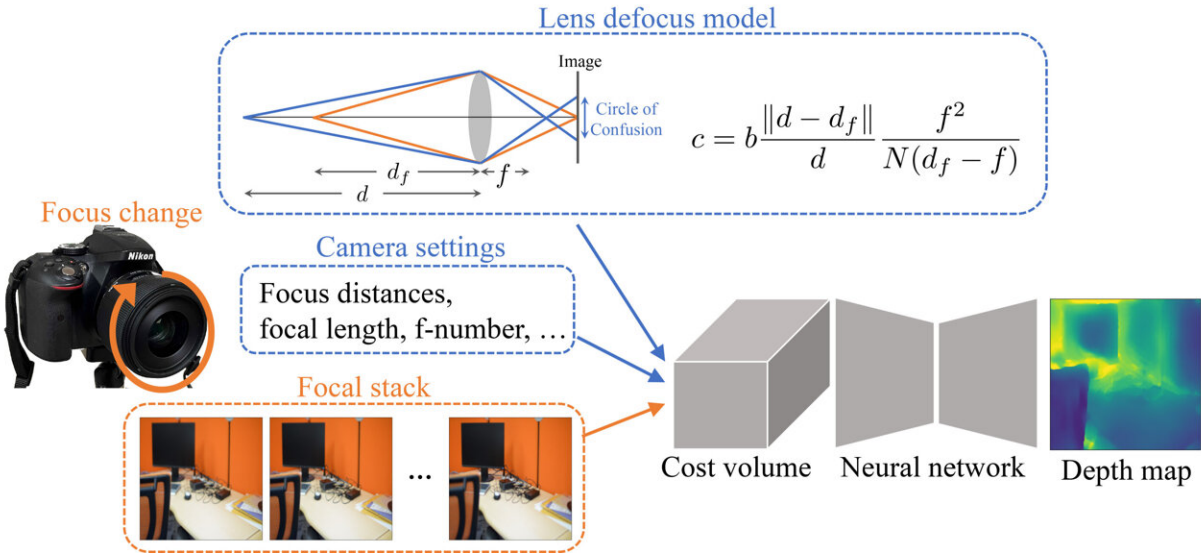


Innovations in depth from focus/defocus pave the way to more capable computer vision systems

February 9 2024



The proposed method takes as input the focal stack and camera settings, and establishes a cost volume based on a lens defocus model. This design enables depth estimation with different camera settings at training and test times. Credit: Yuki Fujimura

In several applications of computer vision, such as augmented reality and self-driving cars, estimating the distance between objects and the camera is an essential task. Depth from focus/defocus is one of the techniques that achieve such a process using the blur in the images as a clue. Depth

from focus/defocus usually requires a stack of images of the same scene taken with different focus distances, a technique known as "focal stack."

Over the past decade or so, scientists have proposed many different methods for depth from focus/defocus, most of which can be divided into two categories. The first category includes model-based methods, which use mathematical and optics models to estimate scene depth based on sharpness or blur. The main problem with such methods, however, is that they fail for texture-less surfaces, which look virtually the same across the entire focal stack.

The second category includes learning-based methods, which can be trained to perform depth from focus/defocus efficiently, even for texture-less surfaces. However, these approaches fail if the camera settings used for an input focal stack are different from those used in the training dataset.

Overcoming these limitations now, a team of researchers from Japan has developed an innovative method for depth from focus/defocus that simultaneously addresses the abovementioned issues. Their study, [published](#) in the *International Journal of Computer Vision*, was led by Yasuhiro Mukaigawa and Yuki Fujimura from Nara Institute of Science and Technology (NAIST), Japan.

The proposed technique, dubbed "deep depth from focal stack" (DDFS), combines model-based depth estimation with a learning framework to get the best of both worlds. Inspired by a strategy used in stereo vision, DDFS involves establishing a "cost volume" based on the input focal stack, the camera settings, and a lens defocus model.

Simply put, the cost volume represents a set of depth hypotheses—potential depth values for each pixel—and an associated cost value calculated on the basis of consistency between images in the

focal stack. "The cost volume imposes a constraint between the defocus images and scene depth, serving as an intermediate representation that enables depth estimation with different camera settings at training and test times," explains Mukaigawa.

The DDFS method also employs an encoder–decoder network, a commonly used machine learning architecture. This network estimates the scene depth progressively in a coarse-to-fine fashion, using "cost aggregation" at each stage for learning localized structures in the images adaptively.

The researchers compared the performance of DDFS with that of other state-of-the-art depth from focus/defocus methods. Notably, the proposed approach outperformed most methods in various metrics for several image datasets. Additional experiments on focal stacks captured with the research team's camera further proved the potential of DDFS, making it useful even with only a few input images in the input stacks, unlike other techniques.

Overall, DDFS could serve as a promising approach for applications where depth estimation is required, including robotics, autonomous vehicles, 3D image reconstruction, virtual and augmented reality, and surveillance. "Our method with [camera](#)-setting invariance can help extend the applicability of learning-based depth estimation techniques," concludes Mukaigawa.

More information: Yuki Fujimura et al, Deep Depth from Focal Stack with Defocus Model for Camera-Setting Invariance, *International Journal of Computer Vision* (2023). [DOI: 10.1007/s11263-023-01964-x](https://doi.org/10.1007/s11263-023-01964-x)

Provided by Nara Institute of Science and Technology

Citation: Innovations in depth from focus/defocus pave the way to more capable computer vision systems (2024, February 9) retrieved 1 March 2024 from <https://techxplore.com/news/2024-02-depth-focusdefocus-pave-capable-vision.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.