New interactive method synchronizes multiple videos

August 8 2014

Disney Research Zurich has developed a new tool to help video editors synchronize multiple video clips based on the visual content of the videos, rather than relying on timecodes or other external markers. Current editing tools include a "snapping" interface that aligns video clips based on start-and-end times; by contrast, Disney Research's VideoSnapping method is based on an analysis of the content of the video. This makes it easier to synchronize multiple clips without such cues as global timecodes or audio, and even when clips are shot at different trajectories and different formats.

"Being able to synchronize multiple videos based only on visual similarities makes possible some exciting new applications," said Oliver Wang, research scientist at DRZ. "For instance, popular events today increasingly are recorded by a large number of cameras, such as smartphones; with VideoSnapping, we can efficiently create a single video from these crowdsourced video clips, even though they are shot from different positions on devices with different capture rates and may have little or no overlap between them."

Wang and the rest of the Disney Research Zurich team will present their findings at ACM SIGGRAPH 2014, the International Conference on Computer Graphics and Interactive Techniques, Aug. 10-14, in Vancouver, Canada.

"Our video alignment technology was originally inspired by an experimental short film we produced last year in Switzerland. In one shot

we had to combine multiple takes of a scene from hand-guided cameras and there was no obvious way to do this easily," added Markus Gross, the director of Disney Research in Zurich and the film's producer.

Alignment of video clips is a task essential to a number of applications, including the compositing of multiple takes in movie production and visual effects, stitching together of video mosaics and combining multiple videos at different exposure levels to create high-dynamic range (HDR) videos.

Without timecodes or audio to guide temporal alignment of clips, editors now must manually align them, usually only with the help of start and endpoints or pre-specified markers. By contrast, with the VideoSnapping method, an editor has the option to drag a pair of video clips in a timeline-style interface, where an algorithm analyzes features in the video frames and determines the best match, causing the videos to snap into alignment. Alternately, temporal synchronization can be automatically computed, allowing for the method to scale to large numbers of clips without need for interaction.

"A major benefit of our temporal snapping is that it can greatly reduce the difficulty of finding good spatial alignments," said Alexander Sorkine-Hornung, Disney Research Zurich senior research scientist. "In fact, we have shown that by applying a simple, out-of-the-box method for spatial alignment after synchronization, we can achieve similar or better quality when compared to state-of-the-art spatiotemporal approaches."

This method doesn't work if the videos lack usable common features. But the researchers have identified possible workarounds. For instance, in a case where videos show two different people repeating a series of actions, synchronization is possible by using skeleton data extracted from Kinect sensors.

More information: In addition to Sorkine-Hornung and Wang, the research team included Disney Research Zurich's Christopher Schroers and Henning Zimmer and DRZ director Markus Gross. For more information on the project and explanatory video, visit the web site at www.disneyresearch.com/project/videosnapping

Provided by Disney Research