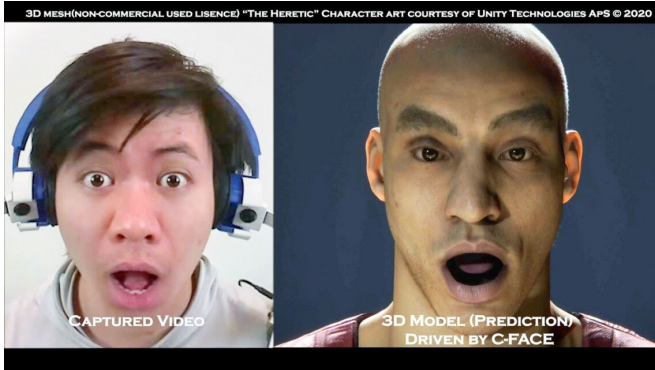


Earphone tracks facial expressions, even with a face mask

13 October 2020, by Melanie Lefkowitz



Captured video of a user's facial expression (left), with a 3-D model predicted by C-Face. Credit: Cornell University

Cornell researchers have invented an earphone that can continuously track full facial expressions by observing the contour of the cheeks—and can then translate expressions into emojis or silent speech commands.

With the ear-mounted device, called C-Face, users could express emotions to online collaborators without holding cameras in front of their faces—an especially useful communication tool as much of the world engages in remote work or learning.

"This device is simpler, less obtrusive and more capable than any existing ear-mounted wearable technologies for tracking [facial expressions](#)," said Cheng Zhang, assistant professor of information science and senior author of "C-Face: Continuously Reconstructing Facial Expressions by Deep Learning Contours of the Face With Ear-Mounted Miniature Cameras."

The paper will be presented at the Association for Computing Machinery Symposium on User Interface Software and Technology, to be held virtually Oct. 20-23.

"In previous wearable technology aiming to recognize facial expressions, most solutions needed to attach sensors on the face," said Zhang, director of Cornell's SciFi Lab, "and even with so much instrumentation, they could only recognize a limited set of discrete facial expressions."

With C-Face, avatars in virtual reality environments could express how their users are actually feeling, and instructors could get valuable information about student engagement during online lessons. It could also be used to direct a computer system, such as a [music player](#), using only facial cues.

Because it works by detecting muscle movement, C-Face can capture facial expressions even when users are wearing masks, Zhang said.

The device consists of two miniature RGB cameras—digital cameras that capture red, green and bands of light—positioned below each ear with headphones or earphones. The cameras record changes in facial contours caused when facial muscles move.

"The most exciting finding is that facial contours are highly informative of facial expressions," the researchers wrote. "When we perform a facial expression, our facial muscles stretch and contract. They push and pull the skin and affect the tension of nearby facial muscles. This effect causes the outline of the cheeks (contours) to alter from the point of view of the ear."

Once the images are captured, they're reconstructed using computer vision and a deep learning model. Since the [raw data](#) is in 2-D, a convolutional neural network—a kind of artificial intelligence model that is good at classifying, detecting and retrieving images—helps reconstruct the contours into expressions.

The model translates the images of cheeks to 42 facial feature points, or landmarks, representing the

shapes and positions of the mouth, eyes and eyebrows, since those features are the most affected by changes in expression.

Because of restrictions caused by the COVID-19 pandemic, the researchers could test the device on only nine participants, including two of the study's authors. They compared its performance with a state-of-art computer vision library, which extracts facial landmarks the image of full face captured by frontal cameras. The average error of the reconstructed landmarks was under 0.8 mm.

These reconstructed facial expressions represented by 42 feature points can also be translated to eight emojis, including "natural," "angry" and "kissy-face," as well as eight silent speech commands designed to control a music device, such as "play," "next song" and "volume up."

Among the nine participants, they found that emoji recognition was more than 88% accurate, and silent speech was nearly 85% accurate.

The ability to direct devices using facial expressions could be useful for working in libraries or other shared workspaces, for example, where people might not want to disturb others by speaking out loud. Translating expressions into emojis could help those in virtual reality collaborations communicate more seamlessly, said Francois Guimbretière, professor of information science and a co-author of the C-Face paper.

"Having a [virtual reality headset](#) allows your collaborators to move around and show you the spaces where they are, but it's very difficult in that situation to capture their faces," Guimbretière said. "What is very exciting about C-Face is that it gives you the opportunity to wear a VR set, and also to be able to translate your emotions directly to others."

One limitation to C-Face is the earphones' limited battery capacity, Zhang said. As its next step, the team plans to work on a sensing technology that uses less power.

More information: C-Face: Continuously

Reconstructing Facial Expressions by Deep Learning Contours of the Face With Ear-Mounted Miniature Cameras: www.scifilab.org/c-face

Provided by Cornell University

APA citation: Earphone tracks facial expressions, even with a face mask (2020, October 13) retrieved 19 October 2021 from <https://techxplore.com/news/2020-10-earphone-tracks-facial-mask.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.