

行业及产业
电子

OpenAI 推出新一代音视频工具 Sora 2

——人工智能月度跟踪

强于大市

一年内行业指数与沪深 300 指数对比走势：



资料来源：聚源数据，爱建证券研究所

相关研究

- 《电子行业周报：国产示波器实现高端突破》
2025-10-20
- 《电子行业周报：湾芯展：期待新凯来新惊喜》
2025-10-13
- 《电子行业周报：先进封装玻璃基板实现技术突破》
2025-10-09
- 《电子行业周报：AI 服务器+智能手机需求爆发推动 NAND Flash 价格上扬》
2025-09-29
- 《人工智能月度跟踪：CPO&CPC 有望开启新一轮成长周期》
2025-09-29

证券分析师

许亮
S0820525010002
0755-83562506
xuliang@ajzq.com

联系人

朱俊宇
S0820125040021
021-32229888-25520
zhujunyu@ajzq.com

投资要点：

- **引子：**2025 年 9 月 30 日，OpenAI 推出新一代音频、视频生成模型 Sora 2，并同步上线独立 iOS 应用。Sora 2 被 OpenAI 誉为“视频领域的 GPT-3.5 时刻”，此次更新不仅在视频生成质量上实现了显著提升，更首次支持了与画面精准匹配的原生音频生成，标志着 AI 视频生成技术迈入一个全新的发展阶段。
- **Sora 采用 DIT 架构，性能优于同期发布模型。**2024 年 2 月 16 日，OpenAI 发布首个文本生成视频模型 Sora，其核心采用 Diffusion Transformer 深度融合架构，依托自注意力机制等设计，可增强视频帧间连贯性与文本与视觉语义的匹配精准度。相较于 Gen-2、Lumiere 等同期模型，Sora 的优势尤为突出：1) 生成时长达 60 秒，远超 Gen-2 (18 秒)、Lumiere (5 秒)、MoonValley (6 秒) 的上限，能承载更具叙事性的创作需求；2) 生成类型覆盖 T2V (文生视频)、I2V (图生视频)、V2V (视频生视频)，并额外支持 VFI (视频插帧)，多模态创作能力更全面；3) 技术架构上，凭借 Diffusion Transformer 的特性，时序一致性表现更佳，可减少画面闪烁、物体运动不连贯等问题，让动态视觉效果更贴近真实场景。
- **OpenAI Sora 2 相较于初代实现了多维度的卓越升级。**1) 在音视频同步方面，其彻底解决了初代的无声局限，可基于文本指令原声生成贴合场景的完整音效，无论是人物说话时的语音与环境音、奔跑场景的脚步声与风声，还是钢琴演奏的琴键声与空间回响等，均能自动匹配，无需后期额外配音，实现“声画一体”的创作体验；2) 物理模拟精度上，Sora 2 实现了突破性提升：针对初代水流不自然问题，升级水、烟雾等流体模拟使其更贴合真实场景；通过精准模拟重力、惯性与摩擦力，修正了物体飘浮、碰撞反馈不真实等缺陷；同时让人物动作严密符合人体工学，有效解决动作僵硬问题。对比测试显示，其在人物走路、水流动态等场景的真实度提升 36%-70%，其中水流动态提升最为显著。3) 此外，Sora 2 新增“Cameo”功能，用户通过短暂的视频与音频采集，即可将自身或他人的形象与声音植入模型，后续能在任意场景中高度保真地呈现该角色，极大拓展了角色互动可能性，丰富了个性化创作与社交体验。
- **Sora 2 应用十分广泛，已覆盖营销广告、教育培训、产品展示、自媒体创作、艺术创作等多个领域，其中营销广告领域占比最高，达 30%。**其核心优势在于相较传统视频制作模型能显著提升效率并大幅压缩成本：在营销广告领域，美妆广告制作成本从 8000 元降至 25 元、节省比例约 99.70%，汽车广告制作周期从 1 个月缩短至 20 分钟，帮助中小商家打破高预算限制；教育培训场景中，历史教育类内容制作周期从 1 个月缩短至 10 分钟，满足教学内容快速迭代需求；自媒体领域里，美食博主单条内容成本从 200 元降至 20 元、时间从 2 小时缩短至 10 分钟，还能实拍不可行的旅游博主直接生成内容，助力创作者高频更新、响应热点；产品展示与创意艺术领域也实现了成本 97% 以上、时间数倍至数十倍的节省。依托“低成本、高效率、个性化”特性，Sora 2 推动 AI 视频制作规模化、高效化落地，标志着 AI 音视频技术进入“工业化生产阶段”，未来若在视频时长、画面分辨率上进一步突破，应用场景还将向影视制作、游戏开发、虚拟直播等领域拓展。
- **投资建议：**随着 AI 大模型在音视频功能上的持续升级，未来音视频制作的成本将大幅下降，这对于下游广告传媒行业的发展十分有利。同时，AI 生成音视频也将推动数据流量的加速爆发，建议关注存储芯片行业的投资机会。
- **风险提示：**1) 国际贸易摩擦加剧 2) 下游需求不及预期 3) 技术升级进度滞后

目录

引子：OpenAI 推出新一代音视频工具 Sora 2	4
1. Sora 采用 DiT 架构，相较于其他视频模型优势明显	4
2. OpenAI Sora 2 性能实现卓越升级	5
3. Sora 2 应用领域广泛	6
4. 风险提示	8

图表目录

图表 1 : Sora 模型框架	4
图表 2 : Sora 与其他视频生成模型参数性能对比	5
图表 3 : Sora 2 解决初代 Sora 无声的局限	5
图表 4 : Sora2 解决了 Sora 存在的物理问题	6
图表 5 : Sora 2 相较于初代 Sora 产品物理真实度更高	6
图表 6 : Sora 2 应用领域广泛	7
图表 7 : Sora 2 相较于传统视频制作模型: 成本下降且效率提升	7

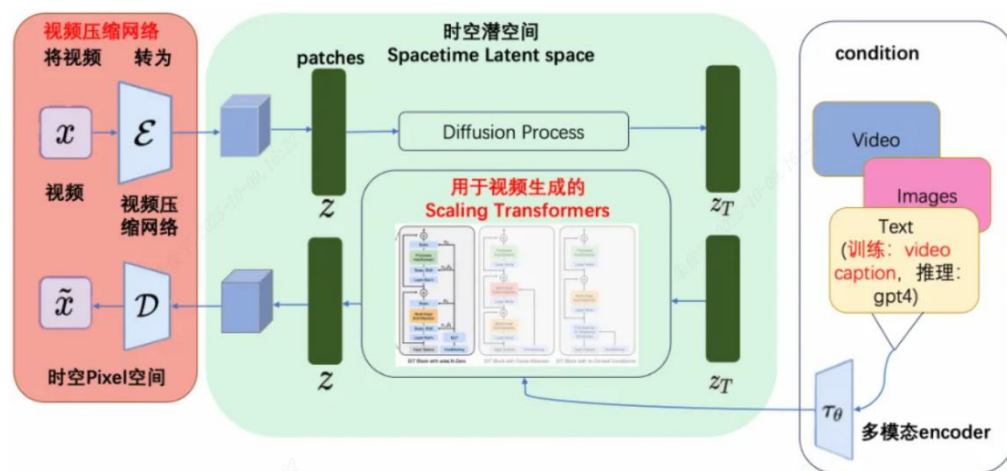
引子：OpenAI 推出新一代音视频工具 Sora 2

2025 年 9 月 30 日，OpenAI 推出新一代音频、视频生成模型 Sora 2，并同步上线独立 iOS 应用。Sora 2 被 OpenAI 誉为“视频领域的 GPT-3.5 时刻”，此次更新不仅在视频生成质量上实现了显著提升，更首次支持了与画面精准匹配的原生音频生成，标志着 AI 视频生成技术迈入一个全新的发展阶段。

1. Sora 采用 DiT 架构，相较于其他视频模型优势明显

2024 年 2 月 16 日，OpenAI 发布首个文本生成视频模型 Sora。Sora 采用 Diffusion Transformer 深度融合架构（将扩散模型与 Transformer 深度结合的创新架构，简称 DiT 架构），旨在提升从文本到视频的生成效果。其依托自注意力机制、动态学习策略及面向视频生成任务优化的架构设计，可增强生成视频的帧间连贯性，同时提升视觉内容与文本语义的匹配精准度。

图表 1：Sora 模型框架



资料来源：阿里云，爱建证券研究所

相较于 Gen-2、Lumiere、MoonValley 等视频生成模型，Sora 具备在视频时长、多模态生成覆盖度、技术架构的时序一致性等方面展现出显著优势。具体而言，1) Sora 视频生成时间长，远超 Gen-2（18 秒）、Lumiere（5 秒）、MoonValley（6 秒）的时长上限，能承载更具叙事性或完整性的创作需求；2) 生成类型上，它不仅覆盖 T2V（文生视频）、I2V（图生视频）、V2V（视频生视频），还额外支持 VFI（视频插帧），多模态创作的丰富度更全面。3) 技术架构层面，Sora 依托 Diffusion Transformer 架构，时序一致性表现更佳，可减少画面闪烁、物体运动不连贯等问题，使动态视觉效果更贴近真实场景。

图表 2：Sora 与其他视频生成模型参数性能对比

模型	企业或团队	发布时间	主要生成类型	最大时长	机位控制	其他特点	是否开源
Sora	Open AI	2024/02	T2V、I2V、V2V、VFI	60s	有	Diffuse Transformer 架构，时序一致性较好	否
Gen-2	Runway	2023/06	T2V、I2V、V2V	18s	有		否
Lumiere	Google	2024/01	T2V、I2V	5s	无	STUNet 架构	否
MoonValley	MoonValley	2023/11	T2V、I2V	6s	有		否
Emu Video Emu Edit	Meta	2023/09	T2V	4s	无	Latent Diffusion	否
Morph	Morph Studio	2023/11	T2V	7s	无		否
Stable Video Diffusion	Stability AI	2023/11	I2V	4s	无		是
MagicVideo-V2	字节跳动	2024/01	T2V、I2V、V2V、VFI	-	无	U-Net 核心	否
Pika 1.0	PIKA Labs	2023/11	T2V、I2V、V2V	7s	有		否

资料来源：陈巍，DeepSeek 技术社区，爱建证券研究所

注：仅选取与 Sora 发布时间接近模型，未覆盖全部

T2V：文生视频；I2V：图生视频；V2V：视频生视频；VFI：视频插帧

但作为初代文本生成视频模型，Sora 仅能生成完全无声的视频，用户需在后期手动为其添加音频。这一额外操作不仅会打断连贯的创作流程，还大幅增加了内容制作的时间成本与人力成本。为解决这一核心痛点，OpenAI 于 2025 年 9 月 30 日推出新一代模型 Sora 2。

2. OpenAI Sora 2 性能实现卓越升级

相较于初代 Sora，Sora 2 首先解决无声局限，实现原生音视频同步。创作者只需输入关键词，即可生成包含音效的完整影片，无需后期额外配音。从具体场景来看，初代 Sora 在人物说话、奔跑、海浪拍岸等场景中均无音效，而 Sora 2 能针对不同场景自动生成贴合场景的声音。

图表 3：Sora 2 解决初代 Sora 无声的局限

画面内容	Sora	Sora 2 自动生成音效
人物说话	无声	语音、呼吸声、环境音
奔跑场景	无声	脚步声、喘息声、风声
海浪拍岸	无声	海浪声、海鸥叫声、风声
咖啡制作	无声	研磨声、蒸汽声、杯子碰撞声
汽车行驶	无声	引擎声、轮胎声、风阻声
钢琴演奏	无声	琴键声、踏板声、空间回响

资料来源：API 易，爱建证券研究所

Sora 2 在物理模拟精度上实现了对初代 Sora 的突破性升级。针对初代“水流方向不自然”的问题，Sora 2 升级了水、烟雾等流体的模拟精度，使其运动轨迹与形态更贴近现实物理规律；针对初代存在的重力效果错误（如物体飘浮）与碰撞反馈不真实问题，它能够通过精准模拟重力、惯性、摩擦力进行修正；针对人物动作僵硬、不流畅的缺陷，它还能让人物动作严格符合人体工学。

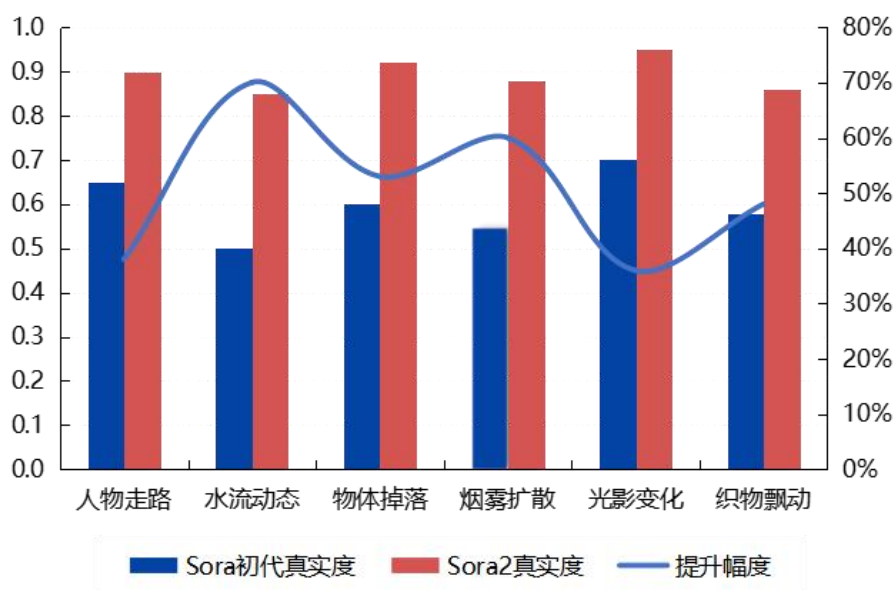
图表 4: Sora2 解决了 Sora 存在的物理问题

Sora 存在的物理问题	Sora2 对应的改进/升级
水流方向不自然	流体（水、烟雾）模拟更真实
重力效果不正确（物体飘浮）	精准模拟重力、惯性、摩擦力
碰撞反馈不真实	
人物动作僵硬、不流畅	人物动作符合人体工学
光影变化不符合逻辑	光影变化符合光学原理

资料来源：API 易，爱建证券研究所

从对比测试结果来看，Sora 2 在人物走路、水流动态、物体掉落等场景的真实度，相较于初代均有明显提升，提升区间为 36%-70%，其中水流动态场景提升最显著，达 70%。

图表 5: Sora 2 相较于初代 Sora 产品物理真实度更高



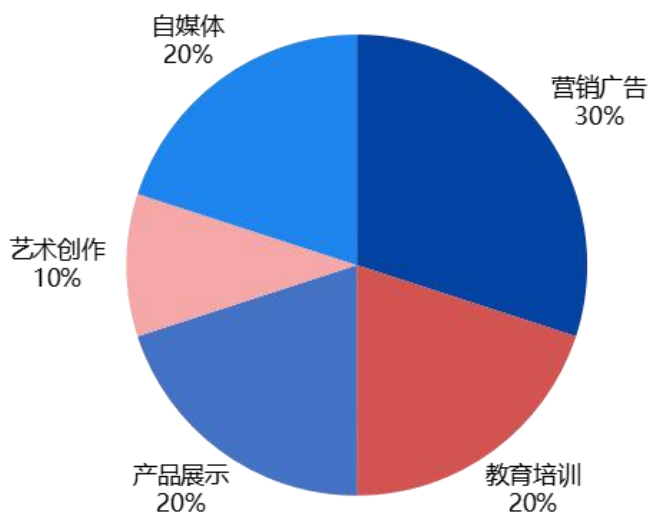
资料来源：API 易，爱建证券研究所

此外，Sora 2 还引入了“Cameo”功能。借助该功能，用户只需完成一次短暂的视频与音频采集，就能将自己或朋友的形象与声音注入模型；后续在任意场景中，这一角色都能以高度保真的方式呈现。这一设计有效拓展了角色互动的可能性，丰富了用户的使用体验。

3. Sora 2 应用领域广泛

Sora 2 应用领域广泛，涵盖营销广告、教育培训、产品展示、自媒体创作以及艺术创作等领域，其中营销广告领域的占比最高，达 30%。

图表 6: Sora 2 应用领域广泛



资料来源: API 易, 爱建证券研究所

Sora 2 相较于传统视频制作模型, 核心优势在于能在显著提升效率的同时大幅压缩成本。 1) 在营销广告领域, Sora 2 凭借极低的制作成本, 帮助中小商家打破“高预算才能做广告”的限制(如美妆广告制作成本从 8000 元降至 25 元, 节省比例约 99.70%), 让中小商家也能轻松开展广告宣传; 2) 在教育培训场景中, Sora 2 可大幅压缩课程制作时间(如历史教育类内容, 它将原本 1 个月的制作周期缩短至 10 分钟), 有效满足了培训机构对教学内容快速迭代的需求; 3) 在自媒体领域中, Sora 2 的低成本与高效率特性, 支持创作者和商家更频繁地更新内容、快速响应市场热点, 进一步提升了内容运营的竞争力。

图表 7: Sora 2 相较于传统视频制作模型: 成本下降且效率提升

模型	案例类型	传统制作成本	Sora2 成本	节省比例	时间节省
营销广告	美妆广告	8,000 元	25 元	99.70%	3 天→8 分钟
	房地产宣传	50,000 元	60 元	99.90%	2 周→15 分钟
	汽车广告	200,000 元	80 元	99.96%	1 月→20 分钟
教育培训	历史教育	100,000 元	30 元	99.97%	1 月→10 分钟
	科学教育	50,000 元	35 元	99.93%	2 周→12 分钟
自媒体内容	美食博主	200 元/条	20 元/条	90.00%	2 小时→10 分钟
	旅游博主	实拍不可行	40 元/条	-	直接生成
产品展示	电商产品	500 元/条	15 元/条	97.00%	1 天→5 分钟
	SaaS 演示	8,000 元	50 元	99.40%	3 天→15 分钟
创意艺术	艺术 MV	100,000 元	600 元	99.40%	1 月→2 小时

资料来源: API 易, 爱建证券研究所

在 Sora 2 技术持续迭代的背景下, OpenAI 不仅攻克 AI 视频生成领域的核心痛点, 更依托“低成本、高效率、个性化”特性, 推动 AI 视频制作迈向规模化、高效化落地。从行业视角看, Sora 2 的发布标志 AI 音视频技术正式进入“工业化生产阶段”; 后续模型若在更长视频时长、更高画面分辨率上实现进一步突破, 其应用场景将向影视制作、游戏开发、虚拟直播等领域持续拓展。

4. 风险提示

- 1) 国际贸易摩擦加剧
- 2) 下游需求不及预期
- 3) 技术升级进度滞

爱建证券有限责任公司

上海市浦东新区前滩大道 199 弄 5 号

电话: 021-32229888

传真: 021-68728700

服务热线: 956021

邮政编码: 200124

邮箱: ajzq@ajzq.com

网址: <http://www.ajzq.com>

评级说明

投资建议的评级标准

报告中投资建议所涉及的评级分为股票评级和行业评级（另有说明的除外）。评级标准为报告发布日后 6 个月内的相对市场表现，也即以报告发布日后的 6 个月内的公司股价（或行业指数）相对同期相关证券市场代表性指数的涨跌幅作为基准。其中：A 股市场：沪深 300 指数（000300.SH）；新三板市场：三板成指（899001.CSI）（针对协议转让标的）或三板做市指数（899002.CSI）（针对做市转让标的）；北交所市场：北证 50 指数（899050.BJ）；香港市场：恒生指数（HIS.HI）；美国市场：标普 500 指数（SPX.GI）或纳斯达克指数（IXIC.GI）。

股票评级

买入	相对同期相关证券市场代表性指数涨幅大于 15%
增持	相对同期相关证券市场代表性指数涨幅在 5% ~ 15% 之间
持有	相对同期相关证券市场代表性指数涨幅在 -5% ~ 5% 之间
卖出	相对同期相关证券市场代表性指数涨幅小于 -5%

行业评级

强于大市	相对表现优于同期相关证券市场代表性指数
中性	相对表现与同期相关证券市场代表性指数持平
弱于大市	相对表现弱于同期相关证券市场代表性指数

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，本报告采用信息和数据来自公开、合规渠道，所表述的观点均准确地反映了我们对标的证券和发行人的独立看法。研究报告对所涉及的证券或发行人的评价是分析师本人通过财务分析预测、数量化方法、或行业比较分析所得出的结论，但使用以上信息和分析方法可能存在局限性，请谨慎参考。

法律主体声明

本报告由爱建证券有限责任公司（以下统称为“爱建证券”）证券研究所制作，爱建证券具备中国证监会批复的证券投资咨询业务资格，接受中国证监会监管。

本报告是机密的，仅供我们的签约客户使用，爱建证券不因收件人收到本报告而视其为爱建证券的签约客户。本报告中的信息均来源于我们认为可靠的已公开资料，但爱建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、意见等均仅供签约客户参考，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，爱建证券及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、评估及预测后续可随时更改。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，爱建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。

版权声明

本报告版权归爱建证券所有，未经爱建证券事先书面许可，任何机构或个人不得以任何形式翻版、复制、转载、刊登和引用。否则由此造成的一切不良后果及法律责任由私自翻版、复制、转载、刊登和引用者承担。版权所有，违者必究。