

3.13 Intonation Modeling

Keywords

Intonation modelling; text-to-speech systems; emotion recognition; speech signal processing

Key contact researcher(s)

Dr. Philip N. Garner
Phil.Garner@idiap.ch
Tel.: +41 27 721 77 68

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

Functional description

Intonation refers to the way pitch and energy of a speech signal vary in order to convey meaning, intent or emotion. It is normally ignored in speech recognition, but is essential for natural sounding speech synthesis. Idiap has novel intonation models with a physiological basis, closely matching the human production mechanism. Designed for cross-lingual emotion transfer in speech to speech translation, the models have potential application in emotion recognition and speech modification.

Innovative aspects

- Biologically plausible physiological model
- Cross-lingual adaptation

Commercial application examples

- Speech to speech translation
- Emotion recognition
- Emotional speech synthesis

More information

“Atom Decomposition-based Intonation Modelling”, Pierre-Edouard Honnet, Branislav Gerazov and Philip N. Garner, in: *IEEE 40th International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane, Australia, IEEE, 2015.

Software & IPR status

Much of the underlying software is or will soon be open-source.

- SSP: <https://github.com/idiap/ssp>