Chao Huang

Prepared on Nov. 2023

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 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, Egocentric Video Understanding, 3D Vision,	
WORK EXPERIENCE	Meta Reality Labs Research Research Intern Mentor: Dr. Dejan Markovic Project: We model 3D spatial audio produced by b spatial audio can be rendered at any arbitrary positi	May. 2023 – Nov. 2023 Pittsburgh, PA body motion and speech so that ion in the 3D space. Specifically,
	 we learn explicit sound held representation based on The Chinese University of Hong Kong Research Assistant Mentor: Prof. Chi-Wing Fu Project: We developed novel deep neural network-b point cloud processing tasks like point cloud upsamp 	Jul. 2019 – Dec. 2020 Shatin, Hong Kong ased algorithms for low-level 3D pling and denoising.
RESEARCH EXPERIENCE	 University of Rocehster Jan. 2021 – Present Advisor: Chenliang Xu Rochester, NY Project: Audio-Visual Scene Understanding We formalize a new task: Egocentric Audio-Visual Object Localization, to explore fine-grained audio-visual association in egocentric videos. Concretely, we propose to mitigate out-of-view sounds with visual guidance and address egomotion with a geometry-aware temporal modeling module. We solve the audio-visual separation problem in a generative manner. Specifically, we leverage a generative diffusion model and a Separation U-Net to synthesize separated magnitudes starting from Gaussian noises, conditioned on both the audio mixture and the visual footage. 	
	 Project: Audio-Visual Scene Synthesis We define novel audio-visual scene synthesis as a task to synthesize a target video, including visual frames and the corresponding spatial audio, along an arbitrary camera trajectory from given source videos and trajectories. We introduce a novel acoustic-aware audio generation method to encode prior knowledge of sound propagation and propose a coordinate transformation mechanism for effective direction expression. In addition, we introduce a binaural audio augmentation method to improve acoustic supervision. 	
	 Nanjing University Advisor: Zhan Ma Project: Extreme Image Compression Proposed a novel Multi-Scale AutoEncoder fr global information and local details and adop mization for extreme image compression, whic as in-depth communication and web snapshots 	Sept. 2018 – Jun. 2019 Nanjing, China amework to better preserve the oted generative adversarial opti- h can be used in situations such s.

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

SKILLS Python, C, C++, PyTorch, TensorFlow, NumPy, OpenCV, Slurm