Fast and efficient detection of hand poses could lead to enhanced human-computer interactions

November 4 2015

A*STAR researchers have developed a program that is able to detect three-dimensional hand gestures from low-resolution depth images in real time. Credit: Michael Blann/DigitalVision/Thinkstock

A mobile phone that responds to hand signals rather than the touch of a button may soon be possible thanks to technology developed by A*STAR researchers that efficiently detects three-dimensional human hand movements from two-dimensional images in real time. Combining

this technology with devices such as laptops or mobile phones can facilitate robot control and enable human-computer interactions.

Extracting correct three-dimensional hand poses from a single image, especially in the presence of complex background signals, is challenging for computers. As the computer has to determine the general position of the hand, as well as each finger, the extraction of hand movements requires the analysis of many parameters. Optimizing these parameters at the same time would be extremely computationally demanding.

To simplify these calculations, Li Cheng and colleagues from the A*STAR Bioinformatics Institute developed a model that breaks this process into two steps. First, the general position of the hand and the wrist is determined. As a second step, the palm and individual finger poses are established using the anatomy of the human hand as a guide for the computer to narrow its options. Separating the task into two steps reduces the overall complexity of the algorithm, and accelerates computations.

Cheng notes that their method is efficient and outperforms competing techniques, "we can estimate three-dimensional hand poses efficiently with processing speeds of more than 67 video frames per second and with an accuracy of around 20 millimeters for finger joints." This speed is faster than usual video signals and therefore allows for real time analysis. The model extracts information for a broad range of poses, under different lighting conditions, and even if a person wears gloves.

Processing video signals in real time may open the door to better human-machine interactions, says Cheng. "For example, instead of the touchscreen technologies used in current mobile phones, future mobile phones might allow a user to access desired apps by simply presenting unique hand poses in front of the phone, or by typing on a 'virtual' keyboard without having to access a physical keyboard."

The field of image processing and the extracting information from hand poses or other types of human motion has become very competitive, adds Cheng, noting that the two-step model developed by the researchers could well lead to more sophisticated ways of rendering three-dimensional human forms and gestures.

Provided by Agency for Science, Technology and Research (A*STAR), Singapore