

周峰

中国, 北京 | 1999.12.25 | 籍贯: 浙江湖州
zhoufeng@bupt.edu.cn | +86-18767299088 (微信)



教育背景

- 北京邮电大学 | 博士研究生 (保送直博) 2022年9月 - 2027年6月 (预计)
- 智能工程与自动化学院, 控制科学与工程专业。
 - 研究方向: 可控视觉生成与三维世界的视觉理解; 导师: 尹建芹教授 (BUPT-COST 实验室)。
- 北京邮电大学 | 本科 2018年9月 - 2022年6月
- 物联网工程专业。GPA: 3.7/4.0 (获保研推免资格)。

实习经历

地平线 (Horizon Robotics) | 三维重建基础模型专项人才计划实习 2025年11月 - 至今

[1] Research | 前馈三维重建基础模型架构探索

- ATSR: 稀疏 Attention 架构** | 将 dense global attention 改造为视图数自适应的 block-level top- k sparse attention, 减少冗余跨视图通信; 在评测集上平均 Pose AUC@30 从 0.879 提升至 0.891, 300 views 下实现 $8.45\times$ 加速。该方法已成功适配业务数据的训练并代替原始 VGGT。[投稿 Nips 26]。
- GeoWeave: 有选择的跨视图通信架构** | 提出基于打分器的 cross-view attention 机制以实现跨视图 token 间有选择的通信, 提升弱重叠、干扰视图下的位姿与点云重建稳定性。[投稿 SIGGRAPH Asia 26]。
- Hybrid-VGGT: 高分辨率 VGGT 架构** | 设计低分辨率全局主干 + 高分辨率 HDE 双分支架构, 通过多尺度局部卷积、线性注意力与层级 DPT 融合, 提高模型的深度、点云的精度表现。[投稿 Nips 26]。

[2] Engineering | 三维重建模型数据, 训练与业务目标支持

- 搭建 mix-17 个公开数据集的训练数据集与统一评测脚本; 在业务数据上适配 pi3 模型的训练; 在业务数据训练集上加入高质量仿真数据, 提高细节物体的重建精度。

实验室研究及学术成果

1. 三维视觉: 三维重建与场景理解

3D Reconstruction, Sparse-view Reconstruction

- 稀疏输入下的三维重建:** 面向 sparse-view 3DGS 在少视角下易过拟合的问题, 系统验证初始化质量是性能上界的主导因素; 提出低频感知 SfM、3DGS self-initialization 与点云一致性过滤流程, 提升稀疏视角下的新视角合成稳定性与几何完整性。[TVCG 大修中]
- 复杂场景中的目标级理解:** 面向大规模三维场景中指定物体难以独立分割、提取与补全的问题, 提出基于 2D Gaussian Splatting 的目标一致分割框架, 并引入个性化 diffusion prior 修补遮挡与不可见区域, 实现目标级 mesh 提取。[TCSVT 2025]
- 单目人体三维重建:** 面向 2D pose lifting 中深度歧义严重、图像信息难以稳定利用的问题, 设计 pose-guided attention 与自适应特征筛选机制, 引入人体相关图像线索并抑制背景过拟合, 在 Human3.6M 与 MPI-INF-3DHP 上取得 SOTA 表现。[AAAI 2024]

2. 可控视觉生成: 分辨率外推机制及可控生成机制

Resolution Extrapolation, Controllable LDM

- 高分辨率生成的机制分析:** 面向 U-Net diffusion 在高分辨率推理时出现布局重复与结构崩坏的问题, 揭示卷积 zero-padding 承担隐式位置编码作用, 并提出 Progressive Boundary Complement, 通过虚拟边界补充位置信息, 实现免训练高分辨率生成。[AAAI 2026 Oral]

- **高分辨率 DiT 外推**: 面向 DiT / FLUX 高分辨率推理中的布局塌缩与纹理模糊问题, 分析位置编码与注意力感受野在全局布局、局部细节中的作用差异, 提出双分支注意力重构与频域融合机制, 激活模型内生高分辨率能力。[CVPR 2026]
- **无标签领域生成适配**: 面向特定领域缺少图文标注、直接微调易损伤原模型可控能力的问题, 提出 guidance-decoupled prior preservation, 将领域知识学习与文本/空间控制能力解耦, 使 diffusion model 仅依赖图像数据完成领域适配。[CVPR 2025]
- **可控生成评价与理论边界**: 面向 Latent Diffusion 中 Autoencoder 评价偏向 gFID、忽略实例级条件保持的问题, 分析其导致的 condition drift 与可控生成上限偏差, 提出更关注实例级重建与条件对齐的评价视角。[TPAMI 在投]
- **可控生成方法体系梳理**: 系统总结 text-to-image diffusion 中条件注入、结构控制、编辑控制与个性化控制等技术路线, 归纳可控生成从 prompt-level 到 spatial-level、instance-level 控制的演进脉络。[TPAMI 2025]

技术能力

- **视觉基础与表征学习**: 掌握成像模型、滤波/边缘/特征点、匹配、光流/跟踪、分类/检测/分割/姿态估计; 理解手工特征、CNN/ViT、自监督与视觉-语言预训练的表征学习范式。
- **三维视觉与场景重建**: 理解相机模型、极线几何、三角化、PnP/BA、SfM/MVS/SLAM; 掌握深度/位姿估计、NeRF/3DGS 场景表示, 以及前馈式 3D foundation models 的跨视角融合建模。
- **扩散生成与可控生成**: 理解 diffusion/score-based generative modeling 的训练与采样机制, 包括 DDPM/DDIM、SDE/ODE sampler、CFG、rectified flow/flow matching; 掌握潜空间扩散、Transformer-based diffusion/flow、图像/视频生成、条件控制与主流生成架构。
- **通用大模型与 AI 工作流**: 了解 LLM/VLM、RAG/Agent、工具调用与多模态交互的应用范式; 能够将大模型作为协作式工具接入日常科研与工程流程, 搭建面向调研、开发、实验迭代与结果复盘的 AI-assisted workflow。

主要论文列表

已发表:

- | | |
|---|-----------------|
| [1] Exploring Position Encoding in Diffusion U-Net for Training-free High-resolution Image Generation | AAAI 2026, Oral |
| Feng Zhou*, Pu Cao*, Yiyang Ma, Lu Yang, Yonghao Dang, Jianqin Yin | |
| [2] Lifting by image – Leveraging Image Cues for Accurate 3D Human Pose Estimation | AAAI 2024 |
| Feng Zhou, Jianqin Yin, Peiyang Li | |
| [3] Image is All You Need to Empower Large-scale Diffusion Models for In-Domain Generation | CVPR 2025 |
| Pu Cao*, Feng Zhou*, Lu Yang, Tianrui Huang, Qing Song | |
| [4] ResDiT: Evoking the Intrinsic Resolution Scalability in Diffusion Transformers | CVPR 2026 |
| Yiyang Ma*, Feng Zhou*, Pu Cao, Yonghao Dang, Jianqin Yin | |
| [5] OMEGAS: Object Mesh Extraction from Large Scenes Guided by Gaussian Segmentation | TCSVT 2025 |
| Lizhi Wang*, Feng Zhou*, Bo Yu, Pu Cao, Jianqin Yin | |
| [6] Controllable Generation with Text-to-Image Diffusion Models: A Survey | TPAMI 2025 |
| Pu Cao, Feng Zhou, Qing Song, Lu Yang | |

审稿中:

- | | |
|--|--------------------|
| [7] Initialize to Generalize: A Stronger Initialization Pipeline for Sparse-View 3DGS | TVCG 大修 |
| Feng Zhou*, Wenkai Guo*, Pu Cao, Zhicheng Zhang, Jianqin Yin | |
| [8] GeoWeave: Learning Reliable Cross-View Dependencies for Feed-Forward 3D Reconstruction | SIGGRAPH Asia 2026 |
| Feng Zhou, Qingfeng Li, Jianqin Yin, Weiqiang Ren, Qian Zhang | |
| [9] ATSR: Adaptive Top-k Sparse Routing for Scalable Multi-View Geometry Transformers | NeurIPS 2026 |
| Qingfeng Li*, Feng Zhou*, Weiqiang Ren, Qian Zhang | |
| [10] A Tilted Seesaw: Revisiting Autoencoder Trade-off for Controllable Diffusion | TPAMI 在投 |
| Pu Cao, Feng Zhou, Yiyang Ma, Xuedan Yin, Qing Song, Lu Yang | |