

Pixel 3: A turn to machine learning for depth estimations

December 4 2018, by Nancy Cohen



The two PDAF images on the left and center look very similar, but in the crop on the right you can see the parallax between them. It is most noticeable on the circular structure in the middle of the crop. Credit: Google blog

TechSpot says what it thinks about the Pixel 3 and it's not an ad: "Pixel 3 is quite possibly the best camera [phone](#) on the market." Tyler Lee in *Ubergizmo* is in full compliment mode. "There is no doubt that all of Google's hard [work](#) and research has paid off as the Pixel 3 does have one of the better cameras around."

OK, we get it. In the ever-jostling vendor struggle to overpower the smartphone market, Pixel 3 is successfully marking its own turf as the phone with very good camera capabilities. And, now, a recent blog on the Google AI blog will only further delight Pixel 3 fans, as it details

how "Portrait Mode" was achieved on Pixel 3.

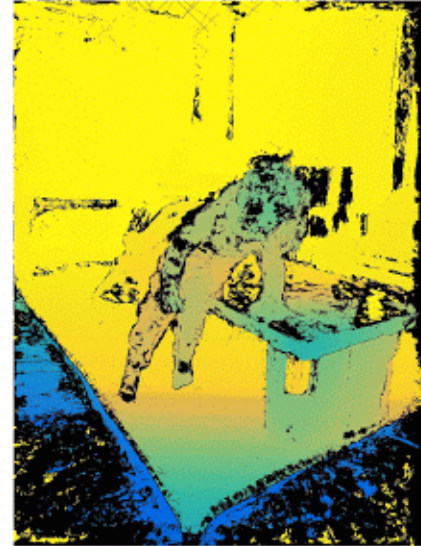
Think of a team turning to machine learning, TensorFlow and depth-perception goals to the nines.

Scott Adam Gordon in *Android Authority* shared with readers what was rather clever about Google's approach and techniques. "With the Google Pixel 3's camera, Google included more depth cues to inform this blur effect for greater accuracy. As well as parallax, Google used sharpness as a depth indicator—more distant objects are less sharp than closer [objects](#)—and real-world object identification. For example, the camera could recognize a person's face in a scene, and work out how near or far it was based on its number of pixels relative to objects around it."

Rahul Garg, research scientist and Neal Wadhwa, software engineer, posted the blog. "This year, on the Pixel 3, we turn to machine learning to improve depth estimation to produce even better Portrait Mode results."

(Google thus far kept the magic to themselves, said Isaiah Mayersen in *TechSpot*. "As Google continues to pave the way in smartphone photography, we'll have to wait and see what competitors bring to the table.")

Wait, what is Portrait Mode? *Webopedia*, thank you:



Left: Custom rig used to collect training data. Middle: An example capture flipping between the five images. Synchronization between the cameras ensures that we can calculate depth for dynamic scenes, such as this one. Right: Ground truth depth. Low confidence points, i.e., points where stereo matches are not reliable due to weak texture, are colored in black and are not used during training. Credit: Sam Ansari and Mike Milne

"In photography and [digital photography](#), [portrait](#) mode is a function of the digital camera that is used when you are taking photos of a single subject. When taking photos in portrait mode, the digital camera will automatically uses a large aperture to help keep the background out of focus by using a narrow depth of field so the subject being [photographed](#) is the only thing in focus."

Not only the single subject but more: "Not only will the person in the portrait be in focus, items near that same plane of [focus](#) will also be sharper, with a realistic increasing blur as items are further in front of and behind that plane," said Ryne Hager in *Android Police*.

Got that? Background intentionally out of focus; the subject carries the maximum impact in focus. This technique is often used in ads and promotional messages to draw maximum attention to a face by blurring the background.

Google needed some photos to train its AI. Enter as 5-phone clump—ok, more politely stated, an assemblage of five phones. They spoke about depth maps, training data and this assemblage, the "Frankenphone," a rig with 5 phones, all Pixel 3, and Wi-Fi based solution.

CNET described Frankenphone. "A quintet of phones sandwiched together generated the data to train the Pixel 3 to judge depth." Stephen Shankland described a "hacked-together [clump](#) of five phones to improve how the feature worked in this year's Pixel 3."

They used TensorFlow Lite, a cross-platform solution for running machine learning models on mobile and embedded devices, and the Pixel 3's GPU to compute depth quickly. Then, they combined resulting depth estimates with masks from their "person segmentation neural network to produce beautiful Portrait Mode results."

Why this matters does not require much analysis: better looking portrait mode shots.

In the bigger picture of picture-taking smartphones, vendors accept that their sales pitch had best include tempting [camera](#) features. Google's computational photography methods can give Pixel a real edge.

"Smartphones have small image sensors that can't compete with traditional cameras for [image quality](#), but Google is ahead of the pack with computational photography methods that can do things like blur backgrounds, increase resolution, tweak exposure, improve shadow details and shoot photos in the dark," said Stephen Shankland, CNET.

More information: [ai.googleblog.com/2018/11/lear ... epth-on-pixel-3.html](https://ai.googleblog.com/2018/11/learning-depth-on-pixel-3.html)

© 2018 Science X Network

Citation: Pixel 3: A turn to machine learning for depth estimations (2018, December 4) retrieved 2 May 2024 from <https://techxplore.com/news/2018-12-pixel-machine-depth.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.