

谷歌 Gemini Omni 横空出世：

AI 视频编辑实现颠覆性跨越，从生成迈向编辑的质变

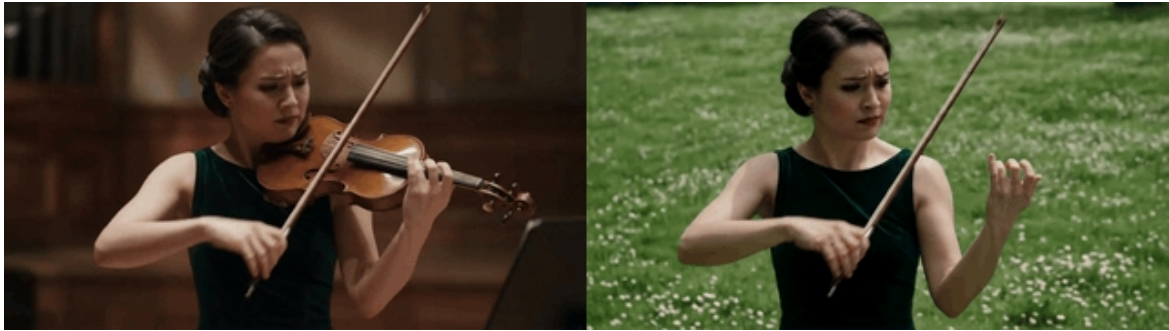
（2026 年 5 月 20 日 北京）在谷歌 I/O 2026 开发者大会上，DeepMind CEO Demis Hassabis 正式发布原生视频生成与编辑模型家族 Gemini Omni，首款成员 Gemini Omni Flash 同步上线，标志着 AI 视频技术完成从“从零生成”到“自然编辑”的关键跨越，为广电视听行业内容生产与制播流程带来革命性变革。



01 核心突破：

对话式语言驱动视频编辑，人物表演无损保留

Gemini Omni 最具颠覆性的创新并非从零生成视频，而是对已有视频进行自然语言驱动的精准确编辑，同时完整保留原有人物的表演细节。正如 Demis Hassabis 在发布会上所言：“正如 Nano Banana 重新定义了图片编辑，Omni 让你能用对话式语言、自然地编辑视频”。



这一技术突破的核心价值在于：

- 背景替换：通过简单 prompt+参考图，可将自拍视频背景替换为演播室、户外场景等，无需绿幕与后期合成。

- 特效添加：一键为视频添加字幕、转场、滤镜等广电级特效，操作门槛大幅降低。

- 角色植入：在不影响原人物表情、肢体语言、说话节奏的前提下，无缝添加虚拟角色或嘉宾。

- 多模态交互：支持文本、图像、语音等多种输入方式，符合广电从业者内容创作习惯。

02 技术升级：

原生音视频同步生成，推理效率革命性提升

Gemini Omni Flash 实现了原生音视频同步生成的技术突破，画面、配乐、音效在同一次推理过程中完成，无需三个模型接力生成，大幅提升内容生产效率。

这一特性对广电行业意义重大：

- 新闻快编：突发新闻可快速生成完整音视频报道，缩短从采集到播出的时间差。

- 短视频生产：适合 IPTV、OTT 平台及短视频矩阵的批量内容生产，降低制作成本。

- 虚拟演播：为虚拟主播、虚拟场景提供实时音视频生成能力，提升直播互动体验。

03 安全保障：

SynthID 水印技术护航，AI 内容可追溯性强化

面对 AI 视频技术带来的内容安全挑战，谷歌同步强化了 SynthID 内容防伪水印技术的应用：

- 已为 1000 亿张图像和视频、6 万年时长的音频打上水印，确保 AI 生成内容可识别。
- 针对高质量深伪视频人类识别率仅约 25% 的行业痛点，提供技术层面的内容溯源解决方案。
- 为广电行业内容审核、版权保护提供技术支撑，符合总局关于 AI 内容规范管理的要求。


04 行业启示：

AI 重构广电内容生产链，新质生产力加速形成

Gemini Omni 的发布对广电视听行业具有深远影响：

- 生产效率质变：从前期拍摄到后期制作的全流程效率提升，降低人力成本，释放创作生产力；
- 制作门槛降低：非专业人员可快速制作广电级内容，助力县级融媒体中心、基层广电机构内容建设；
- 创新场景拓展：为虚拟制作、互动直播、个性化内容推荐等新业态提供技术支撑；
- 安全规范先行：水印技术与内容可追溯体系建设成为 AI 内容生产的标配，符合行业监管要求。

05 后续动态：



Omni Pro 版本即将推出，性能将进一步提升

Demis Hassabis 透露，更强性能的 Gemini Omni Pro 版本将在不久后推出，将提供更高分辨率、更复杂场景的视频编辑能力，为广电行业 4K/8K 超高清内容生产提供更强大的技术支撑。

来源：中国视听科技融媒