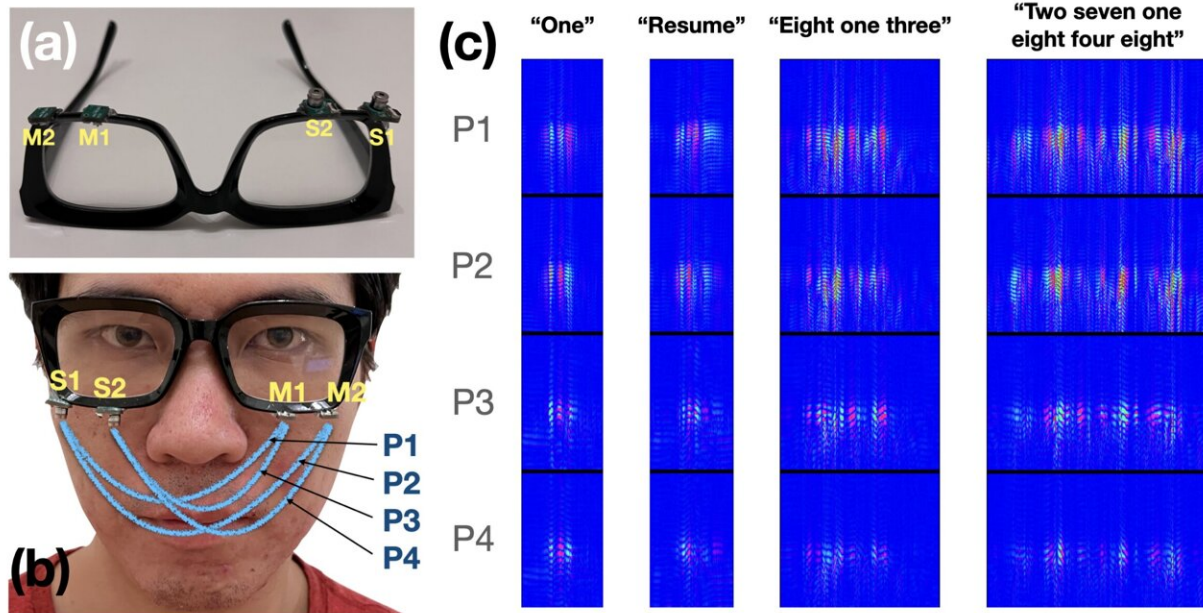# AI-equipped eyeglasses read silent speech

April 6 2023, by Louis DiPietro

System layout and echo profiles. (a-b)Final sensor position and signal paths. S1, S2: speakers; M1, M2: microphones. P1-P4: Paths. Note that each path consists of multiple path reflection and diffraction that originates from the source speaker and ends at the target microphone. The lines in the figure only illustrate the sources and targets. (c) Echo profiles for different utterances. Credit: *EchoSpeech: Continuous Silent Speech Recognition on Minimally-obtrusive Eyewear Powered by Acoustic Sensing* (2023).

Cornell University researchers have developed a silent-speech recognition interface that uses acoustic-sensing and artificial intelligence to continuously recognize up to 31 unvocalized commands, based on lip

and mouth movements.

The low-power, wearable interface—called EchoSpeech—requires just a few minutes of user training data before it will recognize commands and can be run on a smartphone.

Ruidong Zhang, doctoral student of information science, is the lead author of "EchoSpeech: Continuous Silent Speech Recognition on Minimally-obtrusive Eyewear Powered by Acoustic Sensing," which will be presented at the Association for Computing Machinery Conference on Human Factors in Computing Systems (CHI) this month in Hamburg, Germany.

"For people who cannot vocalize sound, this silent speech technology could be an excellent input for a voice synthesizer. It could give patients their voices back," Zhang said of the technology's potential use with further development.

In its present form, EchoSpeech could be used to communicate with others via smartphone in places where speech is inconvenient or inappropriate, like a noisy restaurant or quiet library. The silent speech interface can also be paired with a stylus and used with design software like CAD, all but eliminating the need for a keyboard and a mouse.

Outfitted with a pair of microphones and speakers smaller than pencil erasers, the EchoSpeech glasses become a wearable AI-powered sonar system, sending and receiving soundwaves across the face and sensing mouth movements. A deep learning algorithm then analyzes these echo profiles in real time, with about 95% accuracy.

"We're moving sonar onto the body," said Cheng Zhang, assistant professor of information science and director of Cornell's Smart Computer Interfaces for Future Interactions (SciFi) Lab.

"We're very excited about this system," he said, "because it really pushes the field forward on performance and privacy. It's small, low-power and privacy-sensitive, which are all important features for deploying new, wearable technologies in the real world."

Most technology in silent-speech recognition is limited to a select set of predetermined commands and requires the user to face or wear a camera, which is neither practical nor feasible, Cheng Zhang said. There also are major privacy concerns involving wearable cameras—for both the user and those with whom the user interacts, he said.

Acoustic-sensing technology like EchoSpeech removes the need for wearable video cameras. And because audio data is much smaller than image or video data, it requires less bandwidth to process and can be relayed to a smartphone via Bluetooth in real time, said François Guimbretière, professor in information science.

"And because the data is processed locally on your smartphone instead of uploaded to the cloud," he said, "privacy-sensitive information never leaves your control."

More information: Conference: chi2023.acm.org/

Research: ruidongzhang.com/files/papers/ … _authors_version.pdf

Provided by Cornell University