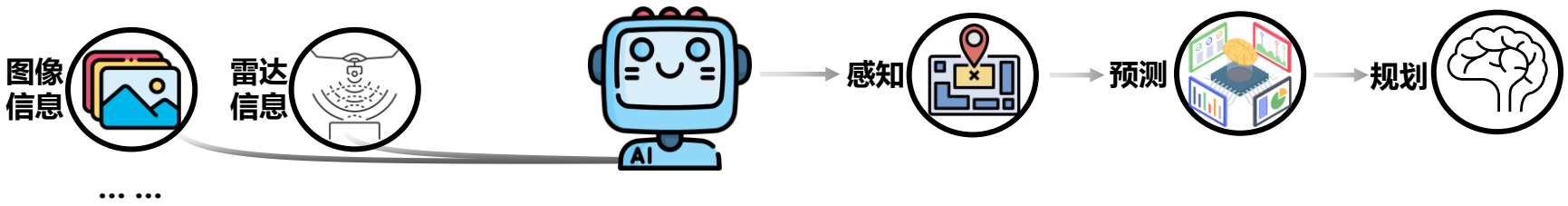
浙江大学
ZHEJIANG UNIVERSITY

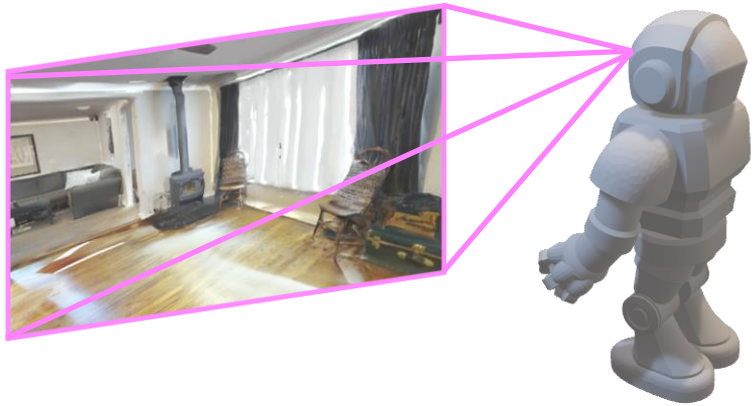
基于神经表征的三维感知、预测与规划

崔兆鹏
浙江大学

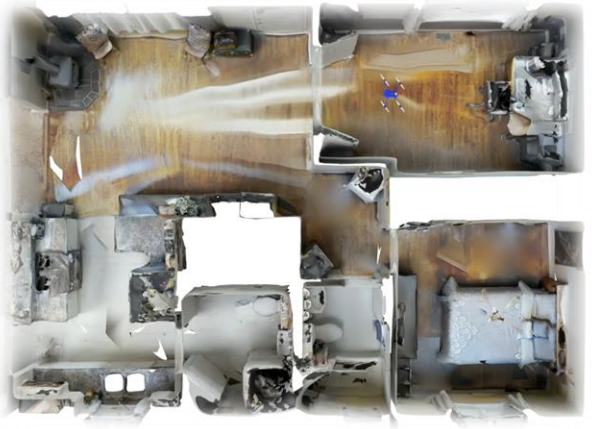


智能无人系统

重建感知



未来预测



精准规划

状态控制

三维神经表征 VS 传统显式表征

传统显式表征

三维神经表征

神经表征是否可以提高三维感知、预测和规划的性能?

常见形式：点云、网格等

简单、高效、可解释性强

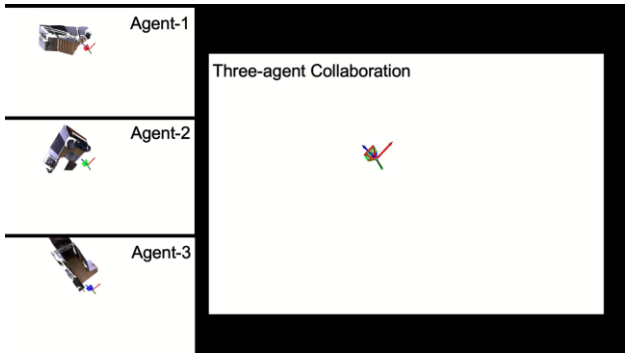
表达能力弱、渲染真实感不足

常见形式：NeRF、3DGS

表达能力强、渲染真实感足

效率低、计算量大

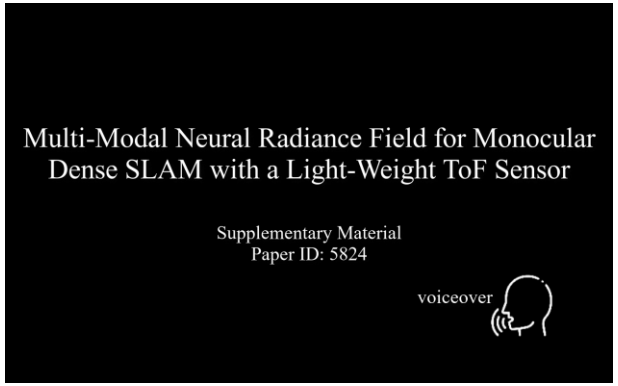
三维神经感知、预测与规划



NICER-SLAM
Neural Implicit Scene Encoding for RGB SLAM

NICE-SLAM
[CVPR'22]

Multi-Modal NeRF SLAM
[ICCV'23]

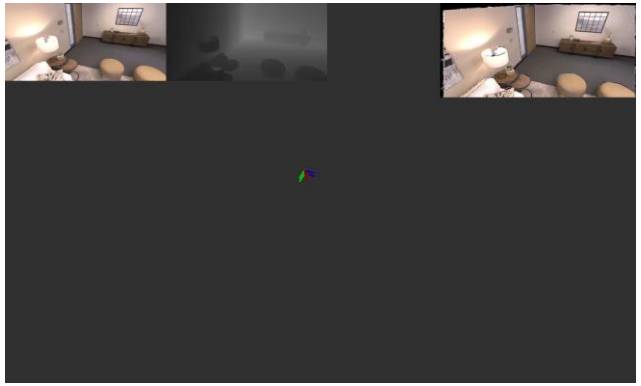


CP-SLAM
[NeurIPS'23]

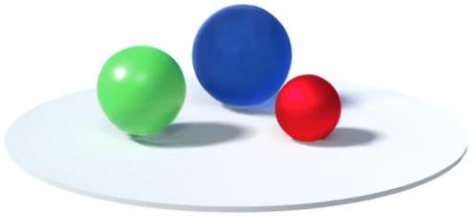
PNeRFLoc
[AAAI'24]

NICER-SLAM
[3DV'24]

CG-SLAM
[ECCV'24]



三维神经感知、预测与规划



GaussianPrediction
[SIG'24]

DreamSpace

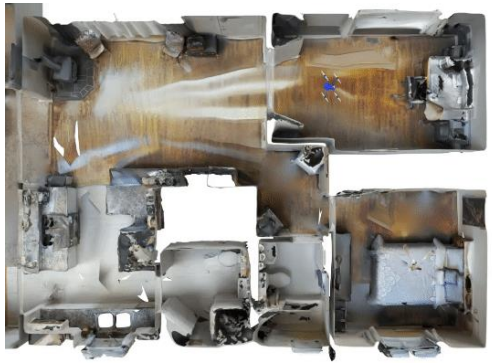
Dreaming Your Room Space with Text-Driven Panoramic Texture Propagation

Bangbang Yang[1] Wenqi Dong[2] Lin Ma[1] Wenbo Hu[1]
Xiao Liu[1] Zhaopeng Cui[2] Yuewen Ma[1†]
[1] ByteDance [2] State Key Lab of CAD&CG, Zhejiang University

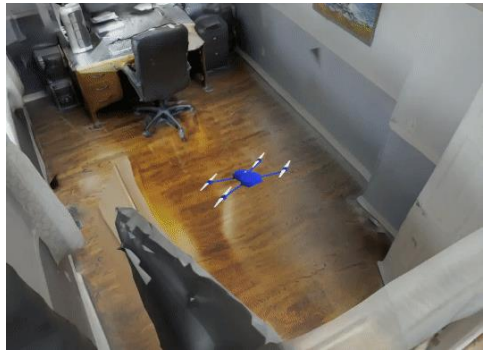
ByteDance Voiceover

DreamSpace
[VR'24]

PCPlanner
[SIGA'24]

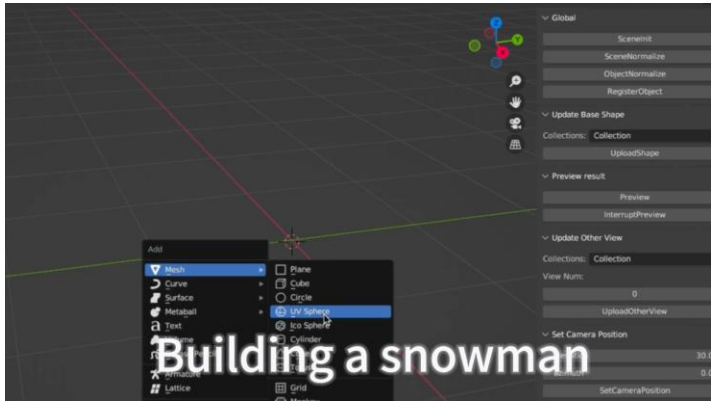


Top view

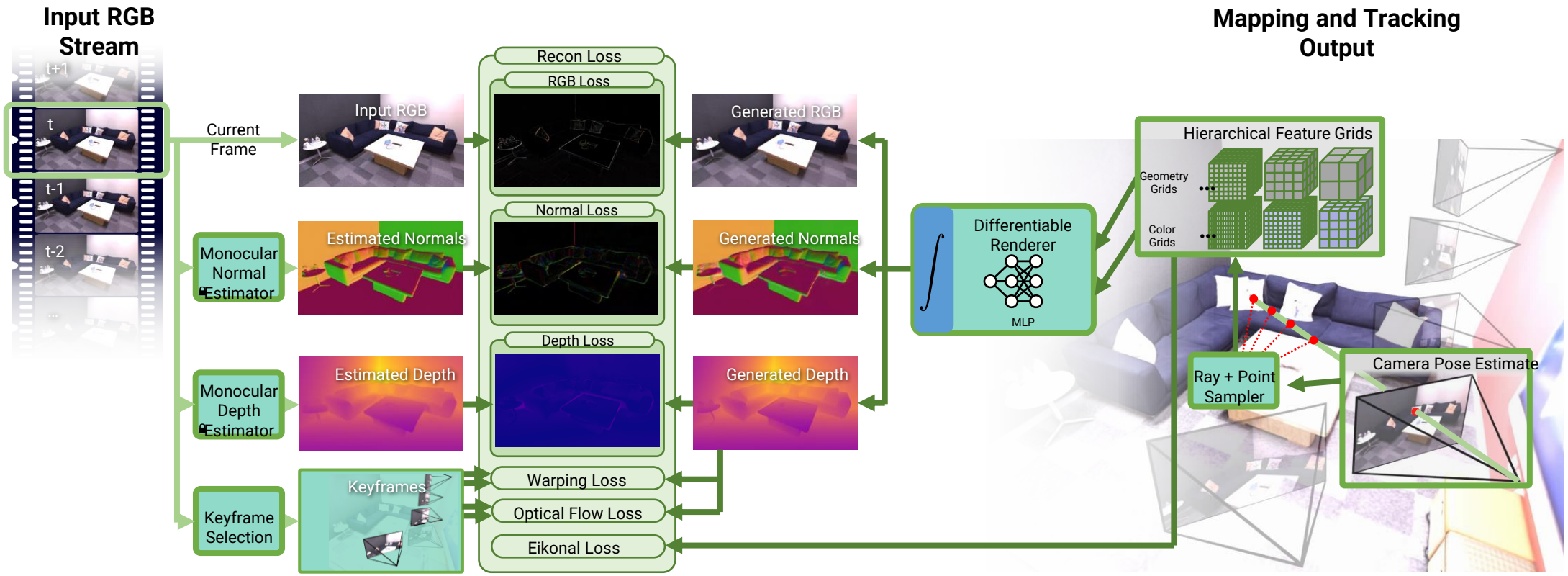


Camera view

Coin3D
[SIG'24]



三维感知：NICER-SLAM



Zhu et al. NICER-SLAM: Neural Implicit Scene Encoding for RGB SLAM. 3DV 2024.

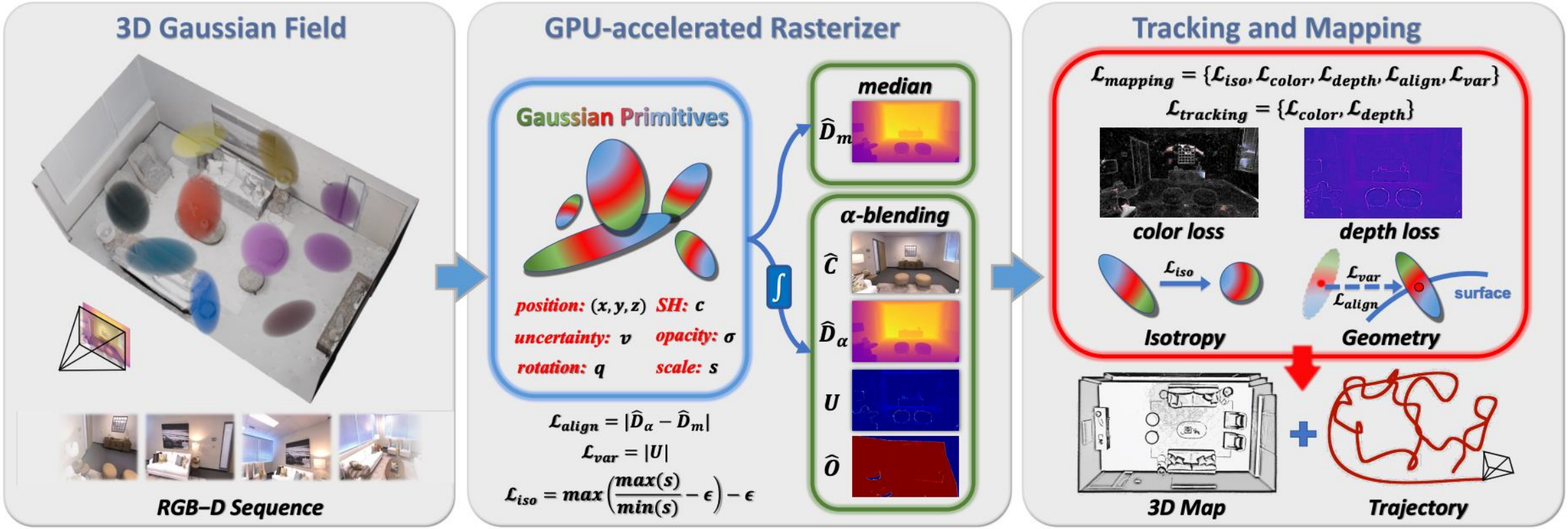
- **RGB-Only：实现了仅依靠单目RGB相机的Neural SLAM框架**
- **高质量稠密重建：将单目几何先验与帧间几何约束融入优化框架，提升系统的稳定性与精度**
- **3DV 2024 Best Paper Honorable Mention （唯一）**

三维感知：NICER-SLAM

真实场景

NICER-SLAM (Ours)
RGB Input

三维感知：CG-SLAM

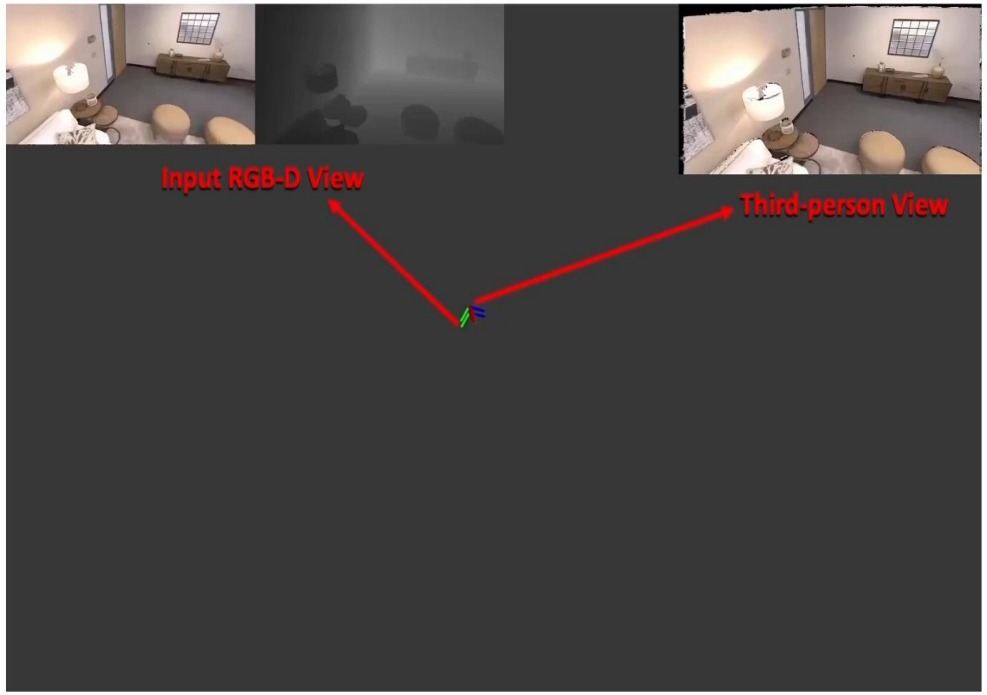


Hu et al. CG-SLAM: Efficient Dense RGB-D SLAM in a Consistent Uncertainty-aware 3D Gaussian Field. ECCV 2024.

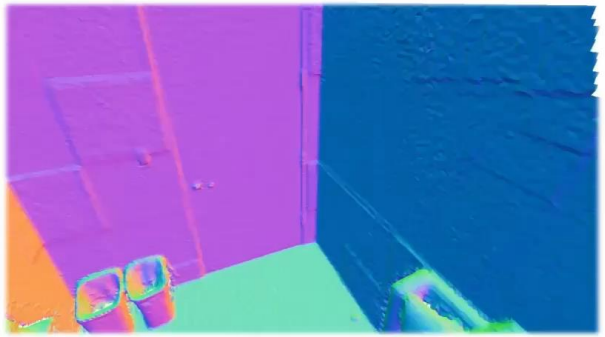
◆ **不确定性建模：** CG-SLAM提出了一种新颖的3DGS不确定性模型，实现了高质量的定位与建图

◆ **实时稠密SLAM：** 挖掘3DGS高效的渲染优化能力，3090显卡实现15Hz以上的同步定位与建图速率

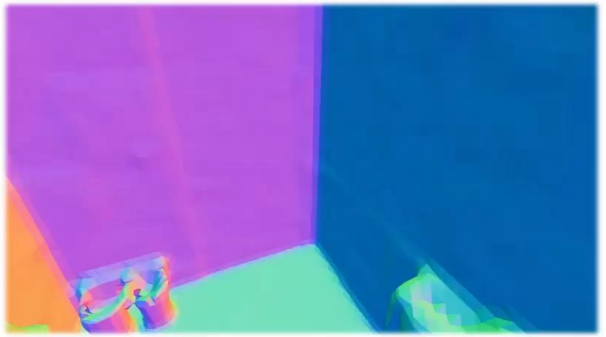
三维感知：CG-SLAM



Replica Office1



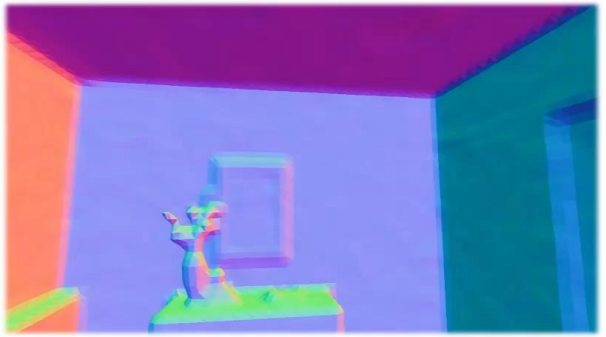
Ours



Co-SLAM

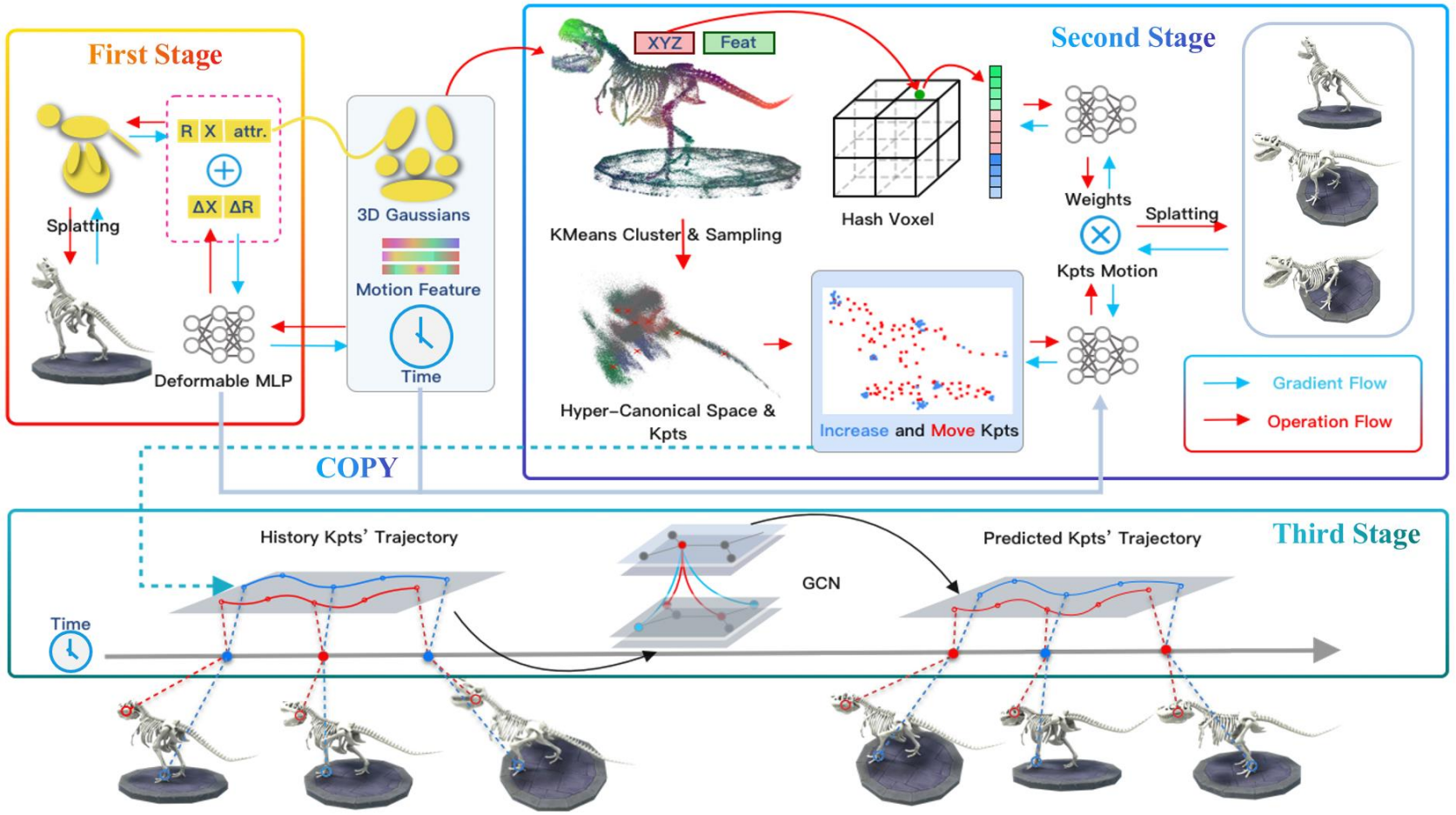
Replica Room1

Ours



Co-SLAM

三维预测：GaussianPrediction



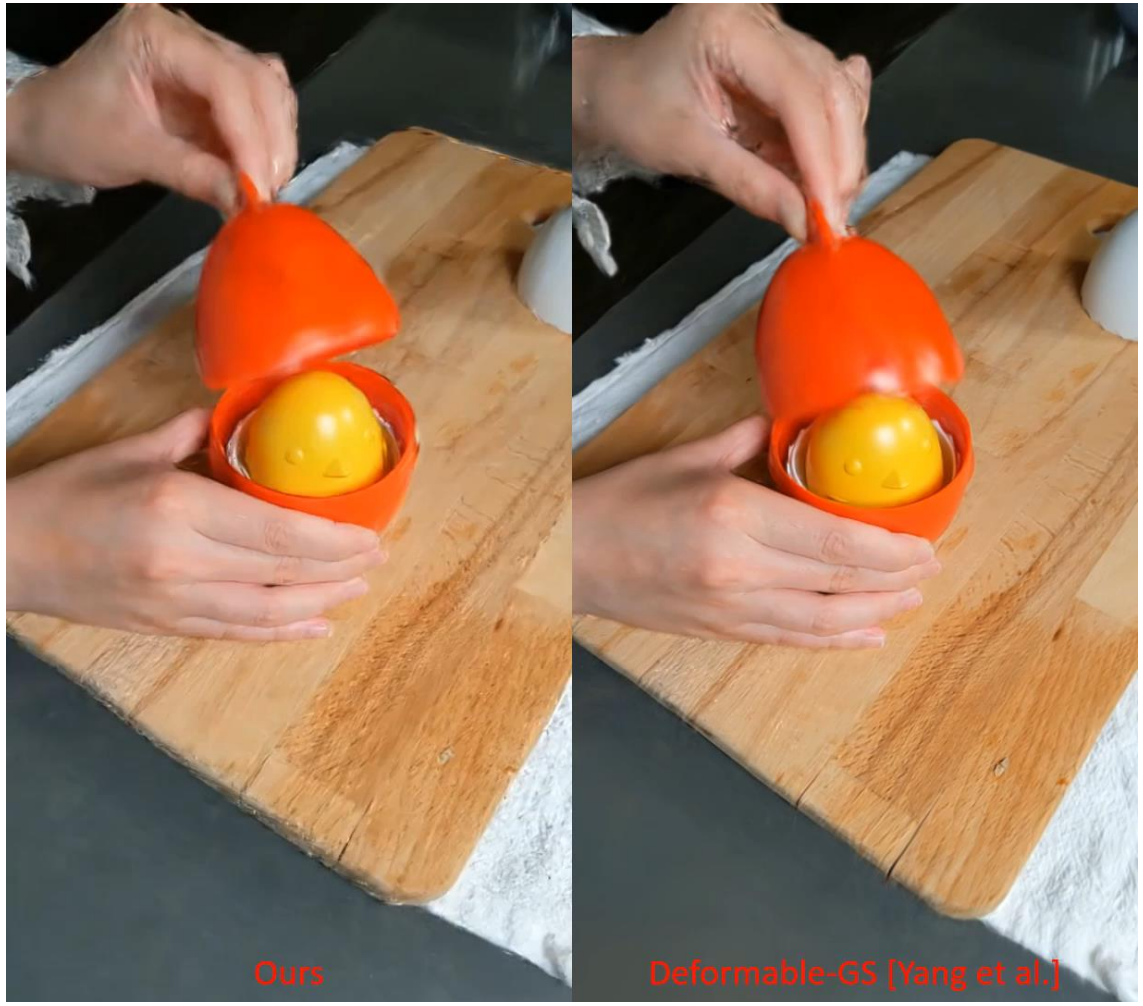
Zhao et al. GaussianPrediction: Dynamic 3D Gaussian Prediction for Motion Extrapolation and Free View Synthesis. SIGGRAPH 2024.

- **同心运动蒸馏：** 降低场景运动预测的复杂度，结合GCN实现了精准的场景动态预测
- **建模预测统一框架：** 提出了基于3DGS的重建与预测统一框架，在动态重建与预测上都达到了SOTA

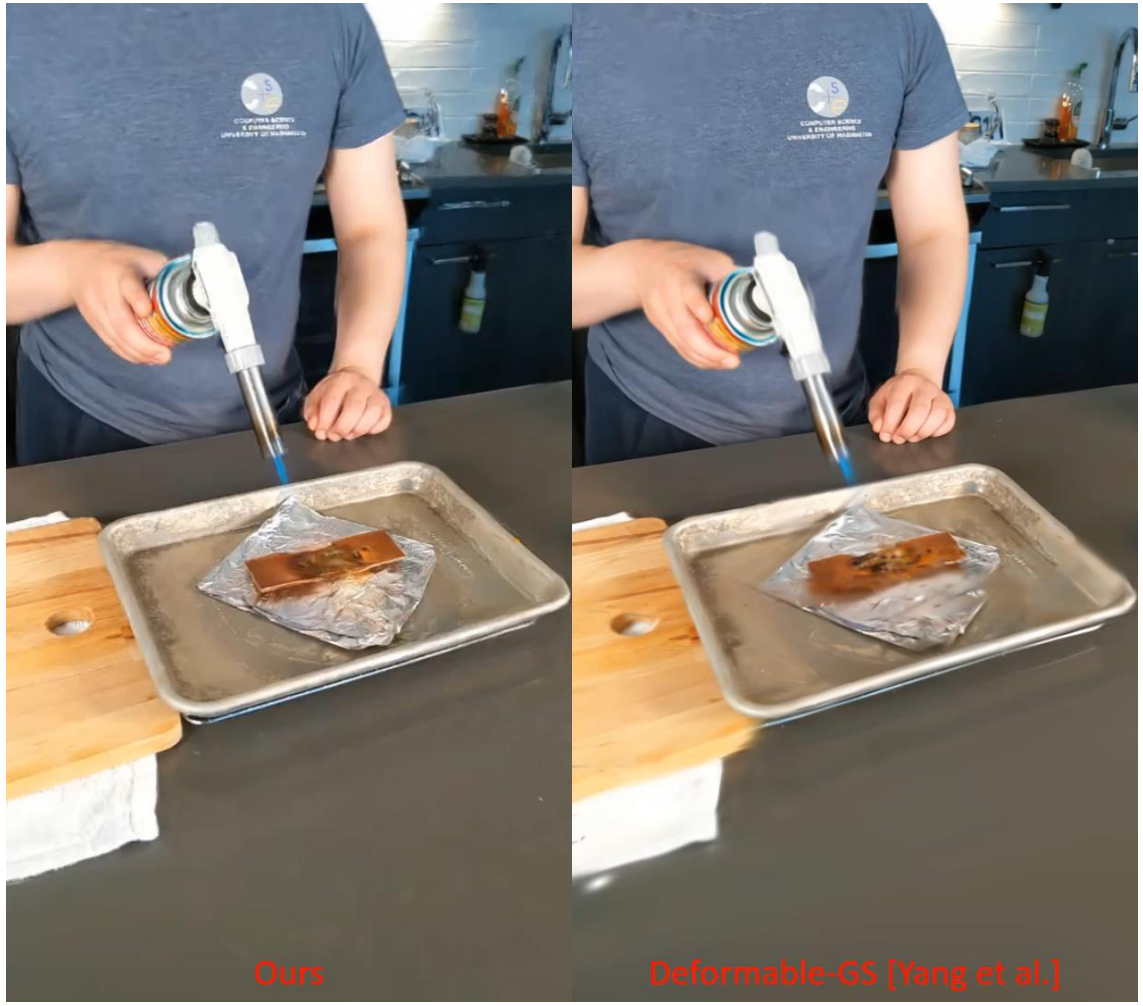
三维预测：GaussianPrediction

浙江大学 ZHEJIANG UNIVERSITY

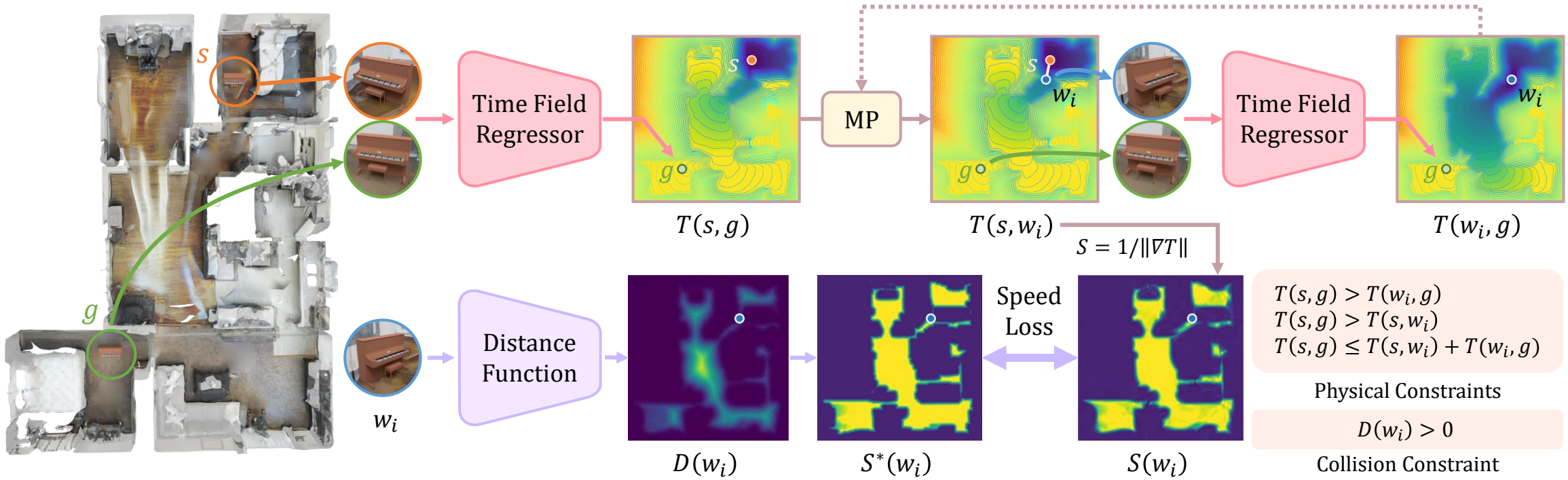
Chicken



Torchocolate



三维规划：PC-Planner

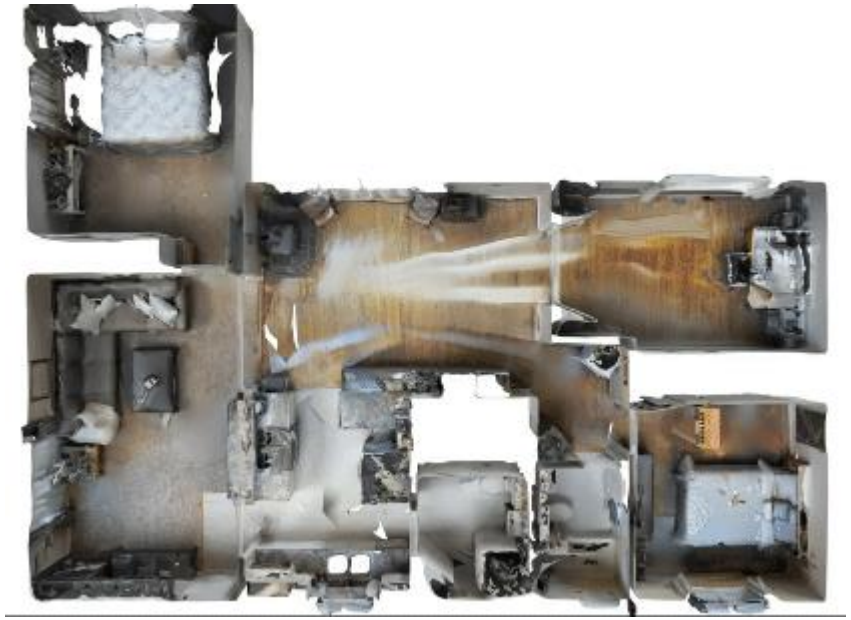


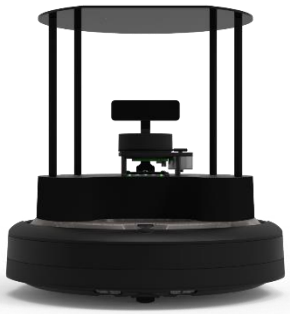
Shen et al. PC-Planner: Physics-Constrained Self-Supervised Learning for Robust Neural Motion Planning with Shape-Aware Distance Function. SIGGRAPH Asia 2024.

- **物理约束：** 将物理时空约束关系引入自监督学习框架，有效解决Eikonal Equation的多解问题
- **任意形状：** 通过新颖的形状距离场表征，支持任意形状机器人的高效运动规划
- **稳定鲁棒：** PC-Planner在3DoF，4DoF和6DoF场景中均达到了SOTA

三维规划：PC-Planner

6DoF合成场景





3DoF真实场景

总结

优势

- ✓ 三维神经感知：实现高精度的实时三维稠密建模
- ✓ 三维神经预测：支持短时未来场景的任意视角渲染
- ✓ 三维神经规划：实现基于物理信息引导的全局运动规划自监督学习

挑战

- ➢ 三维神经感知：需进一步降低计算量，实现端侧的实时感知
- ➢ 三维神经预测：需处理复杂运动，实现准确的长时预测
- ➢ 三维神经规划：需解决高维空间难题，实现动态环境下的鲁棒规划

浙江大学
ZHEJIANG UNIVERSITY

请各位专家批评指正