输出上训练一个条件扩散模型。

相比基于点云的显式生成模型Point-E，Shap-E直接生成隐函数的参数来渲染纹理网格和神经辐射场，收敛速度更快，在更高维的多表示输出空间中实现了更好的样本质量。

图：使用给定的文本提示由Shap-E生成的条件网格



图：Point-E与Shap-E对比图



资料来源：Point · E: A System for Generating 3D Point Clouds from Complex Prompts（Alex Nichol等），Shap · E: Generating Conditional 3D Implicit Functions（Heewoo Jun&Alex Nichol），天风证券研究所

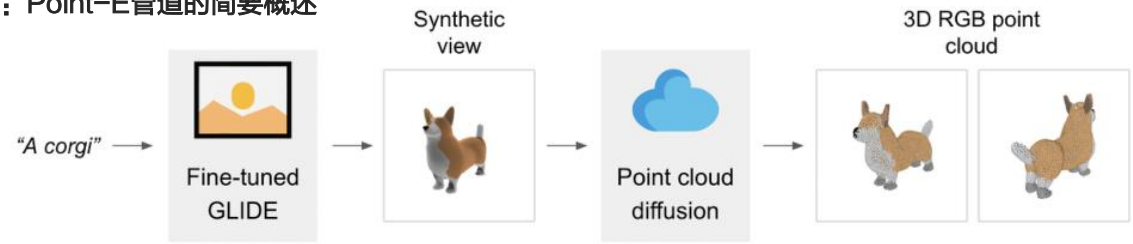
1.5 Open AI技术路径剖析：Point-E-从复杂提示生成3D点云的系统

Point-E：基于目前主流的两种text-to-3D模型进行融合尝试，结合了文本-图像模型与图像-三维模型这两类方法的优点。使用文本到图像的扩散模型生成一个单一的合成视图，然后使用第二个扩散模型生成一个三维点云，该模型以生成的图像为条件。Point-E在采样质量方面达不到最先进的水平，但它的采样速度要快一到两个数量级，为一些用例提供了实用的权衡。

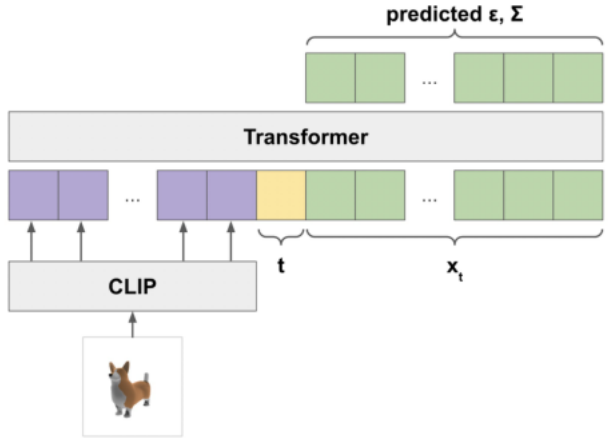
生成步骤：

- 首先，对来自数据集的渲染3D模型进行了微调，使用30亿参数的GLIDE模型，生成一个以文本标题为条件的合成视图。
- 接下来，使用一个有条件的、排列不变的扩散模型，在合成视图的基础上生成一个低分辨率点云(1024个点)。
- 最后，在低分辨率点云和合成视图的条件下，生成一个精细点云(4096个点)。

图：Point-E管道的简要概述



图：Point-E的点云扩散模型架构



Point-E模型的技术优势：

- Point-E可以在短时间内产生样本，较此前的先进技术快了数个数量级，具有较高的实用性。
- Point-E的两个步骤可以在几秒钟内完成，不需要昂贵的优化程序。
- Point-E假设图像包含来自文本的相关信息，无需明确地限制文本上的点云。

Point-E模型的技术缺陷：

- 目前，管道需要合成渲染。
- 当Point-E产生彩色的三维形状，它以相对较低的分辨率以3D格式(点云)完成，不能捕获细粒度的形状或纹理。
- 模型具有一定的偏差，如DALL · E2系统，其中的许多偏差继承自数据集。

图：Point-E模型被误用的例子

"a 3D printable gear, a single gear 3 inches in diameter and half inch thick"

1.5 Open AI技术路径剖析：Shap-E-隐式3D生成模型

在Shap-E的模型架构设计中，首先训练一个编码器来生成隐式表征（implicit representation），然后在编码器产生的潜表征（latent representation）上训练扩散模型。生成步骤：

➢ **3D编码器：给定一个已知三维资产的稠密显式表征，训练一个编码器来生成隐式函数的参数。**

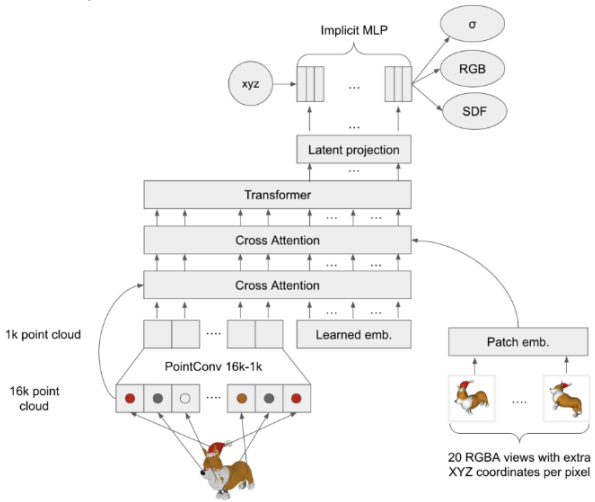
如下图编码器架构所示，给编码器提供点云和三维资产的渲染视图，输出多层感知器（MLP）的参数，将资产表示为一个隐式函数。Shap-E只用NeRF渲染目标对编码器进行预训练，因为研究人员在实验中发现这个操作比基于网格（mesh-based）的目标更稳定，而且可以进行优化。

➢ **潜扩散（Latent Diffusion）：对于生成部分，采用Point-E的基于Transformer的扩散结构，并用潜向量序列替代点云。**

潜表征为shape 1024×1024的序列，可以将其作为1024个token的序列送入Transformer，其中每个token对应于MLP权重矩阵的不同行。因此，该模型在计算上与基本的Point-E模型大致相当（即具有相同的上下文长度和宽度），同时由于输入和输出通道的增加，在一个更高的维度空间中生成样本。

文本条件下，Shap-E在CLIP R-precision和CLIP分数两个指标上都比Point-E模型有所提高。图像条件下，Shap-E和Point-E模型达到了大致相同的最终评估性能，Shap-E在CLIP R-precision方面略有优势，在CLIP分数方面略有劣势。**这表明显式和隐式建模可以从相同的数据和模型架构中学习不同的特征。**

图：Shap-E的编码器架构



图：图像条件下Point-E与Shap-E的比较

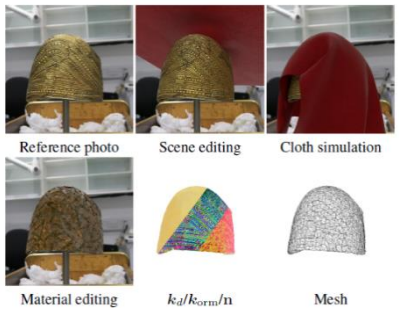
1.6 NVIDIA：3D MoMa、Magic3D、NVIDIA Picasso与Neuralangelo

3DMoMa：从二维图像中提取三维物体。2022年6月，NVIDIA推出3D MoMa，可通过图像输入生成三角网格组成的3D模型，并可直接导入图形引擎。这项方案的重点是，可直接导入支持三角形建模的3D建模引擎、游戏引擎、电影渲染器，可以在手机、浏览器上运行。3D MoMa生成的3D模型自带三角形网格，将3D模型生成自动化，将有望加速艺术、游戏、影视等内容创作。

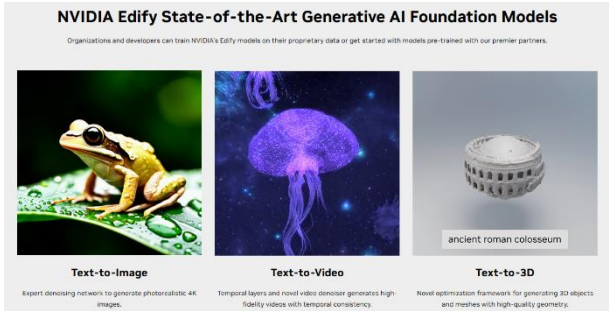
Magic3D：高分辨率的文本到3D内容创建技术。2022年11月，英伟达推出Magic3D，采用了与DreamFusion类似的两段式生成技术路线，但使用不同的生成模型。Magic3D可以在40分钟内创建高质量的三维网格模型，比DreamFusion快2倍，并实现了更高的分辨率

NVIDIA Picasso：用于构建生成式AI视觉应用程序的云服务。2023年3月，英伟达推出NVIDIA Picasso，企业、软件创建者和服务提供商可以在其模型上运行推理，在专有数据上训练NVIDIA Edify基础模型，或者从预训练的模型开始，从文本提示生成图像、视频和3D内容。Picasso服务针对GPU进行了全面优化，并在NVIDIA DGX Cloud上简化了训练、优化和推理。此外，NVIDIA也与Adobe、Getty Images、Shutterstock等企业进行了合作，共同开发NVIDIA Picasso模型。

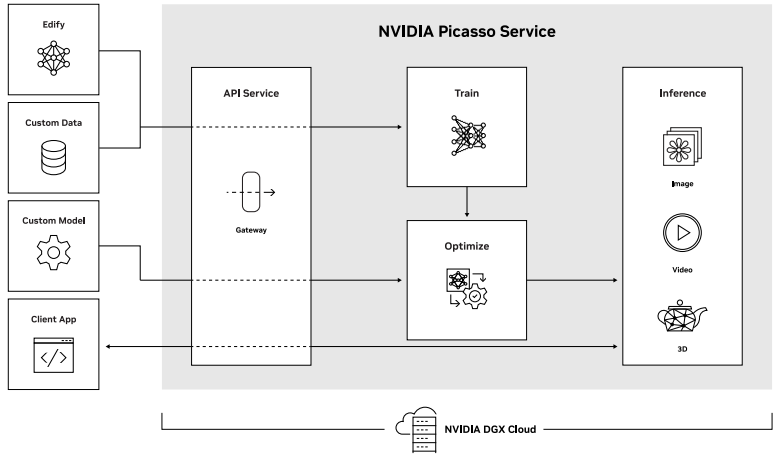
图：3D MoMa建模效果图



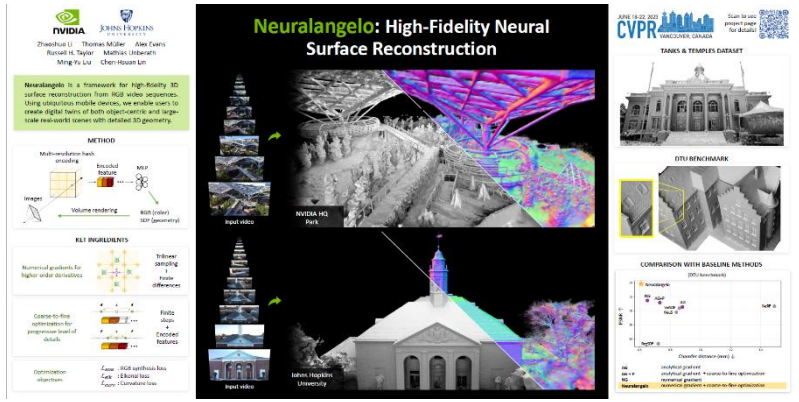
图：NVIDIA Picasso的功能



图：NVIDIA Picasso运行机制示意图



图：Neuralangelo宣传海报



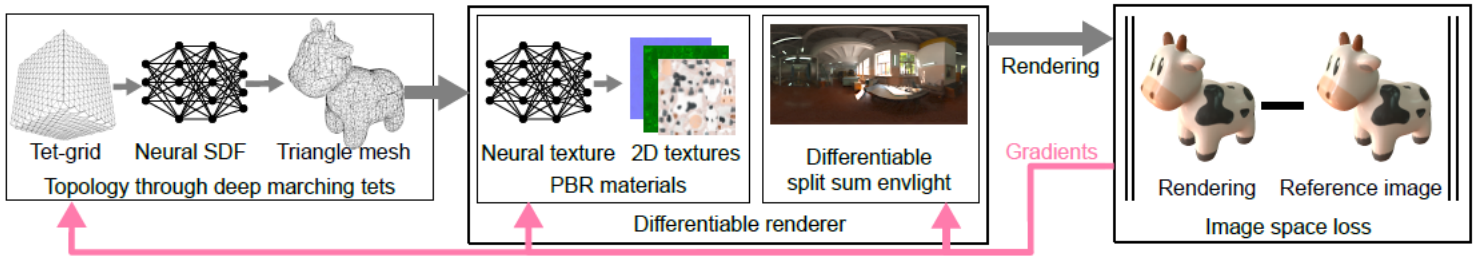
Neuralangelo：高保真的神经表面重建技术。2023年6月，英伟达提出了Neuralangelo，将多分辨率3D哈希网格的表示能力与神经表面渲染相结合，该方法由两个关键成分实现：1）用于计算高阶导数的数值梯度作为平滑操作；2）对控制不同细节级别的哈希网格进行粗-细优化。**即使没有辅助深度，Neuralangelo也可以有效地从多视图图像中恢复密集的3D表面结构，其保真度显著超过之前的方法，可以从RGB视频捕获中进行详细的大规模场景重建。**

资料来源：NVIDIA官网，Extracting Triangular 3D Models, Materials, and Lighting From Images（J Munkberg等），Neuralangelo: High-Fidelity Neural Surface Reconstruction（Z Li 等），Franks World，青亭网公众号，天风证券研究所

1.6 NVIDIA技术路径剖析：从逆向渲染到Instant NeRF

3D MoMa 采用的是被英伟达图形研究副总裁 David Luebke 称之为“统一计算机视觉和计算机图形的圣杯”的逆向渲染技术。**逆向渲染，即将一系列静态 2D 照片重建成物体或场景的 3D 模型的技术，而 3D MoMa 则是通过将每一个逆向渲染问题都定义为 GPU 加速的可微分组件，使用现代 AI 机器和英伟达 GPU 的原始计算能力来快速生成 3D 对象，使创造者可以在现有工具中不受限制地对其导入、编辑和扩展。**
3D MoMa正是基于逆向渲染流程，可从2D图片中提取3D信息、材质和照明数据。

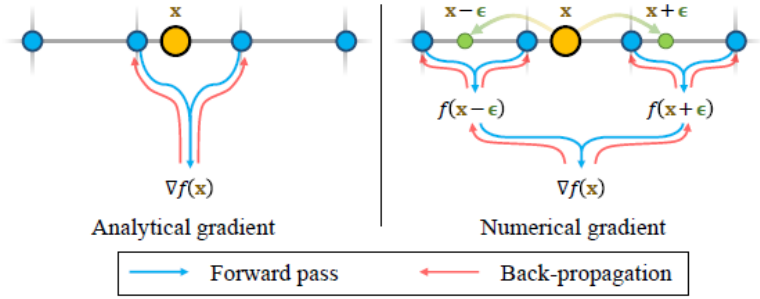
图：3D MoMa的概述



Magic3D利用两阶段优化框架来解决DreamFusion的两大局限性（NeRF的优化极其缓慢；NeRF的低分辨率图像空间监督，导致低质量的三维模型和较长的处理时间）：
1）使用**低分辨率**扩散模型先验获得粗模型，并使用稀疏三维哈希网格结构进行加速
2）粗略表示作为初始，使用高效的可微分的渲染器与**高分辨率**的潜在扩散模型交互，进一步优化纹理三维网格模型

NeuralAngelo建立在3D MoMa的基础上，允许导入更大、更详细的空间和对象。而它特别之处在于，通过采用“即时神经图形基元”，也就是NVIDIA Instant NeRF技术的核心，Neuralangelo由此可以捕捉更细微的细节。
技术步骤：
1）使用数值梯度来计算高阶导数。
2）逐步细化细节层次，逐步减小数值梯度的步长，并启用更高分辨率的哈希网格。
3）使用三个优化目标（RGB合成损失、Eikonal损失、曲率损失）进行优化。

图： NeuralAngelo使用数值梯度来计算高阶导数



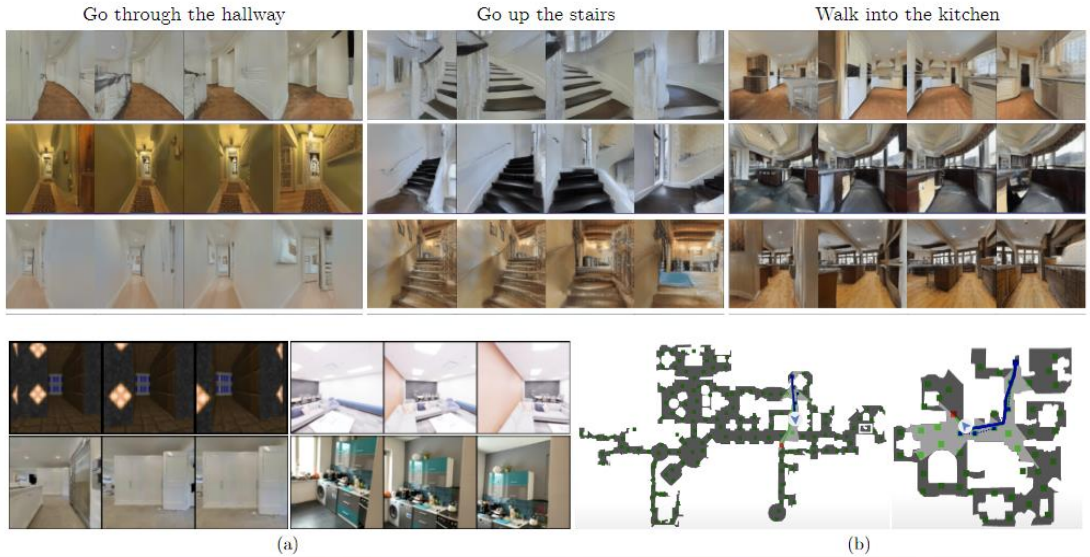
资料来源：Extracting Triangular 3D Models, Materials, and Lighting From Images（J Munkberg等），Neuralangelo: High-Fidelity Neural Surface Reconstruction（Z Li等），天风证券研究所

1.7 Apple：发布3D生成API Object Capture与3D场景生成模型GAUDI

2021年6月，苹果发布了面向Mac的摄影测量API“Object Capture”。Apple Object Capture为Apple设备用户提供了一种相对快速和简单的方法来创建现实世界对象的3D表示——这意味着可以**将物理对象转换为数字对象。**使用iPhone或iPad，可拍摄对象的照片，并使用macOS Monterey上新的对象捕获API将其转换为**针对增强现实(AR)进行优化的3D模型**。物体捕捉功能使用摄影测量技术将 iPhone 或 iPad 上拍摄的一系列照片转换为 USDZ 文件，这些文件可以在“AR 快速查看”中查看，无缝整合到 Xcode 项目中，或在专业的 3D 内容工作流程中使用。

2022年7月，来自苹果的 AI 团队推出了 3D 场景生成的最新神经架构——GAUDI 。GAUDI是一个能够捕捉复杂而真实的三维场景分布的生成模型，可以**从移动的摄像机中进行沉浸式渲染**，采用了一种可扩展但强大的方法来解决这个具有挑战性的问题。研究人员首先优化一个隐表征，将辐射场和摄像机的位置分开，然后将其用于学习生成模型，从而能够以无条件和有条件的方式生成三维场景。GAUDI在多个数据集的无条件生成设置中取得了sota的性能，并允许在给定条件变量（如稀疏的图像观测或描述场景的文本）的情况下有条件地生成三维场景。

图：GAUDI效果演示



图：Object Capture作用机制演示

资料来源：GAUDI: A Neural Architect for Immersive 3D Scene Generation（MA Bautista等），Apple官网，天风证券研究所

1.7 Apple技术路径剖析：发布3D生成API Object Capture与3D场景生成模型GAUDI

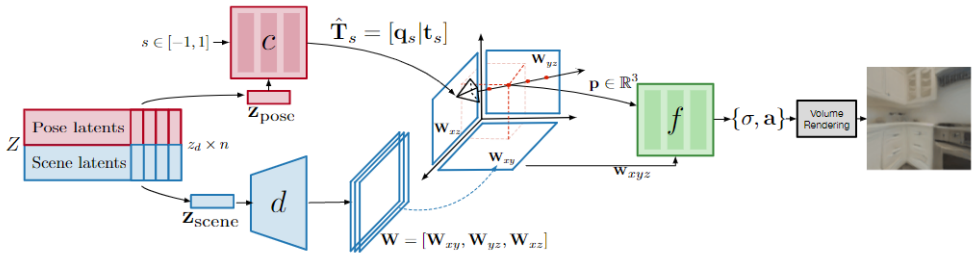
Object Capture技术原理：

- Object Capture 主要**基于 Swift 编程语言构建**，并通过 RealityKit 2（苹果的新一代 AR 引擎）在 macOS Monterey 上运行。
- Object Capture 可以将多张照片拼接在一起，针对照片以创建 3D 模型。
- Object Capture的实现基于多张图片或视频流的数据，通过使用苹果公司的**机器学习框架Core ML和Metal等技术**，进行**深度学习和计算机视觉分析**，从而创建高质量的3D模型。这种技术利用了苹果公司设备的摄像头和传感器，以及强大的硬件和软件性能，能够快速、准确地捕捉并处理图像数据，从而实现高质量的3D模型创建。

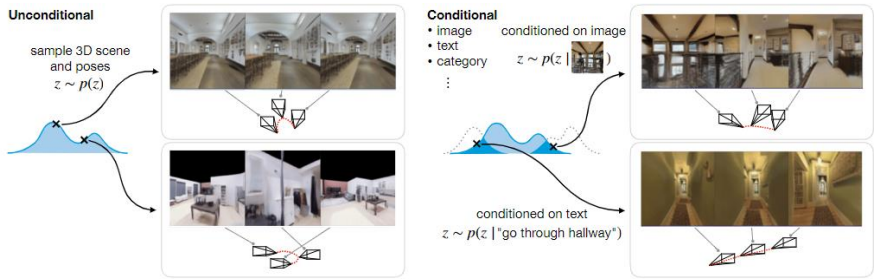
GAUDI基于 3D 场景的神经架构 NeRF，可以根据文字生成 3D 室内场景。根据苹果方面的介绍，GAUDI 的目标是给定 3D 场景轨迹的经验分布时，学习得出生成模型。GAUDI 的 3D 生成包含两个阶段：

- **一是潜在和网络参数的优化**：为每个示例 x ∈ X 获取一个潜在表示 z = [zscene, zpose]，用于表达场景辐射场和在单独的解纠缠向量中的位姿。学习对数千条轨迹的 3D 辐射场和相应相机姿态进行编码的潜在表示。与针对单个对象不同，有效相机姿态随着场景的变化而不同，所以需要对每个场景有效的相机姿态进行编码。
- **二是使用扩散模型在潜在表示上学习生成模型**：给定一组潜在的 Z = {zi∈{0,...,n}}，学习分布 p(Z)，从而能够在有条件和无条件的推理任务中都能很好地建模。前者是根据文本或图像提示来生成 3D 场景，后者则是根据摄像机轨迹来生成 3D场景。

图：GAUDI解码器模型架构



图：GAUDI扩散模型架构



资料来源：Apple官网，GAUDI: A Neural Architect for Immersive 3D Scene Generation（MA Bautista等），青亭网公众号，天风证券研究所

1.8 Google: 开源工具DiscoDiffusion、3D生成模型Dream Fields与DreamFusion

2021年10月，Disco Diffusion上线于 Google Colab 平台。它是一款利用人工智能深度学习进行数字艺术创作的开源工具，可以在 Google Drive 直接运行，也可以部署到本地运行。它利用了一种名为**CLIP-Guided Diffusion的人工智能图像生成技术**，可以让使用者从文本输入中创建详细、逼真的图像。

Dream Fields是由Google团队在2022年所推出的3D AIGC模型。基本原理是将OpenAI的图像分析模型 CLIP 与神经辐射场 (NeRF) 相结合，再利用了Nerf进行3D视图的生成，再通过Clip判断其生成的模型是否达到效果，本质上就是**通过CLIP/DALL-E + NeRF来实现其3D内容的生成。**

图：Dream Fields通过详细的标题来表达特定的艺术风格

DreamFusion同样由Google团队推出，**从模型发展脉络上看，DreamFusion是Dream Fields 的升级演变版本，其将Google的大型AI图像模型Imagen与NeRF的3D功能相结合。不再借助Clip对Nerf做引导，而是直接借助大模型的力量来实现模型的生成。**

图：DreamFusion从文本提示中生成仿真的三维模型

资料来源：Zero-Shot Text-Guided Object Generation with Dream Fields（A Jain等），DREAMFUSION: TEXT-TO-3D USING 2D DIFFUSION（B Poole等），github，天风证券研究所

1.8 Google技术路径剖析：从Dream Fields到DreamFusion的迭代升级

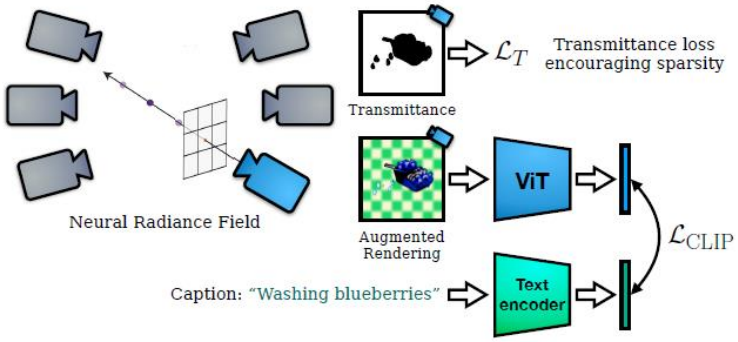
Dream Fields：训练Dream Fields算法时需要多角度2D照片，完成训练后便可生成3D模型、合成新视角。而CLIP的作用，依然是评估文本生成图像的准确性。**文本输入至Dream Fields后，未训练的NeRF模型会从单个视角生成随机视图，然后通过CLIP来评估生成图像的准确性。**也就是说，CLIP可以用来纠正和训练NeRF模型生成图像。这个过程将从不同的视角重复2万次，直到生成符合文本描述的3D模型。

DreamFusion是一种从文本提示生成 3D 模型的新方法，它采用了与Dream Field类似的方法，但模型中的损失函数基于概率密度蒸馏，最小化基于【扩散中前向过程共享的高斯分布族】与【预训练的扩散模型所学习的分数函数】之间的KL散度。技术步骤：

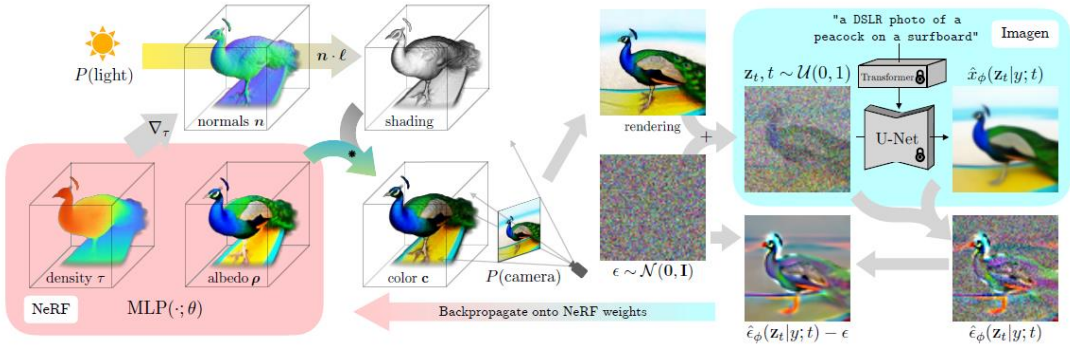
- 先使用一个预训练2D扩散模型基于文本提示生成一张二维图像。
- 然后引入一个基于概率密度蒸馏的损失函数，通过梯度下降法优化一个随机初始化的神经辐射场NeRF模型。

DreamFusion 结合了两种关键方法：神经辐射场和二维扩散。
它逐步改进初始的随机 3D 模型，以匹配从不同角度显示目标对象的 2D 参考图像：现有 AI 模型（如 Nvidia 的Instant NeRF ）使用的方法。然而，与 Instant NeRF 不同，参考不是真实物体的照片，而是由 OpenAI 的DALL-E 2和 Stability.ai 的Stable Diffusion使用的类型的 2D 文本到图像模型生成的合成图像。
在这种情况下，2D 扩散模型是 Google 自己的Imagen，但总体结果是相同的：生成的 3D 模型与原始文本描述生成的 2D 参考图像相匹配。至关重要的是，整个过程既不需要3D训练数据，也无需修改图像扩散模型，完全依赖预训练扩散模型作为先验——这可能为开发实用的、大众市场的基于 AI 的文本到 3D 工具铺平了道路。

图：Dream Fields的训练程序



图：DreamFusion根据自然语言标题生成3D对象



1.9 Meta: Meta MCC实现图像生成3D模型

为了简化AR/VR内容开发方式，Meta于2023年1月研发了一种RGB-D图像生成3D模型方案：MCC。MMC全称是多视图压缩编码，它是一种基于Transformer的编码器-解码器模型，可根据一帧RGB-D图像合成/重建3D模型，潜在应用场景包括AR/VR、3D视觉重建、机器人导航、数字孪生/虚拟仿真等等。

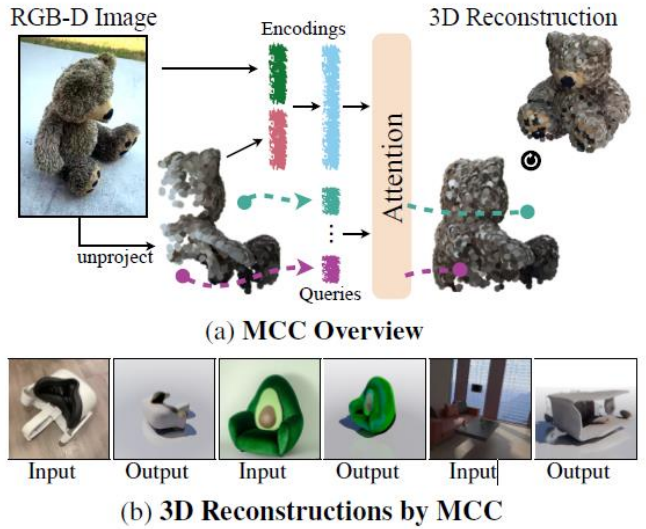
Transformer：一种采用自注意力机制的深度学习模型，谷歌曾使用它来增强搜索引擎，而近期热门的ChatGPT模型也是基于Transformer。**起初，Transformer更常用与自然语言处理领域，而随着它与大规模、通用类别的学习模型结合，便也开始被用于语言处理之外的领域，比如图像合成、图像分析。**

RGB-D：与普通彩色2D图像不同，RGB-D是具有深度的彩色图像，相当于普通RGB三通道彩色图像加上深度图（Depth Map）。

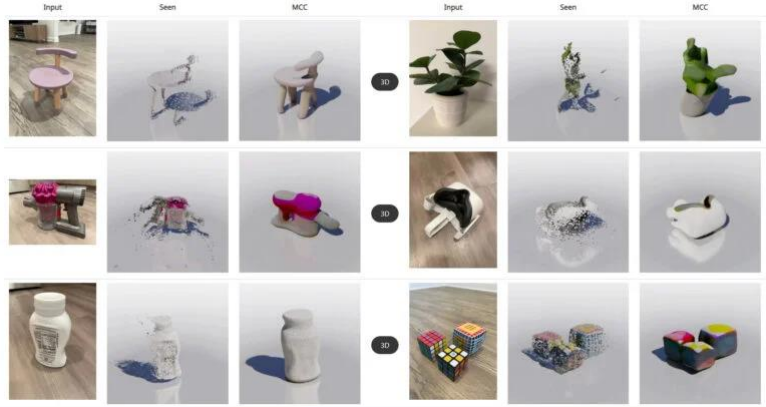
在2018年的F8大会上，Meta就曾公布3D照片研究，可通过双摄手机拍摄出具有3D效果的照片，其中包含一定的深度信息。其甚至还研发了将2D图像转3D的CNN模型，特点是支持单摄手机。这意味着，它如果结合MCC方案，或许可以将单摄手机捕捉的2D图像合成为3D模型。

利用MCC方案，3D开发/合成将有望实现规模化。随着深度传感器、深度捕捉AI模型在手机上普及，具有深度信息的图像越来越容易获得，因此MCC可使用的数据规模足够大。

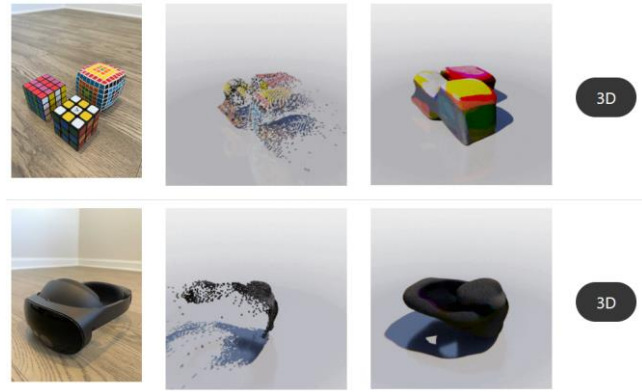
图：MCC的概述及其3D重建效果



图：Meta MCC可从单张图像合成完整的3D模型



图：Meta研究表明可通过双摄手机拍摄出具有3D效果的照片



资料来源：Multiview Compressive Coding for 3D Reconstruction（CY Wu等），the-decoder，Meta MCC官网，天风证券研究所

1.9 Meta技术路径剖析：MCC-基于Transformer的编码器-解码器模型

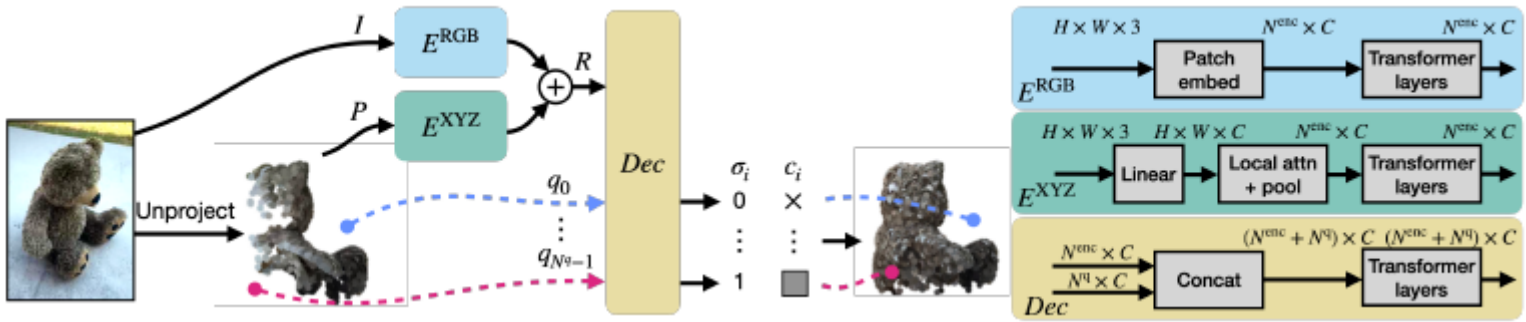
MCC采用简单的解码器-编码器架构，将RGB-D图像输入到MCC中会产生输入编码，然后解码器将在输入编码中访问3D点数据，以预测该点的占用率和RGB色彩（将3D重建定义为二元分类问题）。简单来讲，MCC只需要处理3D点云数据，而3D点可以捕捉任何对象或场景，通用性比网格和立体像素更好，因此用大规模RGB-D图像数据就能训练模型。
另外，RGB-D图像可通过手机的LiDAR传感器来捕捉，或是由深度模型来计算（比如MiDas、COLMAP）。

科研人员利用来自不同数据集的深度图像/视频来训练MCC，这些数据部分未包含3D场景、3D对象的全部角度，而这将需要AI重新构建。此外，MCC也可以将AI合成的图像转化为3D模型。

因此，MCC最大的特点是可预测RGB-D图像中看不见、被遮挡的3D几何形状。科研人员表示：MCC模型与基于图像的自监督学习、掩码自动编码器（MAE）的最新进展有直接关系，MAE也是通过预测图像中看不见的形状来学习图像表示。

MCC的技术优势：
- 无需具有注释的3D数据，成本更低、数据更容易收集
- 普适性好，对于未见过的新对象类别，支持零样本学习，可直接处理成3D模型
- 易于扩展，且将来可以轻松生成大型数据集，为3D重建带来规模化处理

图：MCC将输入RGB图像的像素解投影到相应的3D点



资料来源：Multiview Compressive Coding for 3D Reconstruction（CY Wu等），天风证券研究所

目录

1、生成式AI在视频/3D/游戏等领域的渗透加速

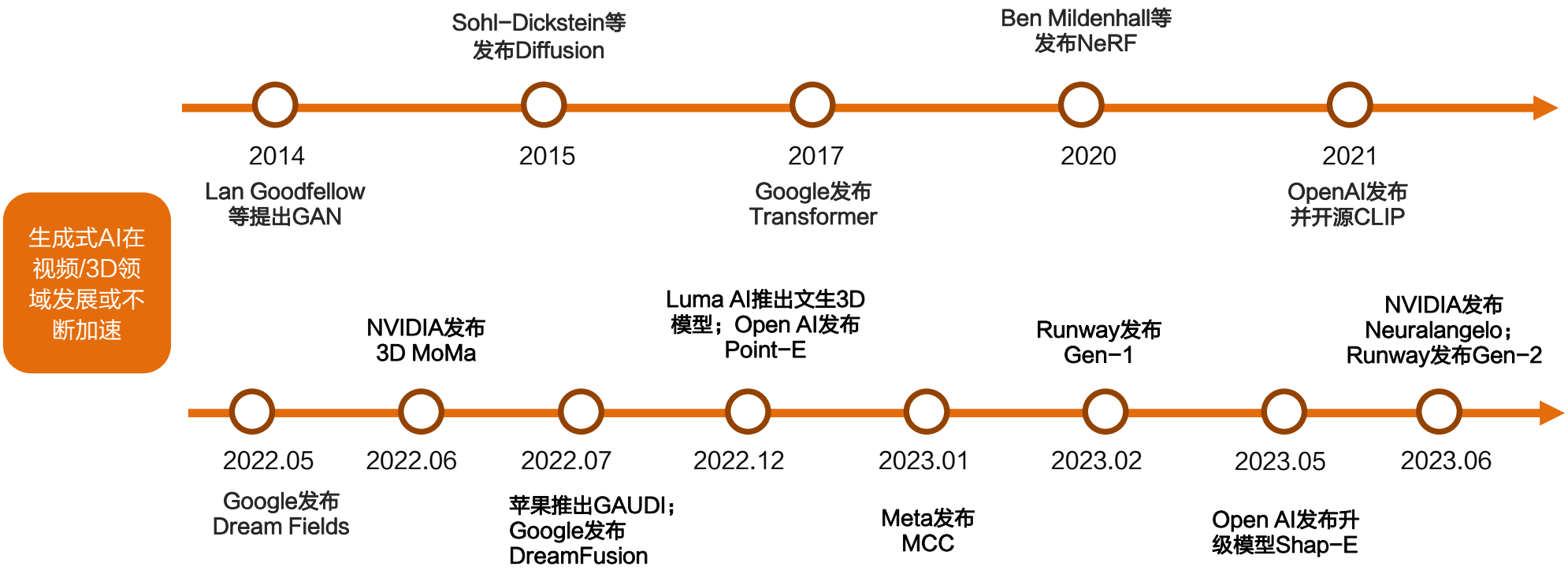
2、生成式AI下游应用场景展望

3、风险提示

2.1 生成式AI将实现对视频/3D/游戏等下游应用场景的渗透

今年以来，我们已经看到生成式AI在文本领域、代码生成、图像领域的快速发展，如GPT-4的推出，Midjourney、Stable Diffusion的版本迭代，Github Copilot X升级等等。
生成式AI在视频和3D模型领域的发展相对较慢，但随着海外科技大厂及初创公司纷纷布局并推出基础的3D和视频生成模型和产品，未来在创意领域（如电影、游戏、VR、建筑和实体产品设计）的应用潜力广阔。

图：生成式AI模型的进展与相关应用的发展进程时间表



资料来源：Generative Adversarial Nets(IJ Goodfellow等)，Improved Denoising Diffusion Probabilistic Models(A Nichol等)，Attention Is All You Need(Vaswani A等)，NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis(B.Mildenhall等)，Open AI官网，Nividia官网，新智元公众号等，天风证券研究所

2.1 生成式AI将实现对视频/3D/游戏等下游应用场景的渗透

我们认为生成式AI将实现对视频/3D/游戏等下游应用场景的渗透。在短视频、创作工具、游戏等下游领域，Runway、Luma AI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、进一步降低使用成本。

表：AI原生产品有望融入现有工具流中实现互补

分类	公司	产品	简介	功能	推出时间	应用场景
互联网平台	字节跳动	剪映	专门为抖音开发的剪辑软件	一键成片；录屏；剪同款；创作脚本；剪辑功能，支持变速，滤镜和美颜等效果，曲库资源。	2019.05	短视频
	快手	云剪	快手云剪”将传统剪辑工具搬到“云”上，实现素材共享、多人协同剪辑	视频剪辑、视频抠像、直播剪辑、文字转视频、数据视频、智能封面、横屏转竖屏、视频去抖、智能去水印、HDR画质增强。	2021.04	短视频
	腾讯	秒剪	微信推出的视频编辑App	一键成片；还有滤镜；转场；字幕等视频编辑功能	2020.09	短视频
	阿里	亲拍	面向淘系商家的短视频拍摄剪辑工具，集合商家“拍-剪-投-优”全链路短视频运营所需功能，提供一站式视频生产解决方案	口播快剪；自由剪辑	2020.10	短视频
	Meta	AI Sandbox	广告工具，包括文本变化、背景生成和图像突出等工具，使广告文本更吸引人并改进部分创意	文本变化：生成多版本文本；背景生成：根据文本输入创建背景图片；图像裁剪：调整创意资产以适应多个界面（如 Stories 或 Reels）的不同长宽比	2023.05	图像
软件工具公司	Adobe	Firefly	创意生成式AI 模型集，将成为跨Adobe云端产品的全新Adobe Sensei生成式AI服务的一部分	文生图；去除/填充；文本特效；更换logo着色等	2023.03	图像
	Midjourney	Midjourney 5.2	生成式AI，从简单的文本提示中创造出高质量的图像，通过Discord聊天应用程序工作	文生图；Zoom out	2023.06	图像
游戏引擎	Unity	Muse	在创作过程中提供ai驱动辅助的扩展平台	输入文本创建角色动画；快速创建2Dsprite资产；生成逼真的纹理；与ChatGPT一样生成文案	2023.06	游戏
	Epic	RealityScan	一款可以将智能手机照片转换为高保真3D模型的免费3D扫描应用	3D模型生成	2022.04	游戏
		RealityCapture 1.2.2	适用于Windows的摄影测量软件，能够从一组图像和/或激光扫描中创建超现实的3D模型	3D模型生成；3D渲染器增强；更快重建；纹理增强	2023.06	游戏
	Tafi	Daz3D	一款突破性的文本到3D角色引擎	通过输入文本来快速创建附带UV和拓扑的3D模型，输出到游戏引擎和3D软件中	2023.06	游戏

资料来源：各产品官网，腾讯应用宝，Apple Store，Adobe官方公众号，科技狐公众号，极客网，Meta官网等，天风证券研究所

2.1 视频与建模工具或与传统工具/工作流结合，进一步拉动生成式AI的需求

我们的观点：

内容创作工具的重要性如何？我们认为关键在于拉动远比工具本身更大的市场，类似于短视频时代的前身GIF创作工具，如剪映这种创作工具和抖音这种内容生态，Maya这种创作工具和动画的内容生态，视频与建模工具将进一步大幅拉动生成式AI的需求。

模型能力或出现明显分化。我们认为当前Diffusion开源模型下各公司生成图片的能力尚未明显出现分化，但建模和视频更重要的在于和传统工具与工作流的结合 。

海外接下来半年关注什么？我们认为从能力来看，图片生成的可控性快速提高或将出现，矢量图、结构、视频、3D模型生成也将提速。尤其关注Unity与Apple的合作，在AI生成内容/建模/App适配上将会如何塑造空间计算的内容与应用的标准生态。

建议关注：

AI+工具：A股【万兴科技】（计算机覆盖）；港股【美图】【腾讯】；美股【Adobe】【Unity】【Nvidia】

AI+游戏：A股【神州泰岳】【恺英网络】【掌趣科技】【虹软科技】

AI+影视：A股【光线传媒】【中国电影】【上海电影】

目录

1、生成式AI在视频/3D/游戏等领域的渗透加速

2、生成式AI下游应用场景展望

3、风险提示

4 风险提示

1）生成式AI发展不及预期：生成式AI在图像/视频/3D领域的技术发展不及预期。

2）算力成本及硬件发展不及预期：模型推理和训练的算力成本下降不及预期；算力硬件发展出现瓶颈。

3）相关应用产品上线后效果不及预期：生成式AI在图像/视频/3D领域的相关产品推出后效果不及预期，商业化进展不及预期。

分析师声明

本报告署名分析师在此声明：我们具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，本报告所表述的所有观点均准确地反映了我们对标的证券和发行人的个人看法。我们所得报酬的任何部分不曾与，不与，也将不会与本报告中的具体投资建议或观点有直接或间接联系。

一般声明

本报告是机密的，仅供我们的客户使用，天风证券不因收件人收到本报告而视其为天风证券的客户。本报告中的信息均来源于我们认为可靠的已公开资料，但天风证券对这些信息的准确性及完整性不作任何保证。本报告中的信息、意见等均仅供客户参考，不构成所述证券买卖的出价或征价邀请或要约。该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。客户应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专家的意见。对依据或者使用本报告所造成的一切后果，天风证券及/或其关联人员均不承担任何法律责任。

本报告所载的意见、评估及预测仅为本报告出具日的观点和判断。该等意见、评估及预测无需通知即可随时更改。过往的表现亦不应作为日后表现的预示和担保。在不同时期，天风证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。

天风证券的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。天风证券没有将此意见及建议向报告所有接收者进行更新的义务。天风证券的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

特别声明

在法律许可的情况下，天风证券可能会持有本报告中提及公司所发行的证券并进行交易，也可能为这些公司提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。因此，投资者应当考虑到天风证券及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突，投资者请勿将本报告视为投资或其他决定的唯一参考依据。

投资评级声明

类别	说明	评级	体系
股票投资评级	自报告日后的6个月内，相对同期沪深300指数的涨跌幅	买入	预期股价相对收益20%以上
		增持	预期股价相对收益10%-20%
		持有	预期股价相对收益-10%-10%
		卖出	预期股价相对收益-10%以下
行业投资评级	自报告日后的6个月内，相对同期沪深300指数的涨跌幅	强于大市	预期行业指数涨幅5%以上
		中性	预期行业指数涨幅-5%-5%
		弱于大市	预期行业指数涨幅-5%以下