Supervisor: Naylor,P.A. 2nd marker: Brookes,D.M.

Title: Audio Signal Zoom for Small Microphone Arrays

Description:

Beamforming is an established technique for performing audio zoom (making the sound appear closer) but requires large microphone arrays typically. For small devices like phones, alternative approaches are required. This project will investigate algorithms exploiting up to 3 microphones in the context of smartphones. The aim will be to capture some example audio containing more than one sound source (person talking) and develop an algorithm that can zoom (in audio terms) towards one selected source. The implementation will be in Matlab.

**Aims & background material (student)**

The aim of the project is to develop an algorithm to zoom towards one selected audio source that can be adopted in a mobile phone. The core part of the project involves speech processing, which requires background knowledge on beamforming, blind source separation/extraction, and spatio-temporal processing using masks. Also, knowledge on programming on MATLAB is required.

**Student Summary of project deliverables, fallbacks & extensions (student)**

In terms of project deliverables, it consists of practical and theoretical work. On the practical aspect, it is expected that by the end of the project, a successful demonstration of the algorithm using MATLAB on simulated data to separate or zoom towards a specific sound source would be performed. Real-life captured audio data would be a further step if possible. Theoretically, a demonstration of understand and knowledge regarding topics such as beamforming, blind source separation, and spatio-temporal processing using masks would be expected.

On the timeline of the project, it could be divided into three phases. Phase I is generating audio data through simulation in MATLAB. To simulate room acoustics, room impulse response (RIR) has to be generated. Here based on some available RIR generator like AudioLabs (by Prof. E. Habets) and MCRoomSim (by University of Sydney), which adopts the method of images (Allen & Berkley, 1978). The outcome of this phase would be a simulator that allows manual control of details of microphones and sound sources. This would allow a more rapid verification environment for the algorithm, as one can easily modify details such as position, amplitude of the sound source and generate a new set of test data, instead of real-life recording.

Phase II, which is the core part of the project, is the audio data processing. There are three common approaches: Beamforming, Blind-source separation/extraction, and Spatio-temporal processing. In this phase, based on researches on the aforementioned aspects, the main task is to develop the source separation/extraction algorithm and implement them on the simulated data.

Phase III would be evaluation of the algorithm performance. There are currently a few existing techniques for measuring audio quality or audio distortion, such as Bark Spectral Distortion (BSD), Polqa, Pesq. However, as they are not tailored to evaluate the specific problem of source separation, the evaluation standard would be decided upon the completion of Phase II.

In parallel to the three phases, another deliverable is reading relevant literature, such as research journals and papers – not only to build up necessary knowledge database, but also to justify key decisions, such as the unlikeliness to adopt beamforming due to size constraint of microphones in mobile phones.

**Summary of Risks (student)**

In terms of the core part of the project, audio processing, the source separation algorithm is novel to a large extent. As the algorithm can be pushed for in many different directions and approaches as aforementioned, there are risks that the approach adopted may not be optimal. Besides, despite a comprehensive collection of research on Beamforming and Blind Source Separation, the research on spatio-temporal processing using Oracle masks is on an ad-hoc basis. It increases the difficulty to find related information regarding this approach in the context of this project.