Digital Renaissance: NVIDIA Neuralangelo Research Reconstructs 3D Scenes

Author: Isha Salian

Editor's note: Neuralangelo is now available on GitHub .

Neuralangelo, a new AI model by NVIDIA Research for 3D reconstruction using neural networks, turns 2D video clips into detailed 3D structures — generating lifelike virtual replicas of buildings, sculptures and other real-world objects.

Like Michelangelo sculpting stunning, life-like visions from blocks of marble, Neuralangelo generates 3D structures with intricate details and textures. Creative professionals can then import these 3D objects into design applications, editing them further for use in art, video game development, robotics and industrial digital twins .

Neuralangelo's ability to translate the textures of complex materials — including roof shingles, panes of glass and smooth marble — from 2D videos to 3D assets significantly surpasses prior methods. The high fidelity makes its 3D reconstructions easier for developers and creative professionals to rapidly create usable virtual objects for their projects using footage captured by smartphones.

"The 3D reconstruction capabilities Neuralangelo offers will be a huge benefit to creators, helping them recreate the real world in the digital world," said Ming-Yu Liu, senior director of research and co-author on the paper. "This tool will eventually enable developers to import detailed objects — whether small statues or massive buildings — into virtual environments for video games or industrial digital twins."

In a demo, NVIDIA researchers showcased how the model could recreate objects as iconic as Michelangelo's David and as commonplace as a flatbed truck. Neuralangelo can also reconstruct building interiors and exteriors — demonstrated with a detailed 3D model of the park at NVIDIA's Bay Area campus.

Prior AI models to reconstruct 3D scenes have struggled to accurately capture repetitive texture patterns, homogenous colors and strong color variations. Neuralangelo adopts instant neural graphics primitives, the technology behind NVIDIA Instant NeRF , to help capture these finer details.

Using a 2D video of an object or scene filmed from various angles, the model selects several frames that capture different viewpoints — like an artist considering a subject from multiple sides to get a sense of depth, size and shape.

Once it's determined the camera position of each frame, Neuralangelo's AI creates a rough 3D representation of the scene, like a sculptor starting to chisel the subject's shape.

The model then optimizes the render to sharpen the details, just as a sculptor painstakingly hews stone to mimic the texture of fabric or a human figure.

The final result is a 3D object or large-scale scene that can be used in virtual reality applications, digital twins or robotics development.

Neuralangelo is one of nearly 30 projects by NVIDIA Research to be presented at the Conference on Computer Vision and Pattern Recognition (CVPR), taking place June 18-22 in Vancouver. The papers span topics including pose estimation, 3D reconstruction and video generation.

One of these projects, DiffCollage , is a diffusion method that creates large-scale content — including long landscape orientation, 360-degree panorama and looped-motion images. When fed a training dataset of images with a standard aspect ratio, DiffCollage treats these smaller images as sections of a

larger visual — like pieces of a collage. This enables diffusion models to generate cohesive-looking large content without being trained on images of the same scale.

The technique can also transform text prompts into video sequences, demonstrated using a pretrained diffusion model that captures human motion:

■ ■ ■ ■

Learn more about NVIDIA Research at CVPR .

Original URL: https://blogs.nvidia.com/blog/2023/06/01/neuralangelo-ai-research-3d-reconstruction/