Researchers use innovative data collection method -- A video by Dutch band C-Mon & Kypski

May 17 2011

Researchers at New York University's Courant Institute of Mathematical Sciences have adopted an innovative data collection method for their latest work in the area of computer vision—a music video created by the Dutch progressive-electro band C-Mon & Kypski. Individual frames from the band's recent video for its song "More is Less" served as a unique visual database for the Courant researchers' work to develop computer vision technology.

Computer vision, a developing technology, aims to give eyesight to machines and is currently used in a range of applications. These include Microsoft's Kinect, which detects poses in order for game play to be controlled using only the body, and cell-phone technology that allows users to cash checks by merely snapping a picture.

However, for computer vision to truly mimic the human vision system, it must be able to reliably detect specific objects or individuals under a variety of conditions—poor lighting, cluttered backgrounds, unusual clothing, and other sources of variation. In building such a system, developers have sought to implement an algorithm to perform "pose estimation"—computer recognition of individuals or objects based on their positioning. However, in order for a computer to succeed at pose estimation it must draw from a large database of people or objects in a variety of poses—after detecting a certain pose in its field of vision, it draws on its vast database of images to find a match.

"If we had many examples of people in similar pose, but under differing conditions, we could construct an algorithm that matches based on pose and ignores the distracting information—lighting, clothing, and background," explained Graham Taylor, a post-doctoral fellow at the Courant Institute and one of the project's researchers. "But how do we collect such data?"

Departing from traditional data-collection methods, the team turned to Dutch progressive-electro band C-Mon & Kypski and, specifically, its video crowd-sourcing project--"One Frame of Fame" (http://oneframeoffame.com/)--which asks fans to replace one frame of the band's music video for the song "More or Less" with a capture from their webcams. In the project, a visitor to the band's website is shown a single frame of the video and asked to perform an imitation in front of the camera. The new contribution is spliced into the video that updates once an hour.

"This turned out to be the perfect data source for developing an algorithm that learns to compute similarity based on pose," explained Taylor, who obtained his doctorate in computer science from the University of Toronto. "Armed with the band's data and a few machine learning tricks up our sleeves, we built a system that is highly effective at matching people in similar pose but under widely different settings."

More information: The research team, which also includes NYU doctoral student Ian Spiro as well as Courant Professors Chris Bregler and Rob Fergus, will present its findings in at the 24th IEEE Conference on Computer Vision and Pattern Recognition (June 21-23) in Colorado Springs. The paper is available here: movement.nyu.edu/imitation

Provided by New York University