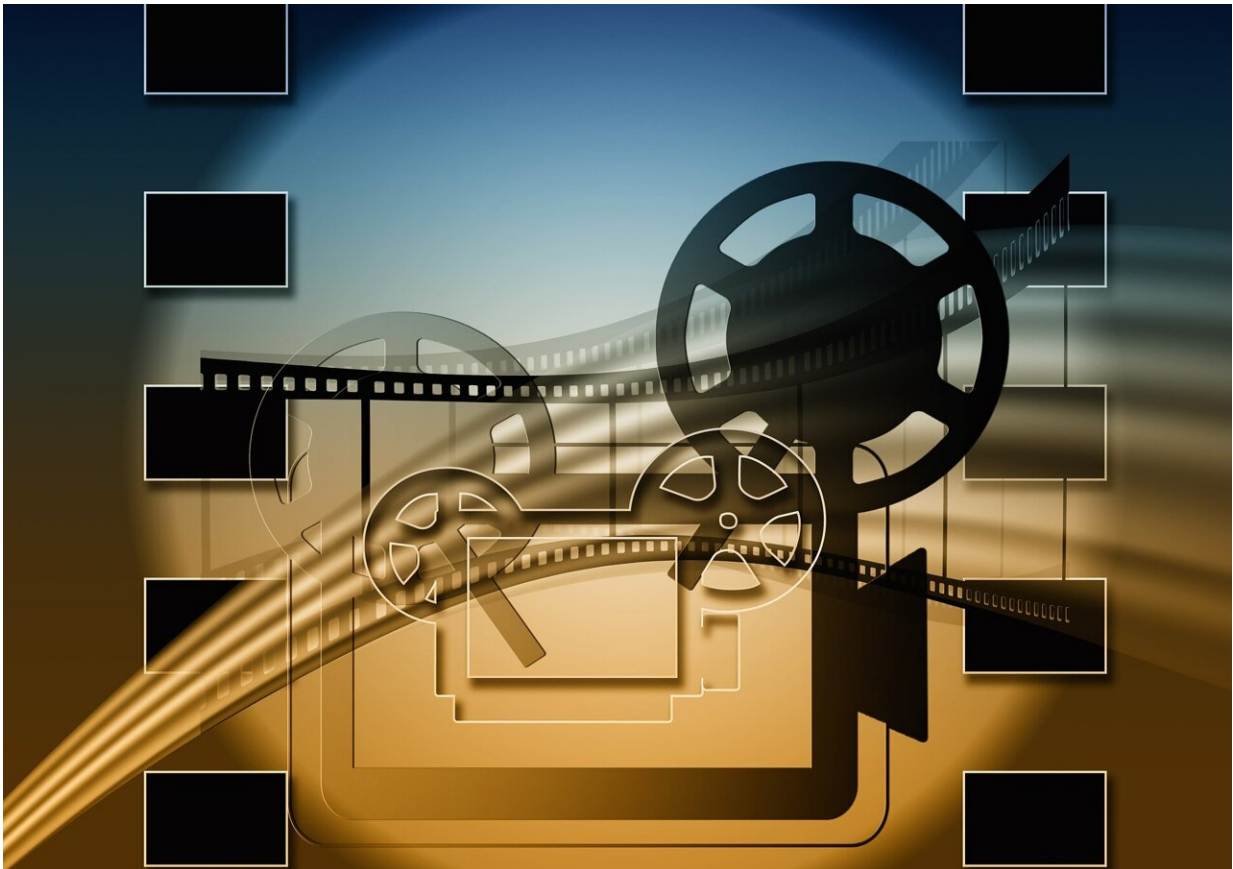# Software creates entirely new views from existing video

July 13 2023

Credit: Pixabay/CC0 Public Domain

Filmmakers may soon be able to stabilize shaky video, change viewpoints and create freeze-frame, zoom and slow-motion

effects—without shooting any new footage—thanks to an algorithm developed by researchers at Cornell University and Google Research.

The software, called [DynIBar](link), synthesizes new views using pixel information from the original video, and even works with moving objects and unstable camerawork. The work is a major advance over previous efforts, which yielded only a few seconds of video, and often rendered moving subjects as blurry or glitchy.

The code for this research effort is [freely available](link), though the project is at an early stage and not yet integrated into commercial video editing tools.

"While this research is still in its early days, I'm really excited about potential future applications for both personal and professional use," said Noah Snavely, a research scientist at Google Research and associate professor of computer science at Cornell Tech and in the Cornell Ann S. Bowers College of Computing and Information Science.

Snavely presented this work, "DynIBaR: Neural Dynamic Image-Based Rendering," at the [2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition](link), on June 20, where it received an honorable mention for the best paper award. Zhengqi Li, Ph.D, of Google Research was the lead author on the study.

"Over the last few years, we've seen major progress in view synthesis methods—algorithms that can take a collection of images capturing a scene from a discrete set of viewpoints, and can render new views of that scene," said Snavely. "However, most of these methods fail on scenes with moving people or pets, swaying trees and so on. This is a big problem because many interesting things in the world are things that move."

Existing methods to render new views of still scenes, such as ones that make a photo appear 3D, take the 2D grid of pixels from an image and reconstruct the 3D shape and appearance of each object in the photo. DynIBar takes this a step further by also estimating how the objects move over time. But considering all four dimensions creates an incredibly difficult math problem.

The researchers simplified this problem by using a computer graphics approach developed in the 1990s called image-based rendering. At the time, it was difficult for traditional computer graphics methods to render complex scenes with many small parts—such as a leafy tree—so graphics researchers developed methods that take images of a scene and then alter and recombine the parts to generate new images. In this way, most of the complexity was stored within the source image and could load faster.

"We incorporated the classic idea of image-based rendering and that makes our method able to handle really complex scenes and longer videos," said co-author Qianqian Wang, a doctoral student in the field of computer science at Cornell Tech. Wang developed a method to use image-based rendering to synthesize new views of still images, which the new software builds on.

Despite the advance, these features may not be coming to your smartphone any time soon. The software takes several hours to process just 10 or 20 seconds of video, even on a powerful computer. In the near-term, the technology may be more appropriate for use in offline video editing software, Snavely said.

The next hurdle will be figuring out how to render new images when pixel information is lacking from the original video, such as when the subject moves too fast or the user wants to rotate the viewpoint 180 degrees. Snavely and Wang envision that soon it may be possible to

incorporate generative AI techniques, such as text-to-image generators, to help fill in those gaps.

Provided by Cornell University