Chao Huang

Prepared on Aug. 2025

CONTACT

Wegmans Hall, 2207
University of Rochester
Rochester, NY, 14627

chuang65@cs.rochester.edu
+1 (585) 910-9360
http://wikichao.github.io

EDUCATION

University of Rochester — Jan. 2021 – Present
Ph.D. in Computer Science — Rochester, NY

Nanjing University — Sept. 2015 – Jun. 2019
B.Eng. in Electronic Science and Engineering — Nanjing, China

RESEARCH AREA

Multimodal Learning and Generation, Audio-Visual Learning, Video Understanding

WORK EXPERIENCE

Advanced Micro Devices Gen AI — May. 2025 – Aug. 2025
Research Intern — Remote
Mentor: *Dr.* Jiang Liu

Meta Reality Labs Research — May. 2024 – Aug. 2024
Research Intern — Cambridge, United Kingdom
Mentor: *Dr.* Sanjeel Parekh, *Dr.* Ruohan Gao, *Dr.* Anurag Kumar

Meta Reality Labs Research — May. 2023 – Nov. 2023
Research Intern — Pittsburgh, PA
Mentor: *Dr.* Dejan Markovic, *Dr.* Alexander Richard

The Chinese University of Hong Kong — Jul. 2019 – Dec. 2020
Research Assistant — Shatin, Hong Kong
Mentor: *Prof.* Chi-Wing Fu

PUBLICATIONS

- **Chao Huang**, Susan Liang, Yapeng Tian, Anurag Kumar, and Chenliang Xu. High-Quality Sound Separation Across Diverse Categories via Visually-Guided Generative Modeling. Submitted to *International Journal on Computer Vision*, IJCV.
- Susan Liang, **Chao Huang**, Yunlong Tang, Zeliang Zhang, and Chenliang Xu. π-AVAS: Can Physics-Integrated Audio-Visual Modeling Boost Neural Acoustic Synthesis? In *International Conference on Computer Vision*, 2025. (**ICCV**)
- Yunlong Tang, Junjia Guo, Pinxin Liu, ..., **Chao Huang**, ..., Zeliang Zhang, and Chenliang Xu. Generative AI for Cel-Animation: A Survey. In *International Conference on Computer Vision Workshops*, 2025. (**ICCVW**)
- **Chao Huang**, Ruohan Gao, J. M. F. Tsang, Jan Kurcius, Cagdas Bilen, Chenliang Xu, Anurag Kumar, Sanjeel Parekh. Learning to Highlight Audio by Watching Movies. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2025. (**CVPR**)

- **Chao Huang**, Susan Liang, Yunlong Tang, Li Ma, Yapeng Tian, and Chenliang Xu. FreSca: Scaling in Frequency Space Enhances Diffusion Models. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2025. (**CVPRW**)
- Yunlong Tang, Junjia Guo, ..., **Chao Huang**, Jing Bi, Zeliang Zhang, Pooyan Fazli, and Chenliang Xu. VIDCOMPOSITION: Can MLLMs Analyze Compositions in Compiled Videos? In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2025. (**CVPR**)
- Yunlong Tang, Jing Bi, Siting Xu, ..., **Chao Huang**, ..., Ping Luo, Jiebo Luo, and Chenliang Xu. Video Understanding with Large Language Models: A Survey. In *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, 2025. (**TCSVT**)
- **Chao Huang**, Susan Liang, Yapeng Tian, Anurag Kumar, and Chenliang Xu. High-Quality Audio-Visual Separation with Generative Diffusion Models. In *17th Asian Conference on Computer Vision*, 2024. **(ACCV)** ***Best Paper, Runner-Up***
- Susan Liang, **Chao Huang**, Yapeng Tian, Anurag Kumar, and Chenliang Xu. Language-Guided Joint Audio-Visual Editing Via One-Shot Adaptation. In *17th Asian Conference on Computer Vision*, 2024. **(ACCV)**
- **Chao Huang**, Dejan Markovic, Chenliang Xu, Alexander Richard. Modeling and Driving Human Body Soundfields through Acoustic Primitives. In *The 18th European Conference on Computer Vision*, 2024. **(ECCV)**
- Susan Liang, **Chao Huang**, Yapeng Tian, Anurag Kumar, and Chenliang Xu. Neural Acoustic Context Field: Rendering Realistic Room Impulse Response With Neural Fields. In *IEEE International Conference on Computer Vision Workshops*, 2023. (**ICCVW**)
- Susan Liang, **Chao Huang**, Yapeng Tian, Anurag Kumar, and Chenliang Xu. AV-NeRF: Learning Neural Fields for Real-World Audio-Visual Scene Synthesis. In *Thirty-seventh Conference on Neural Information Processing Systems* , 2023. **(NeurIPS)**
- **Chao Huang**, Yapeng Tian, Anurag Kumar, and Chenliang Xu. Egocentric Audio-Visual Object Localization. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023. (**CVPR**)
- Luchuan Song, Jing Bi, **Chao Huang**, and Chenliang Xu. Audio-visual action prediction with soft-boundary in egocentric videos. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2023. (**CVPRW**)
- **Chao Huang**, Yapeng Tian, Anurag Kumar, and Chenliang Xu. Audio-Visual Object Localization in Egocentric Videos. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2022. (**CVPRW**)
- **Chao Huang**, Haojie Liu, Tong Chen, Qiu Shen, and Zhan Ma. Extreme Image Compression via Multiscale Autoencoders with Generative Adversarial Optimization. In *IEEE Visual Communications and Image Processing*, 2019. ***Oral Presentation*** (**VCIP**)
- Xuefei Yan, David J Brady, Weiping Zhang, Changzhi Yu, Yulin Jiang, Jianqiang Wang, **Chao Huang**, Zian Li, Zhan Ma. Compressive Sampling for Array Cameras. *SIAM Journal on Imaging Science*, 2021. (**SIIMS**)

ARXIV PREPRINTS

- **Chao Huang**, Yuesheng Ma, Junxuan Huang, Susan Liang, Yunlong Tang, Jing Bi, Wenqiang Liu, Nima Mesgarani , and Chenliang Xu. ZeroSep: Separate Anything in Audio with Zero Training. In *arXiv preprint*, 2025.
- **Chao Huang**, Susan Liang, Yunlong Tang, Yapeng Tian, Anurag Kumar, and Chenliang Xu. Scaling Concept with Text-Guided Diffusion Models. In *arXiv preprint*, 2024.
- **Chao Huang***, Ruihui Li*, Xianzhi Li, and Chi-wing Fu. Non-local Part-Aware Point Cloud Denoising. *arXiv preprint*, 2020. (* joint 1st authors)

PATENTS

- "First-person audio-visual object localization systems and methods." XU Chenliang, **Chao Huang**, Yapeng Tian, FNU Anurag Kumar - US Patent App. 18/599,398, 2024.
- "Neural radiance field systems and methods for synthesis of audio-visual scenes." XU Chenliang, Susan Liang, **Chao Huang**, Yapeng Tian, FNU Anurag Kumar - US Patent App. 18/431,491, 2024.

AWARDS

Best Paper Award, Honorable Mention at ACCV 2024
for the paper "High-Quality Visually-Guided Sound Separation from Diverse Categories"

INVITED TALKS

Egocentric Audio-Visual Localization
Invited paper talk at The Joint International 3rd Ego4D and 11th EPIC Workshop, CVPR 2023

SERVICES

Reviewer: AAAI (2023-2025), CVPR (2022-2025), ICCV (2025), ACM MM (2023-2025), NeurIPS (2024-2025), TMM, TIP, SIGGRAPH