Chao Huang

Prepared on Nov. 2024

CONTACT

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

chuang65@ur.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, Video Understanding

WORK EXPERIENCE

Meta Reality Labs Research — May. 2024 – Aug. 2024
Research Intern — Cambridge, United Kingdom
Mentor: *Dr.* Sanjeel Parekh, *Dr.* Ruohan Gao, *Dr.* Anurag Kumar
Project: We presented a new task visually-guided acoustic highlighting to bridge the gap between visual and acoustic saliency in video content. We leverage movies for free supervision and propose a transformer-based multimodal framework.

Meta Reality Labs Research — May. 2023 – Nov. 2023
Research Intern — Pittsburgh, PA
Mentor: *Dr.* Dejan Markovic, *Dr.* Alexander Richard
Project: We model 3D spatial audio produced by body motion and speech so that spatial audio can be rendered at any arbitrary position in the 3D space. Specifically, we learn explicit sound field representation based on low-order spherical harmonics.

The Chinese University of Hong Kong — Jul. 2019 – Dec. 2020
Research Assistant — Shatin, Hong Kong
Mentor: *Prof.* Chi-Wing Fu
Project: We developed novel deep neural network-based algorithms for low-level 3D point cloud processing tasks like point cloud upsampling and denoising.

RESEARCH EXPERIENCE

University of Rocehster — Jan. 2021 – Present
Advisor: Chenliang Xu — Rochester, NY
Project: **Multi-modal Scene Understanding and Generation**

- A series of work integrating multimodal cues, particularly video, audio, and text for various downstream tasks such as audio-visual localization, audio-visual separation, text-guided image/audio editing, joint audio-visual editing through text, and visually-guided spatial audio generation.

Nanjing University — Sept. 2018 – Jun. 2019
Advisor: Zhan Ma — Nanjing, China
Project: **Extreme Image Compression**

- Proposed a novel Multi-Scale AutoEncoder framework to better preserve the global information and local details and adopted generative adversarial optimization for extreme image compression, which can be used in situations such as in-depth communication and web snapshots.

PUBLICATIONS

- **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**, Susan Liang, Yapeng Tian, Anurag Kumar, and Chenliang Xu. DAVIS: High-Quality Audio-Visual Separation with Generative Diffusion

Models. In *17th Asian Conference on Computer Vision **(ACCV)***, 2024. (oral presentation)

- 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 **(ACCV)***, 2024.
- **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 **(ECCV)***, 2024.
- Susan Liang, **Chao Huang**, Yapeng Tian, Anurag Kumar, and Chenliang Xu. Neural Acoustic Context Field: Rendering Realistic Room Impulse Response With Neural Fields. *arXiv preprint*, 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***, Ruihui Li*, Xianzhi Li, and Chi-wing Fu. Non-local Part-Aware Point Cloud Denoising. *arXiv preprint*, 2020. (* joint 1st authors)
- **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)

SERVICES **Reviewer**: AAAI, CVPR, ACM MM, TMM, TIP