Chang ZENG

PhD Candidate

Tokyo, Japan
07085400924
zengchang@nii.ac.jp
zengchang233.github.io/
researchgate.net/profile/Chang-Zeng-3

Education

Doctor, Informatics, National Institute of Informatics & SOKENDAI 2020.10 - present

Major in singing voice synthesis, speaker verification, and antispoofing

Sponsored by MEXT scholarship and supervised by Professor Junichi Yamagishi

Master,Electrical Engineering and Information Systems, The University of Tokyo 2017.10 - 2020.3

Major in speaker verification

Laboratory focuses on speech synthesis and speech assessment. I major in speaker verification, supervised by Professor Nobuaki Minematsu

Bachelor, Tianjin University 2012.9 - 2016.7

Major in Measurement and Control Technology and Instruments

Publications

1. **Zeng, C.**, Wang, X., Cooper, E., Miao, X., & Yamagishi, J. (2022). Attention back-end for automatic speaker verification with multiple enrollment utterances. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 6717-6721). IEEE. *https://doi.org/10.1109/ICASSP43922.2022.9746688*

2. **Zeng, C.**, Zhang, L., Liu, M., & Yamagishi, J. (2022). Spoofing-Aware Attention based ASV Back-end with Multiple Enrollment Utterances and a Sampling Strategy for the SASV Challenge 2022. *InterSpeech 2022.* *https://doi.org/10.21437/Interspeech.2022-10495*

3. **Zeng, C.**, Miao, X., Wang, X., Cooper, E., & Yamagishi, J. (2022). Joint Speaker Encoder and Neural Back-end Model for Fully End-to-End Automatic Speaker Verification with Multiple Enrollment Utterances. *arXiv.* *https://doi.org/10.48550/arXiv.2209.00485*.

4. Wang, C., **Zeng, C (co-first author).**, & He, X. (2022). HiFi-WaveGAN: Generative Adversarial Network with Auxiliary Spectrogram-Phase Loss for High-Fidelity Singing Voice Generation. *arXiv.* *https://arxiv.org/abs/2210.12740*. *Project page: https://wavelandspeech.github.io/xiaoice2*. *Project page: https://wavelandspeech.github.io/hifi-wavegan*

5. Wang, C., **Zeng, C (co-first author).**, & He, X. (2022). Xiaoicesing 2: A High-Fidelity Singing Voice Synthesizer Based on Generative Adversarial Network. *arXiv.* *https://arxiv.org/abs/2210.14666*. *Project page: https://wavelandspeech.github.io/xiaoice2*

6. Liu, M., Wang, L., Lee, K. A., Zhang, H., **Zeng, C.**, & Dang, J. (2021). DeepLip: A Benchmark for Deep Learning-Based Audio-Visual Lip Biometrics. In *2021 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)* (pp. 122-129). IEEE. *https://doi.org/10.1109/ASRU51503.2021.9688240*

7. Li, K., Li, S., Lu, X., Masato, A., Liu, M., Zhang, L., **Zeng, C.**, Wang, L., Dang, J., & Unoki, M. "Data Augmentation Using McAdams-Coefficient-Based Speaker Anonymization for Fake Audio Detection". *InterSpeech 2022*. *https://doi.org/10.21437/Interspeech.2022-10088*

8. Liu, X., Liu, M., Zhang, L., Zhang, L., **Zeng, C.**, Li, K., Li, N., Lee, K., Wang, L., & Dang, J. "Deep Spectro-temporal Artifacts for Detecting Synthesized Speech". *In Proceedings of the 1st International Workshop on Deepfake Detection for Audio Multimedia (DDAM '22). 69–75.* *https://doi.org/10.1145/3552466.3556527*

Skills

Programming: shell, python, cpp

Deep learning library: PyTorch, PyTorch-lightning

Speech toolkits: KALDI, ESPNet, WeNet, Speechbrain

Work Experience

Avatar Researcher 2022.07 - present

Xiaobing.ai

Singing voice synthesis, Vocoder, Singing voice conversion

Algorithm Engineer 2020.04 - 2020.10

Alibaba

Spoken term detection (QbE), End-to-End ASR, Speaker verification

Intern Algorithm Engineer 2019.7 - 2019.10

Alibaba

Speaker recognition

Intern Algorithm Engineer 2018.10 - 2018.11

Hikivision

Speaker recognition

Proficiency

Japanese JLPT-N2; TOFEL score 85 points.

Voxceleb Speaker Recognition Challenge 2019 **rank #4**

Zhijiang Cup ASR for uncertainty scenarios **rank #2**

Rank #4 of Audio Deep Synthesis Detection Challenge 2022 track1 (Low-quality Fake Audio Detection, LF)

Rank #5 of Audio Deep Synthesis Detection Challenge 2022 track2 (Partially Fake Audio Detection, PF)

Research peference

Currently I am interesting in generative task in speech such as TTS, VC or music task. I am very appreciate if you can provide such a position for me.