CYCLE-CONSISTENT ADVERSARIAL NETWORKS FOR NON-PARALLEL VOCAL EFFORT BASED SPEAKING STYLE CONVERSION

Shreys Seshadri1, Lauri Juvela1, Junichi Yamagishi2,3, Okko Räsänen1,4 and Paavo Alku1

1Department of Signal Processing and Acoustics, Aalto University, Finland, 2Digital Content and Media Sciences Research Division, National Institute of Informatics, Japan, 3The Centre for Speech Technology Research, University of Edinburgh, United Kingdom, 4Unit of Computing Sciences, Tampere University, Finland

Goal of the study
- Speaking style conversion (SSC) is the technology of converting natural speech signals from one style to another.
- This study focuses on SSC for speech with varying vocal effort, focused on conversion between normal and Lombard.
- We use CycleGANs as a mapping model with PML vocoder features.
- The CycleGAN was compared in subjective listening tests with other standard mapping methods used in conversion.

Data
- Read and conversational speech recordings from 20 Finnish speakers (10 female, 10 male) in normal and Lombard styles.
- Lombard and normal speech pairs were read from 20 different speakers to a total of 24 different listeners.
- Lombardness and quality subjective tests are conducted in the listening tests.
- Norrand and Lombard to normal style conversion, and Lombard to normal style conversion.

Results
- Lombardness Test
  - Lombard to Normal
  - Normal to Lombard
- Quality Test
  - Lombard to Normal
  - Normal to Lombard

Conclusions
- This work studied the use of non-parallel learning schemes to the task of vocal effort speech style conversion, in this case between normal and Lombard speech.
- CycleGAN produces encouraging results compared to the baseline methods, producing the largest Lombard effort in normal-to-Lombard conversion while having indistinguishable quality from the INCA-based approach.
- In Lombard-to-normal conversion, CycleGAN achieves superior speech quality to the other methods.
- CycleGANs seem like a promising candidate for SSC problems, as they appear to provide a strong alternative for non-parallel training on problems where parallel data scarcity is a real challenge.
- The implementation of the CycleGAN is available on GitHub:
  https://github.com/shreyas53/CycleGAN-ncCNN.

References