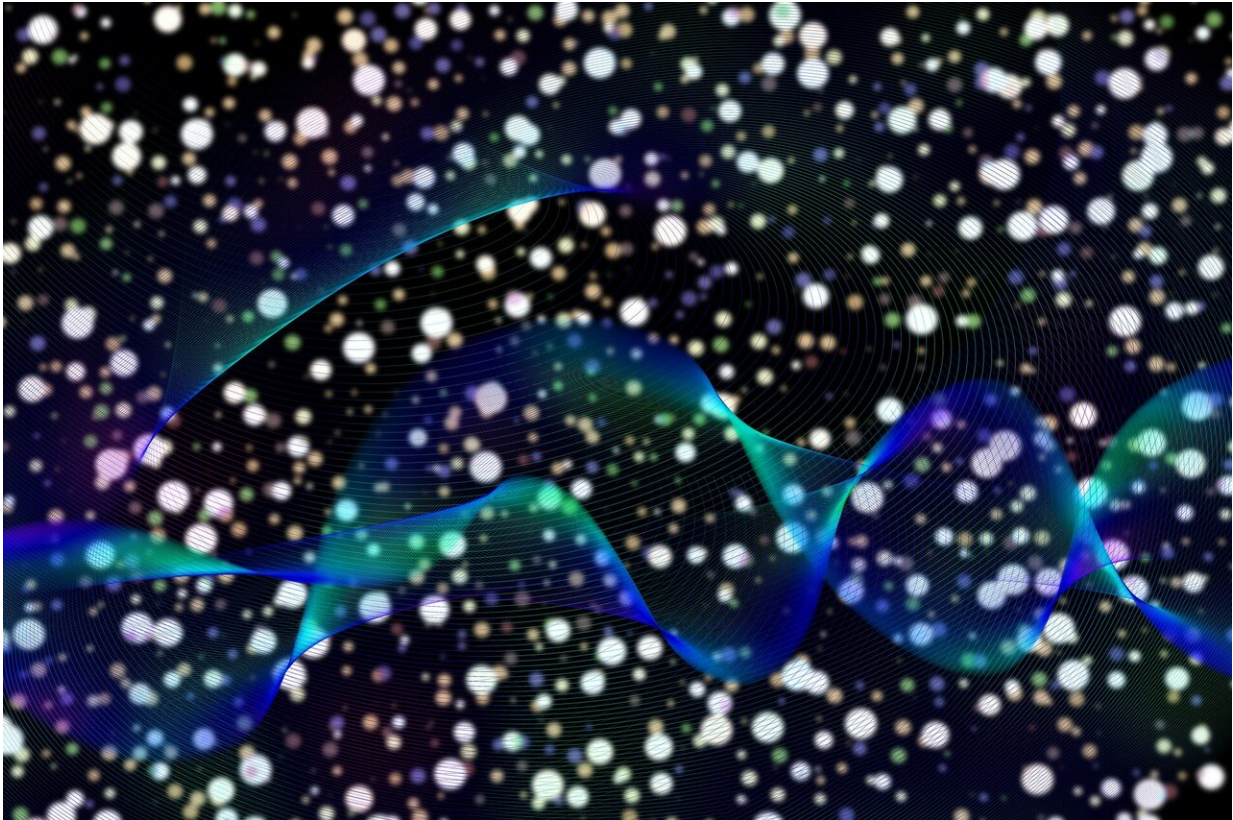


Turning a single photo into a video

June 15 2021, by Sarah McQuate



Credit: CC0 Public Domain

Sometimes photos cannot truly capture a scene. How much more epic would that vacation photo of Niagara Falls be if the water were moving?

Researchers at the University of Washington have developed a deep learning method that can do just that: If given a single photo of a

waterfall, the system creates a [video](#) showing that water cascading down. All that's missing is the roar of the water and the feeling of the spray on your face.

The team's method can animate any flowing material, including smoke and clouds. This technique produces a short video that loops seamlessly, giving the impression of endless movement. The researchers will present this approach June 22 at the Conference on Computer Vision and Pattern Recognition.

"A picture captures a moment frozen in time. But a lot of information is lost in a static image. What led to this moment, and how are things changing? Think about the last time you found yourself fixated on something really interesting—chances are, it wasn't totally static," said lead author Aleksander Holynski, a doctoral student in the Paul G. Allen School of Computer Science & Engineering.

"What's special about our method is that it doesn't require any user input or extra information," Holynski said. "All you need is a picture. And it produces as output a high-resolution, seamlessly looping video that quite often looks like a real video."

Developing a method that turns a single photo into a believable video has been a challenge for the field.

"It effectively requires you to predict the future," Holynski said. "And in the real world, there are nearly infinite possibilities of what might happen next."

The team's system consists of two parts: First, it predicts how things were moving when a photo was taken, and then uses that information to create the animation.

To estimate motion, the team trained a [neural network](#) with thousands of videos of waterfalls, rivers, oceans and other material with [fluid motion](#). The training process consisted of asking the network to guess the motion of a video when only given the first frame. After comparing its prediction with the actual video, the network learned to identify clues—ripples in a stream, for example—to help it predict what happened next. Then the team's system uses that information to determine if and how each pixel should move.

The researchers tried to use a technique called "splatting" to animate the photo. This method moves each pixel according to its predicted motion. But this created a problem.

"Think about a flowing waterfall," Holynski said. "If you just move the pixels down the waterfall, after a few frames of the video, you'll have no pixels at the top!"

So the team created "symmetric splatting." Essentially, the method predicts both the future and the past for an image and then combines them into one animation.

"Looking back at the waterfall example, if we move into the past, the pixels will move up the waterfall. So we will start to see a hole near the bottom," Holynski said. "We integrate information from both of these animations so there are never any glaringly large holes in our warped images."

Finally, the researchers wanted their animation to loop seamlessly to create a look of continuous movement. The animation network follows a few tricks to keep things clean, including transitioning different parts of the frame at different times and deciding how quickly or slowly to blend each pixel depending on its surroundings.

The team's method works best for objects with predictable fluid motion. Currently, the technology struggles to predict how reflections should move or how water distorts the appearance of objects beneath it.

"When we see a [waterfall](#), we know how the water should behave. The same is true for fire or smoke. These types of motions obey the same set of physical laws, and there are usually cues in the image that tell us how things should be moving," Holynski said. "We'd love to extend our work to operate on a wider range of objects, like animating a person's hair blowing in the wind. I'm hoping that eventually the pictures that we share with our friends and family won't be static images. Instead, they'll all be dynamic animations like the ones our method produces."

More information: Conference on Computer Vision and Pattern Recognition: cvpr2021.thecvf.com/

Provided by University of Washington

Citation: Turning a single photo into a video (2021, June 15) retrieved 19 April 2024 from <https://techxplore.com/news/2021-06-photo-video.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.