

Decoding and Synthesizing Speech from ECoG Using Transformer-Based Models

Toshihisa Tanaka

Tokyo University of Agriculture and Technology

Abstract: In this talk, we present the latest research findings on decoding and synthesizing speech from electrocorticogram (ECoG) signals. For speech decoding, we successfully decoded sentences from covert speech using a Transformer model trained on overt speech. The performance evaluation revealed that the model trained on overt speech demonstrated equivalent performance in decoding covert speech. This suggests the possibility of addressing the challenge of collecting training data for covert speech by utilizing overt speech. Furthermore, we report on a novel BCI paradigm that successfully reconstructs spoken sentence waveforms from ECoG by combining a Transformer-based encoder and a pre-trained neural vocoder. The performance evaluation showed that the Transformer model significantly outperformed the BLSTM. In this talk, we will discuss the details of these research findings and future perspectives. We will explore the potential of deep learning techniques, particularly Transformer-based models, towards realizing speech interfaces using ECoG signals.

Biography



Toshihisa Tanaka received the B.E., the M.E., and the Ph.D. degrees from the Tokyo Institute of Technology in 1997, 2000, and 2002, respectively. From 2000 to 2002, he was a JSPS Research Fellow. From October 2002 to March 2004, he was a Research Scientist at RIKEN Brain Science Institute. In April 2004, he joined the Department of Electrical and Electronic Engineering, at the Tokyo University of Agriculture and Technology, where he is currently a Professor. In 2005, he was a Royal Society visiting fellow at the Communications and Signal Processing Group, Imperial College London, U.K. From June 2011 to October 2011, he was a visiting faculty member in the Department of Electrical Engineering, the University of Hawaii at Manoa.

His research interests include a broad area of signal processing and machine learning, including brain and biomedical signal processing, brain-machine interfaces, and adaptive systems. He is a co-editor of *Signal Processing Techniques for Knowledge Extraction and Information Fusion* (with Mandic, Springer), 2008, and a leading co-editor of *Signal Processing and Machine Learning for Brain-Machine Interfaces* (with Arvaneh, IET, U.K.), 2018.

He served as an associate editor and a guest editor of special issues in journals, including *IEEE Access*, *Neurocomputing*, *IEICE Transactions on Fundamentals*, *Computational Intelligence and Neuroscience* (Hindawi), *IEEE Transactions on Neural Networks and Learning Systems*, *Applied Sciences* (MDPI), and *Advances in Data Science and Adaptive Analysis* (World Scientific). He served as editor-in-chief of *Signals* (MDPI). Currently, he serves as an associate editor of *Neural Networks* (Elsevier). He was a General Co-Chair of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) Tokyo in 2021. Furthermore,

he serves as a Vice-President of the Asia-Pacific Signal and Information Processing Association (APSIPA). He is a senior member of IEEE, and a member of IEICE, APSIPA, the Society for Neuroscience, and the Japan Epilepsy Society. He is the co-founder and CTO of Sigron, Inc.