

IWT VoxSort™ Auto Speaker Diarization

VoxSort™ Technology Achievement

STATUS QUO

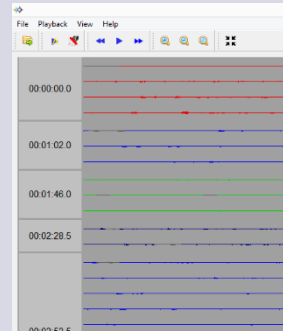
- Transcribers seeking to parse multi-speaker conversations do this manually or with slow, processor-intensive software.
- Other software requires significant computing resources to sort multi-speaker audio recordings.
- Current diarization software computing requirements prohibit real-time analysis or distributed use on small platforms.

NEW INSIGHTS

- IWT's research into speaker-dependent recognition led to new methods of speaker-specific sorting
- IWT's techniques that achieved unprecedented speech recognition accuracy and miniaturization provide the basis for better diarization.

MAIN ACHIEVEMENT:

VoxSort™ achieves higher-accuracy diarization using in some cases one percent of the computing resources of competing software while achieving greater accuracy.



HOW IT WORKS:

The approach is a higher-performance alternative to Gaussian Mixture Modelling (GMM) and Hidden Markov Modelling (HMM) methods used by competing systems. VoxSort™ leverages IWT's continuing research into ultra-miniaturization of its speech recognition technology. Key identification features are extracted with speed and efficiency up to 100X better than current approaches.

ASSUMPTIONS AND LIMITATIONS:

1. Volume of sound is normal and approximately equal for all speakers;
2. Limited non-speech sound;
3. Limited or no speaker overlap;
4. Utterance duration > 3 seconds;
5. Combined duration of each speaker should be longer than 2 minutes;
6. Not more than 7 total speakers

QUANTITATIVE IMPACT

- Provides a capability for monitoring, transcribing and otherwise parsing/processing multi-speaker conversations.
- Enables this capability to be loaded or embedded on any mobile device.
- Enables parallel analysis of large numbers of conversations on larger platforms.

END-OF-PHASE GOAL

- Real-time, mobile diarization with fully accurate segmentation and non-speech exclusion.
- Increased noise immunity for processing less-than-optimal inputs.
- Increased speech overlap capability

Highly miniaturized/optimized automatic speaker diarization/sorting software