



Development of a Meeting Transcription System

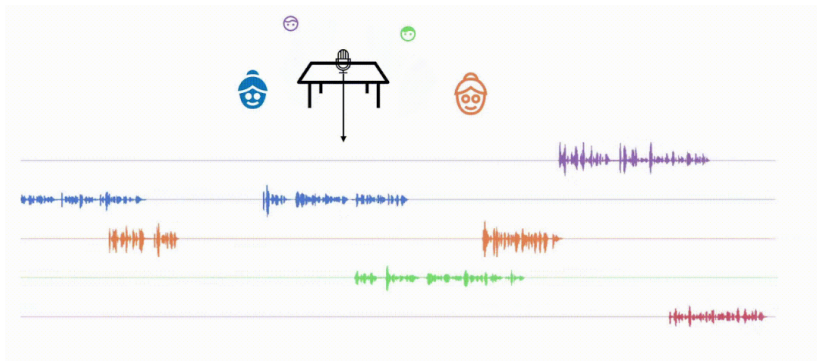
NT project group winter semester 22/23

Tobias Cord-Landwehr



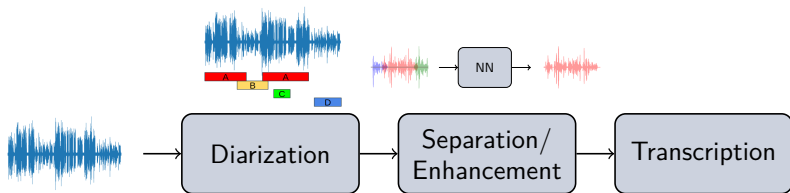
Department of Communications Engineering – Paderborn University
Prof. Dr.-Ing. Reinhold Haeb-Umbach
July 18, 2022

Scope of the Project



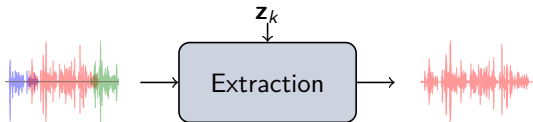
- Development of a functional meeting transcription system
- Provide the transcription for each speaker in a scenario with
 - ▶ varying numbers of speakers
 - ▶ segments of overlapping speech
 - ▶ silence segments or background noise

Meeting Transcription



- Components are implemented over the course of multiple semesters
- Components:
 - ▶ Diarization: Who spoke when? (Previous two project groups)
 - ▶ Enhancement/Extraction: Separate overlapping speech & enhance the signal quality (Upcoming project group)
 - ▶ Transcription: Map the enhanced speech signals to text (Future)

Speech Enhancement/Extraction of meeting data



- Implement a state of the art model for neural network-based source extraction
- Exploit the information provided by a previous diarization component
 - ▶ Use knowledge about speaker identity
 - ▶ Use activity information
- Ensure robustness against environmental distortions (e.g. reverberation, noise)

Requirements/Tools

Requirements

- Experience in Python coding
- Familiarity with Linux
- Knowledge of statistical signal processing
- Basic understanding of neural networks (beneficial)

Tools

- PyTorch: neural network framework
 - NumPy: scientific computing & signal processing
 - Toolboxes for audio signal processing
-
- **Kick-Off meeting: October 12, 13:00h, P7.4.02**