

Teaching AI to see depth in photographs and paintings

August 12 2021



Researchers in Simon Fraser University's Computational Photography Lab at SFU are successfully teaching artificial intelligence how to determine depth from a single photograph. Credit: SFU

Researchers in SFU's Computational Photography Lab hope to give computers a visual advantage that we humans take for granted—the ability to see depth in photographs. While humans naturally can determine how close or far objects are from a single point of view, like a photograph or a painting, it's a challenge for computers—but one they may soon overcome.

Researchers recently published their work improving a process called

monocular depth estimation, a technique that teaches computers how to see depth using machine learning.

"When we look at a picture, we can tell the relative distance of objects by looking at their size, position, and relation to each other," says Mahdi Miangoleh, an MSc student working in the lab. "This requires recognizing the objects in a scene and knowing what size the objects are in real life. This task alone is an active research topic for [neural networks](#)."

Despite progress in recent years, existing efforts to provide high resolution results that can transform an image into a 3-dimensional (3D) space have failed.

To counter this, the lab recognized the untapped potential of existing neural [network](#) models in the literature. The proposed research explains the lack of high-resolution results in current methods through the limitations of convolutional neural networks. Despite major advancements in recent years, the neural networks still have a relatively small capacity to generate many details at once.

Another limitation is how much of the scene these networks can 'look at' at once, which determines how much information the neural network can make use of to understand complex scenes. By working to increase the resolution of their visual estimations, the researchers are now making it possible to create detailed 3D renderings that look realistic to a human eye. These so-called "depth maps" are used to create 3D renderings of scenes and simulate camera motion in computer graphics.

"Our method analyzes an image and optimizes the process by looking at the image content according to the limitations of current architectures," explains Ph.D. student Sebastian Dille. "We give our input image to our neural network in many different forms, to create as many details as the

model allows while preserving a realistic geometry."

The team also published a friendly explainer for the theory behind the method, which is available on YouTube.

"With the high-resolution depth maps that the team is able to develop for real-world photographs, artists and [content creators](#) can now immediately transfer their photograph or artwork into a rich 3D world," says computing science professor and lab director, Yağız Aksoy, whose team collaborated with researchers Sylvain Paris and Long Mai, from Adobe Research.

Tools enable artists to turn 2D art into 3D worlds

Global artists are already utilizing the applications enabled by Aksoy's lab's research. Akira Saito, a visual [artist](#) based in Japan, is creating videos that take viewers into fantastic 3D worlds dreamed up in 2D artwork. To do this he combines tools such as Houdini, a computer animation software, with the depth map generated by Aksoy and his team.

Creative content creators on TikTok are using the research to express themselves in new ways.

"It's a great pleasure to see independent artists make use of our technology in their own way," says Aksoy, whose lab has plans to extend this work to videos and develop new tools that will make depth maps more useful for artists.

"We have made great leaps in computer vision and [computer](#) graphics in recent years, but the adoption of these new AI technologies by the artist community needs to be an organic process, and that takes time."

More information: S. Mahdi et al, Boosting Monocular Depth Estimation Models to High-Resolution via Content-Adaptive Multi-Resolution Merging, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2021):
[openaccess.thecvf.com/content/ ... CVPR 2021 paper.html](https://openaccess.thecvf.com/content/CVPR2021/papers/Mahdi_Boosting_Monocular_Depth_Estimation_Models_to_High-Resolution_via_Content-Adaptive_Multi-Resolution_Merging_CVPR2021_paper.html)

Project Github: yaksoy.github.io/highresdepth/

Provided by Simon Fraser University

Citation: Teaching AI to see depth in photographs and paintings (2021, August 12) retrieved 7 May 2024 from <https://techxplore.com/news/2021-08-ai-depth.html>

<p>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.</p>
--