Researchers leverage redundancy in casually shot videos to enable scene-space effects

August 5 2015

The same sort of video processing effects that usually require video to be shot in controlled environments where 3-D positions of cameras and objects are precisely known can be achieved with real-world, handheld video shot from consumer-grade cameras using a new approach pioneered by Disney Research.

The technique, developed with Braunschweig University of Technology, compensates for the lack of exact 3-D information about a scene by taking advantage of the fact that most elements of a scene are visible many times in a video. The researchers found they could sample pixels of scene structure from multiple frames of a video and add a filtering process that compensates for inaccuracies in 3-D positions.

Using this method, which the researchers will describe at ACM SIGGRAPH 2015, the International Conference on Computer Graphics and Interactive Techniques, in Los Angeles Aug. 9-13, they showed they could eliminate flickering and other "noise" from videos, correct the blurring of lettering caused by camera movement, and remove objects from the foreground or background of scenes.

These and other effects were demonstrated using video recorded on an iPhone 5s, a GoPro Hero 3, a Canon point-and-shoot camera, or a Sony DSLR camera.

"We believe that our novel processing approach will enable new video applications that were previously impossible, limited, or could not be fully explored because of inevitably unreliable depth information," said Oliver Wang, research scientist at Disney Research.

Processing video based on the 3-D position of pixels - known as scene-space processing - has a number of advantages over traditional, 2-D "image-space" processing. Its utility will only increase as more hardware for recording 3-D information, such as depth-enabled smartphones and Kinect game controllers, reaches the mass market.

Even so, obtaining exact 3-D information will remain elusive for the foreseeable future for video recorded under uncontrolled conditions, said Felix Klose, a researcher at Disney and at UT Braunschweig.

To leverage the high degree of redundancy in most videos and compensate for 3-D inaccuracies, the Disney team developed an efficient method for collecting all of the video pixels sampled from multiple frames that conceivably might represent the same point in a scene. They then developed a filtering process that discards outliers from this sample and computes the output pixel color as a weighted combination of the collected samples.

This framework is compatible with parallel computer processing and, despite the large amount of data accessed, reasonable runtimes can be achieved using standard desktop computers.

More information: www.disneyresearch.com/publication/scenespace/

Provided by Disney Research