3.18 Sparse Phonological Vocoding

Keywords

Deep phonological representation learning, Structured sparse coding, Structured compressive sensing, Linguistic parsing

Key contact researcher(s)

Prof. Hervé Bourlard
Dr. Afsaneh Asaei
Dr. Milos Cernak
herve.bourlard@idiap.ch
afsaneh.asaei@idiap.ch
milos.cernak@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 11851

Functional description

Phonological representation of speech signal and multimedia in general, are sparse and their coefficients are highly structured. The underlying structured sparsity patterns are related to segmental and supra-segmental speech parameters, associated to the production and perception of speech at different time scales correlated with physiology of articulation as well as linguistics. The structured sparsity can thus be utilized for innovative technologies in the context of speech production and ineligibility assessment. They also enable highly efficient speech representation. Our developed computational platform can achieve ultra low bit speech coding and it is applicable for automatic speech segmentation or linguistic parsing as well as assessment of speech production and ineligibility.

Innovative aspects

- Deep learning based phonological vocoding
- Structured compressive sensing of phonological representations
- Structured sparse coding of phonological representation
- Class-specific codebook of diverse linguistic structures

Commercial application examples

- Low bit rate speech coding
- Automatic speech segmentation and alignment
- Speech production and ineligibility assessment
- Robust recognition of articulatory distorted speech

More information

“A. Asaei, M. Cernak, H. Bourlard, On Compressibility of Neural Network phonological Features for Low Bit Rate Speech Coding, proceeding of Interspeech, Dresden, Germany, September 2015.

Software & IPR status

Some databases associated with certain languages may require separate agreements.
The rights are covered by a patent application: US2015846036: Signal processing method and apparatus based on structured sparsity of phonological features.