

Zhanpeng Luo

✉ ZhanpengLuo@pitt.edu | ↗ (+1) 412-224-8668
/github | 🌐 Personal Website

EDUCATION

University of Pittsburgh

Bachelor of Computer Sciene GPA:3.90

Pittsburgh, U.S.

Aug. 24 - Jun. 26

Sichuan University

Bachelor of Computer Sciene GPA:3.94

Chengdu, China

Aug. 22 - Jun. 24

PUBLICATIONS & PATENTS

Zhanpeng Luo, Haoxi Ran, Li Lu. *Instant4D: 4D Gaussian Splatting in Minutes*. Accepted by **NeurIPS 2025**. [\[Project Page\]](#)

Zhanpeng Luo, Ce Zhang, ..., Guanya Shi, Katia Sycara, Yaqi Xie. *pySpatial: Generating 3D Visual Programs for Zero-Shot Spatial Reasoning*. Accepted by **ICLR 2026**. [\[Project Page\]](#)

PREPRINTS

Zhanpeng Luo, Linna Wang, Guangwu Qian, Li Lu. *Imagine with the Teacher: Complete Shape in a Multi-View Distillation Way*. [\[arXiv\]](#)

Qianwei Wang*, Bowen Li*, **Zhanpeng Luo**, Yifan Xu, Alexander Gray, Tom Silver, Sebastian Scherer, Katia Sycara, Yaqi Xie. In submission to *IEEE Transactions on Robotics (T-RO)*.

EXPERIENCE & AWARD

Academic:

Poster Presentation:

Robotics Institute Summer Scholar ([RISS 2025](#))

Pittsburgh, Aug. 25

39th Annual Conference on Neural Information Processing Systems

San Diego, Dec. 25

Conference Reviewer: NeurIPS 25 Workshop, ICLR 26;

Award:

39th NeurIPS Scholar Award

Oct. 25

Third Prize of China Nation Olympic Biology Competition (Jiangsu Province)

Aug. 21

Dean List: 6 semester/6

Internship:

25 Summer, RISS, on emboidied agent, supervised by [Katia Sycara](#)

24 Summer, Research Internship, on real-time graphics, supervised by [Bernhard Kerbl](#)

23 Summer, Research Internship, on medical image analysis, West China Hospital

ACTIVE RESEARCH INTEREST

Post-Training for Generative View Synthesis Model:

I am interested in 3D asset generation, feed-forward reconstructions, and world modeling. I am now working to improve modeling the dynamic and complex worlds. Existing 3D-based pipeline can not support this goal in a scalable way based on several observation:

1) 3D autoregressive approach *e.g.* (LVSM, Rayzer, AnySplat) interpolate well between input frames but often fail to capture true 3D structure, leading to 2D-like artifacts during scene exploration.

2) Dynamic reconstruction datasets are costly and scarce, while ordinary video data & dataset are abundant and emerging.

3) All current dynamic representation methods *e.g.* (Dynamic Nerf, 4DGS, MoSca, Shape-of-Motion) stuggle to represent a long-term scene in a scalable manner.

Therefore I am thinking of a fully latent, end-to-end, diffusion model that directly regress RGB to reconstruct dynamic, complex worlds, and also does generative view synthesis. Considering of resources' constraint, I plan to work on the post-training of video diffusion model based on several points:

- Leveraging the diffusion prior learned from large-scale video data. RecamMaster shows us that with light-weight camera embedding layer, a text-to-video model is able to learn the world's structure.

- With 3D reconstructor as verifier (in GRPO, the reward function), we are distillatiing the 3D prior from *e.g.* (Dust3R, VGGT, π^3). Also it acts a way to moderate reconstruction and generation.
- With proper verifiable reward design, our goal is to enforce 3D consistency, camera controllability and overall visual quality.

SKILL

PyTorch, CUDA Programming, JAX, Swift