Zelin Gao

(+86)188-4314-8545 | jamesgzl@zju.edu.cn / zelingao98@gmail.com | https://jameskuma.github.io

Education

Zhejiang University Hangzhou, Zhejiang Province, China
M.Eng in Control Science and Engineering (GPA: 3.78/4.00) Sep. 2021 - Mar. 2024
- Intelligent and Autonomous System Lab | Co-Advisors: Prof. Yu Zhang & Prof. Weichen Dai
- {2021-2022} Academic Scholarship & First Prize Scholarship

Jilin University Changchun, Jilin Province, China
B.Eng in Instrument Science and Engineering (GPA: 3.66/4.00, Rank: 2/129) Sep. 2017 - Jun. 2021
- Geophysical Exploration Research Center | Advisor: Prof. Chuandong Jiang
- {2020-2021} Outstanding Graduates & Outstanding Graduation Projects
- {2017-2021} **National Scholarship** & Xin Wang Scholarship & First Prize Scholarship
- {2018-2019} First Prize in Mathematical Competition & Second Prize in National Mathematical Modeling Competition

Research Intership

Xiaolong Wang's Lab, University of California, San Diego & NVIDIA Research Apr. 2024 - Nov. 2024
- Research Supervisor: Prof. Xiaolong Wang & Dr. Jiarui Xu & Dr. Shalini De Mello & Dr. Seonwook Park
- Research Result:
  1. One paper has been submitted to CVPR2025 (under review).

Berkeley Artificial Intelligence Research Lab, University of California, Berkeley Apr. 2023 - Nov. 2023
- Research Supervisor: Prof. Angjoo Kanazawa & Dr. Yutong Bai
- Research Result:
  1. One paper has been submitted to CVPR2024 (accepted as spotlight).

Research Experience

CVPR2025 (Under Review & 1st Author) AHA: Expressive Human Animation Driven by Audio
- Research Motivation: Talking upper-body human animation with only audio-driven conditions.
- Research Method:
  1. We propose Hybrid Encoder, which serves as a potential motion dictionary learned from the training dataset.
  2. We propose a talking human dataset, TalkingFOX, with various talking upper bodies and their SMPL-X annotations.

AAAI2025 Adaptive Wavelet-Positional Encoding for High-Frequency Information Learning in INR
- Research Motivation: Represent scene content under various frequency distributions.
- Research Method:
  1. Adaptive Wavelet-Positional Encoding, taking inspiration from APE-BARF.
  2. High-Frequency Perception (HFP) method for decomposing image into different frequency components.

CVPR2024 (Spotlight) SAP3D: More View in 2D, More Perceive in 3D
- Research Motivation: 2D view-consistent generation and 3D reconstruction with arbitrary number of input images.
- Research Method:
  1. Jointly fine-tuning pose and Zero 1-to-3 for object-driven generation, enhancing geometry information of latent space.
  2. Stochastic condition images and stochastic SDS are thus introduced in this paper for generation and reconstruction.

ARXIV2024 HG^3-NeRF for Sparse View Inputs
- Research Motivation: The optimization problem of reconstructing neural radiance fields from sparse view inputs.
- Research Method:
  1. Hierarchical Geometric Guidance (HGG), incorporating depth to the scene representations with local-to-global regions.
  2. Hierarchical Semantic Guidance (HSG), incorporating coarse-to-fine semantic content to the scene representations.

ICCV2023 (Poster) Adaptive Positional Encoding for BA-NeRF
- Research Motivation: The joint optimization problem of reconstructing neural radiance fields from unknown camera parameters.
- Research Method:
  1. Adaptive Positional Encoding with its frequency diversity loss, taking inspiration from Fourier series regression.
  2. Brand new implicit network, consisting of PMLP for fine-grain gradient propagation.

IROS2022 (Oral) Thermal-Inertial SLAM for the Environments with Challenging Illumination
- Research Motivation: The frame-based association of thermal images in the simultaneous localization and mapping.
- Research Method:
  1. The SVD-based image processing method, establishing accurate data association by singular value computation.
  2. Real-Time Optical Flow Tracking on RAFT, a light-weight cost-volume for tracking dense optical flow.

Project Experience

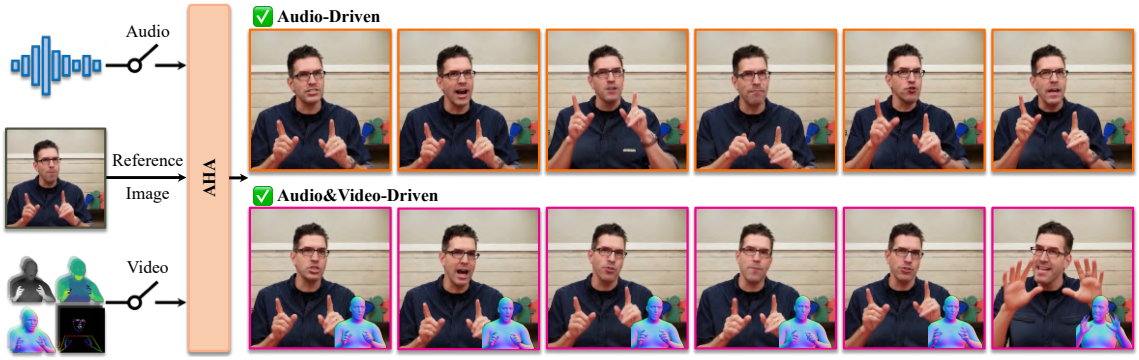
D^4-Dreamer: Text to Non-Rigid Scene Generation May. 2023 - Nov. 2023
- Project Motivation: 4D (XYZ + Temporal) radiance fields generation of fine details.
- Project Method:
  1. We introduce scene content and motion guidance to distill prior from diffusion model to generate non-rigid scenes.
  2. VSD-T of scene motion guidance is proposed to realize smooth motion and further finetune video diffusion model.

APPENDIX/SUPPLYMENTARY

RESEARCH EXPERIENCE

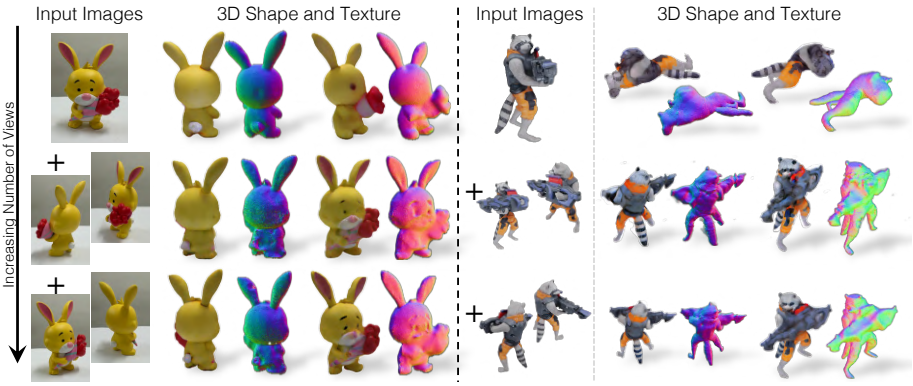
CVPR2025 (Under Review) AHA: Expressive Human Animation Driven by Audio

- Research Highlight:



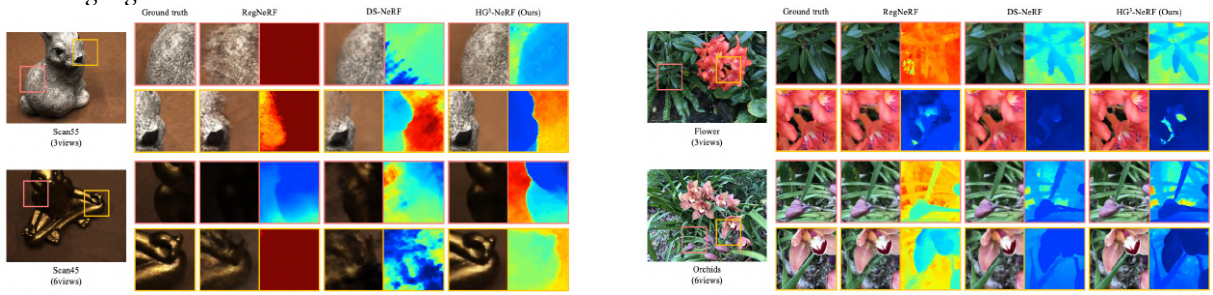
CVPR2024 (Spotlight) SAP3D: More View in 2D, More Perceive in 3D

- Research Highlight:



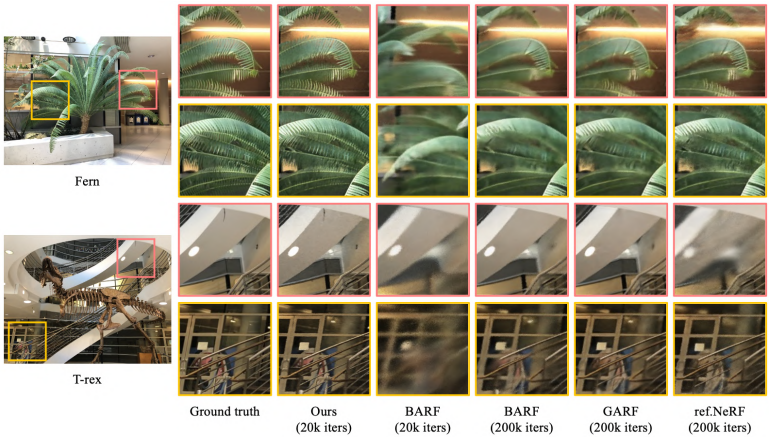
ARXIV2024 HG^3-NeRF for Sparse View Inputs

- Research Highlight:



ICCV2023 (Poster) Adaptive Positional Encoding for BA-NeRF

- Research Highlight:



IROS2022 (Oral) Thermal-Inertial SLAM for the Environments with Challenging Illumination

- Research Highlight: