

ZEYUAN CHEN

☎ 858-625-1886 • ✉ zec016@ucsd.edu • 🏠 Homepage • in LinkedIn

EDUCATION

University of California San Diego, La Jolla, CA Sep. 2022 - Nov. 2026 (expected)

Ph.D. Candidate in Computer Science and Engineering **Advisor:** Prof. [Zhuowen Tu](#)

- Research interests: Video and Image Generation, Unified Understanding and Generation Model.

University of Science and Technology of China, Hefei, China Sep. 2018 - Jun. 2022

B.S. in Data Science and Big Data Technology, School of the Gifted Young

SELECTED PUBLICATIONS

* indicates equal contribution

- **Zeyuan Chen**, Kai Zhang, Zhuowen Tu, Yuanjun Xiong, “[Soft Tail-dropping for Adaptive Visual Tokenization](#),” in ECCV’26.
- **Zeyuan Chen**, Zhikang Dong, Vidit Goel, Bing Zhou, Zhenzhi Wang, Jian Wang, Dhritiman Sagar, Sergei Korolev, Zhuowen Tu, Hao Chen, “StreamTalker: Streaming Talking Person Audio-Video Joint Generation,” in submission.
- Yucheng Mao*, **Zeyuan Chen***, Xiang Zhang, Divyansh Srivastava, Bingnan Li, Haiyang Xu, Zhuowen Tu, “QuadTok: Quad-Tree Visual Tokenizer For Autoregressive Image Generation,” in submission.
- **Zeyuan Chen**, Xiang Zhang, Haiyang Xu, Jianwen Xie, Zhuowen Tu, “[Central-Peripheral Vision-Inspired Multimodality Models for Spatial Reasoning](#)” in WACV’26.
- **Zeyuan Chen**, Hongyi Xu, Guoxian Song, You Xie, Chenxu Zhang, Xin Chen, Chao Wang, Di Chang, Linjie Luo, “[X-Dancer: Expressive music to human dance video generation](#),” in ICCV’25 (**highlight**).
- Xiang Zhang*, **Zeyuan Chen***, Fangyin Wei, Zhuowen Tu, “[Uni-3D: A Universal Model for Panoptic 3D Scene Reconstruction](#),” in ICCV’23.
- Yuefan Wu*, **Zeyuan Chen***, Shaowei Liu, Zhongzheng Ren, Shenlong Wang, “[CASA: Category-agnostic Skeletal Animal Reconstruction](#),” in NeurIPS’22.
- **Zeyuan Chen**, Yinbo Chen, Jingwen Liu, Xingqian Xu, Vidit Goel, Zhangyang Wang, Humphrey Shi, Xiaolong Wang, “[VideoINR: Learning Video Implicit Neural Representation for Continuous Space-Time Super-Resolution](#),” in CVPR’22.
- **Zeyuan Chen**, Yangchao Wang, Yang Yang, Dong Liu, “[PSD: Principled Synthetic-to-Real Dehazing Guided by Physical Priors](#),” in CVPR’21 (**oral**) .
- **Zeyuan Chen**, Yifan Jiang, Dong Liu, Zhangyang Wang, “[CERL: A Unified Optimization Framework for Light Enhancement with Realistic Noise](#),” in IEEE Transactions on Image Processing.

INDUSTRY EXPERIENCE

Research Intern at *Adobe Firefly* Jun. 2026 - Present

Topic: Unified Image and Video Generative Model **Advisor:** Dr. [Yuanjun Xiong](#) and Dr. [Mingze Xu](#)

- Exploring scalable architecture designs for video in Adobe’s next-generation unified generative model.

Research Intern at *Snap Research* Mar. 2026 - Jun. 2026

Topic: Stream Audio-Video Joint Generation **Advisor:** Dr. [Bing Zhou](#), Dr. [Zhikang Dong](#), Dr. [Hao Chen](#)

- Built a streaming autoregressive talking-person generation system that jointly synthesizes speech and video with low latency, at near-real-time speed, matching the performance of state-of-the-arts.
- Constructed a data curation pipeline for Snap’s internal data that enables multi-person conversational video generation and contributes the training of Snap’s internal audio-video generation model.

Research Intern at *Adobe Firefly* Jun. 2025 - Nov. 2025

Topic: Adaptive Visual Tokenization and Generation **Advisor:** Dr. [Yuanjun Xiong](#) and Dr. [Kai Zhang](#)

- Proposed a novel adaptive 1D visual tokenization model for images and videos that achieves improved reconstruction and autoregressive generation quality over state-of-the-arts.

Research Intern at *TikTok*

Jun. 2024 - Nov. 2024

Topic: Music to Human Dance Video Generation

Advisor: Dr. [Hongyi Xu](#) and Dr. [Linjie Luo](#)

- Developed a pipeline that synthesizes human dance motions and generate dance videos with diffusion model.

Research Intern at *TikTok*

Jun. 2023 - Nov. 2023

Topic: Image-based NeRF Style Editing

Advisor: Dr. [Peng Wang](#) and Dr. [Kejie Li](#)

- Proposed a pipeline to fine-tune camera-aware image-based diffusion models on a single NeRF instance.
- Enabled the fine-tuned model to propagate style from a new reference image across viewpoints, yielding more consistent multi-view synthesis than the pre-trained model.
- Achieved NeRF editing quality comparable to existing baselines while significantly improving efficiency.

ACADEMIC EXPERIENCE

Research Assistant at *UIUC*

Jun. 2021 - Aug. 2022

Topic: Non-Rigid 3D Reconstruction

Advisor: Prof. [Shenlong Wang](#)

- Constructed a large-scale dataset for non-rigid 3d construction; Collected over 200 different types of animal models from a video game and use *Blender* to animate these models with randomized actions.
- Proposed an optimization-based framework for reconstructing 3D animal skeletal shape.

Research Assistant at *UC San Diego*

Apr. 2021 - Nov. 2021

Topic: Video Implicit Neural Representation

Advisor: Prof. [Xiaolong Wang](#)

- Proposed a novel Video Implicit Neural Representation (VideoINR) as a continuous video representation.
- VideoINR represents videos in arbitrary space and time resolution with a single network, generalizing to resolutions and frame rates unseen during training without additional fine-tuning on each video.

Research Assistant at *VITA, UT Austin*

Dec 2020 - Mar 2021

Topic: Real-world low-light Enhancement

Advisor: Prof. [Atlas Wang](#) and Prof. [Dong Liu](#)

- Leveraged an iterative optimization framework that disentangles noise from illumination in low-light images, alternating between off-the-shelf denoising and enhancement models to refine components.
- Introduced a self-supervised training scheme grounded in Retinex Theory to fine-tune denoising models directly on real low-light noise, removing the need for paired clean-noisy supervision.

Research Assistant at *VITA, UT Austin*

Jun. 2020 - Nov. 2020

Topic: Self-supervised Image Dehazing.

Advisor: Prof. [Atlas Wang](#) and Prof. [Dong Liu](#)

- Proposed a model-agnostic synthetic-to-real generalization framework for dehazing, consistently achieving state-of-the-art real-world image dehazing performance across multiple backbone models.
- Explored physical and statistical dehazing priors, leveraging them to guide unsupervised model fine-tuning.

SERVICES

Reviewer: CVPR, ICCV, ECCV, NeurIPS, ICLR, AAAI, WACV, TPAMI, TIP

Teaching Assistant: COGS 118, COGS 181, COGS 185 (UCSD Department of Cognitive Science)

SKILLS

Programming & Tools: Python, PyTorch, HuggingFace, Git, Linux, TensorFlow

Techniques: Diffusion Model Training and Distillation, Autoregressive Model Training, LoRA Finetuning