



Research and
Development Center

AI 时代的 3D 内容生产工具

2024 年 3 月 20 日

互联网行业

证券研究报告

行业研究

行业深度研究

行业名称 传媒

投资评级 看好

上次评级 看好

冯翠婷 传媒互联网及海外首席分析师

执业编号: S1500522010001

邮箱: fengcuiting@cindasc.com

信达证券股份有限公司

CINDA SECURITIES CO., LTD

北京市西城区闹市口大街9号院1号楼

邮编: 100031

AI 时代的 3D 内容生产工具

2024 年 3 月 20 日

本期内容提要:

◆近期 AI+3D 新产品产业催化汇总。英伟达 CEO: NVIDIA 3D 平台 Omniverse Cloud 将可以连接到苹果公司混合头显 Vision Pro; 3D 素材赋能 Sora 等文生视频工具深度学习; Stability AI 公司发布全新的 AI 生成 3D 大模型 TripoSR, 效果惊艳。

◆DUST3R 技术储备丰厚, 助力 AI 深度学习 3D 重建, 工业应用未来可期。DUST3R 该框架可根据拍摄的图像重建 3D 模型, 无需相机校准。此外, DUST3R 支持单目和双目视图, 能生成深度图等, 为后续分析和应用提供了丰富数据。

◆谷歌所推出的 DreamFusion 项目, 突破 3D 模型训练样本数量问题。该项目结合 SDS 与 NeRF 技术, 将二维图像扩散模型转化为三维模型, 无需依赖三维训练数据。DreamFusion 生成的 3D 模型几何表现优秀。

◆英伟达发布 GET3D 和 Magic3D 两大创新工具, 重塑元宇宙内容创建流程。GET3D 利用 2D 图像训练, 生成高保真纹理和复杂几何细节的 3D 图形。而 Magic3D 能在 40 分钟内创建带有彩色纹理的 3D 网格模型, 且支持根据提示词编辑已创建的模型。

◆OpenAI 发布了开源的 Shap-E 模型, 能快速创建准确的 3D 模型, 成本效益显著。该模型可生成纹理网格和神经辐射场两种 3D 表示形式, 支持多样化渲染和导入其他 3D 应用, 展现出在游戏开发等领域的广泛应用前景。

◆阿里巴巴发布 Object Drawer, 是首个基于神经渲染的商品三维建模产品。该技术可重建多种商品模型, 低成本高保真。通过优化网络结构和几何先验, 提升推理和训练速度, 实现快速推理高分辨率图像。支持实时高清交互三维展示, 解决光照和纹理问题, 提升渲染效果。

◆商汤科技发布两款 3D 内容生成平台: 格物与琼宇, 满足不同行业对于可交互 3D 实景内容的需求。格物专注于小物体的高精度重建与交互, 适用于商业广告、商品营销等领域; 琼宇则实现城市级大空间的 3D 内容生成, 支持数字孪生、影视创作等。

◆普林斯顿大学发布的 Infinigen AI, 为计算机视觉领域提供了强有力的工具。Infinigen AI 能生成自然界的植物、动物等场景并自定义地形和自然现象, 为计算机视觉领域提供了新的数据生成和场景模拟方法, 适用于 3D 视觉研究。

◆Adobe 联合开发的 DMV3D, 在高效率生成高保真 3D 图像方面表现出色。这是一种基于 transformer 的 3D 大型重建模型, 能在单个 GPU 上快速生成高保真 3D 图像, 避免繁琐的单独训练和优化步骤。

◆Epic Games 推出的 MetaHuman, 大大缩短传统建模的时间及资源。支持云端快速创建高度逼真的 3D 角色, 并允许高度自定义。它简化了 3D 建模的复杂性, 消除了对高端硬件的需求, 可与虚幻引擎 5 紧密集成。

◆Unity 为开发者提供更强大和灵活的创作工具。Unity Muse 利用 NLP 技术

请阅读最后一页免责声明及信息披露 <http://www.cindasc.com> 2

生成 3D 模型、纹理和动画；Sentis 集成 AI 模型实现文字转语音、智能 NPC 等功能；Meshy 提供多种 AI 功能，生成 3D 图形速度快；Unity 6 预计带来高性能渲染与 WebGPU 图形后端。

◆ **风险因素：**宏观经济波动风险、MR 设备市场销量不及预期风险

目 录

一、短期 AI+3D 产业催化汇总.....	4
二、DUSt3R.....	4
三、谷歌：DreamFusion.....	5
四、英伟达：GET3D 和 Magic3D.....	6
五、OpenAI：Shap-E.....	7
六、阿里巴巴：Object Drawer.....	8
七、商汤科技：格物、琼宇.....	9
八、普利斯顿大学：Infinigen AI.....	10
九、Adobe：DMV3D.....	11
十、Epic Games：MetaHuman.....	8
十一、Unity：Muse、Sentis、Meshy（插件）和 Unity6.....	13
十二、风险因素.....	16

图 目 录

图 1：英伟达 GTC 发布会上 Omniverse 适配 Vision Pro 展示.....	4
图 2：Sora 的效果展示.....	4
图 3：TripoSR 的效果展示.....	5
图 4：DUSt3R 的效果展示.....	5
图 5：DUSt3R 的算法结构.....	6
图 6：DreamFusion 的效果展示.....	7
图 7：DreamFusion 算法架构.....	7
图 8：GET3D 的效果展示.....	8
图 9：GET3D 的算法结构.....	8
图 10：Magic3D 的效果展示.....	9
图 11：Magic3D 的算法结构.....	9
图 12：Shap-E 的效果展示.....	10
图 13：Object Drawer 的效果展示.....	11
图 14：NeRF 技术原理示意.....	11
图 15：琼宇的效果展示.....	12
图 16：琼宇的技术储备.....	12
图 17：Infinigen AI 的效果展示.....	13
图 18：Infinigen AI 的技术特点.....	14
图 19：DMV3D 的效果展示.....	14
图 20：DMV3D 的算法结构.....	15
图 21：MetaHuman 的效果展示.....	16
图 22：Unity Muse 的效果展示和特点.....	17
图 23：Unity Sentis 的效果展示和特点.....	18
图 24：Meshy 的效果展示.....	18
图 25：Unity 6 官方预热.....	19

一、短期 AI+3D 产业催化汇总

英伟达 CEO: NVIDIA 3D 平台 Omniverse Cloud 将可以连接到苹果公司混合头显 Vision Pro。在英伟达的 GTC AI 大会上, 公司 CEO 黄仁勋宣布了一项重要的技术进展: NVIDIA 的 3D 平台 Omniverse Cloud, 这一平台将能够与苹果公司的新型混合现实头显 Vision Pro 实现连接。这一举措将使得企业能够将基于 3D 应用的交互式通用场景描述 (OpenUSD) 直接串流至 Apple Vision Pro, 利用英伟达最新的 Omniverse Clouds API 和 Graphics Delivery Network (GDN) 来高效传输数据, 确保用户能够体验到高保真度的视觉效果。

图 1: 英伟达 GTC 发布会上 Omniverse 适配 Vision Pro 展示



资料来源: 英伟达官网, 信达证券研发中心

3D 素材赋能 Sora 等文生视频工具深度学习。Open AI 在 2 月发布的 Sora, 是一款具备卓越视频生成能力的扩散模型。Sora 采用 Transformer 架构并利用 3D 模型素材进行训练, 在文生视频领域表现出色。在生成视频方面, Sora 能够从类似静态噪声的视频开始, 逐步去除噪声, 直至形成清晰的视频画面。此外, Sora 还能一次性生成完整的视频, 或延长已生成的视频。这种能力得益于模型能够同时预览多个帧, 确保即使视频中的主体暂时离开视线, 也能保持其一致性。在 Transformer 架构加持下, Sora 通过使用 3D 模型素材进行训练, 能够生成更真实、更精确的视频内容。

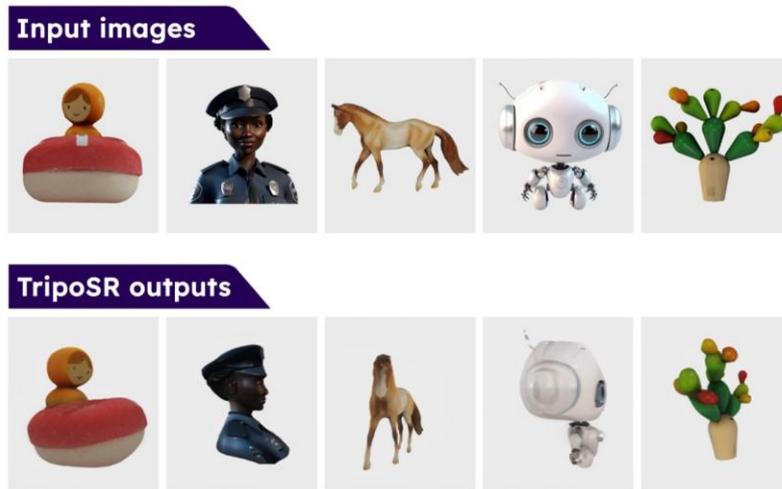
图 2: Sora 的效果展示



资料来源: Open AI 官网, 信达证券研发中心

Stability AI 公司发布全新的 AI 生成 3D 大模型 TripoSR, 效果惊艳。Stability AI 与 Tripo AI 合作推出 TripoSR, 能够在短短一秒内, 凭单个图像即可生成高质量 3D 模型。TripoSR 的独特之处在于, 它即便在低推理预算条件下 (甚至在没有 GPU 的情况下) 也能顺畅运行, 使得这项技术对于广泛的用户和应用场景而言都既易于获取又实用。此外, TripoSR 的模型权重和源代码已按照 MIT 许可证开放下载, 用户可将 TripoSR 用于商业化项目、个人使用和学术研究。Stability AI 和 Tripo AI 致力于打造的这种新型图像转 3D 模型, 旨在满足娱乐、游戏、工业设计和建筑等专业领域对详细 3D 对象可视化日益增长的需求, 并提供一种响应迅速的输出解决方案。

图 3: TripoSR 的效果展示

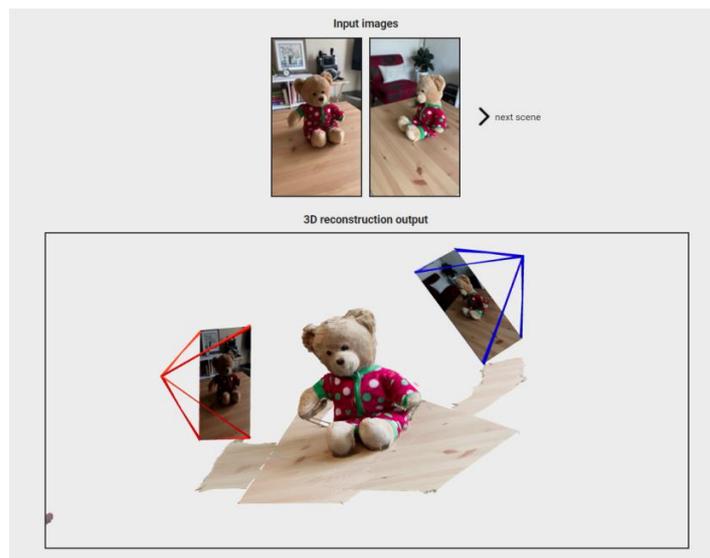


资料来源: Stability AI 官网, 信达证券研发中心

二、DUS_t3R

DUS_t3R (Dense and Unconstrained Stereo 3D Reconstruction) 是一种由芬兰阿尔托大学与 Naver 欧洲实验室联合研发的先进 3D 重建框架。该框架的设计初衷在于简化和优化从拍摄的图像集合中重建三维场景的过程, 免除了对相机校准或视点位置信息的预先要求。

图 4: DUS_t3R 的效果展示

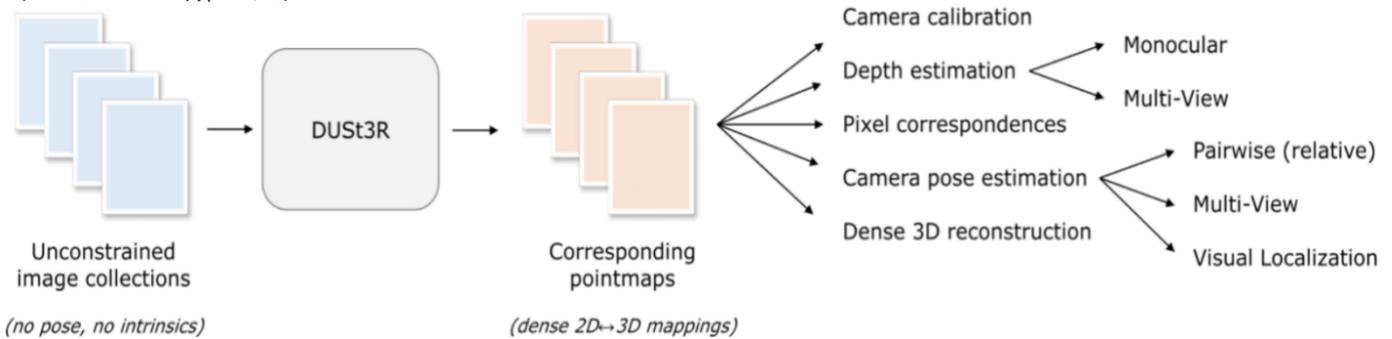


资料来源: naverlabs 官网, Shuzhe Wang 等《DUS_t3R: Geometric 3D Vision Made Easy》, 信达证券研发中心

DUST3R 技术储备丰厚，助力 AI 深度学习 3D 重建。DUST3R 采用点图技术 (Pointmaps) 作为其核心表示机制，该技术通过密集的 2D 场景提供了 3D 点的详尽信息。点图为图像中的每一个像素配备了对应的 3D 点，从而建立了图像像素与 3D 场景点之间的直接联系。除此之外，DUST3R 还采用了 Chat GPT 和 Sora 同款的 Transformer 网络架构。基于标准的 Transformer 编码器和解码器，DUST3R 构建了其网络架构，这使得该模型能够利用预训练的强大模型，在无需显式几何约束的情况下从输入图像中提取丰富的几何和外观信息。

DUST3R 支持端到端训练模式，允许直接从成对图像中学习到点图，无需复杂的多步骤处理流程，如特征匹配、三角测量等。对于多于两张图像的处理，DUST3R 提出了一种全局对齐策略，该策略可以将所有成对点图整合到一个共同的参考框架中，这对于多视图 3D 重建具有重要意义。DUST3R 能够在训练过程中同时学习多个相关任务，如深度估计、相机参数估计、像素对应关系等，通过这种多任务学习策略，模型能够更全面地掌握场景的几何结构。

图 5: DUST3R 的算法结构



资料来源: naverlabs 官网, Shuzhe Wang 等《DUST3R: Geometric 3D Vision Made Easy》, 信达证券研发中心

DUST3R 应用效果出色，工业应用未来可期。DUST3R 能在短时间内从输入图片中重建出高精度的 3D 模型，适用于实时应用和快速原型制作，从而实现快速精确的三维重建。与传统三维重建技术相比，DUST3R 不需要任何相机校准或视点姿势的先验信息，无需相机校准的灵活性大大简化了使用流程。除此之外，DUST3R 能够统一处理单目和双目情况，为用户提供了更多选择和灵活性。除了三维模型，DUST3R 还能生成深度图、置信度图和点云图等，为后续的分析 and 应用提供了丰富的数据支持。通过在多个 3D 视觉任务上设立新的最佳实践，包括单眼和多视图深度估计以及相对姿态估计等，DUST3R 进一步推动未来 AI 生成 3D 模型应用的发展。

三、谷歌：DreamFusion

2021 年末，谷歌与加州大学伯克利分校联合推出了 DreamFusion 项目，旨在实现从二维扩散文本到三维生成的技术。该项目结合了谷歌研究院和 UC Berkeley 研究人员的力量，融合了大型 AI 图像模型 Imagen 和神经辐射场 (NeRF) 技术，以推进 3D 功能的发展。

图 6: DreamFusion 的效果展示

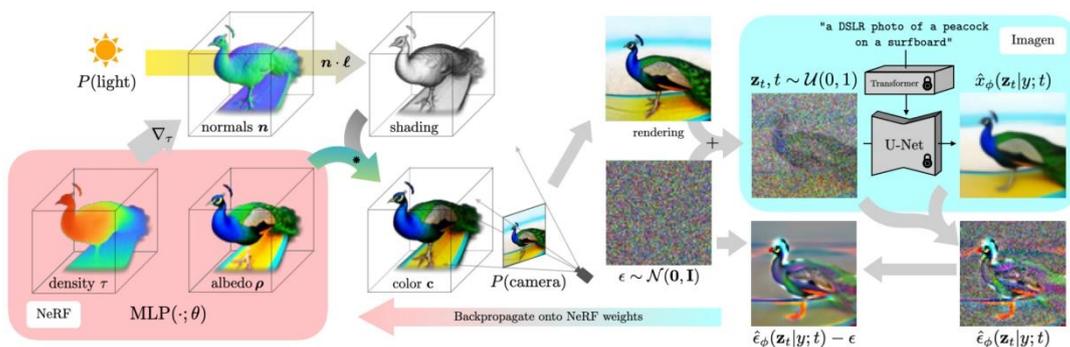


资料来源: DreamFusion 官网, 信达证券研发中心

DreamFusion 依靠 SDS 和 NeRF 技术赋能, 突破 3D 模型训练样本数量问题。DreamFusion 通过采用创新的得分蒸馏采样 (SDS) 方法和 NeRF 相似的渲染引擎, 能够将高质量的二维图像扩散模型转化为三维领域。该技术不依赖于三维或多视图的训练数据, 仅通过预先训练的二维扩散模型进行三维合成, 展示了其在三维生成领域的独特优势和灵活性。SDS 作为 DreamFusion 的关键创新之一, 通过优化损失函数从扩散模型中生成样本, 允许在任意参数空间内 (例如三维空间) 进行样本优化, 使得 DreamFusion 不需要依赖于 3D 训练数据, 也无需对原有的图像扩散模型进行修改, 证明了预训练的图像扩散模型在作为先验知识时的有效性。

图 7: DreamFusion 算法架构

DreamFusion: Text-to-3D using 2D Diffusion



资料来源: DreamFusion 官网, 信达证券研发中心

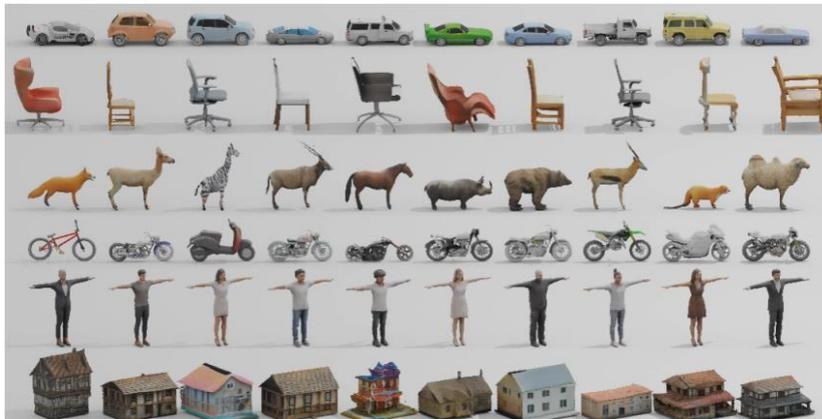
DreamFusion 生成的 3D 模型几何表现优异, 但在分辨率和纹理表现上尚具有进步空间。DreamFusion 生成的

三维模型不仅拥有高质量的法线、表面几何和深度，还能通过朗伯阴影模型进行再照明，使得模型能从多角度观看，并且能够在不同照明条件下重新点亮，甚至可以被合成到其他三维环境中。这一系列特点标志着 DreamFusion 在推进三维内容生成和增强虚拟现实体验方面的重要步骤。尽管 DreamFusion 在 3D 模型生成方面取得了显著进展，但它也存在一些局限性，包括无法生成高分辨率的 3D 几何体或纹理，以及基于 Mip-NeRF 360 的场景渲染模型的可扩展性不佳。

四、英伟达：GET3D 和 Magic3D

2022 年 9 月，英伟达发布了 GET3D 工具，以帮助元宇宙内容创建者更快地生成大量不同的 3D 物体，并为自动生成带贴图的 3D 模型打开了新的可能性。该工具使用 2D 图像训练，能够生成具有高保真纹理和复杂几何细节的三维图形。GET3D 生成的 3D 对象格式与主流图形软件应用相同，用户能够轻松地导入到 3D 渲染器和游戏引擎中进行后续编辑。

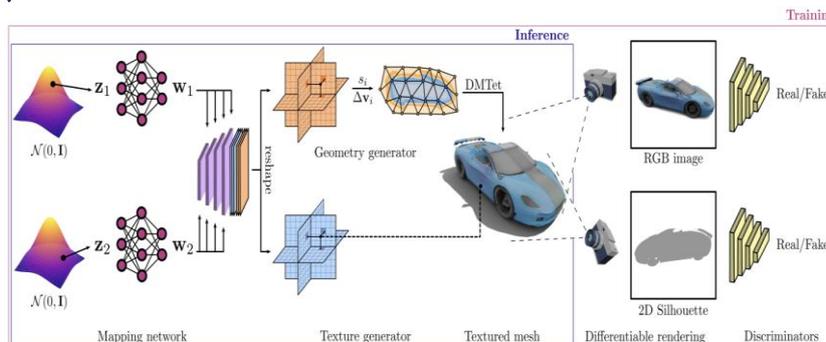
图 8：GET3D 的效果展示



资料来源：英伟达官网，信达证券研发中心

GET3D 通过两个潜码生成 3D 符号距离场 (SDF) 和纹理场。使用 DM Tet 从 SDF 提取 3D 表面网格，并在表面点查询纹理场以获取颜色。除此之外，GET3D 通过定义在 2D 图像上的对抗性损失进行训练，使用基于光栅化的可微分渲染器来获取 RGB 图像和轮廓。在判断方面，模型使用了两个 2D 鉴别器，分别对 RGB 图像和轮廓进行真假分类，整个模型可以端到端训练。GET3D 独特之处在于仅需一块 GPU 就能产出模型，为影视制作、游戏开发和元宇宙建设等领域提供了便利。此外，GET3D 旨在解决以往 3D 生成模型存在的问题，并生成具有丰富几何细节和纹理的 3D 模型，可直接在下游 3D 软件中使用。

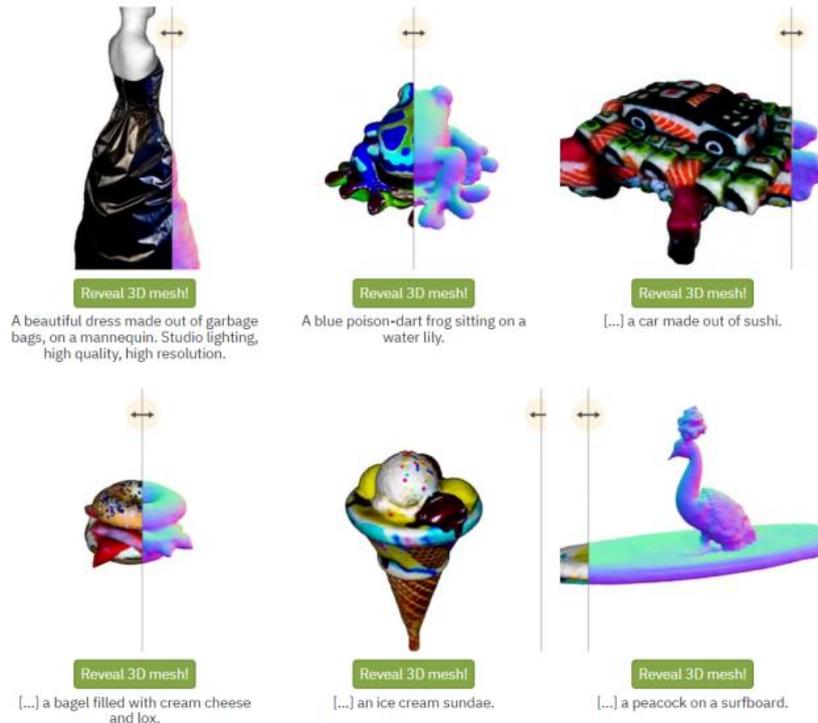
图 9：GET3D 的算法结构



资料来源：英伟达官网，信达证券研发中心

2023年2月，英伟达推出了文生3D模型Magic3D。Magic3D能够在40分钟内创建一个带有彩色纹理的3D网格模型，如“一只坐在睡莲上的蓝色箭毒蛙”。通过高分辨率的文生3D模型和较快的生成速度，Magic3D为游戏和元宇宙世界的开发者提供了快速制作大量3D模型的能力。

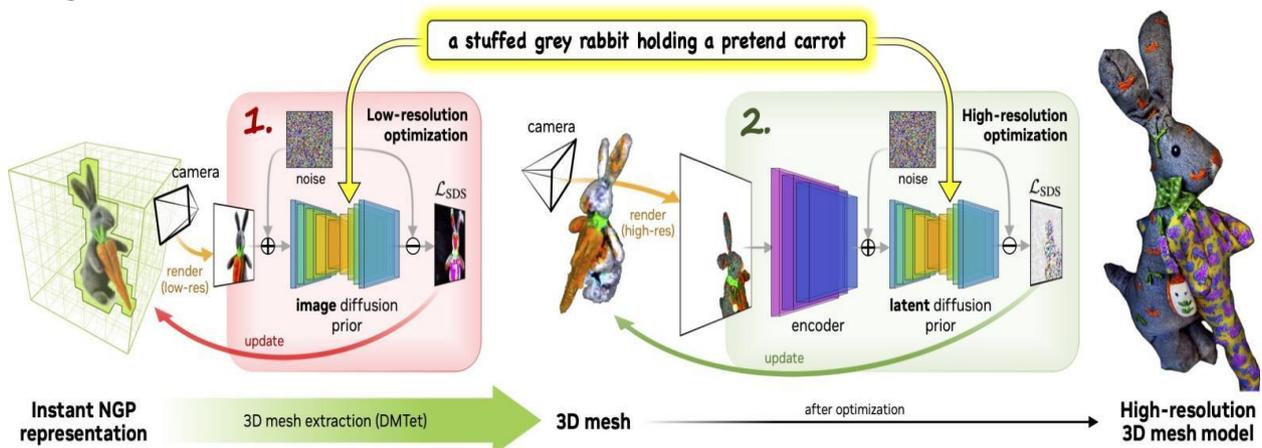
图 10: Magic3D 的效果展示



资料来源：英伟达官网，信达证券研发中心

Magic 3D 采用了一个粗到细的两阶段优化框架，以实现快速且高质量的文本到 3D 内容创造。在第一阶段，我们利用一个低分辨率的扩散先验获取一个粗略模型，并通过哈希网格和稀疏加速结构来加速这一过程。在第二阶段，Magic 3D 使用从粗略神经表征初始化的带纹理的网格模型，通过利用一个高效的微分渲染器与高分辨率潜在扩散模型进行交互，以进行优化，即：Magic 3D 先创建粗糙的低分辨率模型，然后优化为更高分辨率，同时利用文本转图像模型生成 2D 图像，并优化为 NeRF 体积数据。

图 11: Magic3D 的算法结构



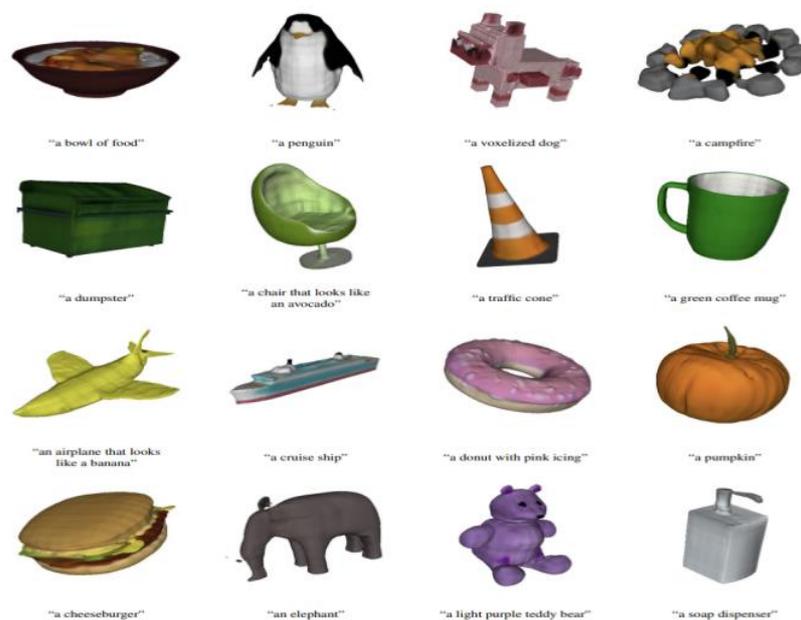
资料来源：英伟达官网，信达证券研发中心

Magic 3D 对比谷歌的 DreamFusion，Magic3D 速度提升了两倍。除了速度优势外，Magic3D 还能根据提示词编辑已创建的 3D 模型，用户可以通过修改基础提示词和低分辨率的 3D 模型来改变生成的模型。

五、OpenAI: Shap-E

2023 年 5 月 6 日，OpenAI 发布了文或图生 3D 的 Shap-E 模型，并在 Github 上开源，它能同时生成两种类型的 3D 表示形式：纹理网格（textured meshes）和神经辐射场（NeRF），使得生成的 3D 模型既可以多样化渲染也可导入到其他 3D 应用程序中。Shap-E 的相关模型权重、推理代码和样本已经向公众开放，这使得开发者和研究人员可以更深入地探索和利用这一模型，同时开源代码也为社区提供了一个共同改进和扩展 Shap-E 功能的平台。

图 12: Shap-E 的效果展示



资料来源: Heewoo Jun, Alex Nicho 《Shap-E: Generating Conditional 3D Implicit Functions》，信达证券研发中心

Shap-E 的工作原理分为编码阶段和生成阶段。在编码阶段，Shap-E 通过一个深度神经网络编码器将 3D 资产映射到隐式函数的参数，该编码器接受一个 3D 资产（如点云或体素）作为输入，并输出隐式函数的参数向量。在生成阶段，它使用一个条件扩散模型来生成这些参数，这种模型基于扩散过程，能够接受文本提示或图像等条件作为输入，并输出隐式函数的参数向量。这种两阶段方法使得 Shap-E 能直接生成纹理网格和神经辐射场，从而创造出具有细腻纹理和复杂形状的高质量 3D 资产。

Shap-E 在生成速度和准确性上具有优势，能快速创建准确的 3D 模型，从而节省设计师和工程师的时间和资源。它的成本效益显著，通过自动化 3D 建模过程，有潜力降低劳动成本并简化工作流程。此外，Shap-E 可以创建具有精细纹理和精美形状的 3D 模型，支持在 CPU 或 GPU 上运行，并且易于通过 Jupyter Notebook 使用。它还可以基于文本提示生成 PLY 文件，使得用户能够为虚拟现实、增强现实和 3D 打印应用创建 3D 模型。Shap-E 在游戏开发、动画制作、虚拟现实和 3D 打印等多个领域都展现出广泛的应用前景。根据目前 Shap-E 的演示效果，其在模型精度和细节呈现上的表现有待提高。

六、阿里巴巴：Object Drawer

2021年阿里巴巴发布了业界首个基于神经渲染的自研商品三维建模产品 Object Drawer，并已在手机天猫 APP 上规模化应用。Object Drawer 技术可重建多个商品模型，实现低成本的商品 3D 建模。它能够对商品环拍照片进行 3D 建模，高保真还原纹理细节，无需人工修模，大幅降低建模成本。此外，模型的训练时间大幅缩短（NeRF 的 10000 倍），模型大小也显著减小（20MB 以内），同时保证任意视角下商品 3D 模型达到实拍照片的展示效果。

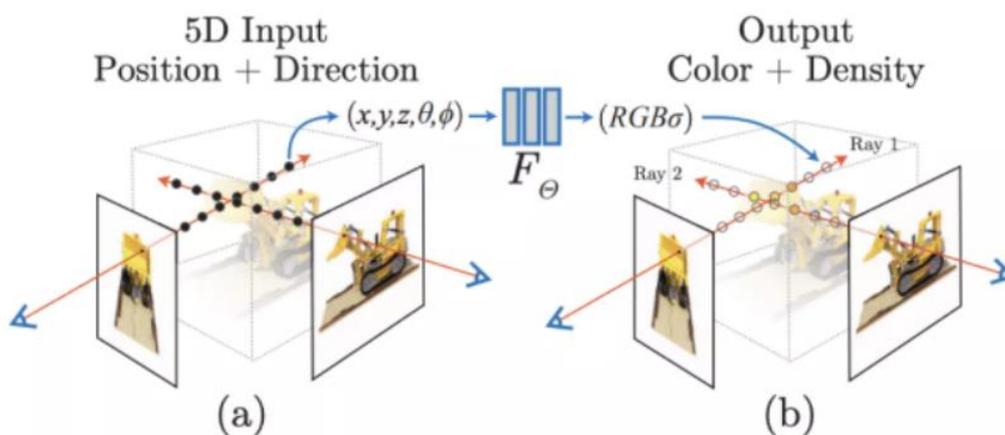
图 13: Object Drawer 的效果展示



资料来源：阿里云，信达证券研发中心

Object Drawer 基于 NeRF 技术发展而来，NeRF 能从多视角图像中重建场景，通过神经网络隐式表示场景的几何和光照信息。Object Drawer 通过优化网络结构、探索几何先验和预存部分信息，显著提升了推理和训练速度。除此之外，Object Drawer 改进了视角鲁棒性效果，能够在视角差异下快速推理出高分辨率图像，支持实时高清可交互的三维模型展示。用户通过手机环绕目标商品拍摄视频，即可自动重建 3D 模型，解决了光照迁移和纹理细节还原的问题，使场景渲染更逼真，输出的 3D 粗模亦可导入 CAD 等图形工具。

图 14: NeRF 技术原理示意



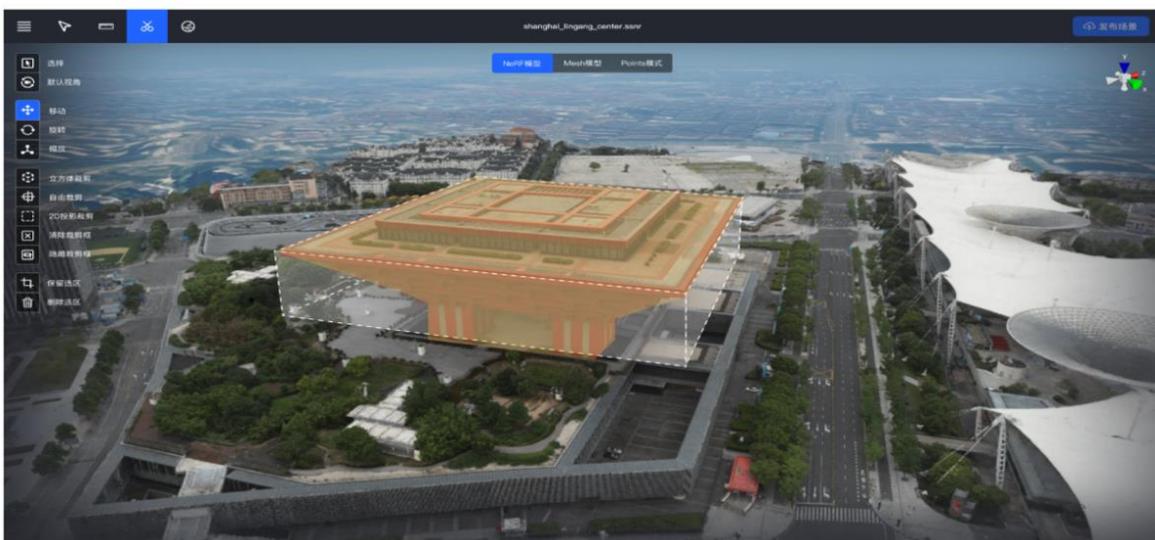
资料来源：阿里云，信达证券研发中心

Object Drawer 已面向部分淘宝和天猫商家开放，用于商品建模，有效降低了 3D 建模的成本和时间，提高了建模的效率和质量。该技术在家居家装、电子商务等行业具有广泛的应用前景。同时，阿里巴巴宣布 Object Drawer 向学术界和普通用户开放，鼓励广大开发者和科研人员交流体验，促进技术的进一步发展和应用。

七、商汤科技：格物、琼宇

2023 年 7 月 7 日，商汤科技在 2023 世界人工智能大会上发布了两款 3D 内容生成平台：格物和琼宇。“格物 SenseThings 2.0”专注于小物体的重建和交互，而“琼宇 SenseSpace 2.0”则专注于大空间的三维重建。这两个平台都基于神经辐射场技术（NeRF），能够实现空间和高精度复刻与交互，适用于从小型物品到城市级大空间的多种 3D 重建需求。

图 15：琼宇的效果展示



资料来源：商汤科技，信达证券研发中心

格物平台专注于小物体的 3D 内容生成，能够实现各类物体，包括高反光和镜面物体如珠宝首饰、小家电和金属物件等的超细节复刻还原。这种精确的还原能力使其适用于商业广告、商品营销、数字孪生管理等领域，提供卓越的复刻效果，并能与线上商城结合，全方位展现商品特性。其技术优势在于，无需改变物品状况下的精确还原，生成的 3D 内容可再编辑和创作，实现 3D 内容生成的全链路。

琼宇平台则是专注于大空间 3D 内容生成，具备城市级大尺度空间重建生成能力，能够复刻和还原超真实感的场景。其应用场景广泛，包括城市及园区的数字孪生、影视创作、文旅、电商等，支持自由漫游和实时交互编辑，能够完成城市级大尺度空间的重建。技术上，琼宇拥有厘米级重建精度和实时渲染与互动能力，通过多源数据融合和超精细化算法，在短时间内完成大规模空间的重建。

图 16: 琼宇的技术储备

大范围、复刻级实景建模

采用多NeRF模型分块训练和拼接技术，实现100km²城市级NeRF模型重建。


多源异构数据融合，保持局部细节

多传感器数据融合支持RGB、RGBD、无人机等多种数据融合重建、地空融合重建以及室内重建。


4K实时、高精度渲染

支持4k分辨率实时渲染，室外：每10000m²精度5cm，室内：每1000m²精度1cm。


训练、渲染、交互一体化系统

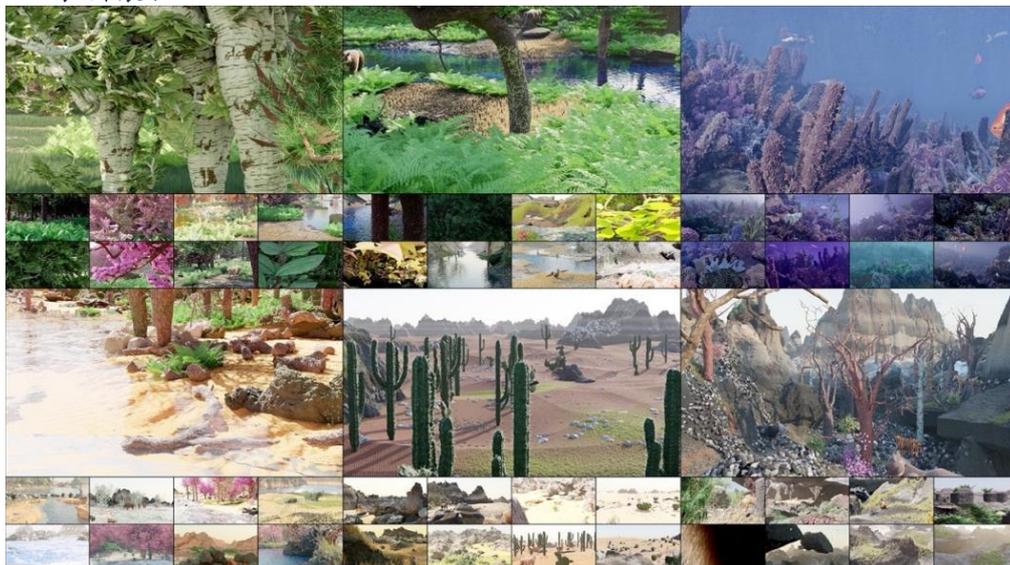
首次实现覆盖算法、算子、计算系统、用户交互的城市级NeRF系统，为训练、渲染、交互和应用提供坚实基础。

资料来源：商汤科技，信达证券研发中心

商汤科技的格物和琼宇平台展现了其在 3D 内容生成和数字孪生技术方面的领先地位，满足了不同行业对于可交互 3D 实景内容的需求。依托于商汤科技的大模型体系“日日新 SenseNova”，这些 AI 生成 3D 模型平台得到了强大的技术支持，推动了 AI 领域的持续创新和发展。

八、普林斯顿大学：Infinigen AI

Infinigen AI 由普林斯顿大学研究人员开发并在 CVPR 会议上发表，是一个基于开源建模软件 Blender 的程序化生成器。它能够通过随机数学规则生成一系列自然界的“植物、动物”场景，并通过输入提示词自定义地形，添加各种自然现象如“云、雨、雪、雷暴、天火”等。Infinigen AI 的发布为计算机视觉领域提供了强大的工具，尤其是在数据生成和场景模拟方面，开辟了新的研究和应用可能性。

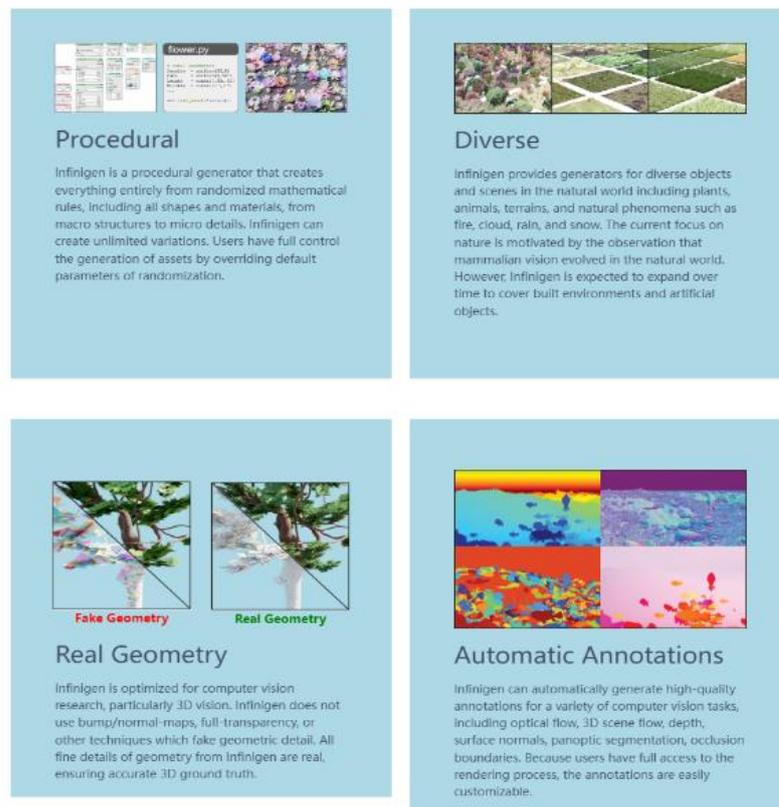
图 17: Infinigen AI 的效果展示


资料来源：Infinigen 官网，信达证券研发中心

技术上，Infinigen AI 是一个基于程序化方法的生成器，完全通过随机化的数学规则来创造内容，包括不同形状和材料，从宏观结构到微观细节，能够创造多样的变化。用户可以通过覆盖随机参数的默认值来完全控制资产的生成。Infinigen 提供了多样化对象和场景的生成器，包括植物、动物、地形和自然现象，如火、云、雨、雪等。当前对自然界的关注是基于哺乳动物视觉在自然世界中演化的观察。研发团队预计在 Infinigen 的未来开发中整合资源，以让 Infinigen 生成更多的环境和对象来涵盖视觉世界的万物。Infinigen 专为计算机视觉研究优化，尤其是 3D 视觉，不使用凹凸/法线贴图、全透明或其他虚假几何细节技术。Infinigen 生成的所有细节都是真实的，确保了精确的 3D 地面真实性。

尽管 Infinigen AI 在生成高质量逼真 3D 场景方面表现出色，其效率仍有待提高。尽管如此，Infinigen AI 的潜在应用范围非常广泛，包括游戏开发、电影制作、虚拟现实、增强现实及计算机视觉研究等领域，其程序化生成方法和开源特性促进了相关技术的进一步发展。Infinigen 还能自动为多种计算机视觉任务生成高质量注释，包括光流、3D 场景流、深度、表面法线、全景分割、遮挡边界等。因为用户可以完全访问渲染过程，这些注释易于自定义。

图 18: Infinigen AI 的技术特点



资料来源: Infinigen 官网, 信达证券研发中心

九、Adobe: DMV3D

Adobe 研究院与澳大利亚国立大学和斯坦福大学的研究人员合作，共同开发了 DMV3D (Diffusion Model for Volumetric 3D)，一个基于 transformer 的 3D 大型重建模型，能在单个 A100 GPU 上仅需 30 秒生成高保真 3D 图像。这 DMV3D 标志着 3D 内容生成领域的重要突破，尤其是在应用程序如 VR、AR、机器人技术和游戏等领域，大大减少了人工成本。

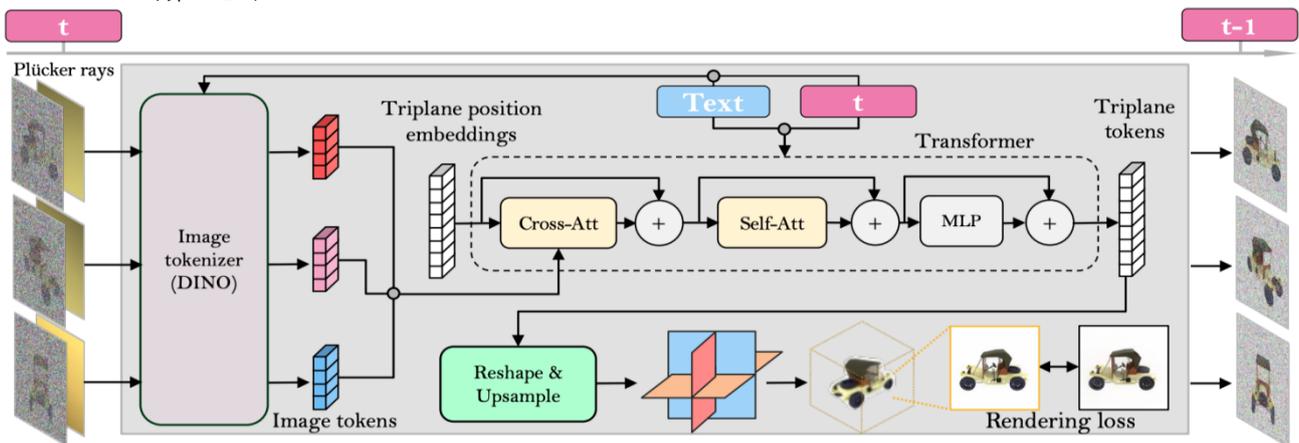
图 19: DMV3D 的效果展示



资料来源: DMV3D. Yinghao Xu 等《DMV3D:Denoising Multi-View Diffusion using 3D Large Reconstruction Model》, 信达证券研发中心

DMV3D 的技术核心在于它是一个单阶段的扩散模型, 将 3D NeRF 重建和渲染集成到其降噪器中, 以端到端的方式进行训练, 无需直接 3D 监督。这种方法避免了单独训练用于潜在空间扩散的 3D NeRF 编码器和繁琐的对每个对象进行优化的方法。在推理过程中, DMV3D 在输入视点渲染去噪后的图像, 并将它们与噪声结合, 以获得下一个去噪步骤的较少噪声的输入。在多视图图像完全去噪后, 模型提供了一个干净的三平面 NeRF, 使得 3D 生成成为可能。除此以外, DMV3D 基于 Transformer 构建, 利用大型 Transformer 模型快速生成高保真 3D 图像, 能够基于单个图像或文本输入生成 3D NeRF, 显著缩短了 3D 对象的创建时间。

图 20: DMV3D 的算法结构



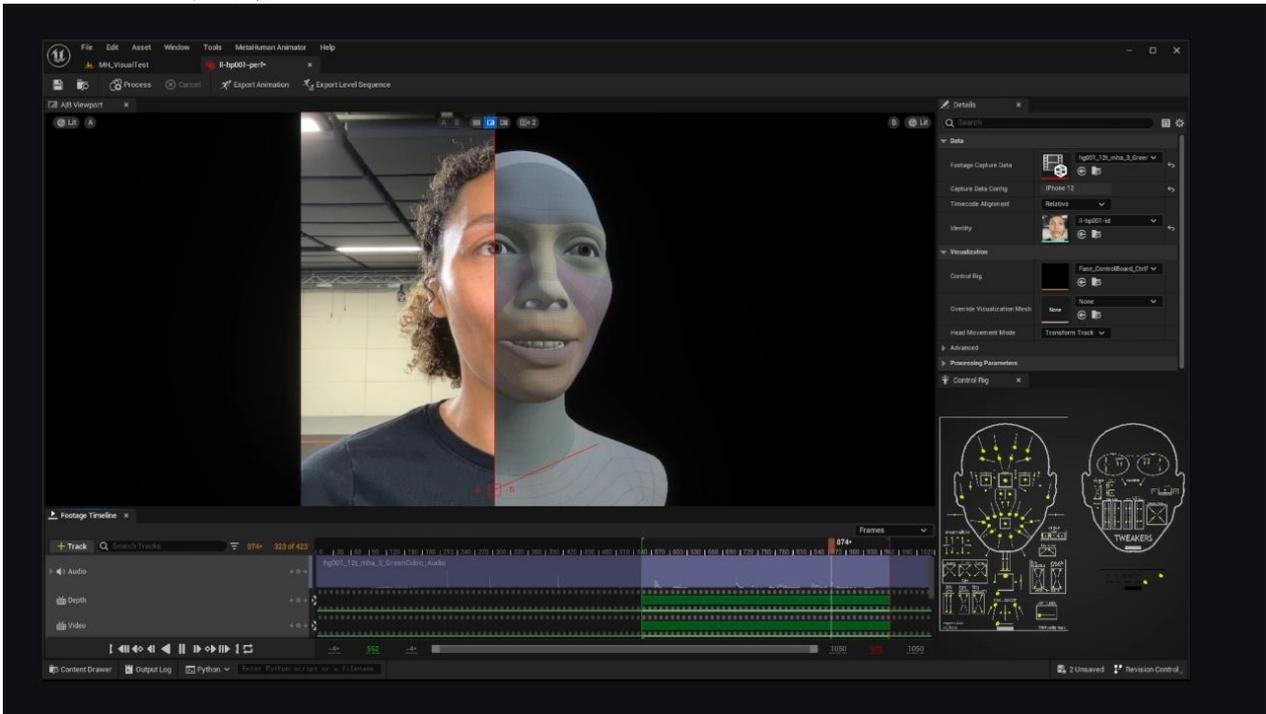
资料来源: DMV3D. Yinghao Xu 等《DMV3D:Denoising Multi-View Diffusion using 3D Large Reconstruction Model》, 信达证券研发中心

尽管 Adobe-DMV3D 在高效率生成高保真 3D 图像方面表现出色, 为 3D 视觉和图形领域带来了新的视角和可能性, 它也面临一些挑战, 特别是在数据集规模小、维数高、非结构化等条件下。截至目前, DMV3D 的实际应用潜力和性能优化仍然是研究和开发的关键方向。

十、Epic Games: MetaHuman

Epic Games 推出的 MetaHuman 是一款 AI 驱动的 3D 人类模型生成工具，旨在为游戏开发、电影、动画以及其他 3D 视觉效果密集型应用提供高度逼真的 3D 人类角色。MetaHuman Creator 作为一个云端应用，使用户能在几分钟内创建复杂的 3D 人类角色，大大减少了传统 3D 建模所需的时间和资源。该工具支持高度的定制化，包括性别、种族、年龄等特征，并且生成的角色可以直接导入到 Unreal Engine 中，也可导出为其他 3D 软件支持的格式。

图 21: MetaHuman 的效果展示



资料来源: MetaHuman 官网, 信达证券研发中心

技术原理上, MetaHuman Creator 是一个基于云的 Web 应用程序, 通过高端服务器完成所有计算工作并通过流式传输到用户设备上。它允许高度自定义, 包括面部特征、肤色、化妆品、头发等, 并集成了 MetaHuman Identity 功能, 使用户能够解算 MetaHuman DNA 数据并在编辑器中生成预览模型。与虚幻引擎 5 紧密集成, 支持在不同平台上保持性能和效果的平衡, 并提供详细的 LOD (Level of Detail) 信息。

MetaHuman Creator 不仅提供了快速创建和高度逼真的 3D 人类角色的能力, 还通过其用户友好的界面和丰富的定制选项, 简化了 3D 人物创作的复杂性。其云端服务消除了对高端本地硬件的需求, 为用户提供了极大的便利。然而, 作为一个高度先进的工具, 它可能需要用户有一定的学习曲线来充分利用其所有功能。此外, Epic Games 提供了大量教程、指南以及详细的文档和支持, 帮助用户解决使用过程中遇到的问题, 同时建立了一个社区, 促进了用户之间的交流和学习。

十一、Unity: Muse、Sentis、Meshy（插件）和 Unity6

Unity 作为全球领先的实时 3D 开发平台，不仅在游戏开发领域占据重要地位，还在 AI 和 3D 内容生成领域进行了积极的探索和创新。它推出了多款与 AI 结合的工具，旨在帮助开发者更高效地创建 3D 内容和提升游戏体验。这些工具包括 Unity Muse、Unity Sentis、Meshy 以及即将到来的 Unity 6 版本，展现了 Unity 在融合 AI 技术以提升 3D 内容创作和游戏互动性方面的创新努力。

Unity Muse 是一个集成在 Unity 开发环境中的生成式 AI 工具，它使用自然语言处理（NLP）技术帮助开发者生成 3D 模型、纹理和动画。通过一个用户友好的聊天机器人界面，开发者可以输入指令来获取资源、生成代码，以及创建 2D 和 3D 纹理。Muse 还计划支持角色动画生成和角色行为定义，使开发者能通过自然语言描述来生成角色的行为代码和触发条件。它的早期测试版本已向开发者开放，提供了一种全新的、以自然语言为基础的交互方式来创建 3D 内容。

图 22: Unity Muse 的效果展示和特点

Ideate and iterate faster with the power of AI

With Unity Muse, your prompts and inputs lead to fast results. Tasks that previously required long hours just to put together a concept, now come together quickly and easily. Spend less time on setup and more time on what matters most to you.



UNITY HDRP DEMO SCENE REIMAGINED WITH UNITY MUSE

Muse features



Texture

Use simple, natural input – a prompt, an image, a doodle – to instantly generate game-ready textures in any style, all inside your project in the Unity Editor.



Sprite

Produce 2D art instantly, directly in the Unity Editor. With just a few simple prompts, you'll receive a selection of modifiable, scene-ready sprites in seconds.



Chat

Solve challenging problems, get ideas for your project, and generate usable code through conversational AI chat that guides you on your creation process.



Animate

Pre-release

Bring humanoid characters to life with just a few text prompts, saving hours of traditional setups – designed for both aspiring and seasoned animators.



Behavior

Pre-release

Instantly set up character interactions with Muse Behavior. Simply describe the desired behaviors and Unity Muse will create behavior trees in the Unity Editor.

资料来源: Unity 官网, 信达证券研发中心

Unity Sentis 是 Unity 推出的工具，它允许开发者将通用的 ONNX 格式 AI 模型接入 Unity Runtime，从而使游戏能够利用 AI 模型来实现文字转语音、智能 NPC 等功能，增强游戏的互动性和智能性。Sentis 为游戏和应用提供了一种简便的方式，使其能够集成最新的 AI 技术，提升用户体验。

图 23: Unity Sentis 的效果展示和特点

High performance on user devices

Sentis leverages the compute power of end-user devices rather than the cloud, which eliminates complex cloud infrastructure, network latency, and recurring inference costs.



Technical features

Documentation

Accessible

An intuitive "plug-and-play" API makes AI model integration accessible via the ONNX model file standard.

No cost

Get free and unlimited AI model inferences with no cloud infrastructure required.

Multiplatform support

Sentis works with all Unity-supported platforms, from mobile to PC to popular game consoles.

Performant

Optimize your AI model for runtime environments, then call inferences at game speed for fluid gameplay.

Hardware optimization

Deploy inferences to the best available hardware (CPU or GPU) and leverage existing Unity compute and job systems.

Secure

AI models are run locally with Unity Sentis, with no data stored or transferred to the cloud.

资料来源: Unity 官网, 信达证券研发中心

Meshy 是可以在 Unity 中使用的 AI 插件，它提供了多种 AI 功能，如 AI Texturing (Beta)、Text to 3D (Beta) 和 Image to 3D (Alpha)。这些功能允许用户通过文本描述、2D 图像或直接文本输入来生成 3D 模型和纹理。Meshy 的生成速度快，支持 PBR 贴图，并提供了 Unity 插件，使得在 Unity 内部直接使用这些 AI 功能成为可能。Meshy 为开发者提供了一种高效且直观的方式，以实现从文本或图像到 3D 内容的快速转换。

图 24: Meshy 的效果展示

Text to 3D

Our text to 3D tool empowers creators, including those with no prior 3D experience, to generate 3D models within just a minute using text input.



Image to 3D

Easily convert your desired object images into stunning 3D models. With a simple image input, our AI transforms 2D into 3D in under a minute.



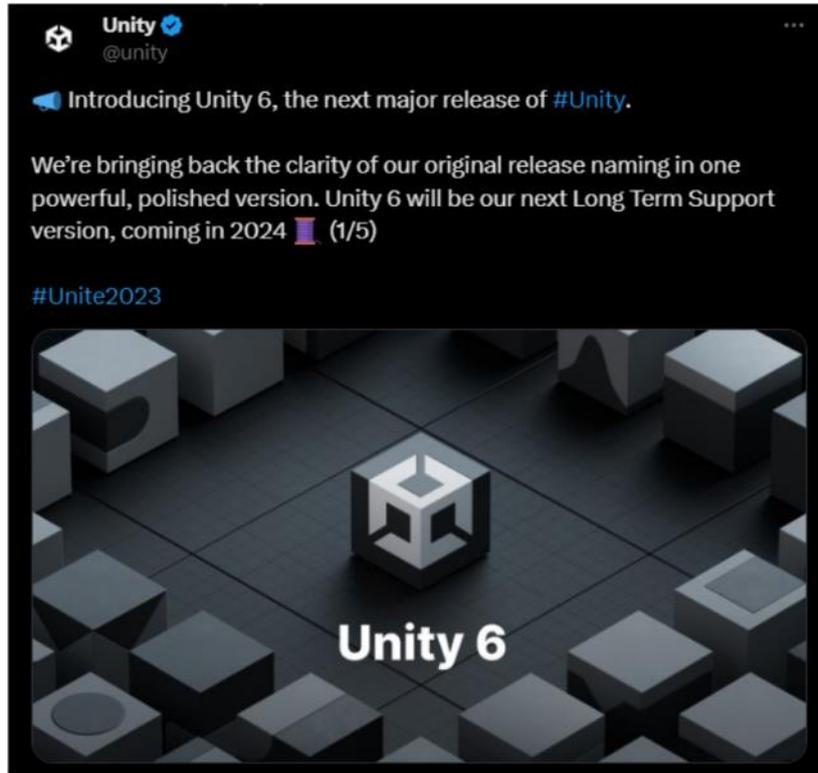
AI Texturing

Our AI texturing tool allows artists to choose either text prompt or 2D concept art, as well as an untextured model as input. AI will do the automatic texturing for your model in less than a minute.

资料来源: Meshy 官网, 信达证券研发中心

Unity 6 是 Unity 的下一个长期更新版本，预计将搭载多项全新产品能力，包括更高性能的渲染、灯光与延展功能。Unity 6 还计划推出一个全新的 WebGPU 图形后端，使得 Unity 游戏能够在各类浏览器中原生运行。这意味着 Unity 6 将为开发者提供更广泛的平台支持，同时带来性能和图形渲染的显著提升，为创造更丰富、更互动的 3D 体验和 AI 应用提供了强大的保障。

图 25: Unity 6 官方预热



资料来源: X-Unity 官网, 信达证券研发中心

通过这些工具，Unity 在 AI 生成 3D 内容方面的探索和创新为开发者提供了更加强大和灵活的创作工具，为未来 3D 数字资产发展和 AI 生 3D 技术进步提供了基础。

我们当前节点建议关注:

- 1) 拥有 3D 内容制作能力，前期储备了 3D 数字资产的公司，有望在前期与苹果的调参测评沟通中占据先发优势，如丝路视觉、凡拓数创、恒信东方、飞天云动、视觉中国；
- 2) 影视、赛事直播、轻度游戏、教育行业有望优先受益，办公、工业、医疗等有望后续应用度提升。海外目前热度较高的苹果 MR 原生应用 JigSpace 对 3D 模型的拆解效果较优有望后续提高市场对教育赛道的认可度。其他方向如 VR 游戏移植、MR 原生游戏开发、赛事直播等方向同样值得关注，如佳创视讯、宝通科技、力盛体育、恺英网络、名臣健康；
- 3) 其他建议关注：风语筑、罗曼股份、岭南股份、锋尚文化等。

十二、风险因素

宏观经济波动风险：宏观经济波动可能致使消费能力及欲望下降，阻碍产业发展。

MR 设备市场销量不及预期风险：MR 设备销量不及预期，可能导致技术无法突破，影响产品渗透，也可能导致产业链厂商研发成本无法收回，行业萎靡。

信达传媒互联网及海外研究团队简介

冯翠婷，信达证券传媒互联网及海外首席分析师，北京大学管理学硕士，香港大学金融学硕士，中山大学管理学学士。2016-2021 年任职于天风证券，覆盖互联网、游戏、广告、电商等多个板块，及元宇宙、体育二级市场研究先行者（首篇报告作者），曾获 21 年东方财富 Choice 金牌分析师第一、Wind 金牌分析师第三、水晶球奖第六、金麒麟第七，20 年 Wind 金牌分析师第一、第一财经第一、金麒麟新锐第三。

刘旺，信达证券传媒互联网及海外团队联系首席分析师。北京大学金融学硕士，北京邮电大学计算机硕士，北京邮电大学计算机学士，曾任职于腾讯，一级市场从业 3 年，创业 5 年（人工智能、虚拟数字人等），拥有人工智能、虚拟数字人、互联网等领域的产业经历。

凤超，信达证券传媒互联网及海外团队高级研究员，本科和研究生分别毕业于清华大学和法国马赛大学，曾在腾讯担任研发工程师，后任职于知名私募机构，担任互联网行业分析师。目前主要负责海外互联网行业的研究，拥有 5 年的行研经验，对港美股市场和互联网行业有长期的跟踪覆盖。主要关注电商、游戏、本地生活、短视频等领域。

李依韩，信达证券传媒互联网及海外团队研究员。中国农业大学金融硕士，2022 年加入信达证券研发中心，覆盖互联网板块。曾任职于华创证券，所在团队曾入围 2021 年新财富传播与文化类最佳分析师评比，2021 年 21 世纪金牌分析师第四名，2021 年金麒麟奖第五名，2021 年水晶球评比入围。

白云汉，信达证券传媒互联网及海外团队研究员。美国康涅狄格大学金融硕士，曾任职于腾讯系创业公司投资部，一级市场从业 2 年。后任职于私募基金担任研究员，二级市场从业 3 年，覆盖传媒互联网赛道。2023 年加入信达证券研发中心，目前主要专注于美股研究以及结合海外映射对 A 股、港股的覆盖。

分析师声明

负责本报告全部或部分内容的每一位分析师在此申明，本人具有证券投资咨询执业资格，并在中国证券业协会注册登记为证券分析师，以勤勉的职业态度，独立、客观地出具本报告；本报告所表述的所有观点准确反映了分析师本人的研究观点；本人薪酬的任何组成部分不曾与，不与，也将不会与本报告中的具体分析意见或观点直接或间接相关。

免责声明

信达证券股份有限公司(以下简称“信达证券”)具有中国证监会批复的证券投资咨询业务资格。本报告由信达证券制作并发布。

本报告是针对与信达证券签署服务协议的签约客户的专属研究产品，为该类客户进行投资决策时提供辅助和参考，双方对权利与义务均有严格约定。本报告仅提供给上述特定客户，并不面向公众发布。信达证券不会因接收人收到本报告而视其为本公司的当然客户。客户应当认识到有关本报告的电话、短信、邮件提示仅为研究观点的简要沟通，对本报告的参考使用须以本报告的完整版本为准。

本报告是基于信达证券认为可靠的已公开信息编制，但信达证券不保证所载信息的准确性和完整性。本报告所载的意见、评估及预测仅为本报告最初出具日的观点和判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会出现不同程度的波动，涉及证券或投资标的的历史表现不应作为日后表现的保证。在不同时期，或因使用不同假设和标准，采用不同观点和分析方法，致使信达证券发出与本报告所载意见、评估及预测不一致的研究报告，对此信达证券可不发出特别通知。

在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议，也没有考虑到客户特殊的投资目标、财务状况或需求。客户应考虑本报告中的任何意见或建议是否符合其特定状况，若有必要应寻求专家意见。本报告所载的资料、工具、意见及推测仅供参考，并非作为或被视为出售或购买证券或其他投资标的的邀请或向人做出邀请。

在法律允许的情况下，信达证券或其关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并可能会为这些公司正在提供或争取提供投资银行业务服务。

本报告版权仅为信达证券所有。未经信达证券书面同意，任何机构和个人不得以任何形式翻版、复制、发布、转发或引用本报告的任何部分。若信达证券以外的机构向其客户发放本报告，则由该机构独自为此发送行为负责，信达证券对此等行为不承担任何责任。本报告同时不构成信达证券向发送本报告的机构之客户提供的投资建议。

如未经信达证券授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。信达证券将保留随时追究其法律责任的权利。

评级说明

投资建议的比较标准	股票投资评级	行业投资评级
本报告采用的基准指数：沪深 300 指数（以下简称基准）； 时间段：报告发布之日起 6 个月内。	买入 ：股价相对强于基准 15% 以上；	看好 ：行业指数超越基准；
	增持 ：股价相对强于基准 5%~15%；	中性 ：行业指数与基准基本持平；
	持有 ：股价相对基准波动在±5% 之间；	看淡 ：行业指数弱于基准。
	卖出 ：股价相对弱于基准 5% 以下。	

风险提示

证券市场是一个风险无时不在的市场。投资者在进行证券交易时存在赢利的可能，也存在亏损的风险。建议投资者应当充分深入地了解证券市场蕴含的各项风险并谨慎行事。

本报告中所述证券不一定能在所有的国家和地区向所有类型的投资者销售，投资者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、税收等方面咨询专业顾问的意见。在任何情况下，信达证券不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任，投资者需自行承担风险。