

This commentary is part of a series in TODAY's Science section, in collaboration with the National University of Singapore's (NUS) School of Computing, that explores computer science research projects conducted here.

NG TECK KHIM



Automatically converting videos, photos into 3D

Imagine you are at the scene of a road accident and you need to quickly document the scene. You have a camera in your smartphone, but photographs do not tell you the exact position of objects, their size and shape and how they are laid out. You also do not have much time to measure and

write down all these details that you may need later. What should you do?

In another scenario, you want to measure objects seen on video and view them in 3D. Could the tools for 3D television help? Unfortunately not, because 3D movies are shot with a pair of cameras, while a normal video clip is usually taken using only one camera.

The solution to both scenarios lie in computer-vision technology, which can be used to automatically process these photographs or videos so that the physical measurements of objects in the scene can be computed and their 3D shapes recovered.

This way, you could measure the positions of cars involved in an accident, the skid mark on the road, the angle of the car after the collision, the extent of damage on the car as well as the surrounding objects. The same technology will also allow you to convert some of your old videos into 3D.

The science behind computer-vision technology lies in what is called parallax, which can be used to derive the depth of the points in an image. To illustrate, close your left eye and look with your right eye only. Without moving your head, close your right eye and look with your left eye only. As you switch eyes, you will find that stationary points nearer to you appear to move by larger distances than points further away. This is the parallax effect.

● Dr Ng Teck Khim is associate professor at the School of Computing, National University of Singapore. He and his research team work on geometrical computer vision, signal processing and mobile technologies.

As part of the research in computer-vision technology, computer scientists and engineers have studied how to get 3D from pictures in a topic called “structure-from-motion”, which literally means getting the 3D structure or shape by virtue of moving the camera. The motion of the camera generates the parallax required.

The structure-from-motion algorithm computes the 3D information automatically and is able to attach the relative distances of every point in the photograph.

A research team at the NUS School of Computing, led by myself, has developed such a system based on computer-vision technology to automatically convert videos or photographs into 3D.

For instance, a video or many photos of distant HDB blocks taken by a drone can be uploaded into a computer and the system will compute the 3D points to construct a model.

Viewers can then view the HDB blocks from any angle — even from view directions that have not been used to capture the photograph — without the need for special glasses.

The team has applied this technology to rapid 3D environment modeling, localisation and mapping. Other exciting applications such as advanced augmented reality and 3D printing can also benefit from such technique.