

HAO SHI

PhD Candidate, at [Speech and Audio Processing Lab](#), Kyoto University
✉ shi@sap.ist.i.kyoto-u.ac.jp / hshi.cca@gmail.com

RESEARCH INTERESTS

Automatic Speech Recognition:

- Robust ASR (Under noise conditions, with front-end)
- Knowledge distillation for ASR (Model compression)
- Accents-based ASR (Accent adaptation)

Speech Enhancement:

- Spectrograms fusion (Complementary between different systems, frequency-domain models)
- Front-end for robust ASR (Improve both human hearing and recognition performance)

Speech Separation:

- Blind source separation (for ASR)

EDUCATION

Ph.D. in Computer Science and Technology , Kyoto University, Kyoto, Japan Department of Intelligence Science and Technology, Graduate School of Informatics ◦ Supervisor: Prof. Tatsuya Kawahara	Apr. 2021 – Present.
Master in Computer Science and Technology , Tianjin University, Tianjin, China College of Intelligence and Computing ◦ Supervisor: Prof. Longbiao Wang	Sep. 2018 – Jan. 2021
B.Sc. in Computer Science and Technology , Southwest Jiaotong University, Sichuan, China The School of Information Science and Technology	Sep. 2014 – Jun. 2018

Research Internships

Research Assistant , at Tianjin University	Aug. 2021 – Jan. 2022
---	-----------------------

LANGUAGE SKILL

- * Chinese (native)
- * English (fluent)

PUBLICATIONS

Conference Papers (Reviewed Paper, [First Author](#)):

1. **Hao Shi**, Yuchun Shu, Longbiao Wang, Jianwu Dang, Tatsuya Kawahara, "Fusing Multiple Bandwidth Spectrograms for Improving Speech Enhancement," in Proc. of APSIPA ASC, 2022, pp. xx-xx.
2. **Hao Shi**, Longbiao Wang, Sheng Li, Jianwu Dang, Tatsuya Kawahara, "Subband-Based Spectrogram Fusion for Speech Enhancement by Combining Mapping and Masking Approaches," in Proc. of APSIPA ASC, 2022, pp. xx-xx.
3. **Hao Shi**, Longbiao Wang, Sheng Li, Jianwu Dang, Tatsuya Kawahara, "Monaural speech enhancement based on spectrogram decomposition for convolutional neural network-sensitive feature extraction," in Proc. of Interspeech, 2022, pp. 221-225.
4. **Hao Shi**, Longbiao Wang, Sheng Li, Cunhang Fan, Jianwu Dang, Tatsuya Kawahara, "Spectrograms Fusion-based End-to-end Robust Automatic Speech Recognition," in Proc. of APSIPA ASC, 2021, pp. 438-442.
5. **Hao Shi**, Longbiao Wang, Meng Ge, Sheng Li, Jianwu Dang, "Spectrograms Fusion with Minimum Difference Masks Estimation for Monaural Speech Dereverberation," in Proc. of ICASSP, 2020, pp. 7544-7548.
6. **Hao Shi**, Longbiao Wang, Sheng Li, Chenchen Ding, Meng Ge, Nan Li, Jianwu Dang, Hiroshi Seki, "Singing Voice Extraction with Attention based Spectrograms Fusion," in Proc. of Interspeech, 2020, pp. 2412-2416.

Conference Papers (Reviewed Paper, [Corresponding Author](#)):

1. Tongtong Song, Qiang Xu, Meng Ge, Longbiao Wang, **Hao Shi**, Yongjie Lv, Yuqin Lin, and Jianwu Dang, "Language-specific Characteristic Assistance for Code-switching Speech Recognition," in Proc. of Interspeech, 2022, pp. 3924-3928.

2. Qiang Xu, Tongtong Song, Longbiao Wang, **Hao Shi**, Yuqin Lin, Yongjie Lv, Meng Ge, Qiang Yu, and Jianwu Dang, "Self-Distillation Based on High-level Information Supervision for Compressing End-to-End ASR Model," in Proc. of Interspeech, 2022, pp. 1716–1720.

Conference Papers (Reviewed Paper, [Joint first author, equal contribution](#)):

1. Luya Qiang, **Hao Shi**, Meng Ge, Haoran Yin, Nan Li, Longbiao Wang, Sheng Li, and Jianwu Dang, "Speech Dereverberation Based on Scale-aware Mean Square Error Loss," in Proc. of ICONIP, 2021, pp. 55–63.
2. Haoran Yin, **Hao Shi**, Longbiao Wang, Luya Qiang, Sheng Li, Meng Ge, Gaoyan Zhang, and Jianwu Dang, "Simultaneous Progressive Filtering-based Monaural Speech Enhancement," in Proc. of ICONIP, 2021, pp. 213–221.

Conference Papers (Reviewed Paper, [Co-Author](#)):

1. Yanbing Yang, **Hao Shi**, Yuqin Lin, Meng Ge, Longbiao Wang, Qingzhi Hou and Jianwu Dang, "Adaptive Attention Network with Domain Adversarial Training for Multi-Accent Speech Recognition," in Proc. of ISCSLP, 2022, pp. xx–xx.
2. Meng Ge, Longbiao Wang, Nan Li, **Hao Shi**, Jianwu Dang, Xiangang Li, "Environment-dependent attention-driven recurrent convolutional neural network for robust speech enhancement," in Proc. of Interspeech, 2019, pp. 3153–3157.

MANUSCRIPTS

Conference Papers (First Author):

1. **Hao Shi**, Masato Mimura, Longbiao Wang, Jianwu Dang, Tatsuya Kawahara, "Time-domain Speech Enhancement Assisted by Multi-resolution Frequency Encoder And Decoder," (Under review, ICASSP 2023).

Conference Papers (Corresponding Author):

1. Qiang Xu, Tongtong Song, Longbiao Wang, **Hao Shi**, Yuqin Lin, Qiang Yu, Jianwu Dang, "Inter-layer Soft Knowledge Distillation for Speech Recognition," (Under review, ICASSP 2023).
2. Tongtong Song, Qiang Xu, Haoyu Lu, Longbiao Wang, **Hao Shi**, Yuqin Lin, Yanbing Yang, Jianwu Dang, "Monolingual Recognizers Fusion for Code-switching Speech Recognition," (Under review, ICASSP 2023).

REVIEWER

Secondary Reviewer:

1. ICASSP 2022
2. Interspeech 2022
3. APSIPA 2020
4. ICASSP 2020
5. Interspeech 2020