Zhongweiyang Xu

https://xzwy.github.io/alanweiyang.github.io/

Email : zx21@illinois.edu

Mobile : +1-217-693-1582

EDUCATION

- **University of Illinois Urbana-Champaign** — Champaign, IL
  MS/PhD Student in Electrical and Computer Engineering; GPA: 4.00 — *Aug. 2021 – May 2026*

- **University of Illinois Urbana-Champaign** — Champaign, IL
  Bachelor of Science in Computer Engineering; GPA: 3.97 — *Aug. 2018 – May. 2021*

RESEARCH OVERVIEW AND INTEREST

- **Current**: My research interest is mostly involved with the combination of **classical signal processing** and **modern machine learning** in general. I have worked on single-channel and multi-channel **speech separation/enhancement** a lot and also have decent knowledge about other speech related technologies like **speech recognition**, **voice conversion**, **speech synthesizing** and **audio-visual** related topics.

- **Future**: I plan to step more into **signal** model based **generative** models to solve audio/speech inverse problem. I believe signal models would create a more **explainable** and **efficient** foundation for today's machine learning models. I'm also learning and doing research in the field of **spatial audio**.

INDUSTRY EXPERIENCE

- **Tencent Ethereal Lab** — Shenzhen, China
  Research Scientist Intern — *May 2021 - Aug 2021*
  - **KiNN**: Develop a Knowledge-inspired **computationally efficient** model for **real-time speech enhancement**. This model is designed for deployment in the **Tencent Meeting**'s (most popular video conferecing app in China) speech AI codec.

- **Tencent AI Lab** — Seattle, US
  Research Scientist Intern — *May 2021 - Aug 2021*
  - **SpatialCodec**: Research on multi-channel speech coding. Propose SpatialCodec to code spatial and spectral information independently with Residual VQ-VAE.
  - **uSee**: Research on text-based speech enhancement with score diffusion.
  - **CodecSE**: Research on codec based generative speech enhancement.

- **Meta Reality Lab Research** — Redmond, US
  Research Scientist Intern — *Aug 2021 - Dec 2021*
  - **FoVNet**: Propose a system for configurable Field-of-View speech enhancement on a smart glass with extremely low computation and distortion.

RESEARCH EXPERIENCE

- **UIUC SiNRG Lab** — Champaign, IL
  Graduate Research Assistant Advised by Prof. Romit Roy Choudhury — *May 2021 - current*
  - **Audio/Speech Inverse Problems**:
    1. Combine signal processing and modern machine learning to solve audio inverse problems like source separation, system identification, speech enhancement, DoA estimation, Beamforming...
  - **Generative AI**:
    1. Apply generative models like VQ-VAE, GAN, Flow, Diffusion, EBM, Autoregressive Language Model on signal of interest.

Teaching Experience

- **UIUC CS/ECE374 SP2021**: Introduction to Algorithms & Models of Computation
- **UIUC CS/ECE434 SP2024**: Real-World Algorithms for IoT and Data Science

Publications

- **ACM ICML2022**: *Zhongweiyang Xu, Romit Roy Choudhury*, "Learning to Sparate Voices by Spatial Regions"
- **IEEE ICASSP2023**: *Zhongweiyang Xu*, Xulin Fan*, Mark Hasegawa-Johnson*, "Dual-Path Cross-Modal Attention for Better Audio-Visual Speech Extraction"
- **IEEE ICASSP2024**: *Zhongweiyang Xu, Yong Xu, Vinay Kothapally, Heming Wang, Muqiao Yang, Dong Yu*, "SpatialCodec: Neural Spatial Speech Coding"
- **SPEAR Workshop**: *Zhongweiyang Xu, Debottam Dutta, Xulin Fan ,Mark Hasegawa-Johnson ,Romit Roy Choudhury*, "Multichannel Speech Enhancement for SPEAR Challenge: A Three Stage Approach"
- **Information Fusion**: *Andong Li, Guochen Yu, Zhongweiyang Xu, Cunhang Fan, Xiaodong Li, Chengshi Zheng*, "TaBE: Decoupling spatial and spectral processing with Taylor's unfolding method in the beamspace domain for multi-channel speech enhancement"
- **IEEE ICASSP2024**: *Muqiao Yang, Chunlei Zhang, Yong Xu, Zhongweiyang Xu, Heming Wang, Bhiksha Raj, Dong Yu*, "uSee: Unified Speech Enhancement and Editing with Conditional Diffusion Models"
- **ISCA INTERSPEECH2024**: *Zhongweiyang Xu, Ali Aroudi, Ke Tan, Ashutosh Pandey, Jung-Suk Lee, Buye Xu, Francesco Nesta*, "FoVNet: Configurable Field-of-View Speech Enhancement with Low Computation and Distortion for Smart Glasses"

Preprints

- **arXiv**: *Heming Wang, Meng Yu, Hao Zhang, Chunlei Zhang, Zhongweiyang Xu, Muqiao Yang, Yixuan Zhang, Dong Yu*, "Unifying Robustness and Fidelity: A Comprehensive Study of Pretrained Generative Methods for Speech Enhancement in Adverse Conditions"
- **arXiv**: *Zhongweiyang Xu, Debottam Dutta, Yu-Lin Wei, Romit Roy Choudhury*, "Multi-Source Music Generation with Latent Diffusion"   (ICASSP 2025 in Submission)

Projects

- **Linux Kernel Project**: Develop an OS Kernel for the OS Class.
- **BeatDance on FPGA**: Develop a GuitarHero-like game on a Cyclone-V FPGA.
- **Two Factor Authentication using Earphones**: Use personal earphone's production imperfection as a source for 2FA.
- **RISC-V CPU**: Use SystemVerilog to develop a RISC-V CPU with cache and pipeline.
- **Feature Visualization Survey and Experiment**: Write a survey about feature visualization techniques for vision neural networks and also gives several propositions regarding how to interpret features for audio-related models. Experiment on simple visualization of audio separation networks.

Skills

- **Languages** Python, C, C++, SystemVerilog

- **Coursework and Skills** Machine Learning, Deep Learning, Multimedia Signal Processing, Random Process, Generative AI, Computer Vision, NLP, Optimization, Adaptive Signal Processing, Information Theory, Speech Processing, Advanced DSP, Vector Space Signal Processing, Detection and Estimation Theory, Wireless Sensing, Digital System Design, Operating System.