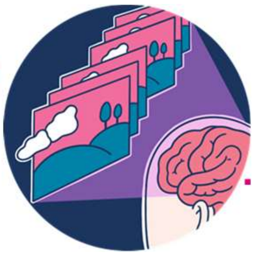
IoB Core Technology

Moonshot Research and Development Program Goal 1 Kanai Project



An ECoG-BCI for Virtual Embodiment

Chang Lab, Weill Institute for Neurosciences - University of California, San Francisco, USA.

Overview

Brain-computer interfaces (BCI) have the potential to facilitate high-bandwidth information transfer between users and technology. As a part of the UCSF BRAVO (BCI Restoration of Arm and Voice) clinical trial, we recorded neural activity directly from the surface of the brain using electrocorticography (ECoG). Here, we show that both speech and facial movements can be decoded directly from the neural activity of individuals with severe vocal-tract and bodily paralysis, and used to subsequently animate a virtual avatar. These findings form a key proof-of-concept demonstrating the potential for naturalistic, embodied communication through BCI-controlled virtual avatars.

ECoG-BCI control of a personalized facial avatar

- Embodied communication is comprised of not only speech outputs, but also intonation, expressiveness, and identity.
- Therefore, in addition to decoding speech output, decoding the accompanying facial expressions and synthesized speech are necessary steps towards a naturalistic speech neuroprosthesis.
- Towards this goal, we designed a three-component, multimodal decoding system capable of decoding speech output as text and synthesized audio, as well accompanying facial and vocal tract movements.
- This articulatory-gesture decoding system used deep, recurrent neural network modeling.
- While speech output as text was incrementally decoded, synthesized audio and facial expressions were decoded after the participant was done speaking, leading to latencies around 8 seconds.
- A combination of high and low frequency features from neural activity from the sensorimotor cortex (SMC) are decoded into flexible but discrete latent articulatory gesture units, which then animate a facial avatar.
- The facial avatar can be used alongside speech attempts or in isolation, such as to convey emotional expressions.
- The avatar (Speech Graphics, Unreal Engine 4.26) is personalized to resemble the participant.
- For further details and figures, see the published paper “A high-performance neuroprosthesis for speech decoding and avatar control” in *Nature* (https://doi.org/10.1038/s41586-023-06443-4).

Towards streaming speech synthesis and articulatory control

- Though our prior work was the first successful demonstration of high accuracy speech synthesis from neural activity during silently attempted speech, the high latencies are prohibitive to natural conversation.
- Ideally, synthesized speech and facial movements would be decoded and played or animated in real time.
- Towards this goal, we developed a speech decoding framework capable of streaming synthesized speech output.
- This model used a recurrent neural network transducer framework, enabling the model to learn both a neural encoder and implicit speech models that can act as a streaming version of a language model over latent units.
- This framework achieved faster decoding speeds (47.4 versus 28.3 words per minute) and lower latencies (1.7 versus 8.0 seconds), compared to prior work, while still maintaining a similar level of accuracy (45.3% versus 45.7% word error rate).
- Combined with our prior work showing that facial movements can be decoded from the same neural signals, future work may apply this same framework to decoding facial movements with low latencies.
- For further details and figures, see the published paper “A streaming brain-to-voice neuroprosthesis to restore naturalistic communication” in *Nature Neuroscience* (https://doi.org/10.1038/s41593-025-01905-6).

Future Prospects

Together, these findings demonstrate that an ECoG-BCI can be used to animate a virtual avatar alongside speech decoding. This work is an important step towards naturalistic, embodied neuroprostheses. Expanding the avatar to larger movement repertoires using continuous control and increasing the speed of the system are important future research directions.

Edward F. Chang, MD

Professor and Chair of Neurosurgery

In addition to being the Joan and Sanford I. Weill Chair of the Department of Neurological Surgery at UCSF, E.F.C. is the Jeanne Robertson Distinguished Professor at the UCSF Weill Institute for Neurosciences, the Co-Director of the UC Berkeley-UCSF Center for Neural Engineering and Prostheses, and a Chan Zuckerberg Biohub Investigator.

BRAVO Research Team: Jessie R. Liu, PhD; Samantha C. Brosler; Irina P. Hallinan; Alexander B. Silva, PhD; Cady M. Kurtz-Miott; Jonah Dunkel Wilker.