3.2 Speech-to-Text Transcription

Keywords

Speech synthesis; text-to-speech systems; Markov models; speech signal processing

Key contact researcher(s)

Prof. Hervé Bourlard
bourlard@idiap.ch
and
Dr. Petr. Motlicek
petr.motlicek@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 2701
Software disclosure 8699

Functional description

Idiap has a growing collection of state of the art software to develop leading-edge speech-to-text systems and to perform speech recognition in various scenarios. Packages include Juicer, Tracter, ISS, and Kaldi. As one of the early pioneers in the development of hybrid systems using Hidden Markov Models (HMM) and Artificial Neural networks (ANN, now referred to DNN), we have leading-edge HMM/DNN technologies and speech recognizers available in several languages. Idiap can provide customized solutions of speech transcription systems for new languages (depending on availability of at least 100 hours of well transcribed audio data). Very large vocabulary (¿ 100 Kwords), continuous speech, speaker independent, HMM/GMM and HMM/DNN-based speech recognition available.

Innovative aspects

- Neural network based acoustic modeling, technology allowing for speaker/channel/sentence-based adaptation.
- Language model – domain customization using data crawling from Internet.
- Offline, near real-time, and real-time solutions, depending on algorithmic delay allowed by customers.

Commercial application examples

- Dictation system customized for specific domain
- Search for specific information in large audio archives
- Audio content data-mining and indexing

More information

Motlicek, Petr, Imseng, David, Potard, Blaise, Garner, Philip N. and Himawan, Ivan, "Exploiting foreign resources for DNN-based ASR", *EURASIP Journal on Audio, Speech, and Music Processing*, 2015.

Software & IPR status

IP lies in the copyright of the code, plus in expertise allowing us to create recognisers from suitable third party databases, dictionaries and tools. Some databases associated with certain languages may require separate agreements.

- Kaldi: http://kaldi.sourceforge.net/about.html