

Extended Cut: NVIDIA Expands Maxine for Video Editing, Showcases 3D Virtual Conferencing Research

Production version of Maxine now available in NVIDIA AI Enterprise; research demo presents how 3D technology can enhance video communications.

Author: Rick Champagne

Professionals, teams, creators and others can tap into the power of AI to create high-quality audio and video effects — even using standard microphones and webcams — with the help of NVIDIA Maxine .

The suite of GPU-accelerated software development kits and cloud-native microservices lets users deploy AI features that enhance audio, video and augmented-reality effects for real-time communications services and platforms. Maxine will also expand features for video editing, enabling teams to reach new heights in video communication.

Plus, an NVIDIA Research demo at this week's SIGGRAPH conference displays how AI can take video conferencing to the next level with 3D features.

Wireless connectivity has enabled people to join virtual meetings from more locations than ever. Typically, audio and video quality are heavily impacted when a caller is on the move or in a location with poor connectivity.

Advanced, real-time Maxine features — such as Background Noise Removal, Super Resolution and Eye Contact — allow remote users to enhance interpersonal communication experiences.

In addition, Maxine can now be used for video editing. NVIDIA partners are transforming this professional workflow with the same Maxine features that elevate video conferencing. The goal when editing a video, whether a sales pitch or a webinar, is to engage the broadest audience possible. Using Maxine, professionals can tap into AI features that enhance audio and video signals.

With Maxine, a spokesperson can look away from the screen to reference notes or a script while their gaze remains as if looking directly into the camera. Users can also film videos in low resolution and enhance the quality later. Plus, Maxine lets people record videos in several different languages and export the video in English.

Maxine features to be released in early access this year include:

Interpreter: Translates from simplified Chinese, Russian, French, German and Spanish to English while animating the user's image to show them speaking English.

Voice Font: Enables users to apply characteristics of a speaker's voice and map it to the audio output.

Audio Super Resolution : Improves audio quality by increasing the temporal resolution of the audio signal and extending bandwidth. It currently supports upsampling from 8,000Hz to 16,000Hz as well as from 16,000Hz to 48,000Hz. This feature is also updated with more than 50% reduction in latency and up to 2x better throughput.

Maxine Client: Brings the AI capabilities of Maxine's microservices to video-conferencing sessions on PCs. The application is optimized for low-latency streaming and will use the cloud for all of its GPU compute requirements. Thin Client will be available on Windows this fall, with additional OS support to follow.

Maxine can be deployed in the cloud, on premises or at the edge, meaning quality communication can be accessible from nearly anywhere.

Many partners and customers are experiencing high-quality video conferencing and editing with Maxine. Two features of Maxine — Eye Contact and Live Portrait — are now available in production releases on the NVIDIA AI Enterprise software platform. Eye Contact simulates direct eye contact with the camera by estimating and aligning the user's gaze with the camera. And Live Portrait animates a person's portrait photo through their live video feed.

Software company Descript aims to make video a staple of every communicator's toolkit, alongside docs and slides. With NVIDIA Maxine, professionals and beginners who use Descript can access AI features that improve their video-content workflows.

"With the NVIDIA Maxine Eye Contact feature, users no longer have to worry about memorizing scripts or doing tedious video retakes," said Jay LeBoeuf, head of business and corporate development at Descript. "They can maintain a perfect on-screen presence while nailing their script every time."

Reincubate's Camo app aims to broaden access to great video by taking advantage of the hardware and devices people already own. It does this by giving users greater control over their image and by implementing a powerful, efficient processing pipeline for video effects and transformation. Using technologies enabled by NVIDIA Maxine, Camo can offer users an easier way to achieve incredible video creation.

"Integrating NVIDIA Maxine into Camo couldn't have been easier, and it's enabled us to get high performance from users' RTX GPUs right out of the box," said Aidan Fitzpatrick, founder and CEO of Reincubate. "With Maxine, the team's been able to move faster and with more confidence."

Quicklink's Cre8 is a powerful video production platform for creating professional, on-brand productions, virtual and hybrid live events. The user-friendly interface combines an intuitive design with all the tools needed to build, edit and customize a professional-looking production. Cre8 incorporates NVIDIA Maxine technology to maximize productivity and the quality of video productions, offering complete control to the operator.

"Quicklink Cre8 now offers the most advanced video production platform on the planet," said Richard Rees, CEO of Quicklink. "With NVIDIA Maxine, we were able to add advanced features, including Auto Framing, Video Noise Removal, Noise and Echo Cancellation, and Eye Contact Simulation."

Los Angeles-based company gemelo.ai provides a platform for creating AI twins that can scale a user's voice, content and interactions. Using Maxine's Live Portrait feature, the gemelo.ai team can unlock new opportunities for scaled, personalized content and one-on-one interactions.

"The realism of Live Portrait has been a game-changer, unlocking new realms of potential for our AI twins," said Paul Jaski, CEO of gemelo.ai. "Our customers can now design and deploy incredibly realistic digital twins with the superpowers of unlimited scalability in content production and interaction across apps, websites and mixed-reality experiences."

In addition to powering the advanced features of Maxine, NVIDIA AI enhances video communication with 3D. NVIDIA Research recently published a paper demonstrating how AI could power a 3D video-conferencing system with minimal capture equipment.

3D telepresence systems are typically expensive, require a large space or production studio, and use high-bandwidth, volumetric video streaming — all of which limits the technology's accessibility. NVIDIA Research shared a new method, which runs on a novel VisionTransformer-based encoder, that takes 2D video input from a standard webcam and turns it into a 3D video representation. Instead of requiring 3D data to be passed back and forth between the participants in a conference, AI enables bandwidth requirements for the call to stay the same as for a 2D conference.

The technology takes a user's 2D video and automatically creates a 3D representation called a neural radiance field, or NeRF, using volumetric rendering. As a result, participants can stream 2D videos, like they would for traditional video conferencing, while decoding high-quality 3D representations that can be rendered in real time. And with Maxine's Live Portrait, users can bring their portraits to life in 3D.

AI-mediated 3D video conferencing could significantly reduce the cost for 3D capture, provide a high-fidelity 3D representation, accommodate photorealistic or stylized avatars, and enable mutual eye contact in video conferencing. Related research projects show how AI can help elevate communications and virtual interactions, as well as inform future NVIDIA technologies for video conferencing.

See the system in action below. SIGGRAPH attendees can visit the Emerging Technologies booth , where groups will be able to simultaneously view the live demo on a 3D display designed by New York-based company Looking Glass .

Learn more about NVIDIA Maxine , which is now available on NVIDIA AI Enterprise .

And see more of the research behind the 3D video conference project.

Featured image courtesy of NVIDIA Research.

Original URL: <https://blogs.nvidia.com/blog/2023/08/08/maxine-3d-video-communications/>