Proposal for a Master thesis

Topic: **Spatial diffuseness features for distant automatic speech recognition**

Description: Hands-free voice control of entertainment devices (e.g., an interactive TV) has become of increasing interest in recent years. However, in distant-talking scenarios, where the speaker is located several meters away from the microphones, the multiple reflections of acoustic sound waves on walls or object inside a room alter the characteristics of the recorded speech signal and limit the performance of automatic speech recognition (ASR) systems.

Since a few years, deep neural networks (DNNs) have been incorporated into the acoustic models of state-of-the-art ASR systems. To increase the robustness against environmental distortions, the so-called spatial diffuseness features have been proposed in [1], being extracted by assuming background and late reverberation to be modelled by a diffuse noise field. It has been shown in this previous work, that spatial diffuseness features improve the recognition accuracy of a DNN-based ASR system without acoustic front-end enhancement of the microphone signals.

This thesis focuses on exploiting spatial diffuseness features for DNN-based ASR systems performing a spatial filtering of the microphone signals. First, different training strategies should be investigated with and without considering the impact of the acoustic front-end signal enhancement on the diffuse noise field. Second, the efficacy of the spatial diffuseness features for improving the recognition accuracy of a DNN-based ASR system should be improved by combining the advantage of different diffuseness estimation schemes.

The implementation is to be done with MATLAB and the open-source Kaldi Toolkit. As prerequisites, the student should be interested in speech recognition and to have basic Shell and MATLAB programming experience.

[1]: Schwarz A., Huemmer C., Maas R., Kellermann W. (ICASSP 2015): *Spatial Diffuseness Features for DNN-Based Speech Recognition in Noisy and Reverberant Environments*

Interactive TV scenario

Supervisor: M.Sc. Christian Hümmer, room 01.178 (Wetterkreuz 15), huemmer@lnt.de

Professor: Prof. Dr.-Ing. Walter Kellermann

Available: Immediately