

Chao Huang

Prepared on Sep. 2025

CONTACT Wegmans Hall, 2207 chuang65@cs.rochester.edu
University of Rochester +1 (585) 910-9360
Rochester, NY, 14627 <http://wikichao.github.io>

EDUCATION **University of Rochester** Jan. 2021 – Present
Ph.D. in Computer Science Rochester, NY, USA

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, USA
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. In *International Journal on Computer Vision*. (**IJCV**)
 - **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 *The Thirty-Ninth Annual Conference on Neural Information Processing Systems*, 2025. (**NeurIPS**)
 - **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, 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

- **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)
- **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)
- **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)
- **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)
- Yunlong Tang, Pinxin Liu, Mingqian Feng, ..., **Chao Huang**, ..., and Chenliang Xu. MMPerspective: Do MLLMs Understand Perspective? A Comprehensive Benchmark for Perspective Perception, Reasoning, and Robustness. In *Thirty-Ninth Conference on Neural Information Processing Systems Datasets And Benchmarks Track*, 2025. (NeurIPS D&B)
- Jiani Liu, Zhiyuan Wang, Zeliang Zhang, **Chao Huang**, Susan Liang, Yunlong Tang, and Chenliang Xu. Harnessing the Computation Redundancy in ViTs to Boost Adversarial Transferability. In *Thirty-Ninth Conference on Neural Information Processing Systems*, 2025. (NeurIPS)
- 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)
- 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)
- 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)
- 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)
- 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)
- 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)

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 & HONORS

Best Paper Award, Honorable Mention at ACCV 2024

for the paper “High-Quality Visually-Guided Sound Separation from Diverse Categories”

Doctoral Consortium at ICCV 2025

for the presentation “Seeing the Sound, Hearing the Scene: Generative and Spatial Audio-Visual Modeling”

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