Hao Zhang

📱 (1)6146152261
✉ h.zhangnwpu@gmail.com
Google Scholar   LinkedIn

RESEARCH INTERESTS

My research centers on speech and audio intelligence across three complementary dimensions: **Listening** (front-end processing and representation learning), **Reasoning** (audio understanding and multimodal inference), and **Speaking** (speech/audio synthesis), with integration of language models, and human-in-the-loop learning.

EDUCATION

08/16 - 05/12 **Ph.D**., Computer Science and Engineering
The Ohio State University (OSU), Columbus, OH, USA
Major: Artificial Intelligence
PhD dissertation: *Deep Learning for Acoustic Echo Cancellation and Active Noise Control*
Advisor: Prof. DeLiang Wang

08/09 - 06/16 **B.E**. and **M.S**., Information and Signal Processing
Northwestern Polytechnical University (NWPU), Xi'an, Shaanxi, China
Bachelor thesis: *Study and Design of Small-Scale Broadband Microphone Arrays*
Master thesis: *Study of Differential Microphone Arrays and Beamforming Algorithms*
Advisor: Prof. Jingdong Chen

PROFESSIONAL EXPERIENCE

06/22 - now **Senior Researcher** (level: T11)
Tencent AI Lab, Bellevue, WA, USA

OTHER EXPERIENCE

05/21 - 08/21 **Amazon** – Sunnyvale, CA, USA – *Applied Scientist Intern*
- Topic: Deep-Adaptive Acoustic Echo Cancellation for Mobile Platforms
- Develop a hybrid echo cancellation method that integrates traditional signal processing into deep learning structure and investigate its performance in situations with continuously changing echo paths.

05/20 - 08/20 **Amazon** – Sunnyvale, CA, USA – *Applied Scientist Intern*
- Topic: Deep Learning Based Low-Latency Speech Enhancement
- Tackle the ultra low-latency speech enhancement problem using a gated temporal convolutional neural network with an asymmetric encoder-decoder structure.

05/19 - 08/19 **Amazon** – Sunnyvale, CA, USA – *Applied Scientist Intern*
- Topic: Deep Learning Based Stereo Music Source Separation
- Separate stereo music into multiple sound objects with a self-attention network and then up-mix the separated stereo audios to generate spatial music.

JOURNAL ARTICLES

[7] **H**.**Zhang**, Y. Zhang, M. Yu and D. Yu, "Enhanced Acoustic Howling Suppression via Hybrid Kalman Filter and Deep Learning Models", in IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2024.

[6] **H**.**Zhang**, A. Pandey, and D. L. Wang, "Low-latency active noise control using attentive recurrent network", in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 31, pp. 1114-1123, 2023.

[5] **H**.**Zhang**, and D. L. Wang, "Deep MCANC: A deep learning approach to multi-channel active noise control", in Neural Networks, vol. 158, pp. 318-327, 2023.

[4] **H**.**Zhang**, and D. L. Wang, "Neural cascade architecture for multi-channel acoustic echo suppression", in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 30, pp. 2326-2336, 2022.

[3] **H**.**Zhang** and D. L. Wang, "Deep ANC: A deep learning approach to active noise control", in Neural Networks, vol. 141, pp. 1-10, 2021.

[2] **H.Zhang**, J. Chen, J. Benesty, "Study of nonuniform linear differential microphone arrays with the minimum-norm filter", in Journal of Applied Acoustics, vol. 98, pp. 62-69, 2015.

[1] J. Benesty, J. Chen, C. Pan, and **H.Zhang**, "Study and design of robust differential beamformers with linear microphone arrays", in The Journal of the Acoustical Society of America, vol. 137, pp. 2238-2238, 2015.

OTHER PAPERS

[26] "Auden-Voice: General-Purpose Voice Encoder for Speech and Language Understanding." (under review).

[25] "Revisiting Audio-language Pretraining for Learning General-purpose Audio Representation." (under review)

[24] S. Wu, C Li, W. Wang, **H.Zhang**, H. Wang, M. Yu, D. Yu. "Audio-thinker: Guiding audio language model when and how to think via reinforcement learning". AAAI 2026.

[23] **H.Zhang**, W. Li, R. Chen, V. Kothapally, M. Yu, D. Yu, "LLM-Enhanced Dialogue Management for Full-Duplex Spoken Dialogue Systems", arXiv preprint arXiv:2502.14145.

[22] J. Hai, Y. Xu, **H.Zhang**, C. Li, H. Wang, M. Elhilali, D. Yu, "Ezaudio: Enhancing text-to-audio generation with efficient diffusion transformer", INTERSPEECH 2025.

[21] J. Shi, C. Zhang, J Tian, J Ni, **H.Zhang**, S. Watanabe, D. Yu, "Balancing Speech Understanding and Generation Using Continual Pre-training for Codec-based Speech LLM", ASRU 2025.

[20] Z. Sun, A. Li, R. Chen, **H.Zhang**, M. Yu, Y. Zhou, D Yu, "SMRU: Split-And-Merge Recurrent-Based UNet For Acoustic Echo Cancellation And Noise Suppression", in IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), 2024.

[19] J. Tian, C. Zhang, J. Shi, **H.Zhang**, J. Yu, S. Watanabe, D. Yu, "Preference Alignment Improves Language Model-Based TTS", in the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2025.

[18] H.T. Chiang, **H.Zhang**, Y. Xu, M. Yu, D. Yu, "Restorative Speech Enhancement: A Progressive Approach Using SE and Codec Modules", arXiv preprint arXiv:2410.01150.

[17] **H.Zhang** and D. L. Wang, "Multi-channel and multi-microphone acoustic echo cancellation using a deep learning based approach", arXiv:2103.02552.

[16] Y. Zhang, **H.Zhang**, M. Yu, and D. Yu, "Neural network augmented kalman filter for robust acoustic howling suppression", in the Conference of the International Speech Communication Association (INTERSPEECH), 2024.

[15] H. Wang, M. Yu, **H.Zhang**, C. Zhang, Z. Xu, M. Yang, Y. Zhang, and D. Yu, "Unifying Robustness and Fidelity: A Comprehensive Study of Pretrained Generative Methods for Speech Enhancement in Adverse Conditions", arXiv preprint arXiv:2309.09028.

[14] **H.Zhang**, Y. Zhang, M. Yu and D. Yu, "Advancing Acoustic Howling Suppression through Recursive Training of Neural Networks", in the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2024.

[13] Y. Zhang, M. Yu, **H.Zhang**, D. Yu and D. L. Wang, "NeuralKalman: A learnable Kalman filter for acoustic echo cancellation", in IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), 2023.

[12] **H.Zhang**, M. Yu, and D. Yu, "Deep Learning for Joint Acoustic Echo and Acoustic Howling Suppression in Hybrid Meetings", in IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), 2023.

[11] **H.Zhang**, M. Yu, and D. Yu, "Hybrid AHS: A Hybrid of Kalman Filter and Deep Learning for Acoustic Howling Suppression", in the Conference of the International Speech Communication Association (INTERSPEECH), 2023.

[10] **H.Zhang**, M. Yu and D. Yu, "Deep AHS: A deep learning approach to acoustic howling suppression", in the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2023.

[9] **H.Zhang**, and D. L. Wang, "Neural Cascade Architecture for Joint Acoustic Echo and Noise Suppression", in the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2022.

[8] **H.Zhang**, S. Kandadai, H. Rao, M. Kim, T. Pruthi, and T Kristjansson, "Deep adaptive AEC: Hybrid of deep learning and adaptive acoustic echo cancellation", in the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2022.

[7] **H.Zhang**, "Deep Learning for Acoustic Echo Cancellation and Active Noise Control", The Ohio State University (Ph.D. dissertation), 2022.

[6] **H.Zhang**, A. Pandey, and D. L. Wang, "Attentive Recurrent Network for Low-Latency Active Noise Control", in the Conference of the International Speech Communication Association (INTERSPEECH), 2022.

[5] **H.Zhang** and D. L. Wang, "A Deep Learning Approach to Multi-Channel and Multi-Microphone Acoustic Echo Cancellation", in the Conference of the International Speech Communication Association (INTERSPEECH), 2021.

[4] **H**.**Zhang** and D. L. Wang, "A Deep Learning Method to Multi-Channel Active Noise Control", in the Conference of the International Speech Communication Association (INTERSPEECH), 2021.

[3] **H**.**Zhang** and D. L. Wang, "A Deep Learning Approach to Active Noise Control", in the Conference of the International Speech Communication Association (INTERSPEECH), pp. 1141-1145, 2020.

[2] **H**.**Zhang**, K. Tan and D. L. Wang, "Deep Learning for Joint Acoustic Echo and Noise Cancellation with Nonlinear Distortions", in the Conference of the International Speech Communication Association (INTERSPEECH), pp. 4255-4259, 2019.

[1] **H**.**Zhang** and D. L. Wang, "Deep learning for acoustic echo cancellation in noisy and double-talk scenarios", in the Conference of the International Speech Communication Association (INTERSPEECH), pp. 3239-3243, 2018.

PATENTS (11 issued & 28 in submission)

[11] **H**.**Zhang**, D. Yu, "Dual-filter kalman method for acoustic feedback cancellation in hands-free karaoke environments", US Patent App. 18/760,813.

[10] **H**.**Zhang**, M. Yu, D. Yu, "Method and apparatus for task-driven speech separation by leveraging speaker distance information", US Patent App. 18/760,823.

[9] **H**.**Zhang**, M. Yu, D. Yu, "Elevating acoustic howling suppression via recursive neural network training", US Patent App. 18/417,540.

[8] M. Yu, **H**.**Zhang**, D. Yu, "Kalmannet: a learnable kalman filter for acoustic echo cancellation", US Patent App. 19/282,341.

[7] M. Yu, **H**.**Zhang**, D. Yu, "Multi-channel acoustic howling suppression using kalman filter with shared parameter estimation", US Patent App. 18/417,540.

[6] M. Yu, **H**.**Zhang**, D. Yu, "Speech codec based generative method for speech enhancement in adverse conditions", US Patent App. 18/494,492.

[5] **H**.**Zhang** M. Yu, D. Yu, "Method and apparatus for neural network augmented kalman filter for acoustic howling suppression", US Patent App. 18/494,302.

[4] **H**.**Zhang** M. Yu, D. Yu, "Deep ahs: a deep learning approach to acoustic howling suppression", US Patent App. 18/327,418.

[3] **H**.**Zhang** M. Yu, D. Yu, "Deep learning for joint acoustic echo and acoustic howling suppression in hybrid meetings", US Patent App. 18/319,039.

[2] **H**.**Zhang** M. Yu, D. Yu, "Hybrid ahs: a hybrid of kalman filter and deep learning for acoustic howling suppression", US Patent App. 18/318,910.

[1] M. Yu, **H**.**Zhang**, D. Yu, "Kalmannet: a learnable Kalman filter for acoustic echo cancellation", US Patent App. 12,406,683.

ACADEMIC SERVICES

• *Journal Reviewer*, IEEE/ACM Transactions on Audio, Speech and Language Procesing, Neural Networks, Neurocomputing Signal Processing, The Journal of the Acoustical Society of America, IEEE Journal of Selected Topics in Signal Processing, Speech Communication, Journal of Supercomputing, Measurement, Journal of Cleaner Production, International Journal of Acoustics and Vibration, Mechanical Systems and Signal Processing, IEEE Signal Processing Letters, Heliyon, Knowledge-Based Systems, Franklin Open

• *Conference Reviewer*, The International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Conference of the International Speech Communication Association (INTERSPEECH), IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)