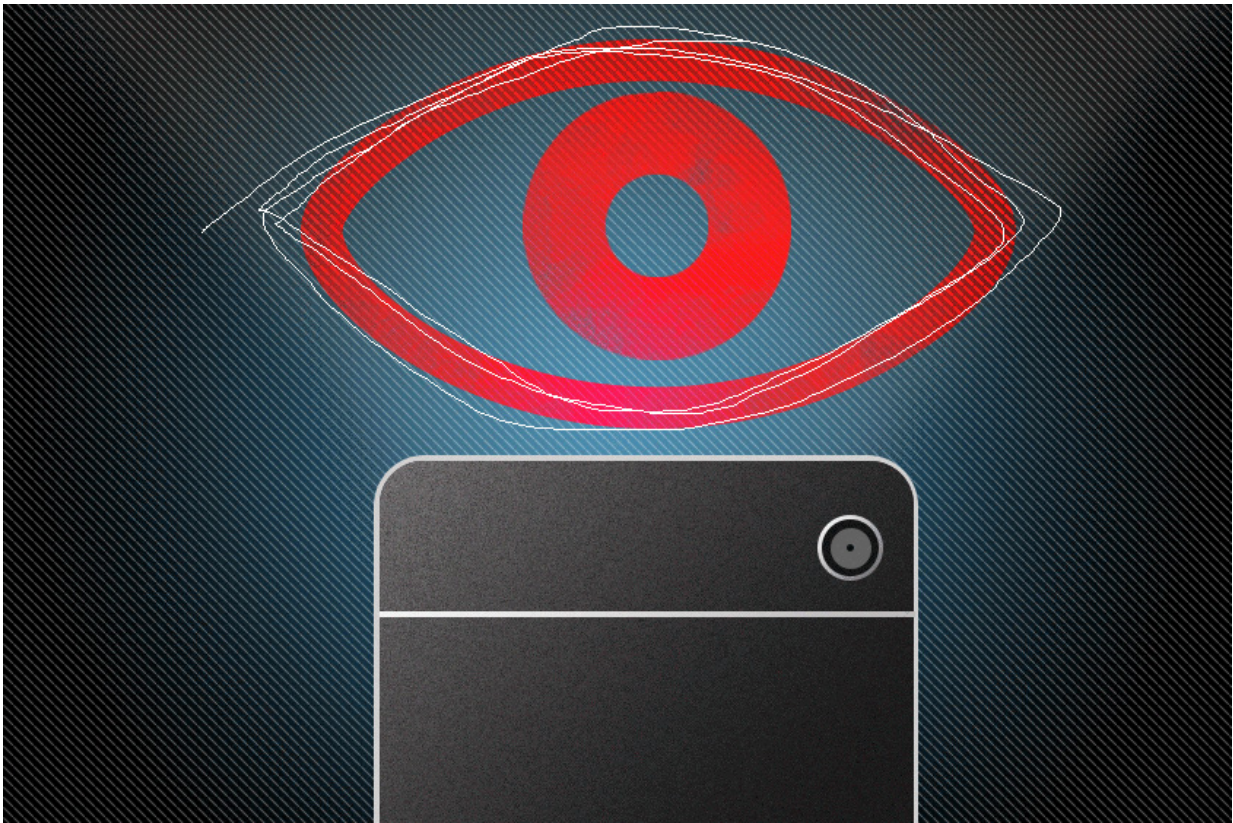# Eye-tracking system uses ordinary cellphone camera

June 16 2016, by Larry Hardesty

Researchers developed a simple application for devices that use Apple's iOS operating system. The application flashes a small dot somewhere on the device's screen, attracting the user's attention, then briefly replaces it with either an "R" or an "L," instructing the user to swipe either the right or left side of the screen. Correctly executing the swipe ensures that the user has actually shifted his or her gaze to the intended location. During this process, the device camera continuously captures images of the user's face. Credit: Massachusetts Institute of Technology

For the past 40 years, eye-tracking technology—which can determine where in a visual scene people are directing their gaze—has been widely used in psychological experiments and marketing research, but it's required pricey hardware that has kept it from finding consumer applications.

Researchers at MIT's Computer Science and Artificial Intelligence Laboratory and the University of Georgia hope to change that, with software that can turn any smartphone into an eye-tracking device. They describe their new system in a paper they're presenting on June 28 at the Computer Vision and Pattern Recognition conference.

In addition to making existing applications of eye-tracking technology more accessible, the system could enable new computer interfaces or help detect signs of incipient neurological disease or mental illness.

"The field is kind of stuck in this chicken-and-egg loop," says Aditya Khosla, an MIT graduate student in electrical engineering and computer science and co-first author on the paper. "Since few people have the external devices, there's no big incentive to develop applications for them. Since there are no applications, there's no incentive for people to buy the devices. We thought we should break this circle and try to make an eye tracker that works on a single mobile device, using just your front-facing camera."

Khosla and his colleagues—co-first author Kyle Krafka of the University of Georgia, MIT professors of electrical engineering and computer science Wojciech Matusik and Antonio Torralba, and three others—built their eye tracker using machine learning, a technique in which computers learn to perform tasks by looking for patterns in large sets of training examples.

Strength in numbers

Khosla and his colleagues' advantage over previous research was the amount of data they had to work with. Currently, Khosla says, their training set includes examples of gaze patterns from 1,500 mobile-device users. Previously, the largest data sets used to train experimental eye-tracking systems had topped out at about 50 users.

To assemble data sets, "most other groups tend to call people into the lab," Khosla says. "It's really hard to scale that up. Calling 50 people in itself is already a fairly tedious process. But we realized we could do this through crowdsourcing."

In the paper, the researchers report an initial round of experiments, using training data drawn from 800 mobile-device users. On that basis, they were able to get the system's margin of error down to 1.5 centimeters, a twofold improvement over previous experimental systems.

Since the paper was submitted, however, they've acquired data on another 700 people, and the additional training data has reduced the margin of error to about a centimeter.

To get a sense of how larger training sets might improve performance, the researchers trained and retrained their system using different-sized subsets of their data. Those experiments suggest that about 10,000 training examples should be enough to lower the margin of error to a half-centimeter, which Khosla estimates will be good enough to make the system commercially viable.

To collect their training examples, the researchers developed a simple application for devices that use Apple's iOS operating system. The application flashes a small dot somewhere on the device's screen, attracting the user's attention, then briefly replaces it with either an "R"

or an "L," instructing the user to tap either the right or left side of the screen. Correctly executing the tap ensures that the user has actually shifted his or her gaze to the intended location. During this process, the device camera continuously captures images of the user's face.

The researchers recruited application users through Amazon's Mechanical Turk crowdsourcing site and paid them a small fee for each successfully executed tap. The data set contains, on average, 1,600 images for each user.

Tightening the net

The researchers' machine-learning system was a neural network, which is a software abstraction but can be thought of as a huge network of very simple information processors arranged into discrete layers. Training modifies the settings of the individual processors so that a data item—in this case, a still image of a mobile-device user—fed to the bottom layer will be processed by the subsequent layers. The output of the top layer will be the solution to a computational problem—in this case, an estimate of the direction of the user's gaze.

Neural networks are large, however, so the MIT and Georgia researchers used a technique called "dark knowledge" to shrink theirs. Dark knowledge involves taking the outputs of a fully trained network, which are generally approximate solutions, and using those as well as the real solutions to train a much smaller network. The technique reduced the size of the researchers' network by roughly 80 percent, enabling it to run much more efficiently on a smartphone. With the reduced network, the eye tracker can operate at about 15 frames per second, which is fast enough to record even brief glances.

"In lots of cases—if you want to do a user study, in computer vision, in marketing, in developing new user interfaces—eye tracking is something

people have been very interested in, but it hasn't really been accessible," says Noah Snavely, an associate professor of [computer science](#) at Cornell University. "You need expensive equipment, or it has to be calibrated very well in order to work. So something that will work on a device everyone has, that seems very compelling. And from what I've seen, the accuracy they get seems like it's in the ballpark that you can do something interesting."

"Part of the excitement is that they've also created this way of collecting data, and also the [data](#) set itself," Snavely adds. "They did all the legwork that will make other people interested in this problem. And the fact that the community will start working on this will lead to fast improvements."

More information: Paper: "Eye tracking for everyone": [drive.google.com/file/d/0B9HHf … aThhbEx1ek9POWM/view](#)

This story is republished courtesy of MIT News ([web.mit.edu/newsoffice/](#)), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology