

3.7 Speaker Segmentation and Linking

Keywords

Speaker segmentation; speaker linking;
Speaker diarization; Information Bottleneck

Key contact researcher(s)

Prof. Hervé Bourlard
herve.bourlard@idiap.ch
Tel.: +41 27 721 77 11

Technology Transfer Office

Dr. Florent Monay
Dr. Hugues Salamin
tto@idiap.ch
Tel.: +41 27 721 77 72

Corporate Sponsorship Program

See Section 4 of the present document

File reference & version number:

Software disclosure 6536

Functional description

The speaker segmentation and linking systems automatically discover speaker boundaries in an audio and link similar speakers across audio recordings in a database. Idiap's speaker segmentation system, also referred to as the speaker diarization system, employs the Information Bottleneck approach for faster-than-real-time, unsupervised segmentation of speakers in an audio recording. The speakers discovered are then linked across audio recordings using the state-of-the-art speaker modelling and comparison approaches. Our systems have been demonstrated to work on a variety of recording conditions such as TV shows, talks, meetings, etc.

Innovative aspects

- Rapid speaker segmentation of audio with the Information Bottleneck approach
- Longitudinal linking of speakers across audio recordings with i-vectors
- Speaker overlap detection to reduce segmentation errors

Commercial application examples

- Pre-processing for speech processing systems (e.g., for speech recognition, speaker recognition, etc.)
- Mining speakers in a large audio collection

More information

“DiarTK: An Open Source Toolkit for Research in Multi-stream Speaker Diarization and its Application to Meetings Recordings”, Deepu Vijayasenan and Fabio Valente, *Proc. of Interspeech*, 2011

Software & IPR status

The Information Bottleneck based speaker diarization system is available as a free and open source software.

- <http://www.idiap.ch/scientific-research/resources/speaker-diarization-toolkit>