ZHENXING MI

The Hong Kong University of Science and Technology, Hong Kong
Email: zmiaa@connect.ust.hk **Homepage:** https://mizhenxing.github.io

RESEARCH INTEREST

My research interests focus on multimodal generation and understanding. My current work integrates VLM capabilities into image and video generation models, bridging in-context reasoning to them. I also worked on NeRF, Multi-view Stereo, 3D generation and Mixture of Experts. My long-term goal is to empower 2D & 3D multimodal generation & understanding by large language models and large vision-language models, and exploring any-to-any unified multimodal models that excel at both reasoning and generation.

EDUCATION

- **The Hong Kong University of Science and Technology** Feb. 2021 - Mar. 2026 (Expected)
  Ph.D. candidate in Computer Science and Engineering
  Advisor: Prof. Dan Xu
- **Huazhong University of Science and Technology** Sept. 2017 - June 2020
  M.Sc. in Automation
  Advisor: Prof. Wenbing Tao
- **Huazhong University of Science and Technology** Sept. 2013 - June 2017
  B.Sc. in Control Science and Engineering
  Advisor: Prof. Wenbing Tao

EXPERIENCE

- **Personalized Generative AI, Snap Research.** July. 2024 - Oct 2024
  Research Intern, Multimodal generation
  Mentor: Kuan-Chieh Wang, Guocheng Qian and Kfir Aberman
- **Applied Research Center (ARC), PCG, Tencent** Jan. 2024 - May 2024
  Research Intern, 3D generation
  Mentor: Xintao Wang, Weihao Cheng

PUBLICATIONS

- **Zhenxing Mi**, Kuan-Chieh Wang, Guocheng Qian, Hanrong Ye, Runtao Liu, Sergey Tulyakov, Kfir Aberman, Dan Xu. "I Think, Therefore I Diffuse: Enabling Multimodal In-Context Reasoning in Diffusion Models.", ***ICML*** 2025.
- **Zhenxing Mi**, and Dan Xu. "LeCO^2-NeRF: Learning Continuous and Compact Large-Scale Occupancy for Urban Scenes.", arXiv:2411.11374, ***Technical report***.
- **Zhenxing Mi**, and Dan Xu. "Switch-NeRF: Learning Scene Decomposition with Mixture of Experts for Large-scale Neural Radiance Fields." ***ICLR*** 2023.
- **Zhenxing Mi**, Di Chang, and Dan Xu. "Generalized Binary Search Network for Highly-Efficient Multi-View Stereo." ***CVPR*** 2022.
- Ganzhangqin Yuan*, Qiancheng Fu*, **Zhenxing Mi***, Yiming Luo*, and Wenbing Tao."SSRNet: Scalable 3D Surface Reconstruction Network." ***IEEE TVCG*** 2022.
- Yiming, Luo*, **Zhenxing Mi***, and Wenbing Tao. "DeepDT: Learning Geometry From Delaunay Triangulation for Surface Reconstruction." ***AAAI*** 2021.
- **Zhenxing Mi***, Yiming Luo*, and Wenbing Tao. "SSRNet: Scalable 3D Surface Reconstruction Network." ***CVPR*** 2020.