

IEEE Distinguished Industry Speaker Dr. Tomohiro Nakatani

The IEEE Signal Processing Society is proud to announce a lecture by Dr. Tomohiro Nakatani:

Enhancing Distant Automatic Speech Recognition via Model-Based Multi-Microphone Front-Ends

The presentation will take place at **Paderborn University, lecture hall L3.204**,
on **February 27, 2026, 11:00h**.

Abstract: Distant Automatic Speech Recognition (DASR) refers to the task of recognizing speech captured by far-field microphones. It supports a wide range of applications, including the recognition of natural human conversations in everyday environments. A major challenge in DASR is maintaining high recognition accuracy in the presence of interfering signals such as background noise, reverberation, and overlapping speech.

This talk will provide an overview of model-based multi-microphone front-end techniques developed to suppress interference in DASR. A key strength of this approach is its ability to decompose signals into individual components using physical and probabilistic signal models, without necessarily requiring prior training. This property enables strong adaptability to unknown and complex environments. Moreover, when combined with neural network approaches, this framework enables highly accurate front-end processing under adverse conditions.

Through challenging DASR scenarios, the talk will demonstrate how dereverberation, denoising, and source separation front-ends can substantially enhance recognition performance.

Biography: Tomohiro Nakatani is a Senior Distinguished Researcher at the Communication Science Laboratories, NTT, Inc., Japan. He received his B.E., M.E., and Ph.D. degrees from Kyoto University in 1989, 1991, and 2002, respectively. Since joining NTT in 1991, he has focused on advancing audio signal processing technologies, including speech enhancement and robust automatic speech recognition (ASR). Together with his colleagues, Dr. Nakatani has developed several influential techniques: the blind dereverberation method Weighted Prediction Error (WPE), the blind source separation method complex Angular Central Gaussian Mixture Model (cACGMM), and the target speech extraction method SpeakerBeam. He also made pioneering contributions to the mask-based beamforming framework. His work has achieved top performance in robust ASR evaluations, including the REVERB Challenge (2014) and the CHiME-1 and CHiME-3 Challenges (2011, 2015). He has served as a member on the IEEE Signal Processing Society Audio and Acoustic Signal Processing Technical Committee (2009–2014) and the Speech and Language Processing Technical Committee (2016–2021). He was elevated to IEEE Fellow in 2021.



Moderators:

Prof. Dr. Reinhold Häb-Umbach (Paderborn University)

Prof. Dr. Gerald Enzner (IEEE SPS Germany Chapter)

Directions:

Building L, Room L3.204, opposite of Südring shopping mall, 33098 Paderborn