

Challenges and Progress in Automatic Speech-to-Speech Translation: Bridging the Gap to Real-Time Interpretation

by

Professor Satoshi Nakamura

The Chinese University of Hong Kong, Shenzhen



Date: 31 March 2025 (Monday)

Time: 10:30 a.m. – 12:00 noon

Venue: Room 222, 2/F, Ho Sin Hang Engineering Building (SHB), CUHK

Abstract

Automatic speech-to-speech translation has long been a dream technology for humanity. Through numerous breakthroughs from years of research, we have now reached the stage where this service can be used on smartphones. However, there are still many challenges remaining before we can achieve translation quality comparable to that of a trained simultaneous interpreter. Key issues include how to translate across languages with different word orders, such as between English and Japanese, without waiting for the end of a sentence or utterance; how to balance latency and content fidelity in translation; and how to extract the speaker's intent from their intonation.

Having conducted research on speech translation over an extended period, I would like to reflect on some of the past progress and introduce ongoing research addressing these challenges.

Biography of the Speaker

Dr. Satoshi Nakamura is Presidential Chair Professor at School of Data Science, The Chinese University of Hong Kong, Shenzhen. He is a Professor Emeritus at the Nara Institute of Science and Technology (NAIST) and an Honorary Professor at the Karlsruhe Institute of Technology, Germany. He is an IEEE Fellow, ISCA Fellow, Information Processing Society of Japan Fellow, and Advanced Telecommunications Research Institute International (ATR) Fellow. He got his Ph.D. by dissertation from Kyoto University in 1992. He was Department Head, Director of Spoken Language Communication Research Laboratories, and Vice President of ATR from 2000 till 2008. From 2009 to 2010, he served as the Director General of Keihanna Research Laboratories and the Executive Director of Knowledge Creating Communication Research Center, National Institute of Information and Communications Technology. He moved to the Nara Institute of Science and Technology as a full professor in 2011. He established the Data Science Center at NAIST and served as a director from 2017 to 2021. He became a professor at the Chinese University of Hong Kong Shenzhen in 2024. His research interests include modeling and systems of spoken language processing, speech processing, spoken language translation, spoken dialog systems, natural language processing, and data science. He is one of the world leaders in speech-to-speech translation research. He has been serving for various speech-to-speech translation research projects, including C-Star, A-Star, and International Workshop on Spoken Language Translation IWSLT. He is chairperson of the International Speech Communication Association Special Interest Group: Spoken Language Translation. He also contributed to standardizing the network-based speech translation at the International Telecommunication Union. He received the Antonio Zampolli Prize in 2012.