Zhanpeng Luo

ZhanpengLuo@pitt.edu | (+1) 412-224-8668
Github | Personal Website

EDUCATION

University of Pittsburgh — Pittsburgh, PA
Bachelor of Computer Sciene — Estimate to graduate at 26 Spring

PUBLICATIONS

Zhanpeng Luo, Haoxi Ran, Li Lu. Instant4D: 4D Gaussian Splatting in Minutes. *Accpted by* ***NeurIPS 25***, [Project Page] [Code] [NeurIPS 25]
Zhanpeng Luo, Ce Zhang … Guanya Shi, Katia Sycara, Yaqi Xie. pySpatial: Generating 3D Visual Programs for Zero-Shot Spatial Reasoning. *Under Review* [Paper]
Linna Wang, Leyi Zhao, **Zhanpeng Luo**, Xuan Wang, Ziliang Feng, and Li Lu "CARE-30: A Causally Driven Multi-Modal Model for Enhanced 30-Day ICU Readmission Predictions", **IEEE, BIBM 23**
Changjing Song, **Zhanpeng Luo**, Li Lu, and Qian Su " MLNet: Enhancing Joint Predictive Modeling of Chronic DiseasesUsing Deep Learning, **IEEE BIBM 23**
Lixian Zhu, **Zhanpeng Luo**, Xiaodong Yang, Bridgeless Power Factor Correction Circuit and Its Control Method. **China Patent** CN 119696357B

EXPERIENCE

Academic:
Poster Presentation: Robotics Institute Summer Scholar (RISS 2025), CMU;
Conference Reviewer: NeurIPS 25 Workshop, ICLR 26, CVPR 26;
Dean List: 6 semester; GPA Major: 4.0 / 4;
Award: NeurIPS Scholar Award;
Internship:
25 Summer, RISS, Supervised by Katia Sycara
24 Summer, Research Internship, Supervised by Bernhard Kerbl
23 Summer, Research Internship, 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.