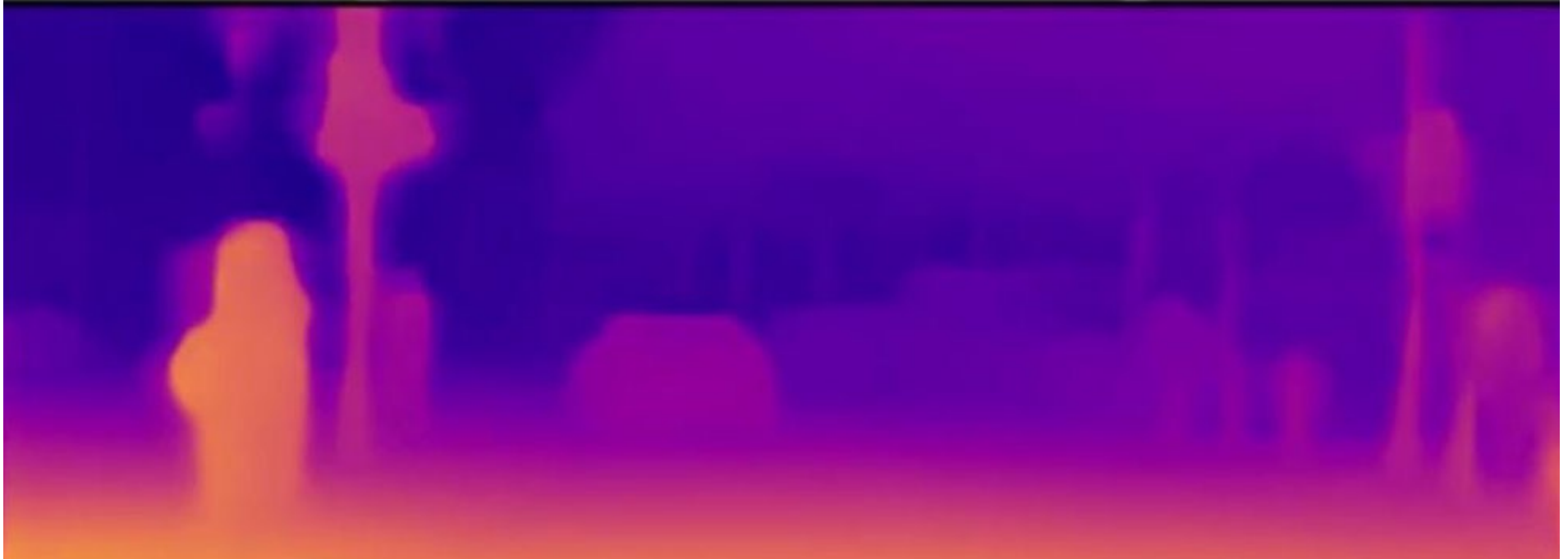


Measuring distance with a single photo

26 July 2017

Most cameras just record colour but now the 3D shapes of objects, captured through only a single lens, can be accurately estimated using new software developed by UCL computer scientists.



- [Video](#)

The method, published today at *CVPR 2017*, gives state-of-the-art results and works with existing photos, allowing any camera to map the depth for every pixel it captures.

The technology has a wide variety of applications, from augmented reality in computer games and apps, to robot interaction, and self-driving cars. Historical images and videos can also be analysed by the software, which is useful for reconstruction of incidents or to automatically convert 2D films into immersive 3D.

“Inferring object-range from a simple image by using real-time software has a whole host of potential uses,” explained supervising researcher, Dr Gabriel Brostow (UCL Computer Science).

“Depth mapping is critical for self-driving cars to avoid collisions, for example. Currently, car manufacturers use a combination of laser-scanners and/or radar sensors, which have limitations. They all use cameras too, but the individual cameras couldn’t provide meaningful depth information. So far, we’ve optimised the software for images of residential areas, and it gives unparalleled depth mapping, even when objects are on the move.”

The new software was developed using machine learning methods and has been trained and tested in outdoor and urban environments. It successfully estimates depths for thin structures such as street signs and poles, as well as people and cars, and quickly predicts a dense depth map for each 512 x 256 pixel image, running at over 25 frames per second.

Currently, depth mapping systems rely on bulky binocular stereo rigs or a single camera paired with a laser or light-pattern projector that don’t work well outdoors because objects move too fast and sunlight dwarfs the projected patterns.

There are other machine-learning based systems also seeking to get depth from single photographs, but those are trained in different ways, with some needing elusive high-quality depth information. The new technology doesn’t need real-life depth datasets, and outperforms all the other systems. Once trained, it runs in the field by processing one normal single-lens photo after another.

“Understanding the shape of a scene from a single image is a fundamental problem. We aren’t the only ones working on it, but we have got the highest quality outdoors results, and are looking to develop it further to work with 360 degree cameras. A 360 degree depth map would be fantastically useful – it could drive wearable tech to assist disabled people with navigation, or to map real-life locations for virtual reality gaming, for example,” added first author and UCL PhD student, Clément Godard (UCL Computer Science).

Co-author, Dr Oisín Mac Aodha, previously at UCL and now at Caltech, added: “At the moment, our software requires a desktop computer to process individual images, but we plan on miniaturising it, so it can be run on hand-held devices such as phones and tablets, making it more accessible to app developers. We’ve also only optimised it for outdoor use, so our next step is to train it on indoor environments.”

The team has patented the technology through UCL Business, but has made the code available free for non-commercial use. Funding for the research was kindly provided by the Engineering and Physical Sciences Research Council.