Speech and Audio Processing Group, Intelligence Media Division

Professor KAWAHARA Tatsuya;

Associate Professor IMOTO Keisuke,   Assistant Professor INOUE Koji;

Program-Specific Assistant Professor OCHI Keiko;

Application Code: IST-9

Description

Speech communication plays a key role in human intelligence. We are studying intelligent processing of speech and audio exchanged by human beings for automatic recognition, understanding, and interaction systems, specifically (1) automatic speech recognition of real-world conversations, (2) analysis of audio scenes composed of multiple sound sources, and (3) humanoid robots that conduct natural dialogue by combining non-verbal information.

Research Topics

1. **Speech** Recognition and Understanding
   Automatic speech recognition (ASR) of real-world conversations, including emotional and speaker attribute information, is studied based on state-of-the-art deep learning models in order to realize intelligent transcription and captioning systems.
2. **Audio** Scene Analysis
   Analysis of the audio environment, where multiple persons and a variety of sound sources exist, is studied based on multi-channel and multi-modal sensing and statistical acoustic signal processing.
3. Human-Robot **Dialogue**
   Spoken dialogue models and systems integrating verbal and non-verbal information, including backchannels, laughter, and emotion, are studied for humanoid robots (androids), which will behave like and naturally interact with human beings.

Applicants are expected to have a background in pattern recognition, machine learning, signal processing, information theory, artificial intelligence, and human interface, as well as some basic knowledge of speech processing.

Reference literature is listed, and review articles are available on our website ("Projects" page).

Contact

Tatsuya Kawahara

E-mail: kawahara@i.kyoto-u.ac.jp

Website: http://sap.ist.i.kyoto-u.ac.jp/EN/