

# New AI technology enables 3D capture and editing of real-life objects

March 12 2024



Credit: Simon Fraser University

Imagine performing a sweep around an object with your smartphone and getting a realistic, fully editable 3D model that you can view from any angle. This is fast becoming reality, thanks to advances in AI.

Researchers at Simon Fraser University (SFU) in Canada have unveiled

new AI technology for doing exactly this. Soon, rather than merely taking 2D photos, everyday consumers will be able to take 3D captures of real-life objects and edit their shapes and appearance as they wish, just as easily as they would with regular 2D photos today.

In a [new paper](#) appearing on the *arXiv* preprint server and presented at the 2023 [Conference on Neural Information Processing Systems](#) (NeurIPS) in New Orleans, Louisiana, researchers demonstrated a new technique called [Proximity Attention Point Rendering \(PAPR\)](#) that can turn a set of 2D photos of an object into a cloud of 3D points that represents the object's shape and appearance.

Each point then gives users a knob to control the object with—dragging a point changes the object's shape, and editing the properties of a point changes the object's appearance. Then in a process known as "rendering," the 3D point cloud can then be viewed from any angle and turned into a 2D photo that shows the edited object as if the photo was taken from that angle in real life.

Using the new AI technology, researchers showed how a statue can be brought to life—the technology automatically converted a set of photos of the statue into a 3D point cloud, which is then animated. The end result is a video of the statue turning its head from side to side as the viewer is guided on a path around it.

"AI and machine learning are really driving a paradigm shift in the reconstruction of 3D objects from 2D images. The remarkable success of machine learning in areas like [computer vision](#) and natural language is inspiring researchers to investigate how traditional 3D graphics pipelines can be re-engineered with the same deep learning-based building blocks that were responsible for the runaway AI success stories of late," said Dr. Ke Li, an assistant professor of computer science at Simon Fraser University (SFU), director of the APEX lab and the senior author on the

paper.

"It turns out that doing so successfully is a lot harder than we anticipated and requires overcoming several technical challenges. What excites me the most is the many possibilities this brings for consumer technology—3D may become as common a medium for visual communication and expression as 2D is today."

One of the biggest challenges in 3D is on how to represent 3D shapes in a way that allows users to edit them easily and intuitively. One previous approach, known as neural radiance fields (NeRFs), does not allow for easy shape editing because it needs the user to provide a description of what happens to every continuous coordinate. A more recent approach, known as 3D Gaussian splatting (3DGS), is also not well-suited for shape editing because the shape surface can get pulverized or torn to pieces after editing.

A key insight came when the researchers realized that instead of considering each 3D point in the point cloud as a discrete splat, they can think of each as a control point in a continuous interpolator. Then when the point is moved, the shape changes automatically in an intuitive way. This is similar to how animators define the motion of objects in animated videos—by specifying the positions of objects at a few points in time, their motion at every point in time is automatically generated by an interpolator.

However, how to mathematically define an interpolator between an arbitrary set of 3D points is not straightforward. The researchers formulated a [machine learning](#) model that can learn the interpolator in an end-to-end fashion using a novel mechanism known as proximity attention.

In recognition of this technological leap, the paper was awarded with a

spotlight at the NeurIPS conference, an honor reserved for the top 3.6% of paper submissions to the conference.

The research team is excited for what's to come. "This opens the way to many applications beyond what we've demonstrated," said Dr. Li. "We are already exploring various ways to leverage PAPR to model moving 3D scenes and the results so far are incredibly promising."

The authors of the paper are Yanshu Zhang, Shichong Peng, Alireza Moazeni and Ke Li. Zhang and Peng are co-first authors, Zhang, Peng and Moazeni are Ph.D. students at the School of Computing Science and all are members of the APEX Lab at Simon Fraser University (SFU).

**More information:** Yanshu Zhang et al, PAPR: Proximity Attention Point Rendering, *arXiv* (2023). [DOI: 10.48550/arxiv.2307.11086](https://doi.org/10.48550/arxiv.2307.11086)

Provided by Simon Fraser University

Citation: New AI technology enables 3D capture and editing of real-life objects (2024, March 12) retrieved 27 April 2024 from <https://techxplore.com/news/2024-03-ai-technology-enables-3d-capture.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.