Disney researchers reconstruct detailed 3D scenes from hundreds of high-resolution 2D images

July 19 2013

Investigators at Disney Research, Zürich have developed a method for using hundreds of photographic images to build 3D computer models of complex, real-life scenes that meet the increasing demands of today's movie, TV and game producers for high-resolution imagery.

Building 3D models from multiple 2D images captured from a variety of viewing positions is nothing new, but doing so for highly detailed or cluttered environments at high resolution has proved difficult because of the large amounts of data involved. The Disney Research, Zürich team, however, developed an algorithm that can effectively leverage these amounts of data, and process them efficiently without the need to keep all of the input data in memory at one time.

The researchers will present their findings at ACM SIGGRAPH 2013, the International Conference on Computer Graphics and Interactive Techniques, July 21-25 in Anaheim, Calif.

Three-dimensional models have become increasingly important for digitizing, visualizing and archiving the real world. In movie production, for instance, creating accurate 3D models of movie sets is often necessary for post-production tasks such as integrating real-world imagery with computer-generated effects.

But Alexander Sorkine-Hornung, a Disney Research, Zürich research

scientist, said the applications for Disney Research, Zürich's new method extend beyond 3D modeling. "Our method could be used for applications such as automatic image segmentation, which would simplify background removal in detailed scenes," he said. "It also would be useful for image-based rendering, in which new 2D images are created by combining real images."

Many 3D models now are obtained using laser scanning. In complex, cluttered environments, however, a single laser scan misses a lot of detail because objects in the foreground can block the laser's view. Photography makes it easier to capture the scene from multiple viewpoints, revealing details that otherwise would be blocked from a single point of view. But performing the computation necessary to combine photographs to build a 3D model is burdensome at high resolutions.

Changil Kim, a Ph.D. student at Disney Research, Zürich and ETH Zürich, said he and his colleagues found a way to make that high density of image data work for them, rather than against them. Their method allows them to use the ample variation of the scene's appearance to calculate depth estimates for individual pixels, rather than patches of pixels. The depth calculations work best at the edges of objects, producing precise silhouettes.

The researchers represent the captured image data in such a way that it can be processed efficiently with a standard graphics processing unit (GPU).

The researchers demonstrated their method by photographing a number of complex outdoor and indoor scenes with a standard DSLR camera, using 100 21-megapixel-resolution images to create each 3D reconstruction. Most existing stereo reconstruction techniques have been tailored for resolutions of just 1 or 2 megapixels.

The photos were captured along a linear path; this geometry provided structure that the researchers could leverage to make processing the data more efficient. However, the researchers also generalized their approach so that it can be applied even to a set of images taken with a hand-held camera.

Provided by Disney Research