

# Voice Conversion Challenge 2016

## General guidelines for the submission of the entries by the participants

### Reminder of training and conversion rules

#### **Training:**

- The use of the test data for training purposes is NOT allowed.
- The use of manual phonetic transcriptions is NOT allowed (in both training and test data).
- The use of automatic speech recognition systems to generate automatic transcriptions is allowed. However, the use of the training or test data to develop the ASR system is not allowed.
- To develop a conversion system for a certain speaker-pair, the use of utterances of the other source and target speakers in the data set provided by the organizers is NOT allowed.
- The use of additional data (different from the database from which the speakers corpora will be provided) for training purposes is allowed.

#### **Conversion:**

- The 54 samples of the test data of each source speaker should be converted for all source-target pairs (25). In total, 1,350 converted samples (54 samples times 25 pairs) should be generated and submitted.
- Multiple entries for a single participant are NOT allowed (the listening test would become unmanageable).
- Manual edition or modification of the converted audios is NOT allowed.
- The conversion systems can be optimized individually (per pair) in the training step but they should not be modified in the conversion step (manual tuning of system parameters is NOT allowed).
- Any acoustic features including suprasegmental and duration features may be transformed. However, transformation of linguistic information is not considered and NOT allowed.

### Data submission

- To submit their entry (converted audios exclusively), participants should send a link to [vcc2016@vc-challenge.org](mailto:vcc2016@vc-challenge.org) to download the data via dropbox or google drive, the deadline for submission is **January 15**.
- All the converted samples of each conversion pair should be submitted with the following format:
  - Filename <source\_speaker>\_<target\_speaker>\_<file identifier> (ex: SF1\_TM1\_100001.wav for the conversion of speakers SF1 to TM1 of the file 100001.wav).
  - All the converted files should be put together in a directory with the name of the team (ex: ./CSTR/SM1\_TM1\_100001.wav).
  - The directory containing all the converted files **should be compressed into a .zip file** and named according to

<TEAM\_NAME>.zip. The information to access this file online (e.g. dropbox) should be sent to the organizers via email with the following line as subject: "VCC2016 submission for <Team name>". Please note that other compression formats **will not be accepted**.

- Please provide the md5 sum values of the zip file.
- Participants need to complete a form after the submission process providing the general technical specification of their conversion system in order to facilitate easy cross-system comparisons (e.g. is it a GMM-based system? does it convert prosodic features?). The organizers will provide further information once the entries submission is completed.

**Please contact the organizers ASAP at [vcc2016@vc-challenge.org](mailto:vcc2016@vc-challenge.org) for any question regarding the rules and entries submission.**

The VCC 2016 organizers:

- Linghui Chen (University of Science and Technology of China)
- Daisuke Saito (Tokyo University)
- Tomoki Toda (Nagoya University)
- Fernando Villavicencio (National Institute of Informatics)
- Mirjam Wester (University of Edinburgh)
- Zhizheng Wu (University of Edinburgh)
- Junichi Yamagishi (National Institute of Informatics)