

## Psychoacoustic Masking Effect for Noise Robust Speech Recognition Robot

Yoshikazu Miyanaga

Hokkaido University, Sapporo, Japan

This topic introduces the design of a noise robust speech recognition system. It is suitable for speech communication robots including an AI-ROBOT. For almost all of communication robots, a strong noisy robust speech recognition has been demanded. For both of a continuous speech dialog-based and a command-based automatic speech recognition (ASR) systems, we can design a strong robust ASR systems against various noise circumstances.

In this presentation, advanced speech analysis techniques named psychoacoustic masking effects have been introduced. In order to develop the robustness under low SNR, Dynamic Range Adjustment (DRA) and Modulation Spectrum Control (MSC) have been developed for the robust speech features and they focus on the speech feature adjustment with an important speech components. DRA normalizes dynamic ranges and MSC eliminates the noise corruption of speech feature parameters.

In addition to DRA and MSC, the psychoacoustic masking effects for speech feature extraction in automatic speech recognition (ASR) is also introduced in this presentation. It is based on the human auditory system. Generally, the mel-frequency cepstral coefficients (MFCC) are the most widely used speech features in ASR systems, and however one of their main drawbacks is background noise, which can affect and hamper the results. This presentation introduces noise robust speech features which improve upon MFCC. A psychoacoustic model-based feature extraction which simulates the perception of sound in the human auditory system is investigated and integrated into the MFCC. In this presentation, this new approach has been useful for noise robust speech recognition embedded into AI-Robots.