

Chao Huang

Prepared on Feb. 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 Ph.D. in Computer Science	Jan. 2021 – Present Rochester, NY
	Nanjing University B.Eng. in Electronic Science and Engineering	Sept. 2015 – Jun. 2019 Nanjing, China
RESEARCH AREA	Multimodal Learning and Generation, Video Understanding	
WORK EXPERIENCE	Meta Reality Labs Research <i>Research Intern</i> Mentor: <i>Dr. Sanjeel Parekh, Dr. Ruohan Gao, Dr. Anurag Kumar</i>	May. 2024 – Aug. 2024 Cambridge, United Kingdom
	Meta Reality Labs Research <i>Research Intern</i> Mentor: <i>Dr. Dejan Markovic, Dr. Alexander Richard</i>	May. 2023 – Nov. 2023 Pittsburgh, PA
	The Chinese University of Hong Kong <i>Research Assistant</i> Mentor: <i>Prof. Chi-Wing Fu</i>	Jul. 2019 – Dec. 2020 Shatin, Hong Kong
RESEARCH EXPERIENCE	University of Rochester Advisor: Chenliang Xu Project: Multi-modal Scene Understanding and Generation	Jan. 2021 – Present Rochester, NY
	Nanjing University Advisor: Zhan Ma Project: Extreme Image Compression	Sept. 2018 – Jun. 2019 Nanjing, China
PUBLICATIONS	<ul style="list-style-type: none">• Chao Huang, Susan Liang, Yapeng Tian, Anurag Kumar, and Chenliang Xu. DAVIS: High-Quality Audio-Visual Separation with Generative Diffusion Models. In <i>17th Asian Conference on Computer Vision (ACCV)</i>, 2024. (Oral Presentation, 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 <i>17th Asian Conference on Computer Vision (ACCV)</i>, 2024.• Chao Huang, Dejan Markovic, Chenliang Xu, Alexander Richard. Modeling and Driving Human Body Soundfields through Acoustic Primitives. In <i>The 18th European Conference on Computer Vision (ECCV)</i>, 2024.	

- 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 (ICCV Workshop)*, 2023.
- 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 (NeurIPS)*, 2023.
- **Chao Huang**, Yapeng Tian, Anurag Kumar, and Chenliang Xu. Egocentric Audio-Visual Object Localization. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.
- 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 (CVPR Workshop)*, 2023.
- **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 (CVPR Workshop)*, 2022.
- 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 (SIIMS)*, 2021.
- **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 (VCIP)*, 2019. (Oral Presentation)

ARXIV PREPRINTS

- Yunlong Tang, Junjia Guo, Pinxin Liu, ..., **Chao Huang**, ..., Zeliang Zhang, and Chenliang Xu. Generative AI for Cel-Animation: A Survey. In *arXiv preprint*, 2025.
- Yunlong Tang, Junjia Guo, Hang Hua, Susan Liang, Mingqian Feng, Xinyang Li, Rui Mao, **Chao Huang**, Jing Bi, Zeliang Zhang, Pooyan Fazli, and Chenliang Xu. VIDCOMPOSITION: Can MLLMs Analyze Compositions in Compiled Videos? In *arXiv preprint*, 2024.
- **Chao Huang**, Susan Liang, Yunlong Tang, Yapeng Tian, Anurag Kumar, and Chenliang Xu. Scaling Concept with Text-Guided Diffusion Models. In *arXiv preprint*, 2024.
- Yunlong Tang, Jing Bi, Siting Xu, ..., **Chao Huang**, ..., Ping Luo, Jiebo Luo, and Chenliang Xu. Video Understanding with Large Language Models: A Survey. In *arXiv preprint*, 2023.
- **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	ACCV Best Paper Honorable Mention Award, 2024
INVITED TALKS	Egocentric Audio-Visual Localization <i>Invited paper talk at The Joint International 3rd Ego4D and 11th EPIC Workshop, CVPR 2023</i>
SERVICES	Reviewer: AAAI, CVPR, ICCV, ACM MM, TMM, TIP, SIGGRAPH