

A computer that reads body language

6 July 2017



Carnegie Mellon University researchers have developed methods to detect the body pose, including facial expressions and hand positions, of multiple individuals. This enables computers to not only identify parts of the body, but to understand how they are moving and positioned. Credit: Carnegie Mellon University

Researchers at Carnegie Mellon University's Robotics Institute have enabled a computer to understand the body poses and movements of multiple people from video in real time—including, for the first time, the pose of each individual's fingers.

This new method was developed with the help of the Panopticon Studio, a two-story dome embedded with 500 video cameras. The insights gained from experiments in that facility now make it possible to detect the pose of a group of people using a single camera and a laptop [computer](#).

Yaser Sheikh, associate professor of robotics, said these methods for tracking 2-D human form and motion open up new ways for people and machines to interact with each other, and for people to use machines to better understand the world around them. The ability to recognize [hand](#) poses, for instance, will make it possible for people to interact with computers in new and more natural ways, such as communicating with computers simply by pointing at things.

Detecting the nuances of nonverbal communication between individuals will allow robots to serve in social spaces, allowing robots to perceive what people around them are doing, what moods they are in and whether they can be interrupted. A self-driving car could get an early warning that a pedestrian is about to step into the street by monitoring body language. Enabling machines to understand human behavior also could enable new approaches to behavioral diagnosis and rehabilitation for conditions such as autism, dyslexia and depression.

"We communicate almost as much with the movement of our bodies as we do with our voice," Sheikh said. "But computers are more or less blind to it."

In sports analytics, real-time pose detection will make it possible for computers not only to track the position of each player on the field of play, as is now the case, but to also know what players are doing with their arms, legs and heads at each point in time. The methods can be used for live events or applied to existing videos.

To encourage more research and applications, the researchers have released their computer code for both multiperson and hand-pose estimation. It already is being widely used by research groups, and more than 20 commercial groups, including automotive companies, have expressed interest in licensing the technology, Sheikh said.

Sheikh and his colleagues will present reports on their multiperson and hand-pose detection methods at CVPR 2017, the Computer Vision and Pattern Recognition Conference, July 21-26 in Honolulu.

Tracking multiple people in real time, particularly in social situations where they may be in contact with each other, presents a number of challenges. Simply using programs that track the pose of an individual does not work well when applied to each individual in a group, particularly when that group gets large. Sheikh and his colleagues took a bottom-

up approach, which first localizes all the body parts in a scene—arms, legs, faces, etc.—and then associates those parts with particular individuals.

The challenges for hand detection are even greater. As people use their hands to hold objects and make gestures, a camera is unlikely to see all parts of the hand at the same time. Unlike the face and body, large datasets do not exist of hand images that have been laboriously annotated with labels of parts and positions.

But for every image that shows only part of the hand, there often exists another image from a different angle with a full or complementary view of the hand, said Hanbyul Joo, a Ph.D. student in robotics. That's where the researchers made use of CMU's multicamera Panoptic Studio.

"A single shot gives you 500 views of a person's hand, plus it automatically annotates the hand position," Joo explained. "Hands are too small to be annotated by most of our cameras, however, so for this study we used just 31 high-definition cameras, but still were able to build a massive data set."

Joo and Tomas Simon, another Ph.D. student, used their hands to generate thousands of views.

"The Panoptic Studio supercharges our research," Sheikh said. It now is being used to improve body, face and hand detectors by jointly training them. Also, as work progresses to move from the 2-D models of humans to 3-D models, the facility's ability to automatically generate annotated images will be crucial.

When the Panoptic Studio was built a decade ago with support from the National Science Foundation, it was not clear what impact it would have, Sheikh said.

"Now, we're able to break through a number of technical barriers primarily as a result of that NSF grant 10 years ago," he added. "We're sharing the code, but we're also sharing all the data captured in the Panoptic Studio."

Provided by Carnegie Mellon University

APA citation: A computer that reads body language (2017, July 6) retrieved 24 January 2022 from <https://techxplore.com/news/2017-07-body-language.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.