

3.4 Grapheme-based and Instance-based Methods for Speech Processing

Functional description

Keywords

Speech processing; Markov models; computational linguistics; graphemes; phonemes

Key contact researcher(s)

Dr. Mathew Magimai Doss
Mathew.MagimaiDoss@idiap.ch
Tel.: +41 27 721 77 88

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 8700
Software disclosure 11457

Phonetic lexicon and acoustic resources (transcribed speech data) are the primary resources needed for development of speech technologies, such as automatic speech recognition, text-to-speech synthesis, spoken term detection. Not all languages or domains may have such well developed resources. To overcome these limitations, Idiap is actively engaged in research on alternate ways to model the words in the language through acoustics. More specifically, (a) by modeling units of written form i.e. graphemes and (b) by modeling spoken instances.

Innovative aspects

- Speech recognition without the need of a phonetically transcribed lexicon
- Development of linguistic resources and speech technology for under-resourced languages and domains

Commercial application examples

- Speech and language technology applications where lexicon with phonetic transcription is not readily available
- Front-end of text-to-speech systems
- Search for specific information in large audio archives
- Audio content data-mining and indexing

More information

“Acoustic and Lexical Resource Constrained ASR using Language-Independent Acoustic Model and Language-Dependent Probabilistic Lexical Model”, Ramya Rasipuram and Mathew Magimai-Doss, *Speech Communication*, Vol. 68, April, 2015.

Software & IPR status

The complete software consists of open source softwares such as HTK, Quicknet, SSP (SSP: <https://github.com/idiap/ssp>) and Idiap proprietary code that interfaces the open source tools, which can be customized to the customers need. Speech databases and lexical resources used for development may require commercial license.