3.12 Very Low Bit-Rate Speech Coding

Keywords

Speech processing; speech coding; very low bit rate real-time speech communication

Key contact researcher(s)

Dr. Philip N. Garner
Phil.Garner@idiap.ch
and
Dr. Milos Cernak
milos.cernak@idiap.ch
Tel.: +41 27 721 77 68/99

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:

N/A

Functional description

Idiap's very low bit-rate (VLBR) speech coding system is a conventional speech coder with an acceptable communication delay, designed for real-time speech communication. Cascaded phone-based speech recognition and synthesis systems transmit phonetic and syllabic information, encoded independently, i.e., asynchronously. It operates at an uncompressed bit rate of 213 bits/sec and achieves an average communication delay of 243 ms. A more recent version also uses phonological speech representation instead of phonetic, and is purely based on artificial neural networks, and no hidden Markov models are then used.

Innovative aspects

- Speech coding inspired by the human speech signal processing, viz., has explicit simultaneous phonetic and syllabic components, which are asynchronously related.
- The recent version transmits binary phonological speech representation, and thus is more ready for a multi-lingual use.

Commercial application examples

End-to-end speech transmission suitable for military and tactical communication systems, and for environments with highly restricted bandwidth, such as under-see communication systems.

More information

M. Cernak, P. N. Garner, A. Lazaridis, P. Motlicek, and X. Na, "Incremental Syllable-Context Phonetic Vocoding," *IEEE/ACM Trans. on Audio, Speech, and Language Processing*, Vol. 23, No. 6, pp. 10191030, 2015.

Software & IPR status

Most of the underlying software is open source. Some databases associated with certain languages may require separate agreements.

- https://gitlab.idiap.ch/milos.cernak/sct-ext
- https://gitlab.idiap.ch/milos.cernak/phonovoc