



Advertisement

Master Thesis

Building a Speaker Clustering Tool by Employing Speaker Recognition Techniques

Berlin, September 2016

Analyzing the similarity of voices is useful in many research fields, such as speaker diarization (dividing an audio recording into speaker-homogeneous regions), speech synthesis (system adaptation in human-computer dialogs), and voice casting (selecting an appropriate voice for media applications). Due to the prohibitive costs of running subjective tests, it is of main interest to be able to automatically detect the speech similarity of voice segments.

In this master thesis, a speech similarity detection tool will be built for speaker clustering based on speaker recognition state-of-the-art techniques, e.g. i-vectors and deep neural networks. The student's tasks involve literature review, the implementation and adaptation of existing code, and the preparation of audio data for system training and testing. The tool's performance is to be compared to that of existing automatic speaker clustering algorithms.

Responsibilities/ Preliminary work plan

- 4 Weeks - Literature review
- 10 Weeks - Tool implementation
- 8 Weeks - Evaluating the clustering performance
- 4 Weeks - Thesis writing

Requirements

- Study subject: computer science, human computer interaction, or similar
- Strong programming skills
- Previous knowledge in / familiar with Deep Neural Networks is a plus

Contact person

Laura Fernández Gallardo, PhD
laura.fernandezgallardo@tu-berlin.de