Jing-Xuan Zhang
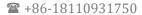
University of Sci. & Tech. of China,

National Engineering Laboratory for Speech and Language Information Processing

☎ +86-18110931750

✉ nosisi@mail.ustc.edu.cn

☰ Objective

Speech synthesis, voice conversion algorithm engineer

🎓 Education

- **2020.2-2020.8, School of Informatics, University of Edinburgh (UoE), UK**
  - ➤ Visiting PhD student supported by China Scholarship Council | IELTS score: 7.0 / 9.0
    Supervisor: Prof Korin Richmond
- **2016.9-2021.6, School of Information Sci. & Tech., University of Sci. & Tech. of China (USTC), China**
  - ➤ PhD of electronic engineering and information science | Avg. score: 90 / 100
    Supervisor: Prof Li-Rong Dai, Associate Prof. Zhen-Hua Ling
- **2012.9-2016.6, School of Gifted Young, USTC, China**
  - ➤ Bachelor of electronic engineering and information science | Avg. score: 88 / 100

🛠 Experience

- **2020.2 - 2020.8, The Centre for Speech Technology Research (CSTR), UoE**
  - ➤ Research on speech generation from lip and tongue videos, design a transfer learning strategy from TTS task and improve the intelligibility 30% over the baseline;
  - ➤ Gave a report of "Sequence-to-Sequence Based Voice Conversion" in CSTR Talk;
  - ➤ Participated in "Voice Conversion Challenge 2020" on behalf of NELSLIP & iFLYTEK team, responsible for acoustic model construction and optimization in both mono- and cross-lingual conversion task. Our team won the first place on both tasks.
- **2016.9 – 2021.6, National Engineering Laboratory for Speech and Language Information Processing (NELSLIP), USTC**
  - ➤ Research on sequence-to-sequence (seq2seq) modeling for speech generation and voice conversion (VC). The seq2seq modeling improves the prosody and duration conversion compared to the conventional method. I also extended the seq2seq VC into non-parallel task by learning disentangled speaker and linguistic representations.
- **2017.11 - 2018.8, TTS Group of iFLYTEK Research, Internship**
  - ➤ Partly participated in "Voice Conversion Challenge 2018" and our system won the first place. Contributed to "Automatic Speaker Verification (ASV2019)" dataset based on our system;
  - ➤ Research on voice conversion based on sequence-to-sequence acoustic model to improve duration and prosody conversion;

For more information, please visit my personal page: https://jxzhanggg.github.io/online-cv_en/

> ➤ Enhanced the voice quality of low-bit quantization WaveNet model using a DNN based speech enhancement model. Improved the quality while kept the stability advantage of the low-bit WaveNet.

- **2016.11 - 2017.5, Project of Speech Enhancement on Mobile Device, SAMSUNG Research & NELSLIP**
  > ➤ Built and optimized a low-latency SE system with progressive learning and multi-task learning of IRM target. Improved PESQ > 0.2 compared to the baseline.

- **2015.5 - 2016.5, National Undergraduate Innovation Projection, USTC**
  > ➤ Built a speech-to-singing system with GUI interface, based on modification of acoustic parameters. Leader of 3 people group, responsible for durational control, code integration and building of the GUI software.

📄 Publications

- **Jing-Xuan Zhang**, Zhen-Hua Ling, Li-Rong Dai, "Non-Parallel Sequence-to-Sequence Voice Conversion with Disentangled Linguistic and Speaker Representations", IEEE/ACM Trans. on Audio, Speech and Lang. Proc., 2020 11 citations

  Open-source project, Github: https://github.com/jxzhanggg/nonparaSeq2seqVC_code 92 stars

- **Jing-Xuan Zhang**, Zhen-Hua Ling, Li-Juan Liu, Yuan Jiang, Li-Rong Dai, "Sequence-to-Sequence Acoustic Modeling for Voice Conversion", IEEE/ACM Trans. on Audio, Speech, and Lang. Proc., 2019 29 citations

- **Jing-Xuan Zhang**, Zhen-Hua Ling, Li-Rong Dai, "Recognition-Synthesis Based Non-Parallel Voice Conversion with Adversarial Learning", accepted by INTERSPEECH, 2020

- **Jing-Xuan Zhang**, Zhen-Hua Ling, Yuan Jiang, Li-Juan Liu, Liang Chen, Li-Rong Dai, "Improving Sequence-to-Sequence Voice Conversion by Adding Text-Supervision", ICASSP, 2019 8 citations

- **Jing-Xuan Zhang**, Zhen-Hua Ling, Li-Rong Dai, "Forward Attention in Sequence-to-Sequence Acoustic Modeling for Speech Synthesis", ICASSP, 2018 27 citations

- Yang Ai, **Jing-Xuan Zhang**, Zhen-Hua Ling, "DNN-Based Spectral Enhancement For Neural Waveform Generators With Low-Bit Quantization", ICASSP, 2019 4 citations

- Xin Wang, *et al.,* **Jing-Xuan Zhang**, Zhen-Hua Ling, "ASVspoof 2019: a large-scale public database of synthetic, converted and replayed speech", Computer Speech and Language, 2019 3 citations

💬 Honors

- 2016-2020，First-Level Postgraduate Student Academic Scholarship, USTC
- 2014-2015，National Endeavor Scholarship, USTC
- 2012-2013，Outstanding Student Scholarship, USTC