

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
- 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
- 2012.9-2016.6, School of Gifted Young, USTC, China
 - MSc of electronic engineering and information science. Avg. score: 88 / 100

Experience

- 2020.2 -2020.8, the Center for Speech Technology Research (CSTR), UoE
 - Speech generation from lip and tongue videos. Supervisor: Prof Korin Richmond;
 - 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 in both mono and cross lingual task.
- 2017.11-2018.8, TTS Group, iFLYTEK Research Internship
 - Partly participated in "Voice Conversion Challenge 2018" and our system won the first place;
 - Research on voice conversion based on sequence-to-sequence acoustic model to improve duration and prosody conversion;
 - Enhanced the voice quality of low-bit WaveNet model using a DNN based SE 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, improving 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](#) [89 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 [28 citations](#)
- **Jing-Xuan Zhang**, Zhen-Hua Ling, Li-Rong Dai, "Recognition-Synthesis Based Non-Parallel Voice Conversion with Adversarial Learning", INTERSPEECH, 2020, submitted
- **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

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

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