

Shu Pu

pushuabc@gmail.com — Homepage — Scholar — Github — Last Updated: Jul 22th 2026

EDUCATION

University of California, Santa Cruz CSE Sep 2026–Jun 2031 (Expected)

Degree: Doctor of Philosophy

Publication

- *Memorization in 3D Shape Generation: An Empirical Study* [PDF]
Shu Pu, Boya Zeng, Kaichen Zhou, Mengyu Wang, Zhuang Liu.
Accepted at **CVPR'26 Findings**. (Preliminary version accepted at **ICCV'25 CEGIS Workshop**).
- *Judge Anything: MLLM as a Judge Across Any Modality* [PDF]
Shu Pu*, Yaochen Wang*, Dongping Chen, Yuhang Chen, Guohao Wang, Qi Qin, Zhongyi Zhang, Zhiyuan Zhang, Zetong Zhou, Shuang Gong, Yi Gui, Yao Wan, Philip S. Yu.
Accepted at **KDD'25 DB Track (Oral)**
- *ISG-Bench: Interleaved Scene Graph for Text-to-Image Generation Assessment* [PDF]
Dongping Chen*, Ruoxi Chen*, **Shu Pu***, Zhaoyi Liu*, Yanru Wu*, Caixi Chen*, Benlin Liu*, Yue Huang, Yao Wan, Pan Zhou, Ranjay Krishna
Accepted at **ICLR'25 (Spotlight 5.1%)**.
- *Thinking Before Looking: Improving Multimodal LLM Reasoning via Mitigating Visual Hallucination* [PDF]
Haojie Zheng *, Tianyang Xu *, Hanchi Sun, **Shu Pu**, Ruoxi Chen, Lichao Sun.

Experience

Research Internship in Princeton University, supervised by Professor Zhuang Liu

Full-time, Remote Research Intern

March 2025 – Present

- Research on 3D generative models and memorization behavior. We propose an analysis that introduces an evaluation framework to measure memorization and uses it to conduct controlled experiments on data and modeling factors that influence and mitigate memorization in 3D generative models.
- Research on building an open-source state-of-the-art image generation model.

Research Internship in Huazhong University of Science and Technology, supervised by Professor Yao Wan

Full-time, Research Intern

June 2024 – March, 2025

- Research on unified generation models and their interleaved generation capabilities. We propose a benchmark, **ISG-Bench**, which focuses on evaluating a model's interleaved image-text generation capabilities and introduce a new evaluation metric for interleaved context based on a scene graph.
- Research on multimodal LLMs for human-aligned evaluation of any-to-any generative tasks. We propose **Judge Anything**, a benchmark with fine-grained human analysis, and an *Omni-Arena* for the development of reward models and any-to-any generative models.

Research Internship in Lehigh University, supervised by Professor Lichao Sun

Part-time, Research Intern

June 2024 – August, 2024

- Research on multi-agent systems. We propose **VIC** (Visual Inference Chain), a novel framework requiring multimodal LLMs to think before processing visual content, effectively mitigating visual hallucinations.

Personal Projects

3D Physical World Perception and Generation Computer Vision, Course Project

- Conducted a self-directed survey on 3D perception and generation. *Survey link* [PDF]
- Explored 3D Gaussian Splatting by studying its codebase and learning CUDA to understand the technique in depth.

Flash Attention and GPU Acceleration

- Implement the assignment about Benchmarking GPUs, Flash Attention and Distributed training in online course CS336, Stanford University. [Code]

Skills

Language: Mandarin (native), English (TOEFL 109)

Research Abilities: Proficient in Python and C++; Knowledge and experience in Multi-modal Perception and Generation, Diffusion models and 3D Vision. Basic knowledge about GPU architecture and acceleration.

Research Interest: Besides unified model and diffusion, I'm also a motivated Computer Graphics learner, especially in 3D vision and geometry. My long-term goal is to help develop a world model which can revolutionize the film, game and animation industry. I believe integrating AI into everyday life is a more worthwhile pursuit.

Others

Research is only one part of who I am.

Hobbies: Football, Swimming, R&B.