从结构化地图到机器人导航：
场景建模、视觉定位与自主规划

From Structured Map to Robot Navigation:
Scene Modeling, Visual Localization, and Autonomous Planning

高 翔

中国科学院自动化研究所

电子邮箱：xiang.gao@ia.ac.cn

个人主页：https://xgao-3dv.github.io/

汇报提纲

地图构建（环境表达）



含义：让机器人把环境“画出来”，形成可用的三维地图或场景模型



室内场景重建



实景三维建模



数字孪生建模



机器人环境建图

视觉定位（位姿解算）



含义：让机器人知道“我在哪里”，估计自身在地图中的位置与姿态



机器人定位



自动驾驶定位



无人机定位



AR/VR空间追踪

自主导航（行为决策）



含义：让机器人知道“我要怎么走”，完成路径规划、避障最终到达目标



仓储物流



无人配送



工业巡检



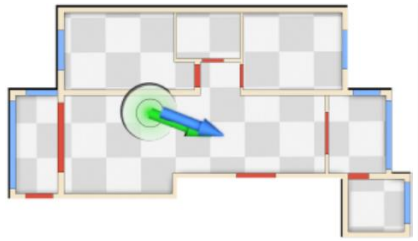
自主飞行

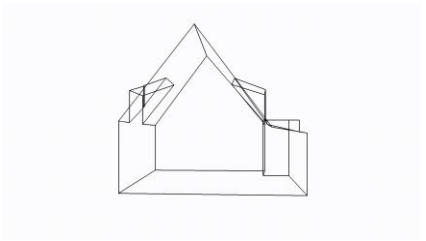
汇报提纲

- 基于结构化基元的建图与定位

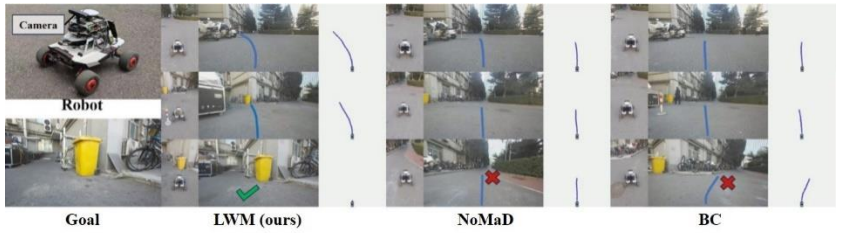


- 室内外场景矢量化建模与定位



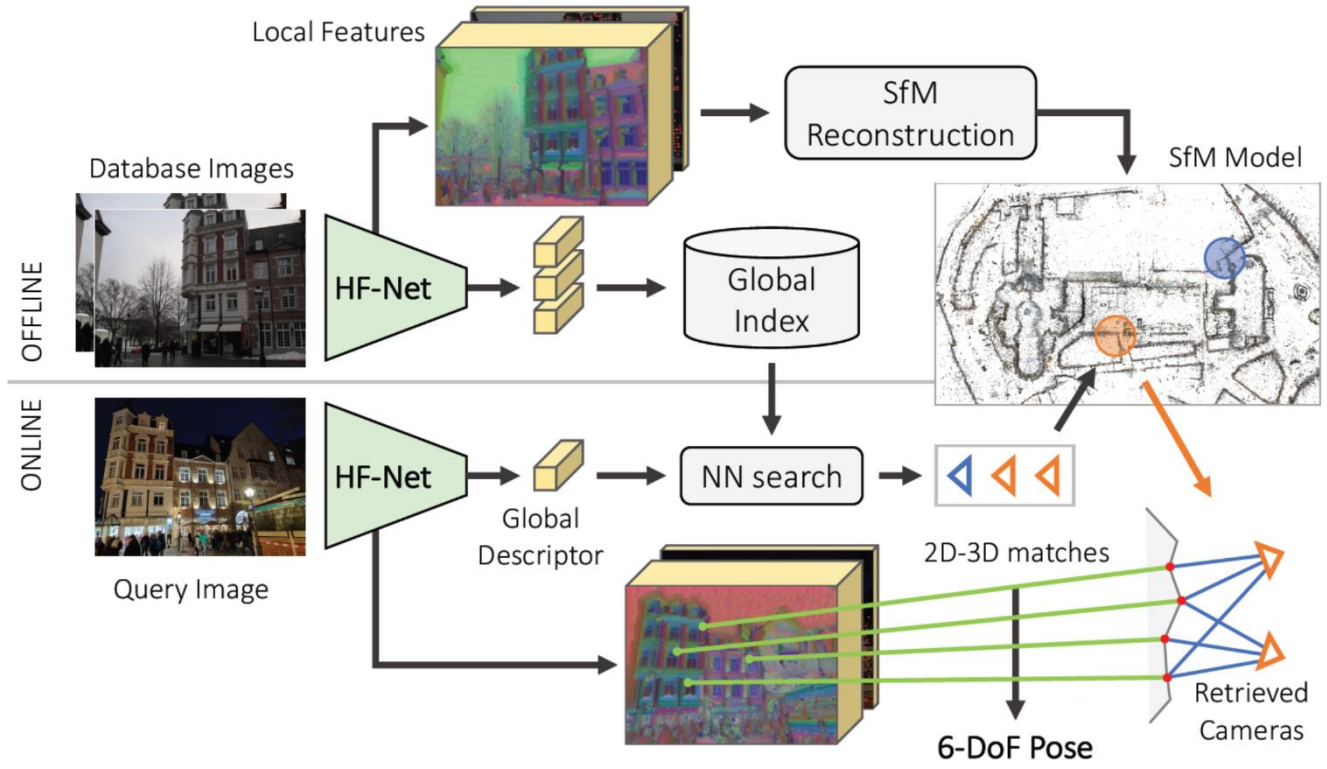


- 潜在世界模型驱动的自主导航



基于结构化基元的建图与定位

- **传统基于点的建图与定位框架（局限）**
  - **点云地图存储成本高**
  - **弱纹理区域表现较差**

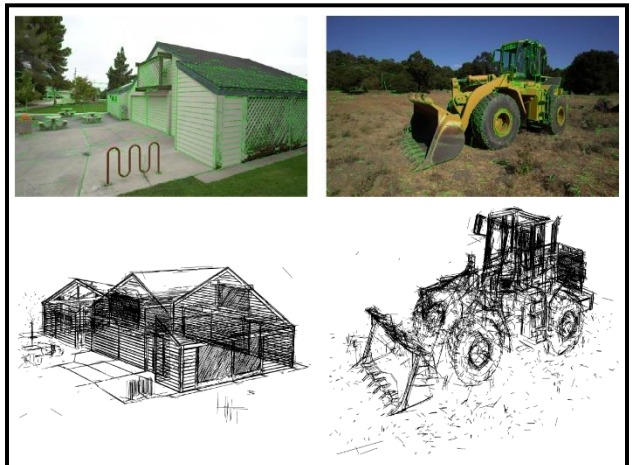


P.-E. Sarlin, et al. From Coarse to Fine: Robust Hierarchical Localization at Large Scale. In Proc. CVPR 2019

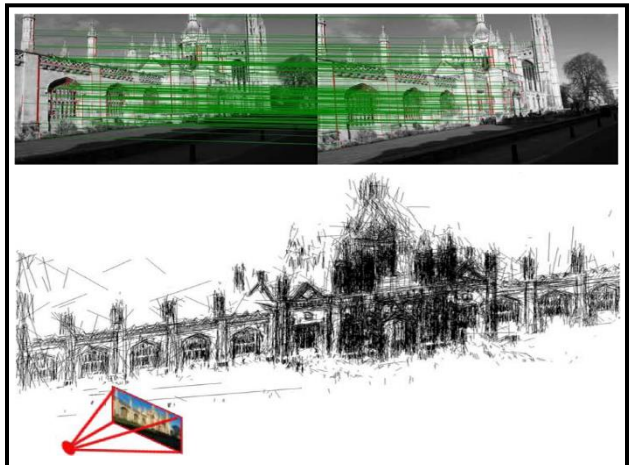
基于结构化基元的建图与定位

- **基于线元与面元的建图与定位**
  - 线元建图：*CLMAP In Proc. ECCV 2024*
  - 线元定位：*LSLM-VLoc IEEE RAL 2024*
  - 面元建图：*NeuralPlane In Proc. ICLR 2025*
  - 面元定位：*PlanaReLoc In Proc. CVPR 2026*

基于线元的建图与定位

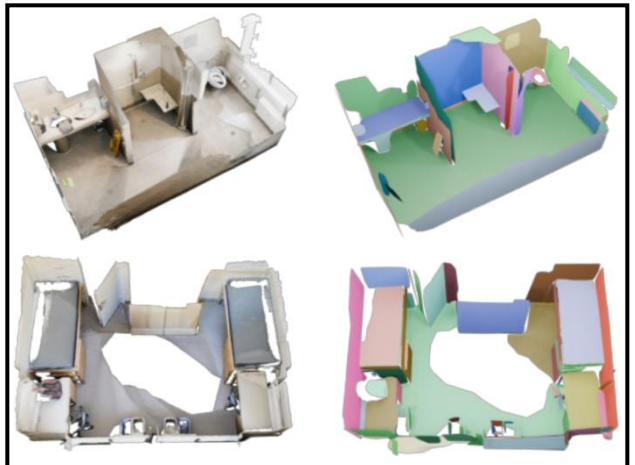


CLMAP

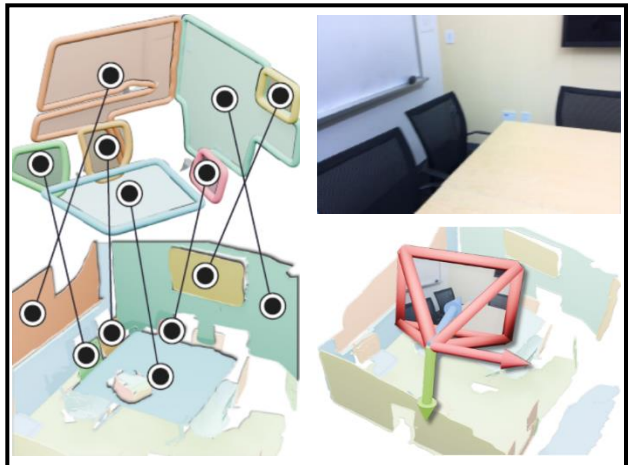


LSLM-VLoc

基于面元的建图与定位



NeuralPlane

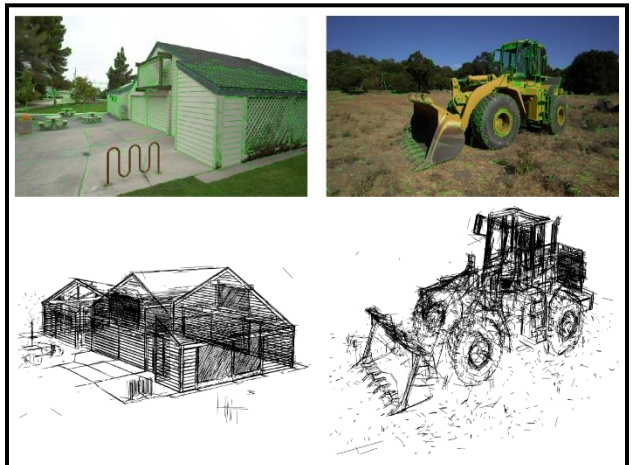


PlanaReLoc

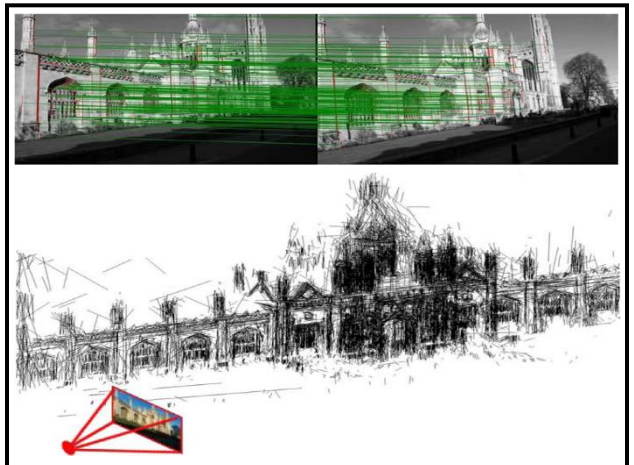
基于结构化基元的建图与定位

- **基于线元与面元的建图与定位**
  - 线元建图：*CLMAP In Proc. ECCV 2024*
  - 线元定位：*LSLM-VLoc IEEE RAL 2024*
  - **面元建图：*NeuralPlane In Proc. ICLR 2025***
  - **面元定位：*PlanaReLoc In Proc. CVPR 2026***

基于线元的建图与定位

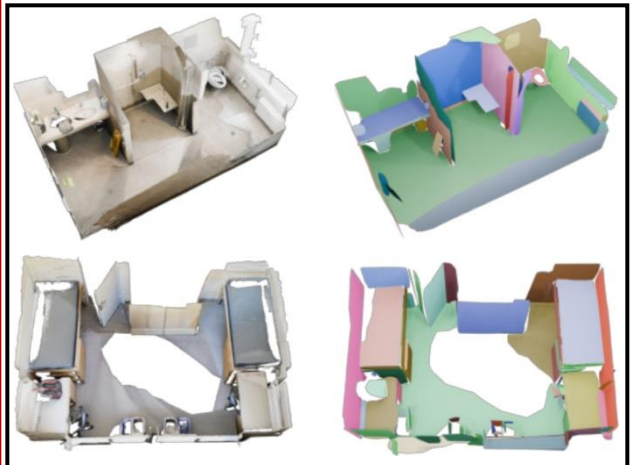


CLMAP

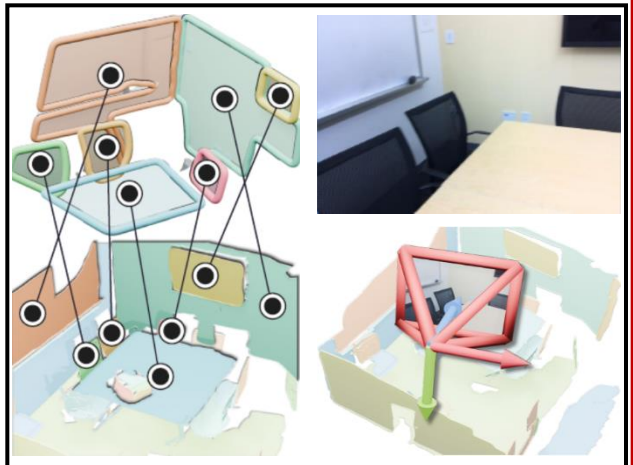


LSLM-VLoc

基于面元的建图与定位



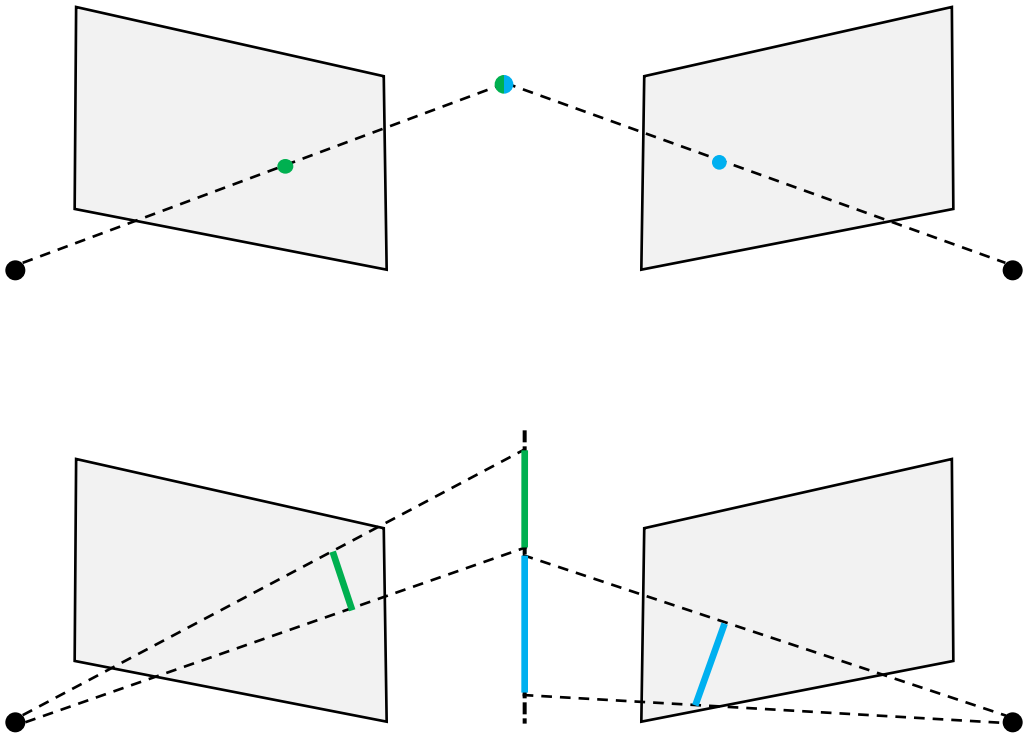
NeuralPlane

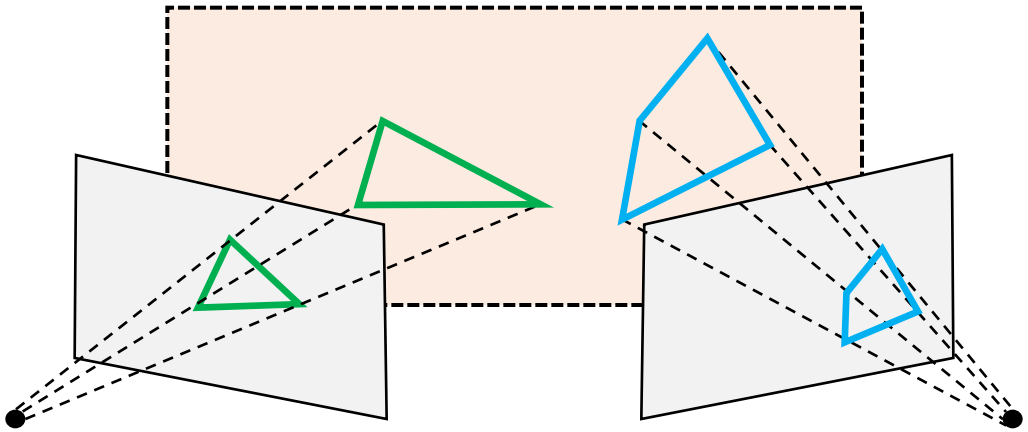


PlanaReLoc

基于结构化基元的建图与定位

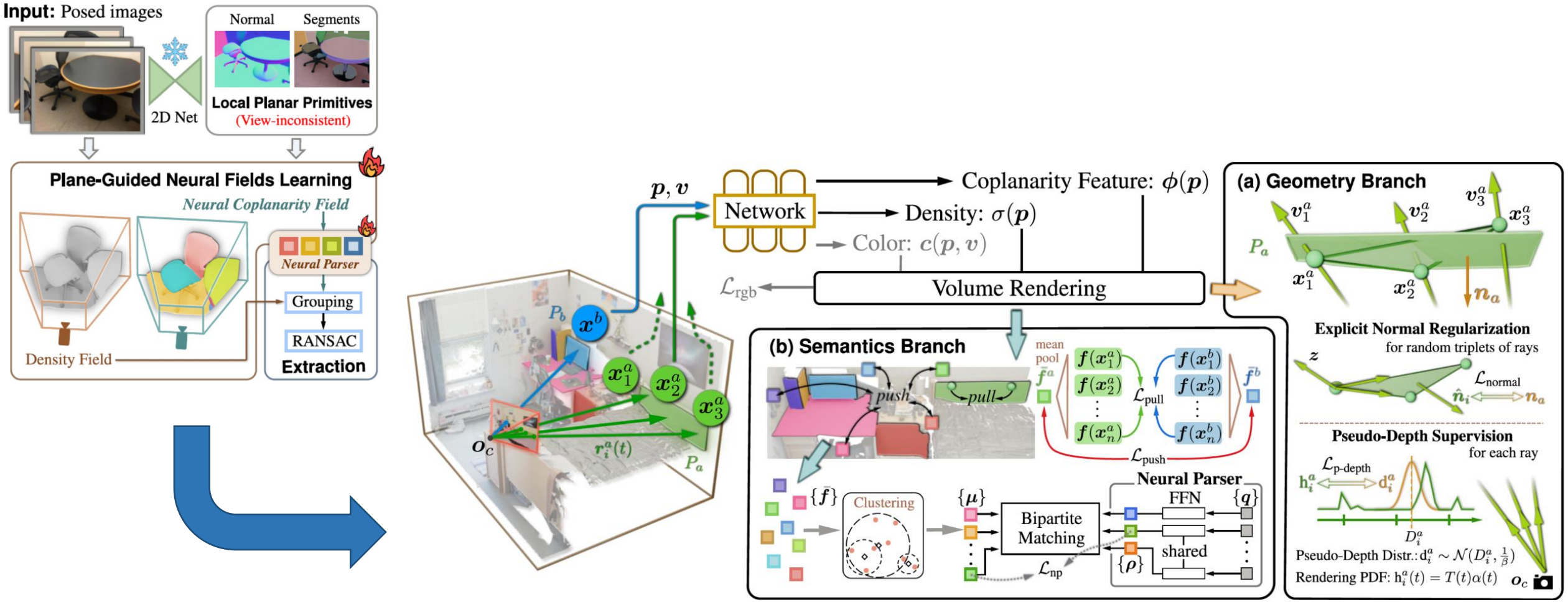
- 基于面元的三维建图：NeuralPlane





基于结构化基元的建图与定位

- **基于面元的三维建图：NeuralPlane**



Hanqiao Ye, et al. Structured 3D Reconstruction in Planar Primitives with Neural Fields. In Proc. ICLR 2025

基于结构化基元的建图与定位

- 基于面元的三维建图：NeuralPlane

ScanNet++定性对比实验结果



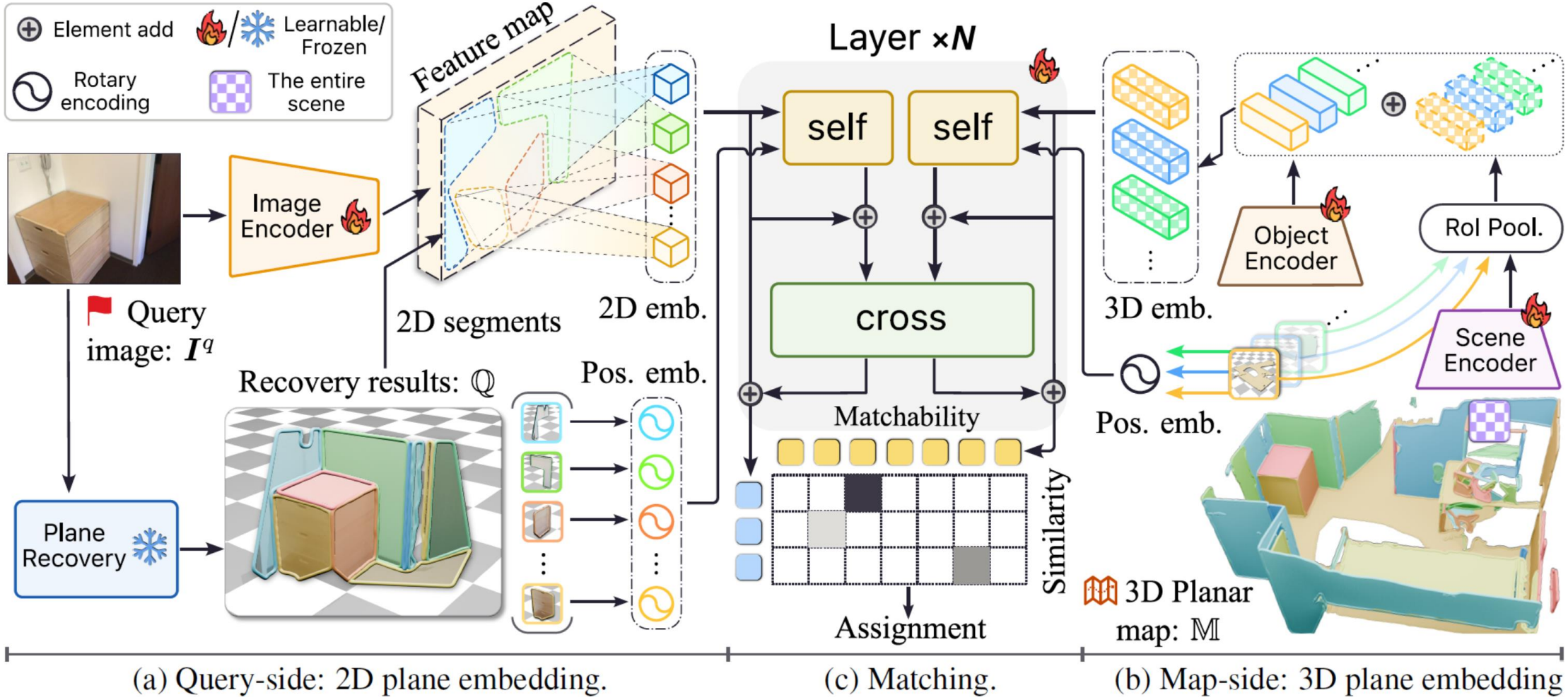
ScanNetv2定性对比实验结果



Hanqiao Ye, et al. Structured 3D Reconstruction in Planar Primitives with Neural Fields. In Proc. ICLR 2025

基于结构化基元的建图与定位

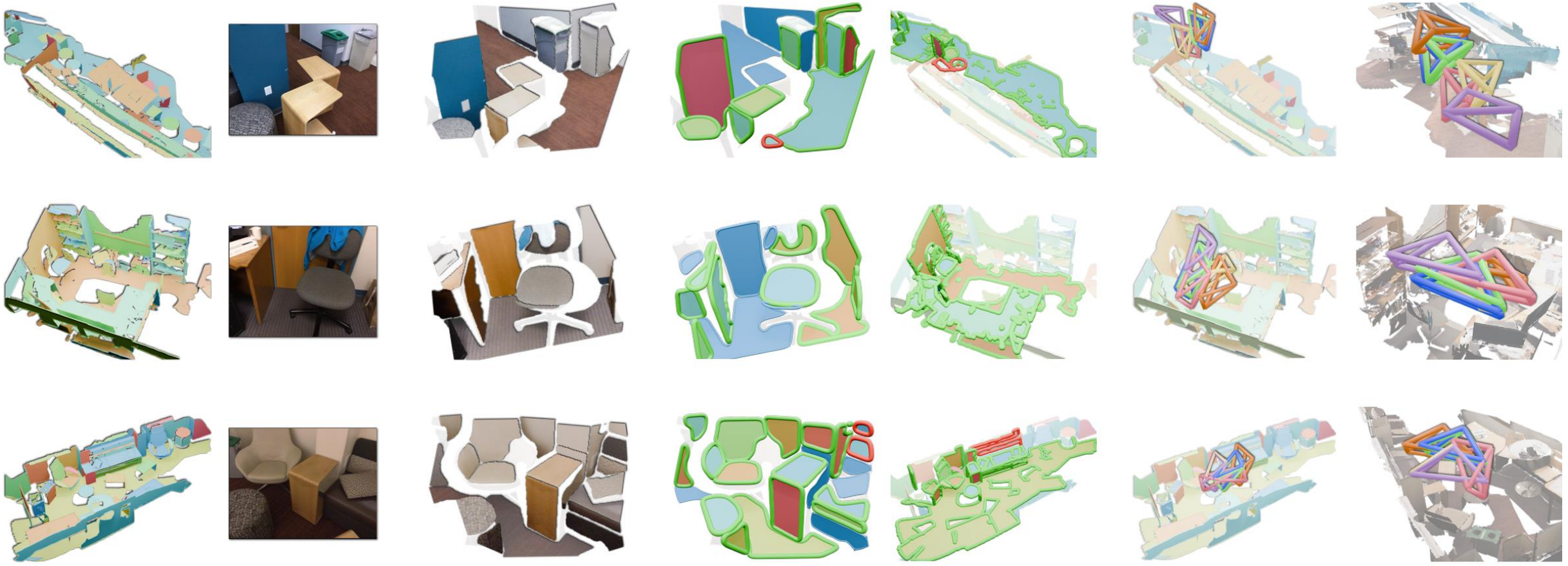
- **基于面元的视觉定位：PlanaReLoc**



Hanqiao Ye, et al. PlanaReLoc: Camera Relocalization in 3D Planar Primitives via Region-Based Structure Matching. ***In Proc. CVPR 2026***

基于结构化基元的建图与定位

- **基于面元的视觉定位：PlanaReLoc**



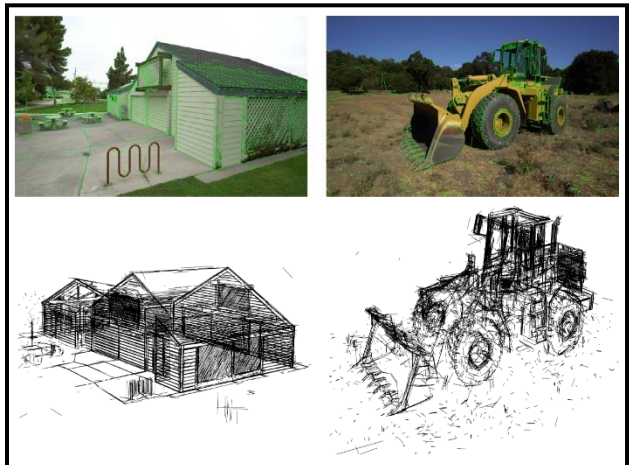
(a) Input: map & query (b) Monocular plane recovery (c) Plane correspondences (d) Poses (viewpoint 1 & 2)

Hanqiao Ye, et al. PlanaReLoc: Camera Relocalization in 3D Planar Primitives via Region-Based Structure Matching. In Proc. CVPR 2026

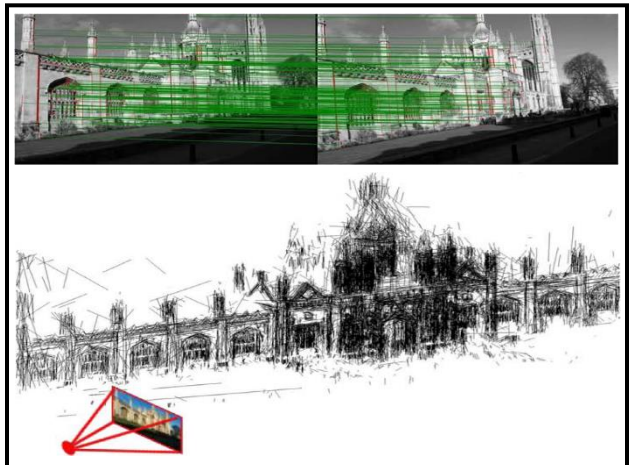
室内外场景矢量化建模与定位

- **基于结构化基元的建图与定位（局限）**
  - **边界不规则**
  - **冗余信息多**
  - **动态易干扰**

基于线元的构图与定位

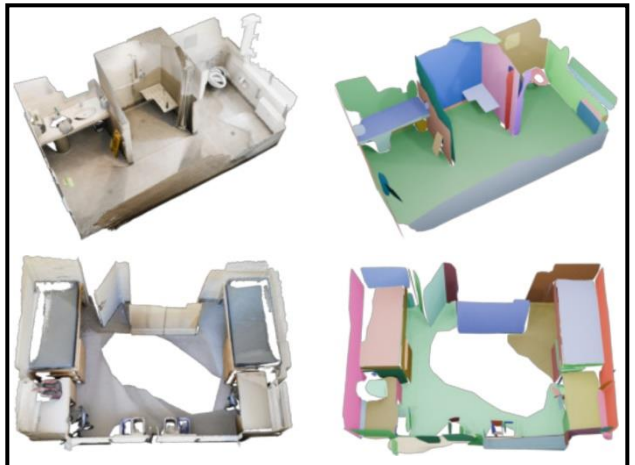


CLMAP

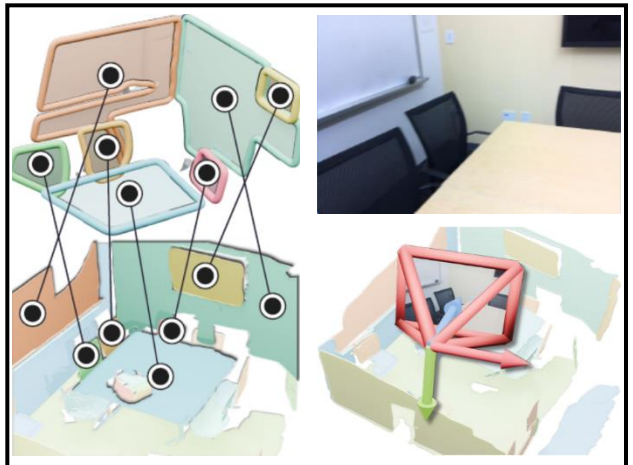


LSLM-VLoc

基于面元的构图与定位



NeuralPlane

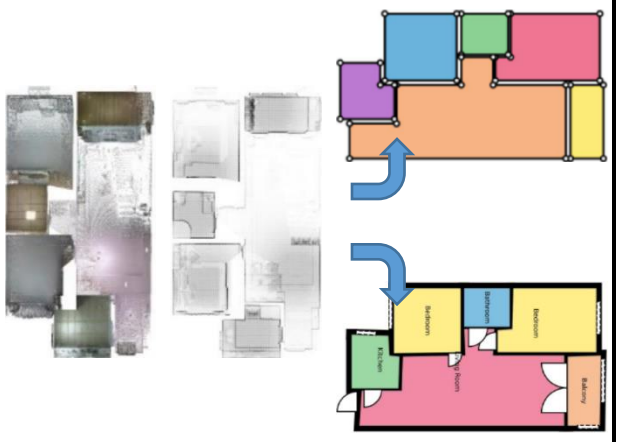


PlanaReLoc

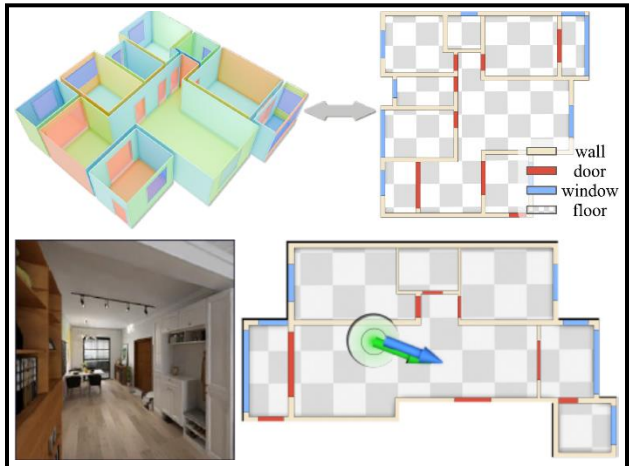
室内外场景矢量化建模与定位

- **室内平面布置图与建筑物线框模型拓扑结构的重建与生成**
  - 室内平面布置图重建：*PolyRoom In Proc. ECCV 2024*
  - 室内平面布置图定位：*PGGT-Loc In Proc. NeurIPS 2026*
  - 建筑物线框模型重建：*BWFormer In Proc. CVPR 2025*
  - 建筑物线框模型生成：*BuildingGPT In Proc. CVPR 2026*

室内平面布置图重建与定位

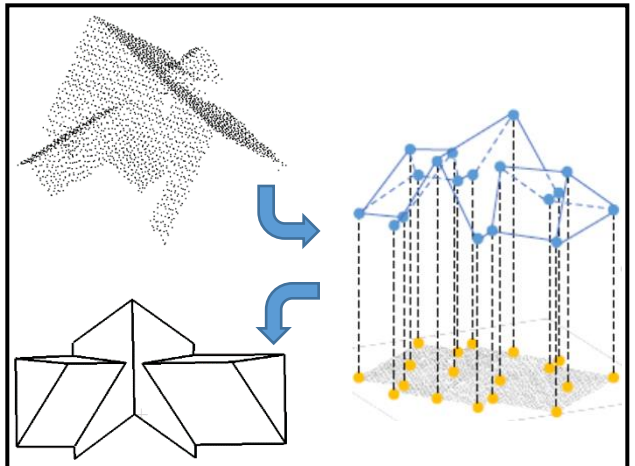


PolyRoom

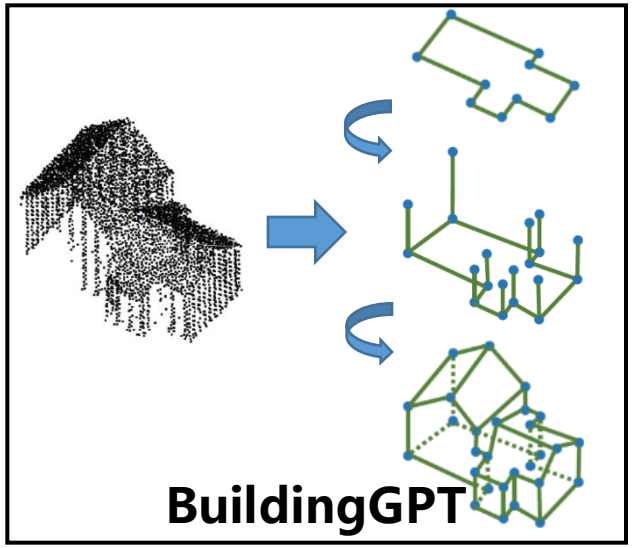


PGGT-Loc

建筑物线框模型重建与生成



BWFormer

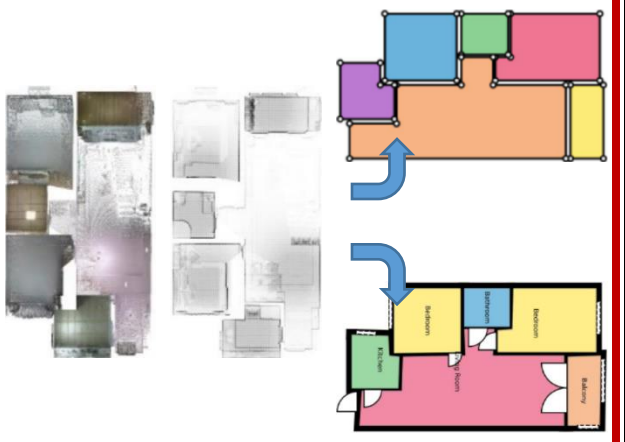


BuildingGPT

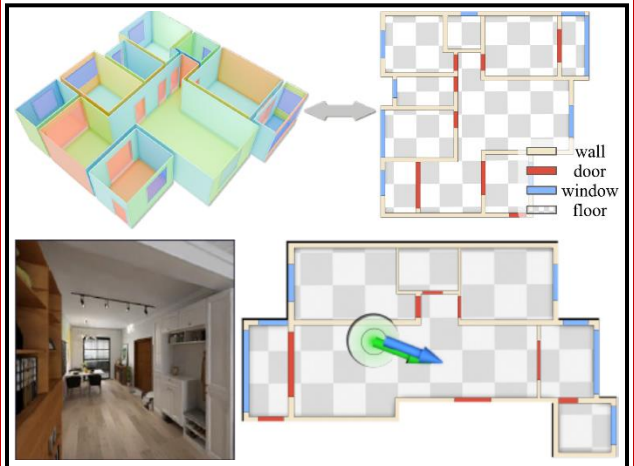
室内外场景矢量化建模与定位

- **室内平面布置图与建筑物线框模型拓扑结构的重建与生成**
  - 室内平面布置图重建：*PolyRoom In Proc. ECCV 2024*
  - **室内平面布置图定位：*PGGT-Loc In Proc. NeurIPS 2026***
  - 建筑物线框模型重建：*BWFormer In Proc. CVPR 2025*
  - **建筑物线框模型生成：*BuildingGPT In Proc. CVPR 2026***

室内平面布置图重建与定位

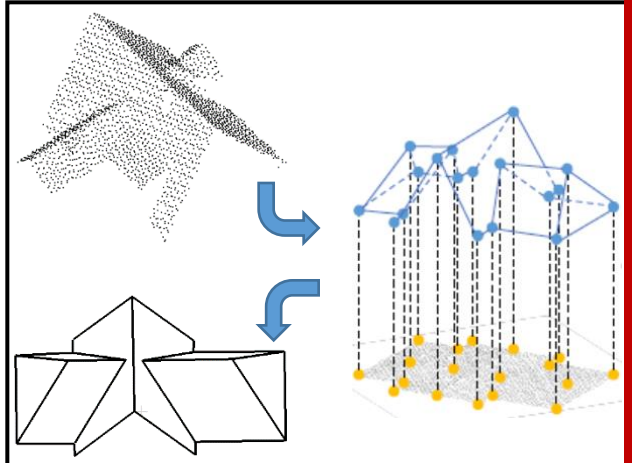


PolyRoom

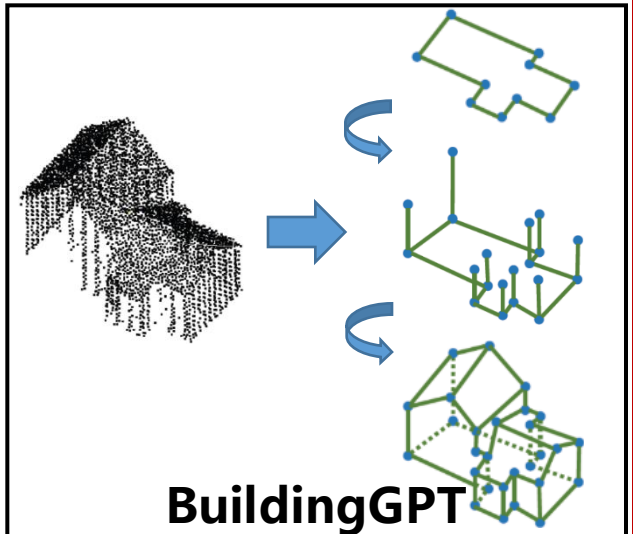


PGGT-Loc

建筑物线框模型重建与生成



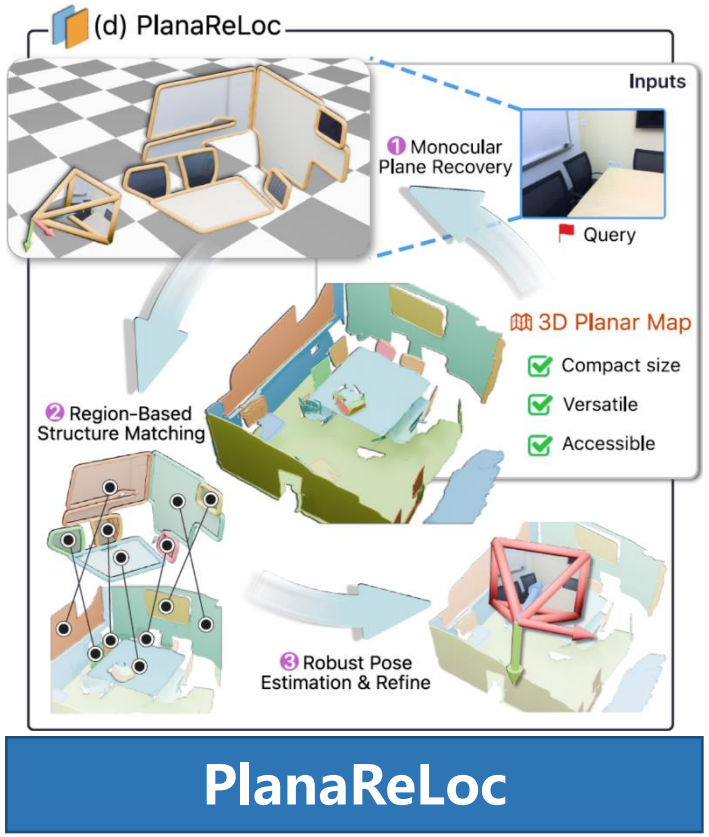
BWFormer



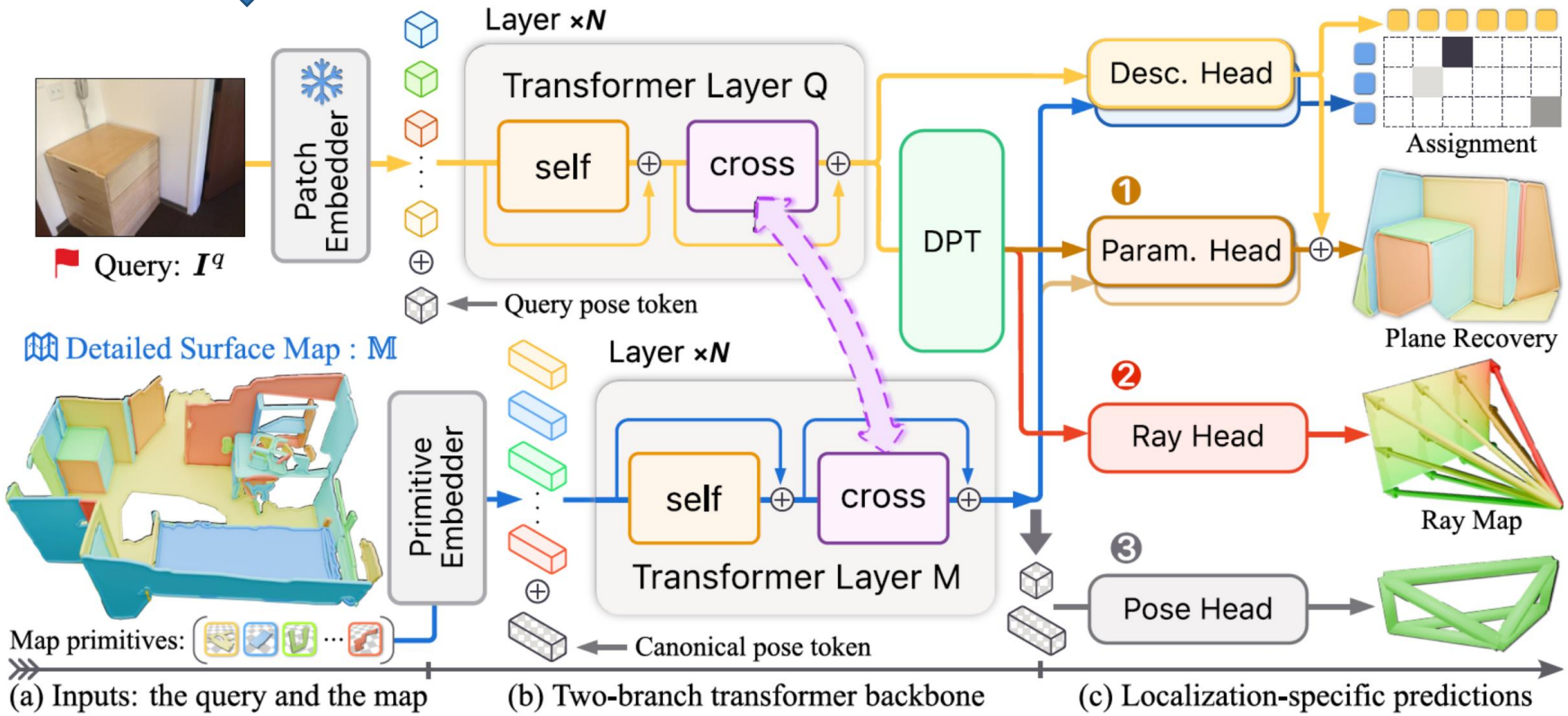
BuildingGPT

室内外场景矢量化建模与定位

- **室内平面布置图定位：PGGT-Loc**



PlanaReLoc

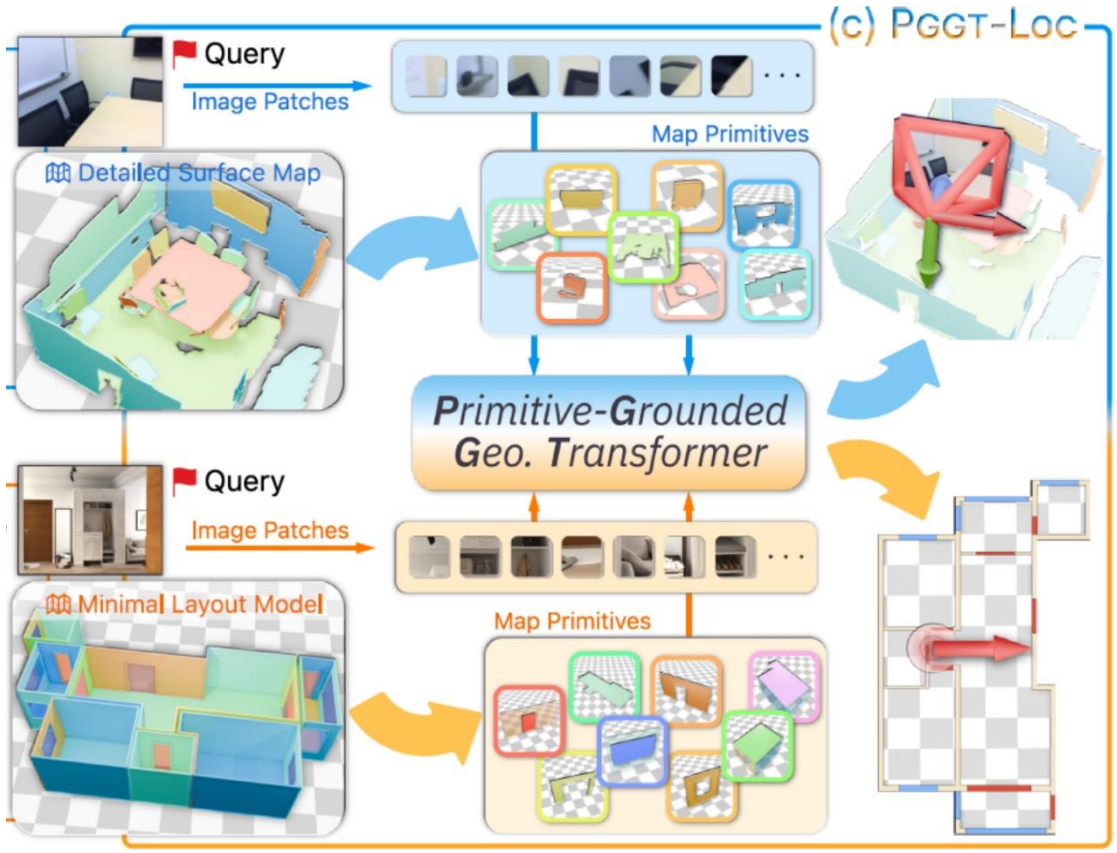


- **预处理：提取查询图像平面区域**
- **后处理：基于平面匹配解算位姿**

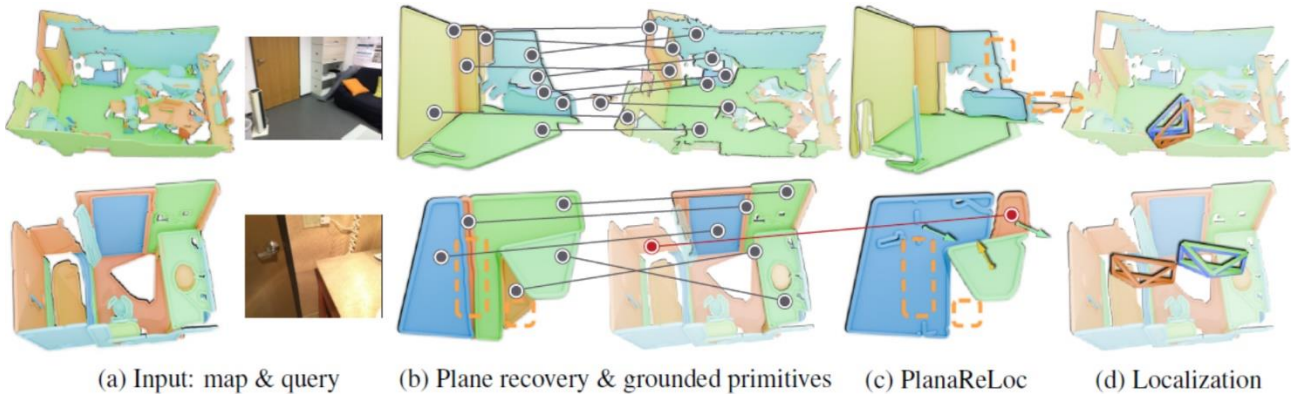
Hanqiao Ye, et al. PGGT-Loc: Primitive-Grounded Geometry Transformer for Feed-Forward Camera Localization. In Proc. NeurIPS 2026

室内外场景矢量化建模与定位

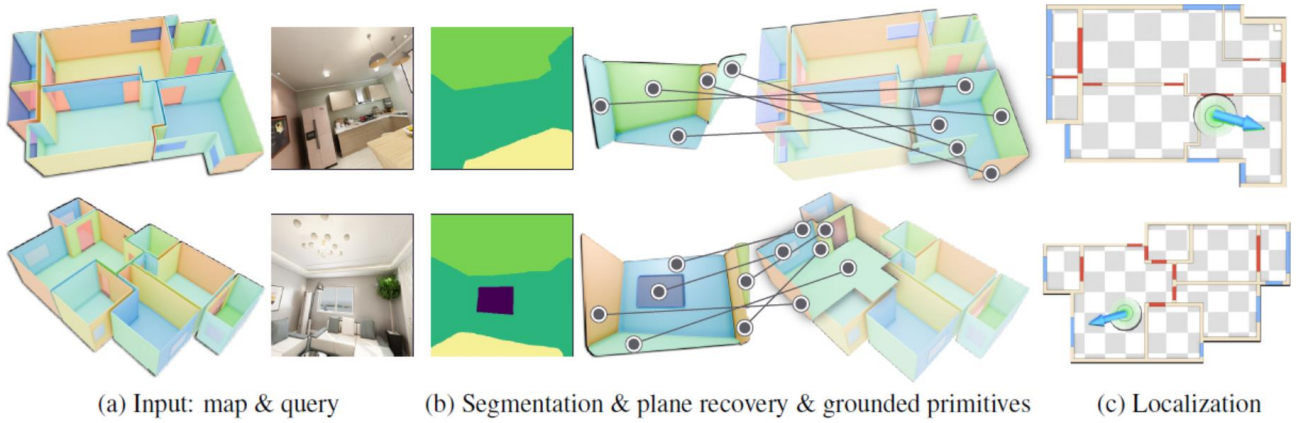
- **室内平面布置图定位：PGGT-Loc**



面元地图定位



平面布置图定位



Hanqiao Ye, et al. PGGT-Loc: Primitive-Grounded Geometry Transformer for Feed-Forward Camera Localization. In Proc. NeurIPS 2026

室内外场景矢量化建模与定位

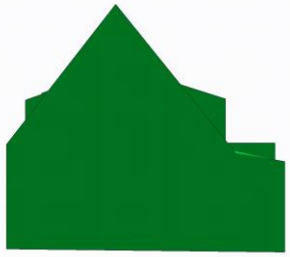
- **建筑物线框模型生成：BuildingGPT**



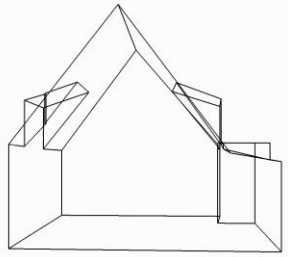
建筑物点云



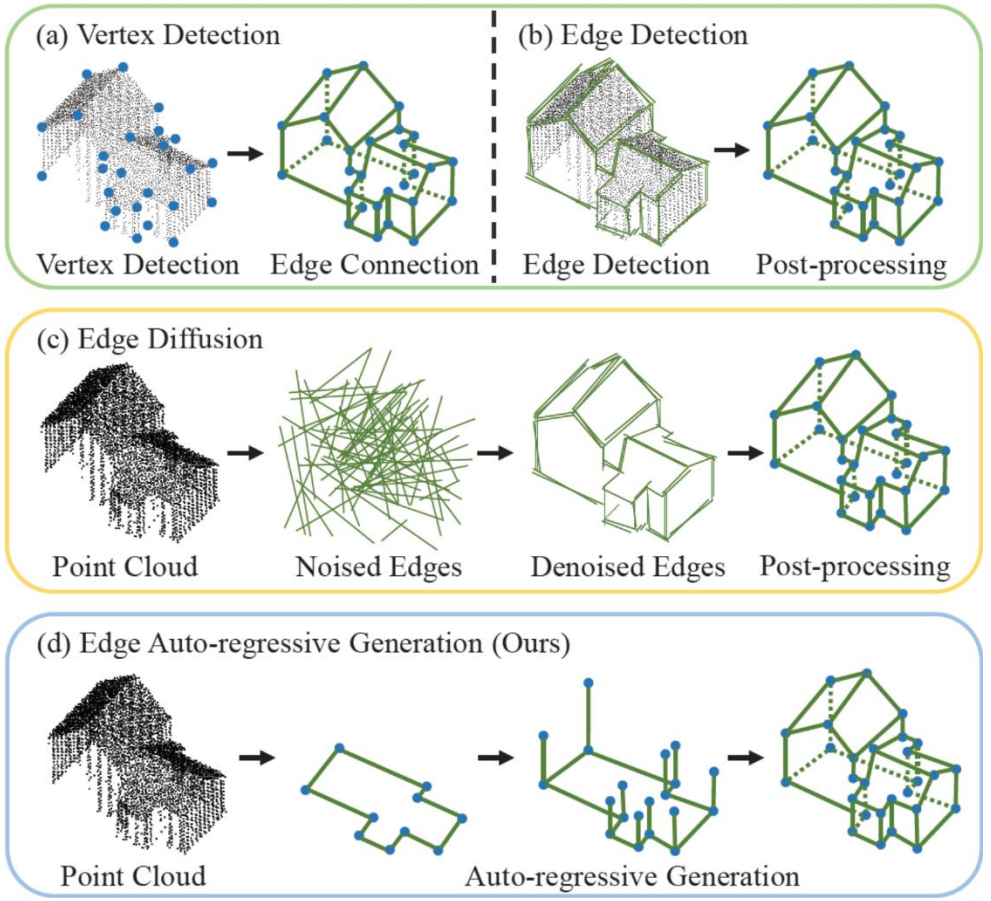
屋顶点云



建筑物网格模型



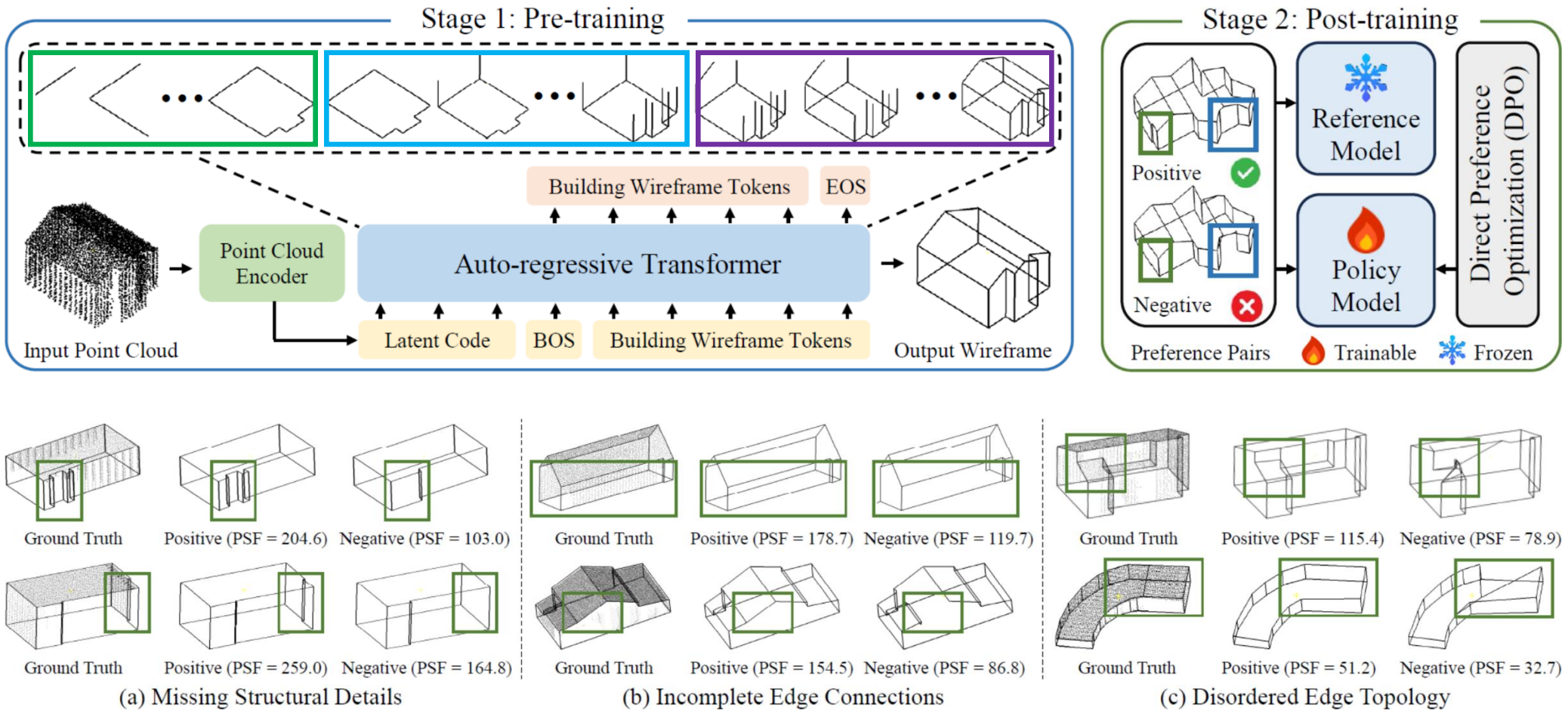
建筑物线框模型



Yuzhou Liu, et al. BuildingGPT: Auto-Regressive Building Wireframe Reconstruction Model with Reinforcement Learning. ***In Proc. CVPR 2026***

室内外场景矢量化建模与定位

- **建筑物线框模型生成：BuildingGPT**



Yuzhou Liu, et al. BuildingGPT: Auto-Regressive Building Wireframe Reconstruction Model with Reinforcement Learning. In Proc. CVPR 2026

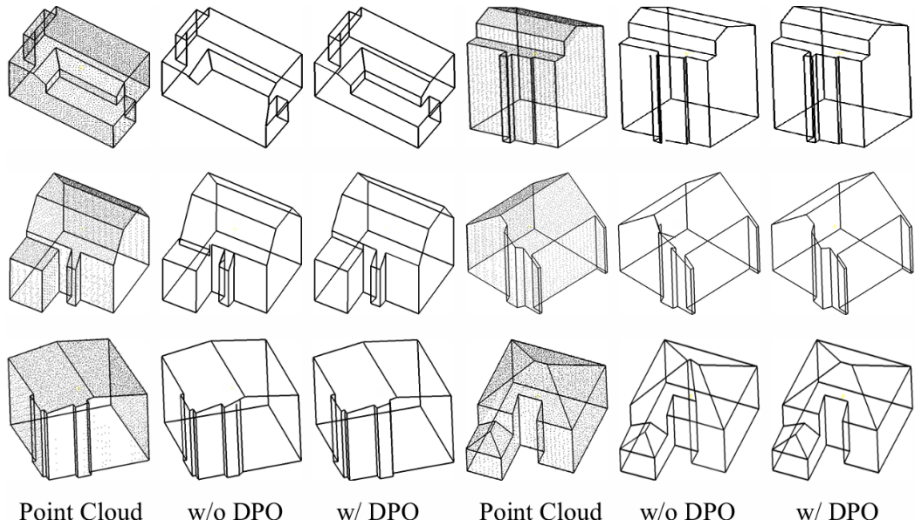
室内外场景矢量化建模与定位

- **建筑物线框模型生成：BuildingGPT**

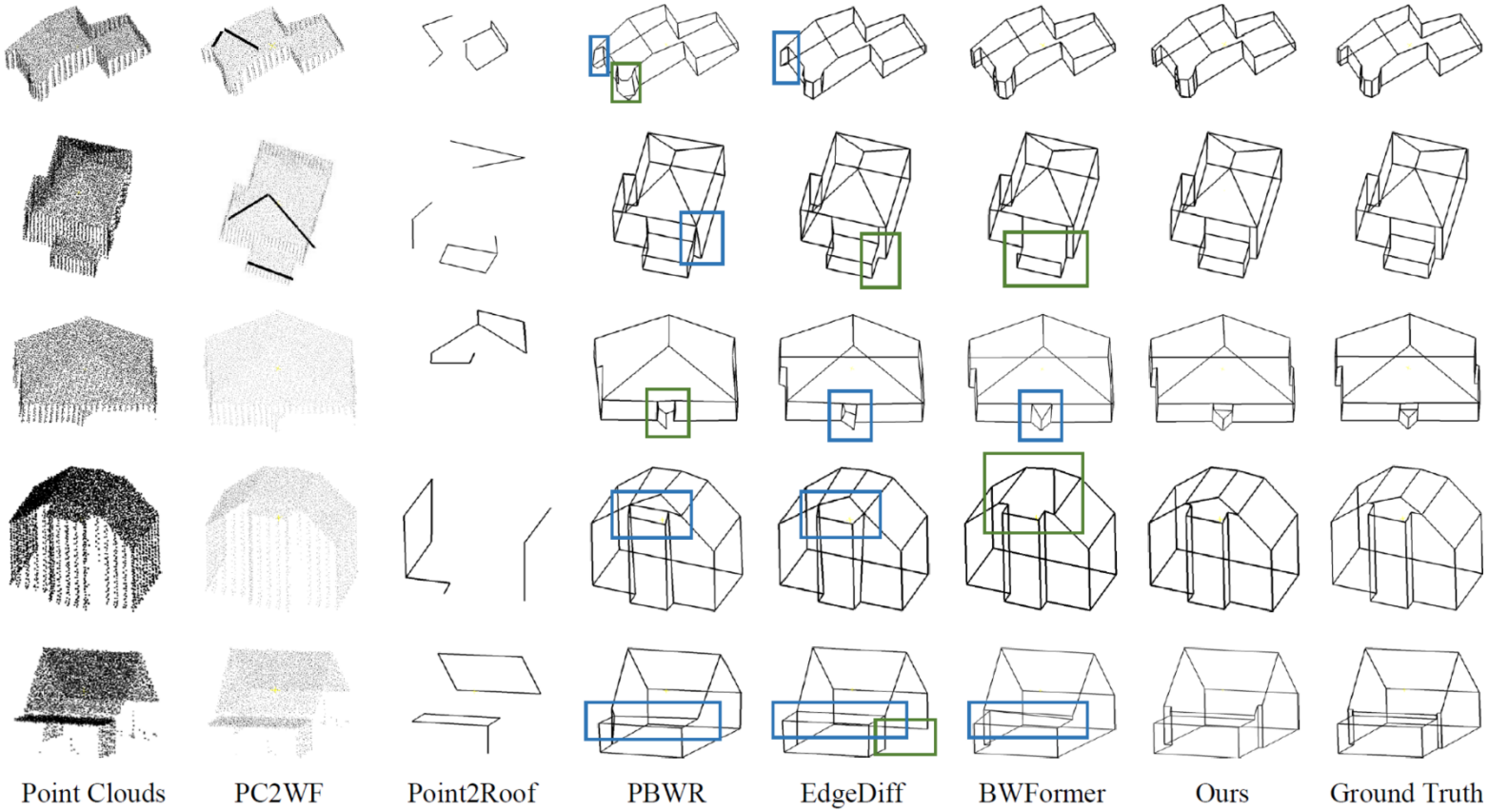
定量消融实验

Method	Distance ($\times 10^{-2}$)		Corner (%)			Edge (%)		
	WED ↓	ACO ↓	CP ↑	CR ↑	CF1 ↑	EP ↑	ER ↑	EF1 ↑
Baseline	1.39	1.25	96.2	96.2	96.0	92.2	91.9	91.9
+ Hierarchical Tokenization	1.19	1.08	96.8	97.1	96.7	93.1	93.3	93.1
+ Post-training with DPO	**0.98**	**0.88**	**97.4**	**97.7**	**97.4**	**94.6**	**94.5**	**94.4**

定性消融实验



定性对比实验



Yuzhou Liu, et al. BuildingGPT: Auto-Regressive Building Wireframe Reconstruction Model with Reinforcement Learning. ***In Proc. CVPR 2026***

潜在世界模型驱动的自主导航

地图构建（环境表达）	视觉定位（位姿解算）	自主导航（行为决策）







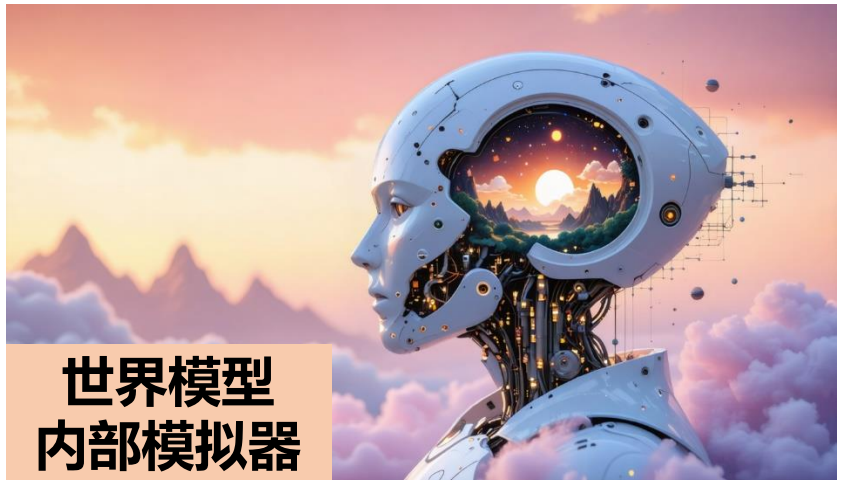


当前状态

动作1

动作2

动作3



未来状态1

未来状态2

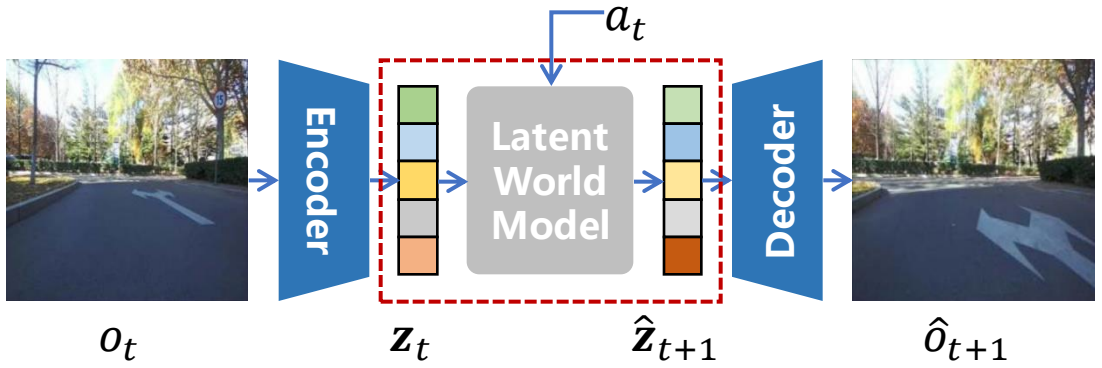
未来状态3



目标状态

潜在世界模型驱动的自主导航

- 面向自主导航的潜在世界模型构建

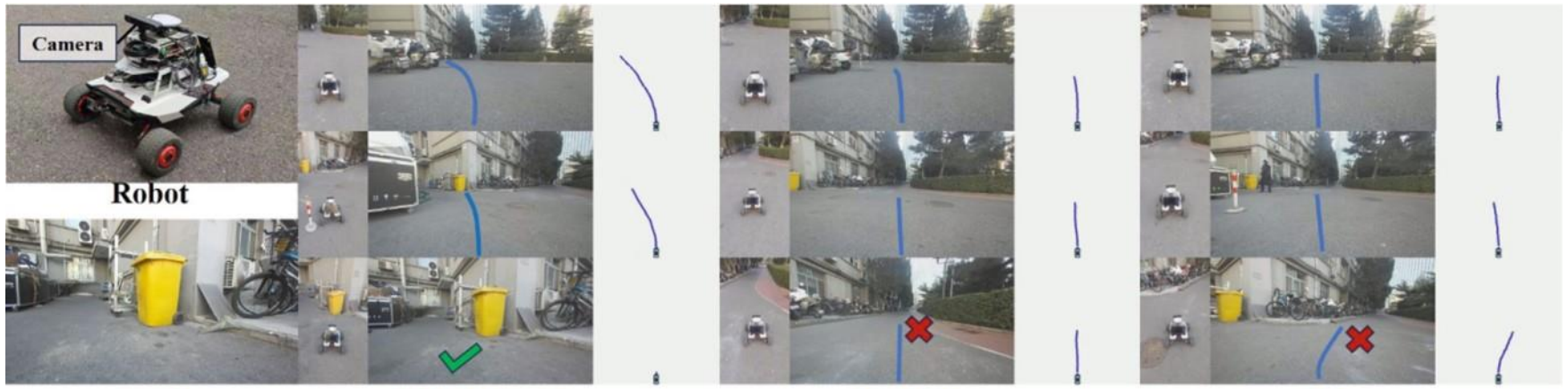


- 潜在世界模型驱动的导航策略学习

Zengmao Wang, et al. Predicting Consequences and Reinforcing Navigation Policies with Latent World Models. In Proc. ECCV 2026

潜在世界模型驱动的自主导航

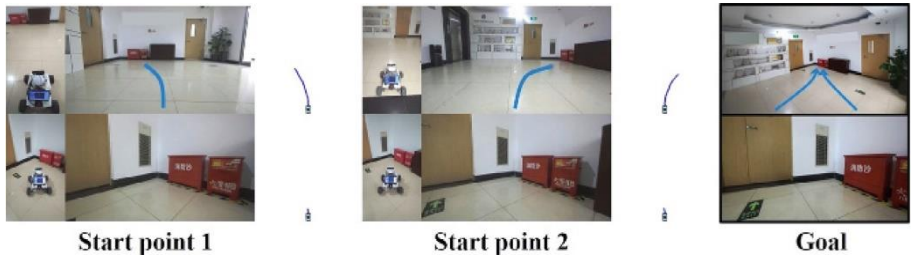
真实场景定性实验



真实场景定量实验

Method	SR ↑	LPIPS ↓
NoMaD	33.3	0.448
BC	33.3	0.566
LWM-B-P (ours)	**66.7**	**0.396**

室内场景零样本泛化实验

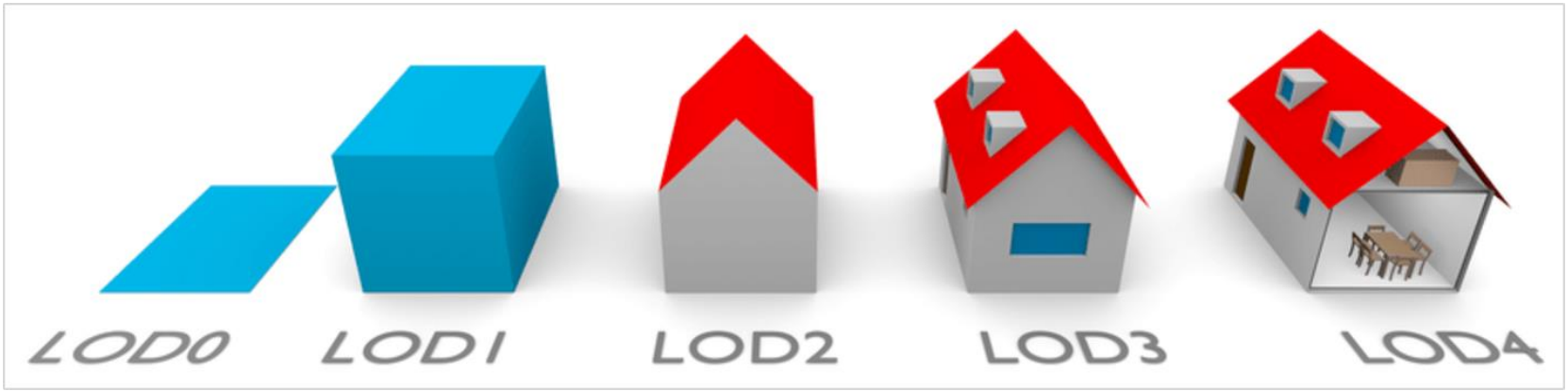


Zengmao Wang, et al. Predicting Consequences and Reinforcing Navigation Policies with Latent World Models. ***In Proc. ECCV 2026***

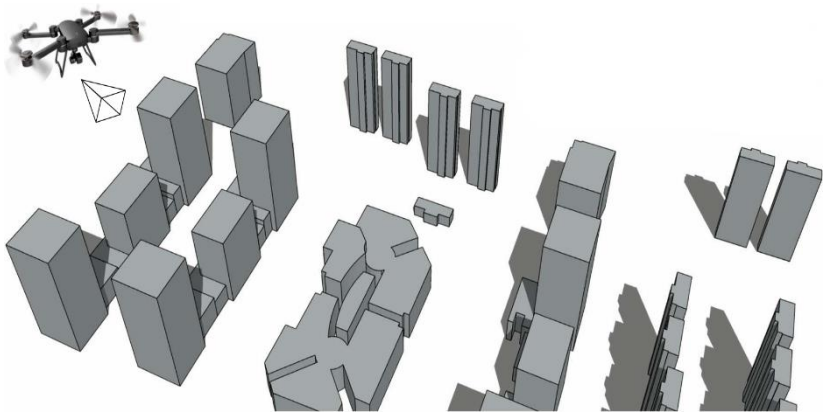
总结

- **基于结构化基元的建图与定位**
  - 线元建图：*CLMAP In Proc. ECCV 2024*
  - 线元定位：*LSLM-VLoc IEEE RAL 2024*
  - **面元建图：*NeuralPlane In Proc. ICLR 2025***
  - **面元定位：*PlanaReLoc In Proc. CVPR 2026***
- **室内外场景矢量化建模与定位**
  - 室内平面布置图重建：*PolyRoom In Proc. ECCV 2024*
  - **室内平面布置图定位：*PGGT-Loc In Proc. NeurIPS 2026***
  - 建筑物线框模型重建：*BWFormer In Proc. CVPR 2025*
  - **建筑物线框模型生成：*BuildingGPT In Proc. CVPR 2026***
- **潜在世界模型驱动的自主导航**
  - **世界模型构建与导航策略学习：*LWM In Proc. ECCV 2026***

展望

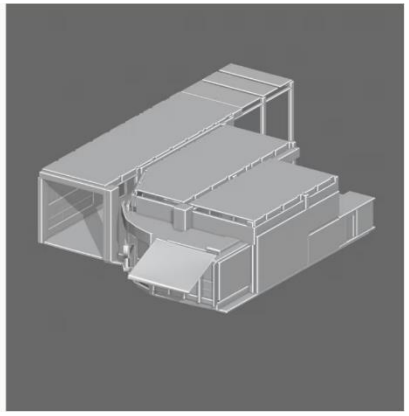






基于LOD模型的视觉定位





基于大模型的高LOD层级建模



显隐结合的世界模型构建

谢　谢

高　翔

中国科学院自动化研究所

电子邮箱：xiang.gao@ia.ac.cn

个人主页：https://xgao-3dv.github.io/