Zhongweiyang Xu

https://zhongweiyangxu.web.illinois.edu/

Email : https://zhongweiyangxu.web.illinois.edu/

Mobile : +1-217-693-1582

Education

- **University of Illinois Urbana-Champaign** — Champaign, IL
  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 **multimedia signal processing** and **machine learning** in general. I have worked on **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 speech technologies and also involve other sensors like **RF** and **IMU** sensors into multimedia applications. I also intend to design more efficient models which are realistic in real-world applications like **hearing aids**, **video conferencing**, and **augmented reality**.

Industry Experience

- **Tencent Ethereal Lab** — Shenzhen, China
  Research Scientist Intern — *May 2021 - Aug. 2021*
  - **KiNN**: Develop a Knowledge-inspired Nerual Netork (KiNN) for **real-time computationally efficient ultra-lightweight** model for **speech enhancement**. This model is designed for deployment in the **Tencent Meeting**'s (most popular video conferecing app in China) speech AI codec and for competition in **ICASSP2023 Deep Noise Suppression Challenge**.

Research Experience

- **UIUC SyNRG Lab** — Champaign, IL
  Graduate Research Assistant Advised by Prof. Romit Roy Choudhury — *May 2021 - current*
  - **Learning to Sparate Voices by Spatial Regions (First Author Paper Accepted to ICML2022)**:
    1. Propose a spatial **region-wise** separation mechanism for **binaural** speech applications like audio augmented reality and hearing aids.
    2. Use traditional clustering as a teacher model to allow **self-supervised** training for binaural spatial separation, so that personalized model (with personalized HRTF encoded) can be trained simply from personal recordings.
    3. Our model shows superior results comparing with using a general model.
- **UIUC Statistical Speech Technology Group Lab** — Champaign, IL
  Graduate Research Assistant Advised by Prof. Mark Hasegawa-Johnson — *May 2021 - current*
  - **Dual-path Attention is All You Need for Audio-Visual Speech Extraction (First Author Paper in Submission to INTERSPEECH2022)**:
    Propose a **dual-path multi-modal attention** model for **audio-visual speech extraction** for better **multi-modal fusion**. Our model is able to achieve **17+ dB** for SI-SNRi on LRS3 dataset (in the wild) for 2-5 number of speakers and achieves SOTA performance which exceeds all previous work by a large margin ($\geq$ 3dB).

Teaching Experience

- **UIUC CS/ECE374 SP2021**: Introduction to Algorithms & Models of Computation

Publications

- **ACM ICML2022**: *Zhongweiyang Xu, Romit Roy Choudhury*, "Learning to Sparate Voices by Spatial Regions", Acceptance rate: 1235/5630 = **21.9%**

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, Wireless Sensing, Digital System Design, Operating System.