Next-Generation **Scene Representations** for **Robot Navigation**

2025. 11. 03

CS580 : Computer Graphics

Team #3 Student Lecture

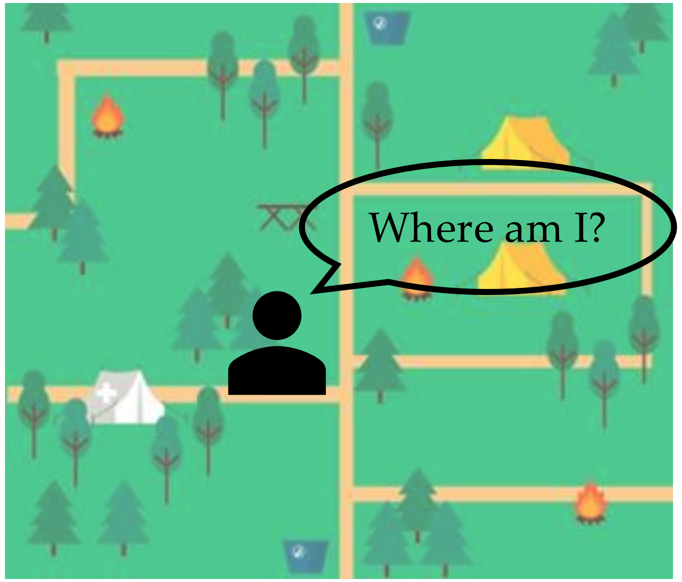
Jiwon Park & Harin Kim

KAIST

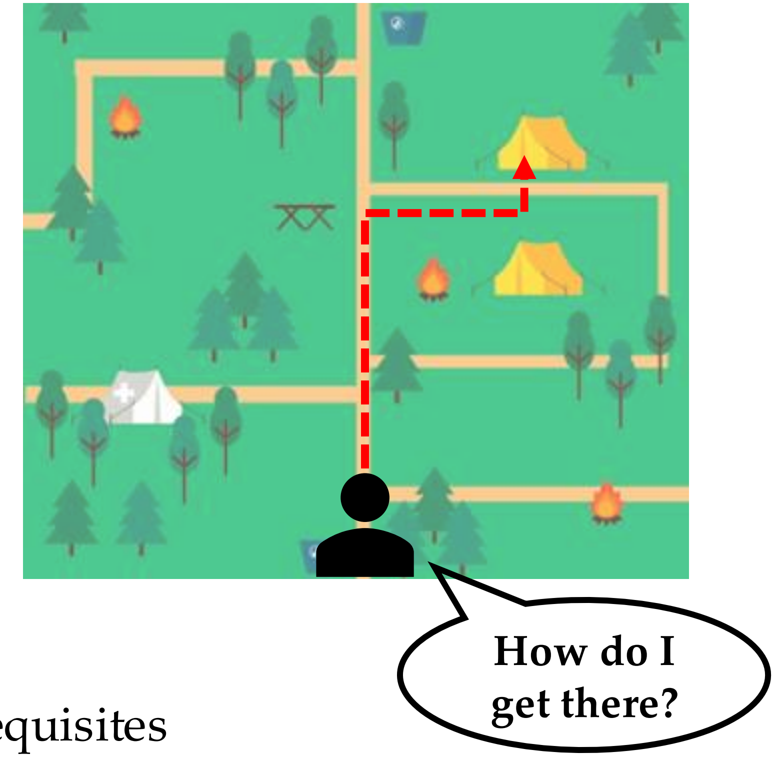
Introduction

- Graphics from **a robot application perspective**
  - Creating representations for robots to understand the world

Localization



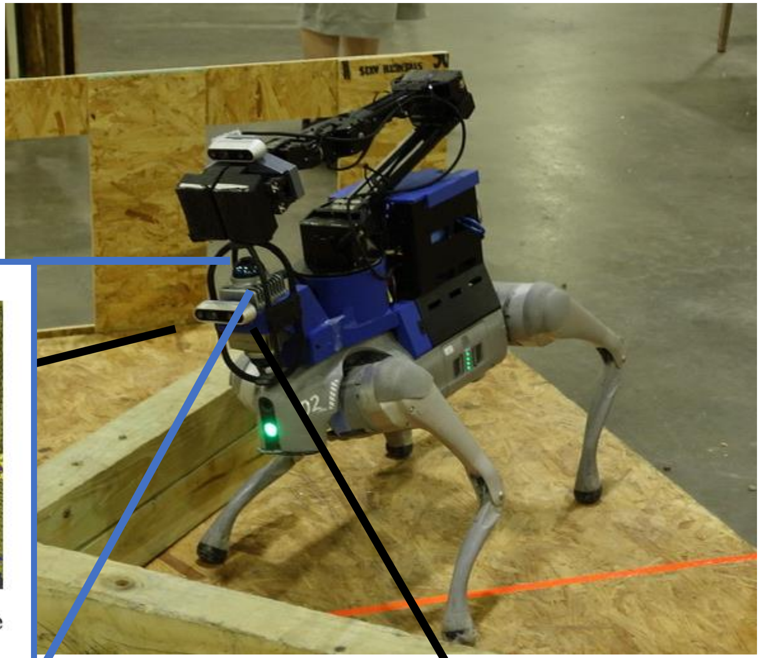
Path planning

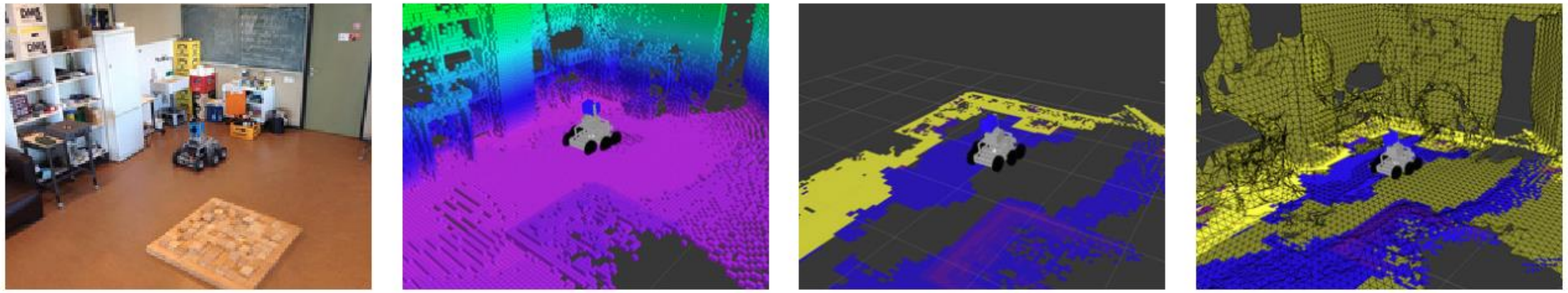


Navigation Prerequisites

Introduction

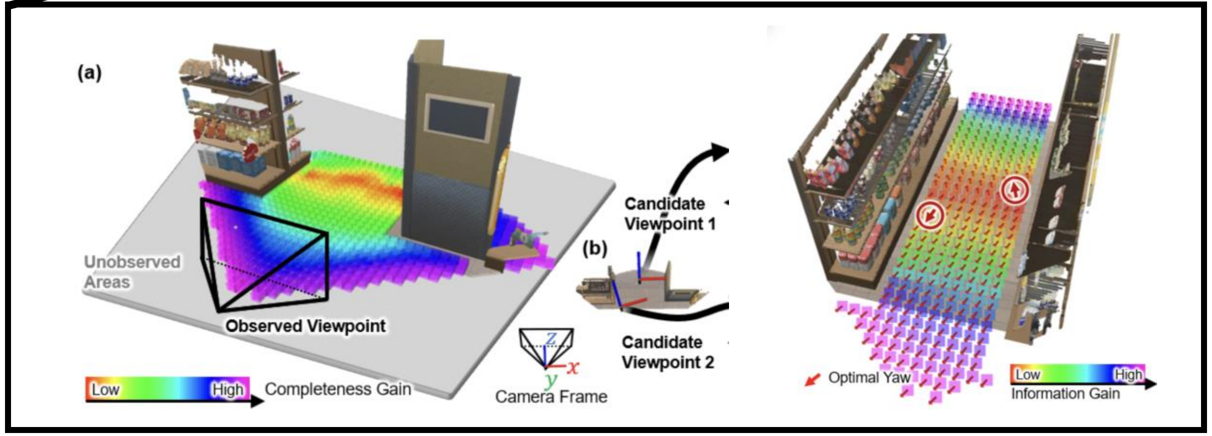
- Graphics from **a robot application perspective**
  - Creating representations for robots to understand the world

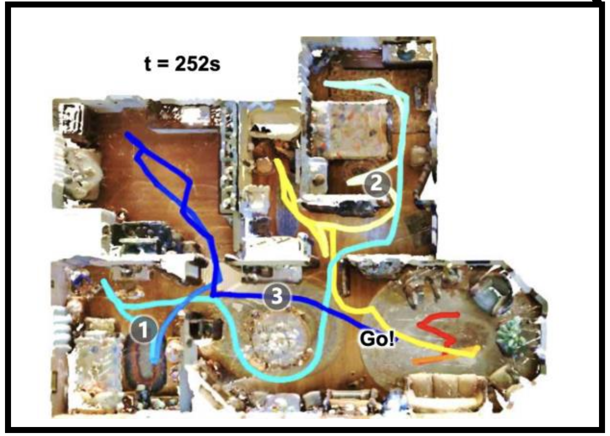




(a) Lab example scene with a step field. (b) Octree representation of the scene. (c) 2.5D Map (DEM) representation of the scene. (d) Mesh representation of the scene.

LiDAR





Camera

KAIST

Contents

- Graphics for Robotics
- Research Trends by Year
- Scene Properties & Representations
- Traditional Scene Representation
  - Navigation examples
- Next-Gen Scene Representations
  - Navigation examples

Paper Lists

NeRF

1. Renderable Neural Radiance Map for Visual Navigation
2. Vision-Only Robot Navigation in a Neural Radiance World

Gaussian

1. RTG-SLAM: Real-time 3D Reconstruction at Scale Using Gaussian Splatting
2. Compressed 3D Gaussian Splatting for Accelerated Novel View Synthesis

KAIST

Graphics for **Robotics?**

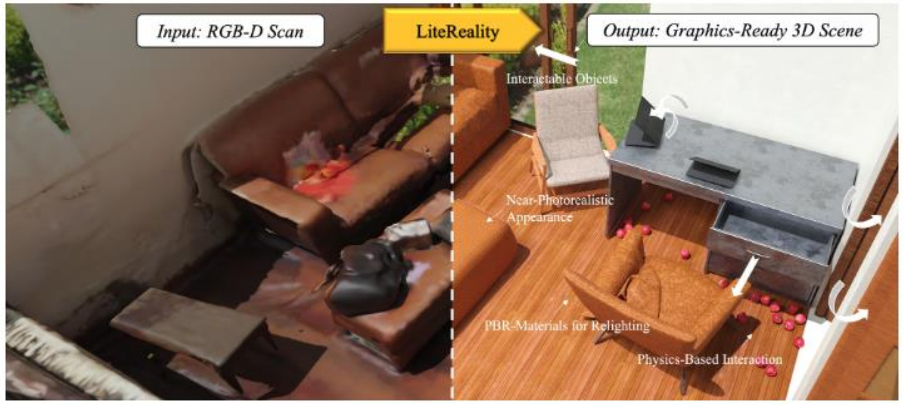
Graphics

- Scene fidelity
- Photorealism
- Geometric consistency

Gaussian Splashing[1]

Graphics for Robotics

- Task-driven
- Robustness
- Sufficiency for action
- Efficiency

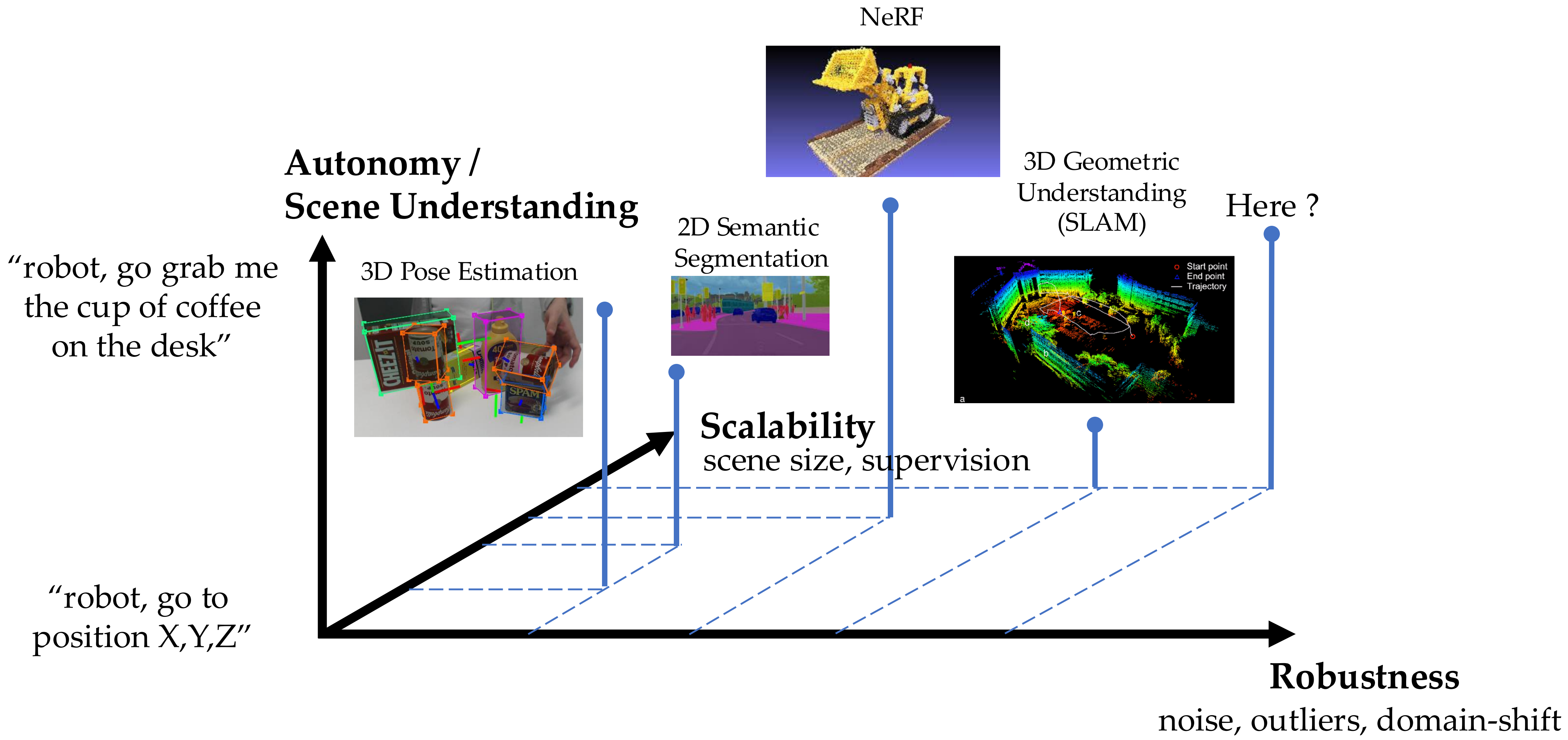


LiteReality[2]

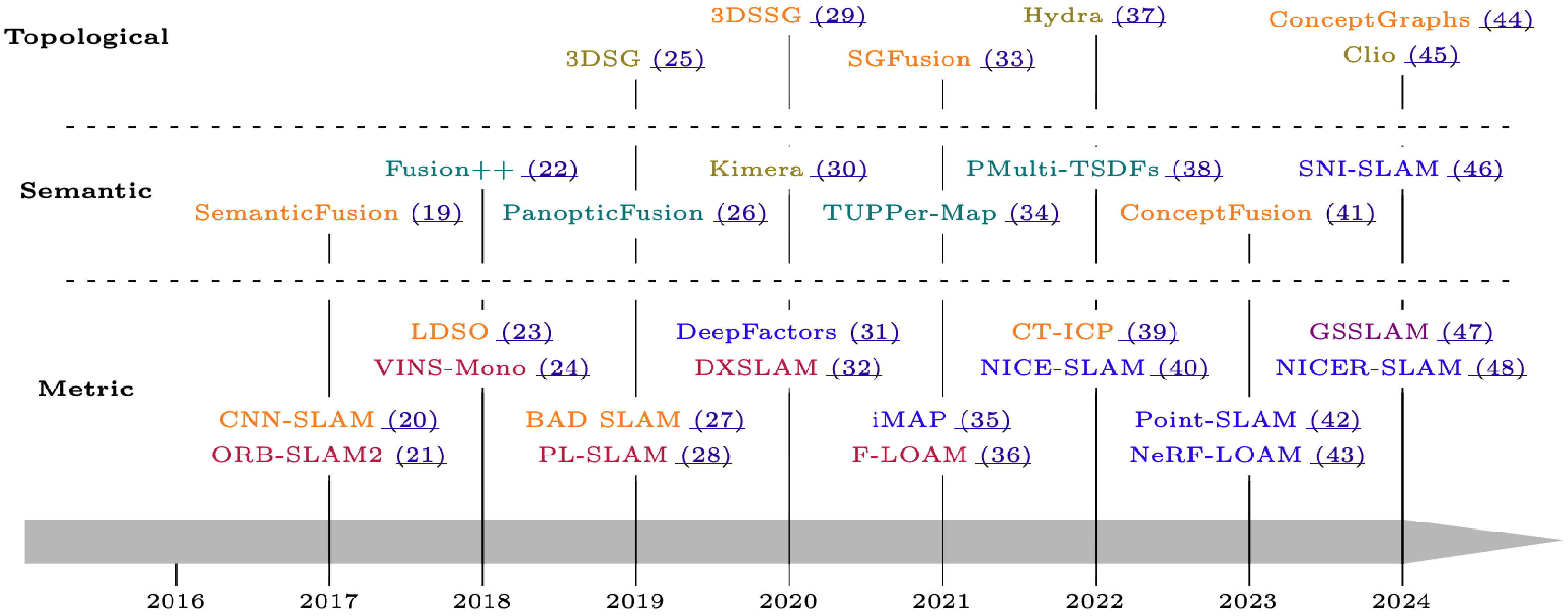
KAIST

[1] Feng, Yutao, et al. "Gaussian splashing: Unified particles for versatile motion synthesis and rendering." Proceedings of the Computer Vision and Pattern Recognition Conference. 2025.
[2] Huang, Zhening, et al. "LiteReality: Graphics-Ready 3D Scene Reconstruction from RGB-D Scans." arXiv preprint arXiv:2507.02861 (2025).

Research Map : Scene Representation for Robotics



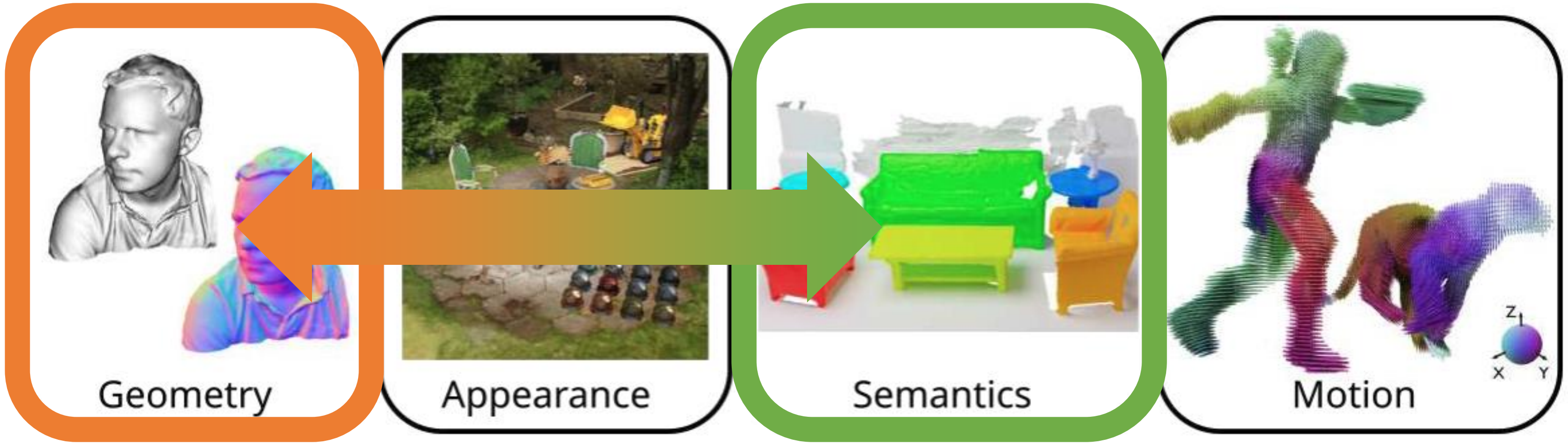
Research Trends by Year



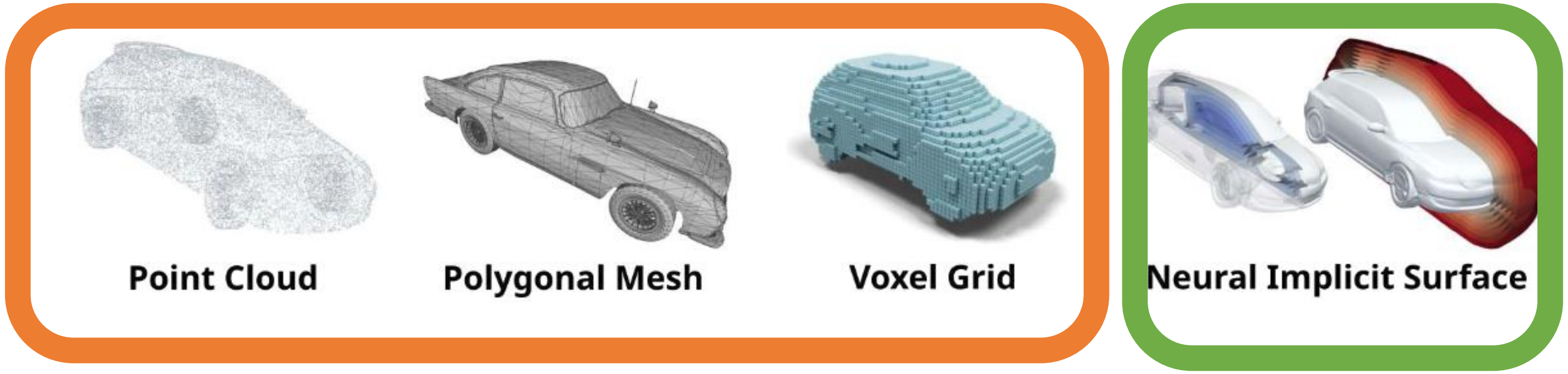
Metric Representation: Feature-based, Point-based, Mesh, Classical volumetric, Neural, Gaussian Splatting

[3] https://www.research-collection.ethz.ch/server/api/core/bitstreams/fbe04935-5a78-4ee1-8b50-e97895b4d29e/content

Challenges



Scene Properties[4]

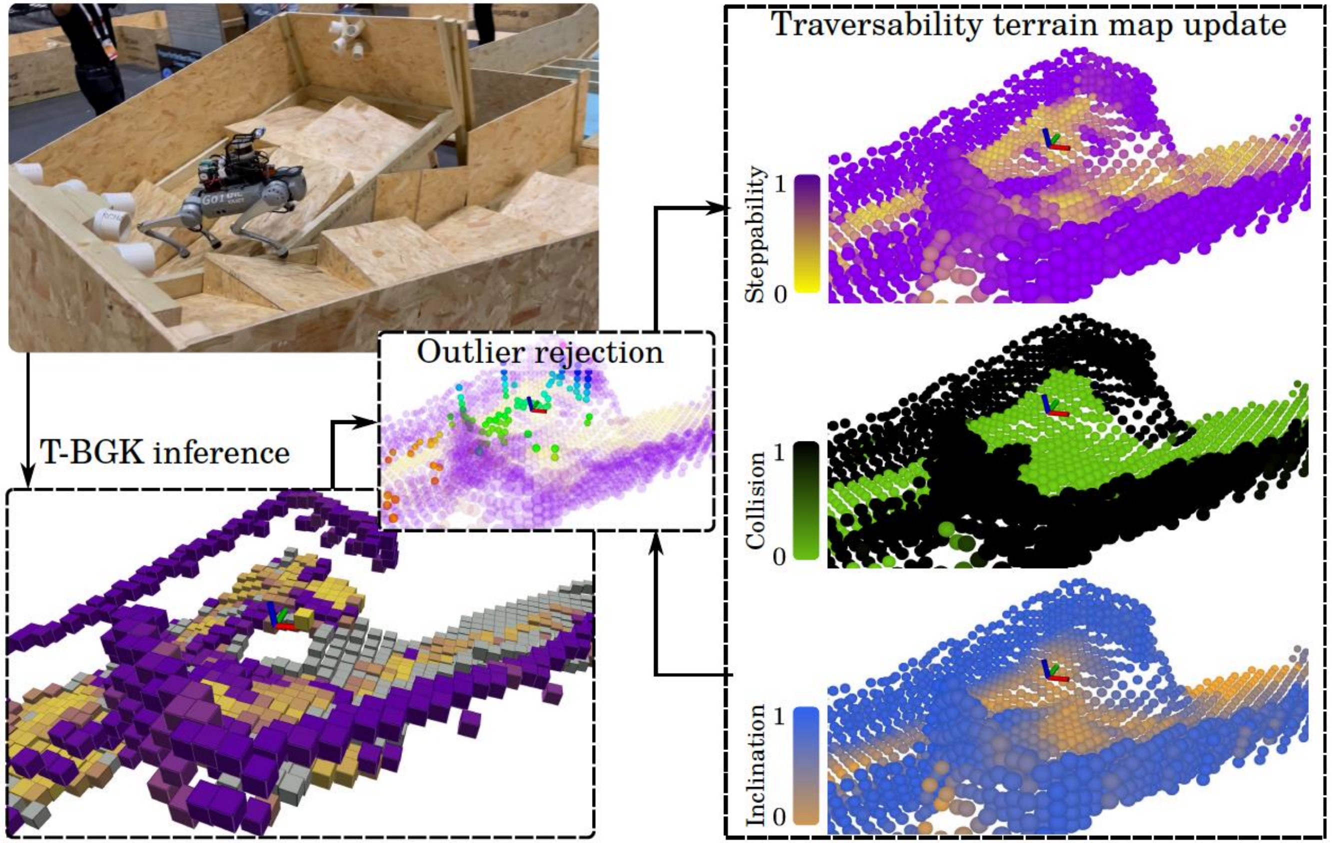


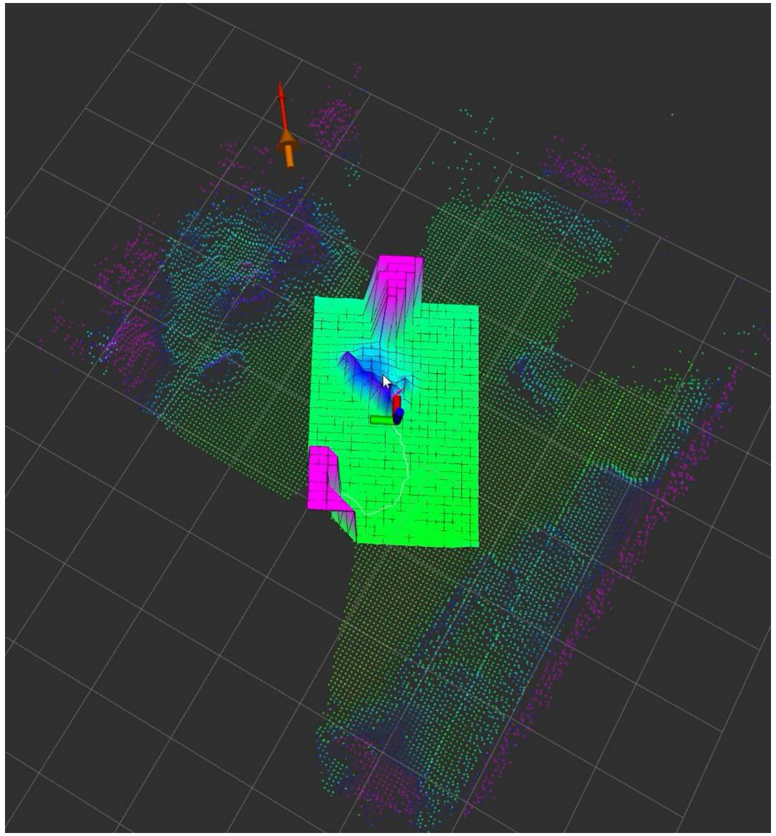
Scene Representations[4]

[4] https://www.research-collection.ethz.ch/server/api/core/bitstreams/ee428acb-f8c4-4c2f-bb6c-932d76f09ae6/content

Traditional Map

- Occupancy Grid



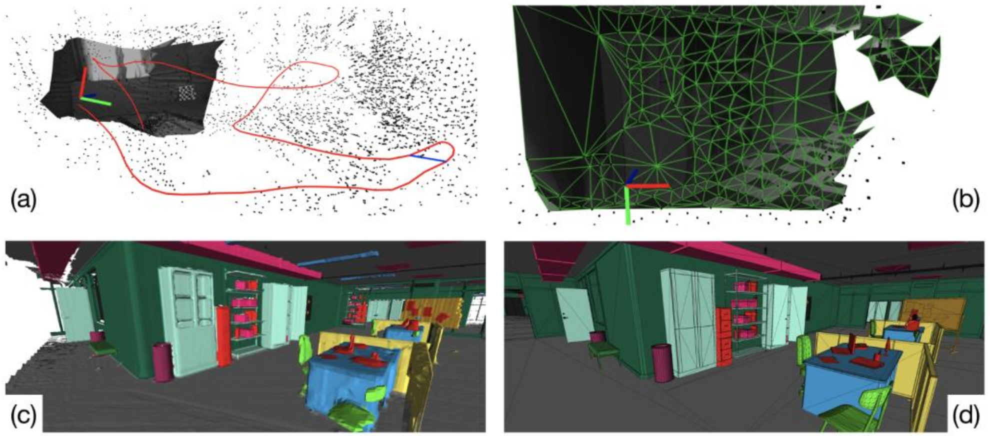


Navigation using 2.5 Dimension Terrain Map[5]

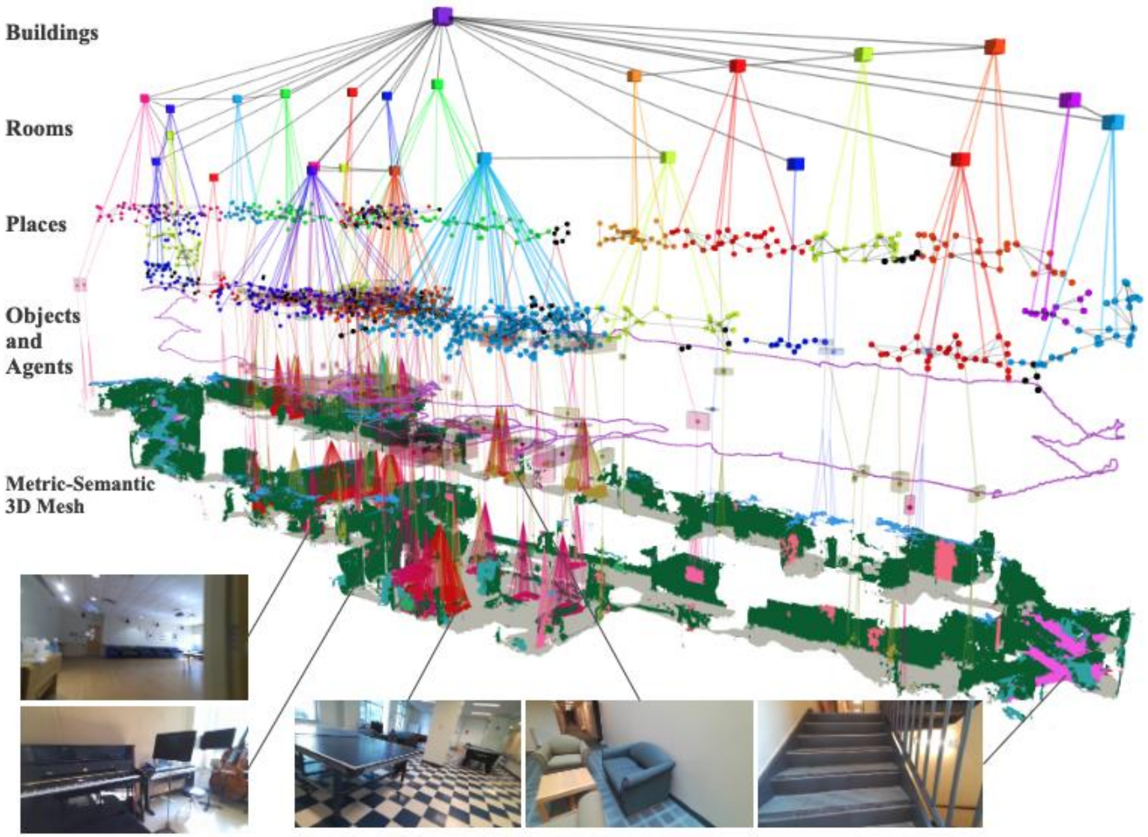
[5] Oh, Minho, et al. "TRIP: Terrain Traversability Mapping With Risk-Aware Prediction for Enhanced Online Quadrupedal Robot Navigation." arXiv preprint arXiv:2411.17134 (2024).

Scene Graphs

- Hierarchical Semantic Representation
  - Nodes : Rooms, Objects
  - Edges : Spatial relationships



Kimera[7]



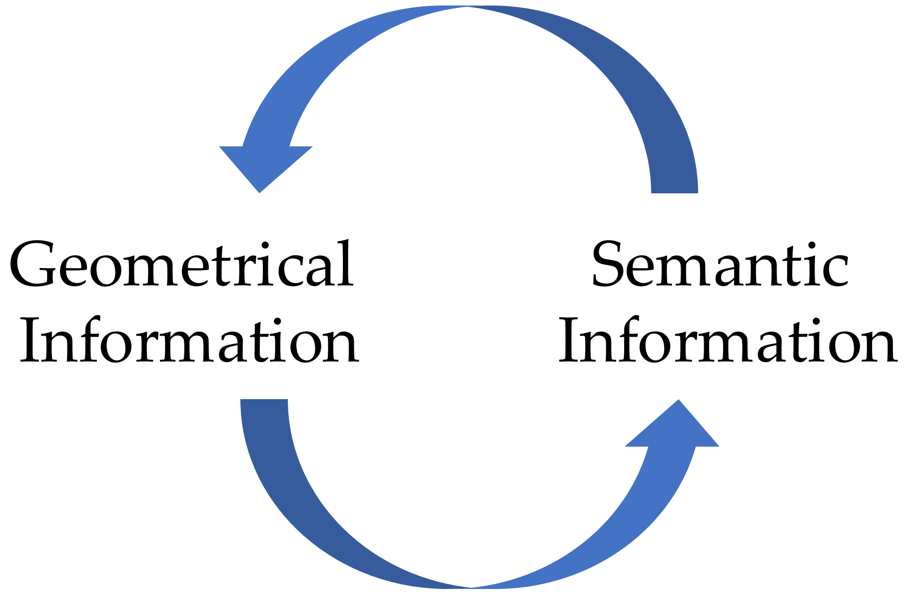
Hydra[8]

KAIST

[7] Rosinol, Antoni, et al. "Kimera: an open-source library for real-time metric-semantic localization and mapping." 2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2020.
[8] Hughes, Nathan, Yun Chang, and Luca Carlone. "Hydra: A Real-time Spatial Perception Engine for 3D Scene Graph Construction and Optimization." CoRR (2022).

Next-Gen Scene Representations

- Scene Graphs
- Representation Learning
- Semantic field, Kinematics field
- Neural Jacobian Fields
- Neural Radiance Fields (NeRF)
- Gaussian Splatting
- ... and so on



KAIST

NeRF-based Navigation (NeRF-Nav, 2022)

- **Assumption** : Pre-trained NeRF of the scene is given.

Idea #1

Position (p)
Direction (d)

NeRF

Color ($\boldsymbol{c}$)
Denstiy (ρ)

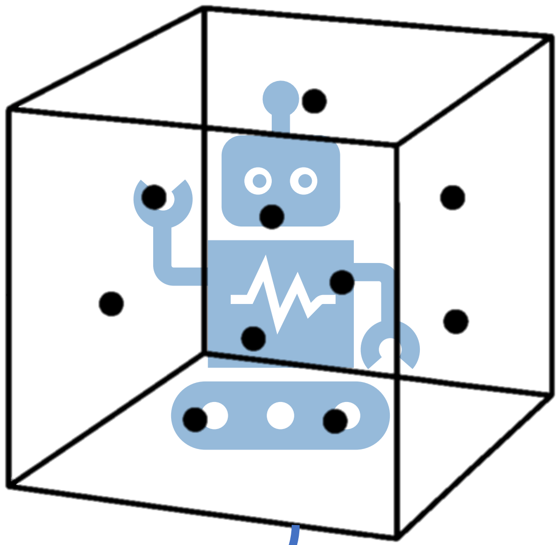
Physical Meaning of NeRF Density

differentiable probability of a given spatial point stopping a ray of light

⇒ Using NeRF Density $\rho(\boldsymbol{p})$ as a Collision Proxy.

Idea #2

Robot body ≒ finite set of points



3D grid of bounding box

Kwon, Obin and Park, Jeongho and Oh, Songhwai. "Renderable Neural Radiance Map for Visual Navigation". CVPR, 2023.

NeRF-based Navigation (NeRF-Nav, 2022)

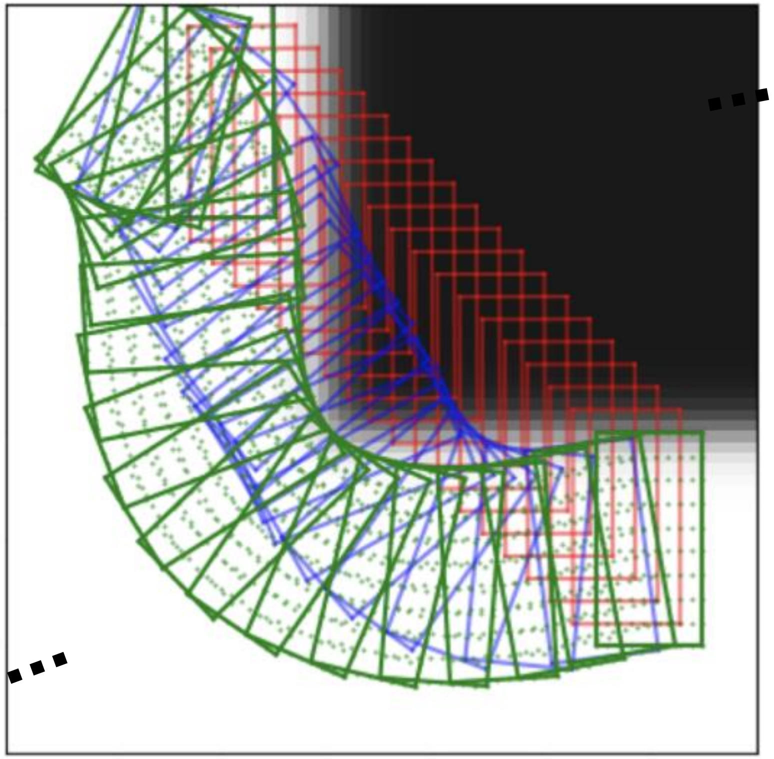
Collision probability | NeRF Density | Traveled distance of point b

$$p_t^{coll} \leq \sum_{b_t \in B} \rho(b_t) \cdot s(b_t)$$

Path cost = accumulated density along the path

Kwon, Obin and Park, Jeongho and Oh, Songhwai. "Renderable Neural Radiance Map for Visual Navigation". CVPR, 2023.

NeRF-based Navigation (NeRF-Nav, 2022)



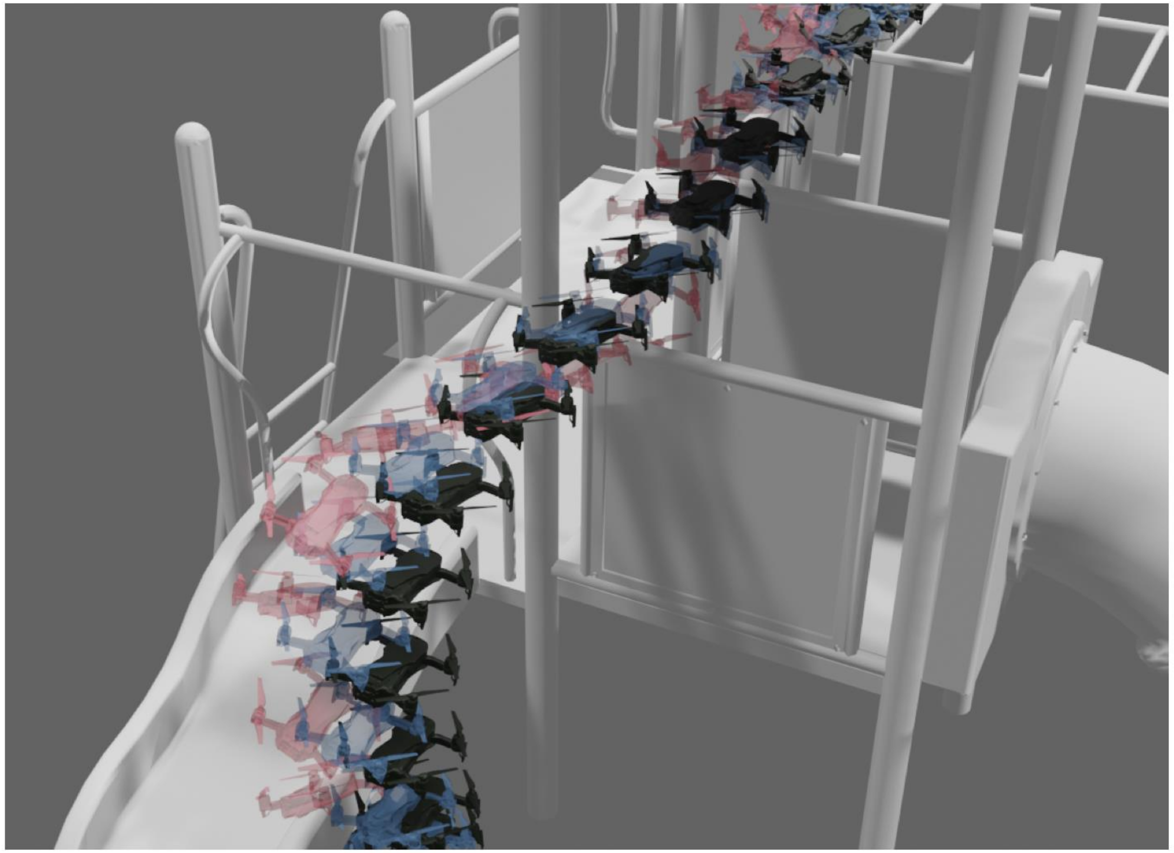
High density region

Low density region

Initial Trajectory
Intermediate Trajectory
Final Trajectory

Kwon, Obin and Park, Jeongho and Oh, Songhwai. "Renderable Neural Radiance Map for Visual Navigation". CVPR, 2023.

NeRF-based Navigation (NeRF-Nav, 2022)



Initial, partially-optimized, fully-optimized

Kwon, Obin and Park, Jeongho and Oh, Songhwai. "Renderable Neural Radiance Map for Visual Navigation". CVPR, 2023.

NeRF-based Navigation (NeRF-Nav, 2022)

- **Limitations**
  - Building a NeRF for a scene is Non-Trivial
  - Lacks explicit uncertainty quantification in geometry and appearance which is important in safe navigation
  - Not directly renderable as a global map

Kwon, Obin and Park, Jeongho and Oh, Songhwai. "Renderable Neural Radiance Map for Visual Navigation". CVPR, 2023.

NeRF-based Navigation (RNR-Map, 2023)

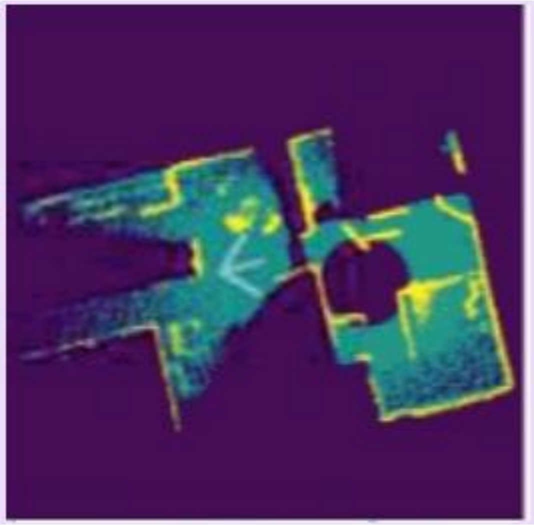
- **Traditional Maps**
  - Only preserves geometry information
  - Loss of visual details
  - Limitations on Image-Goal Navigation

- **Goal**
  - Build a map with both geometry and visual information.
  - How can we embed the visual information from a 3D environment to 2D form?

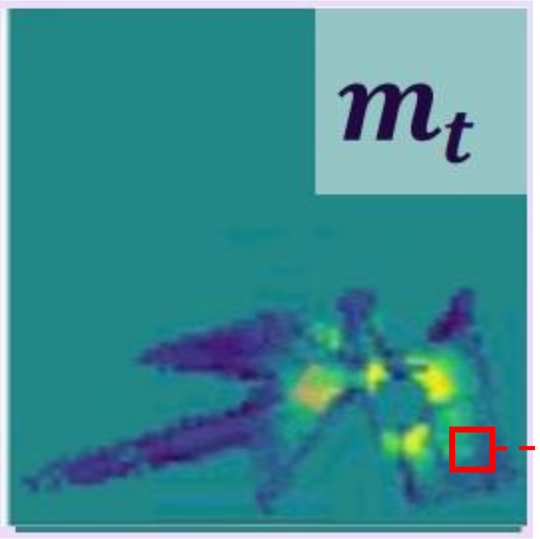
Kwon, Obin and Park, Jeongho and Oh, Songhwai. "Renderable Neural Radiance Map for Visual Navigation". CVPR, 2023.

NeRF-based Navigation (RNR-Map, 2023)

Occupancy Map



RNR-Map

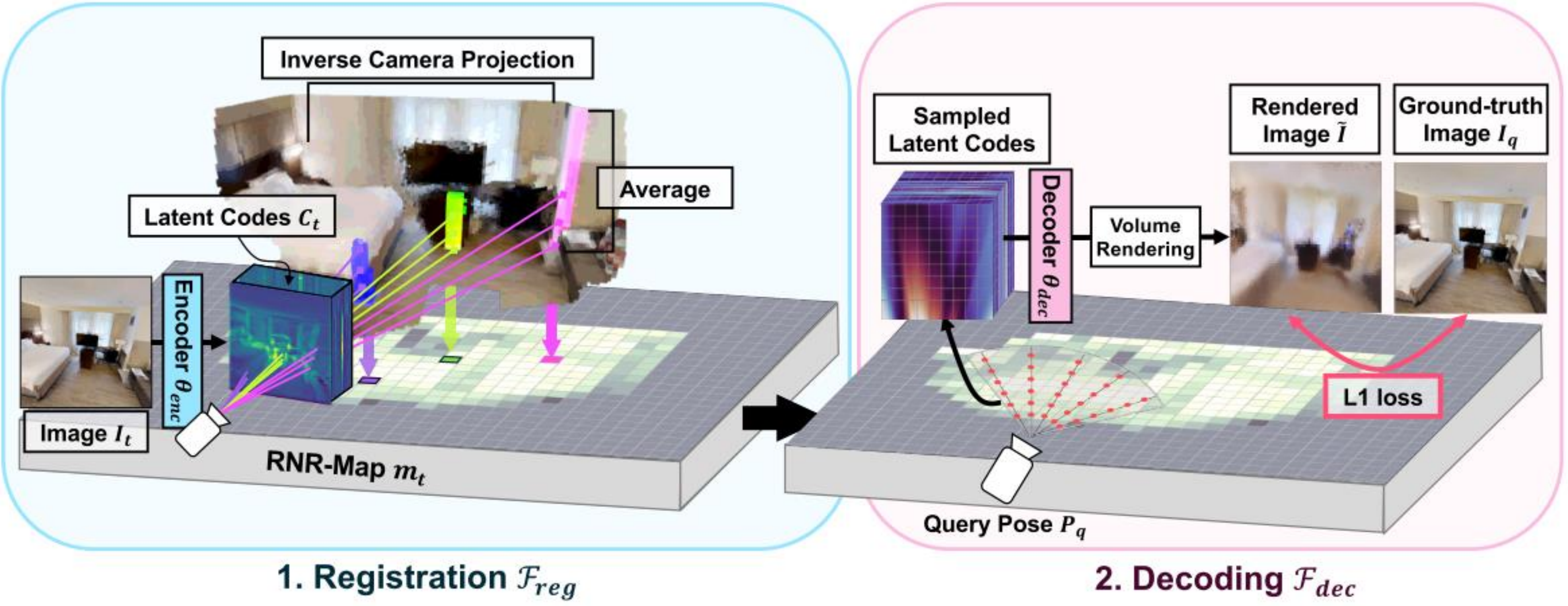


Stores a latent code derived from image observation at that location

Latent code → Conditional NeRF decoder → Rendered view

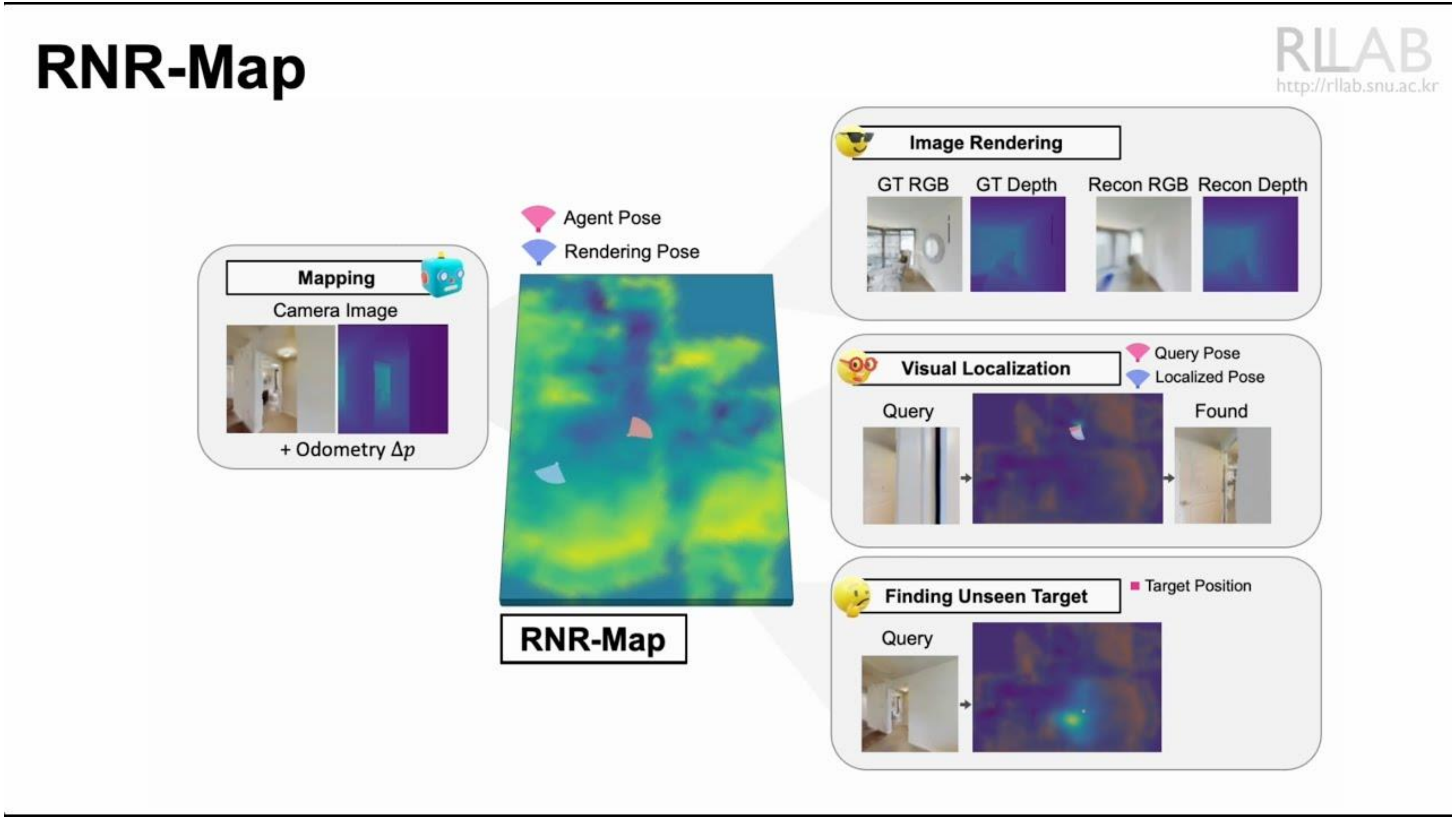
Kwon, Obin and Park, Jeongho and Oh, Songhwai. "Renderable Neural Radiance Map for Visual Navigation". CVPR, 2023.

NeRF-based Navigation (RNR-Map, 2023)



Kwon, Obin and Park, Jeongho and Oh, Songhwai. "Renderable Neural Radiance Map for Visual Navigation". CVPR, 2023.

NeRF-based Navigation (RNR-Map, 2023)



Kwon, Obin and Park, Jeongho and Oh, Songhwai. "Renderable Neural Radiance Map for Visual Navigation". CVPR, 2023.

NeRF-based Navigation (RNR-Map, 2023)

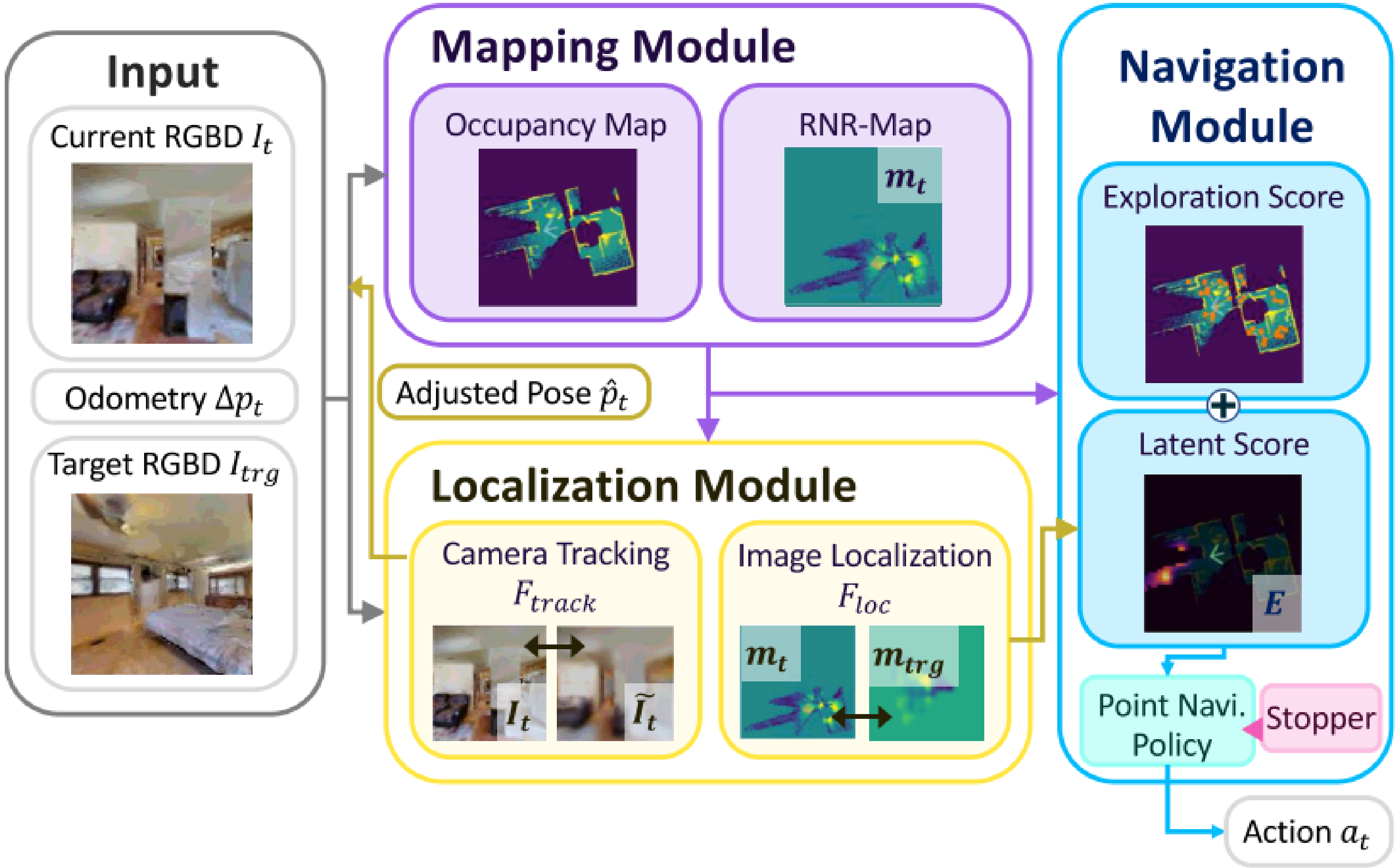


Figure 3. **Navigation System Overview.**

Kwon, Obin and Park, Jeongho and Oh, Songhwai. 「Renderable Neural Radiance Map for Visual Navigation」. CVPR, 2023.

NeRF-based Navigation (RNR-Map, 2023)

- Limitation
  - Implicit geometry & limited occupancy/traversability info
  - Latent code compression / information loss
  - Rendering quality and viewpoint generalisation

Kwon, Obin and Park, Jeongho and Oh, Songhwai. "Renderable Neural Radiance Map for Visual Navigation". CVPR, 2023.

Gaussian Splatting for Robot Navigation

- Challenges

1. Converting **visual information** into **metrically consistent 3D geometry**

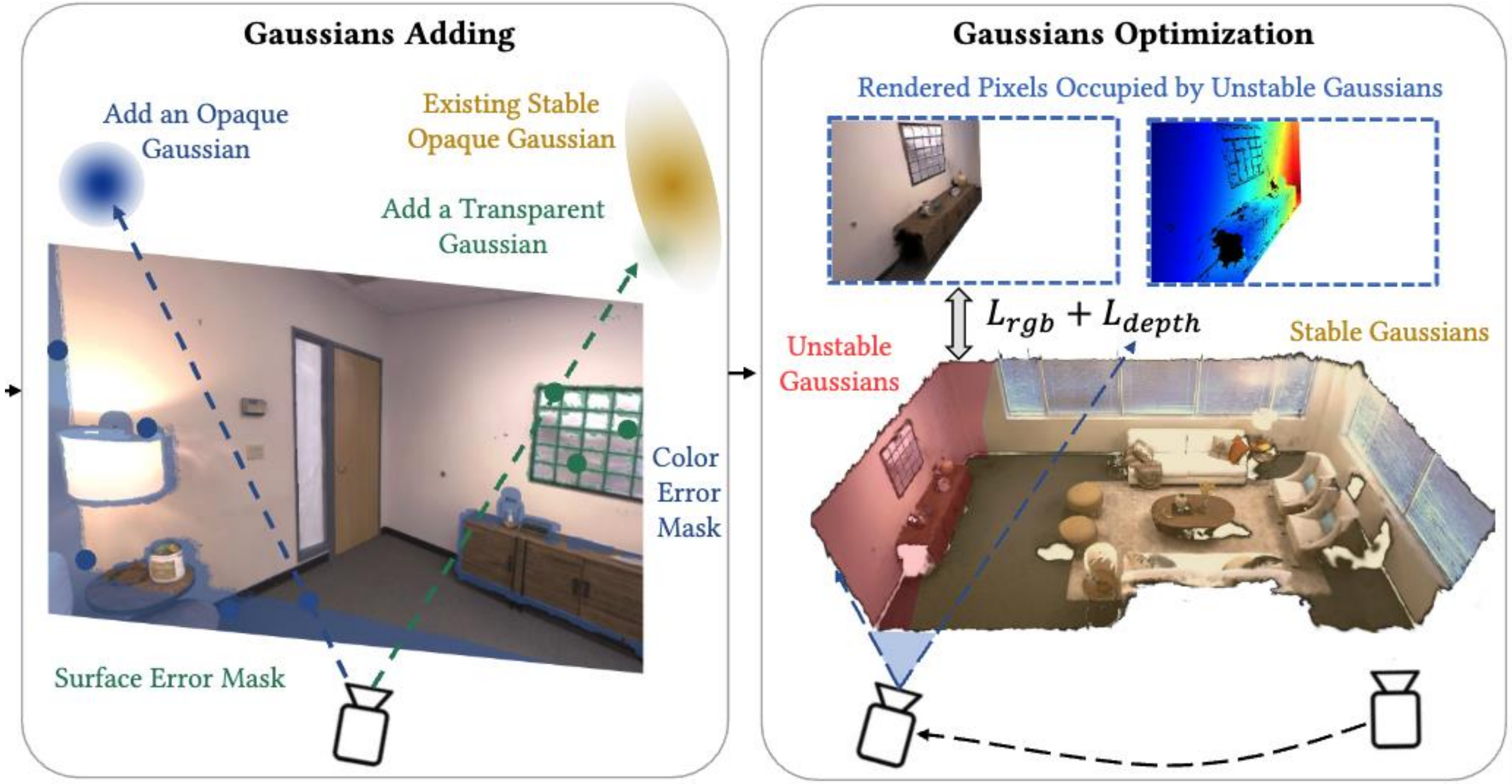
1-1) RTG-SLAM : Real-time 3D Reconstruction at Scale Using Gaussian Splatting

2. **Optimization** for real-time onboard inference with limited compute

2-1) Compressed 3D Gaussian Splatting for Accelerated Novel View Synthesis

Gaussian Splatting for Robot Navigation

- RTG-SLAM
  - 2024 SIGGRAPH
  - Real-time Gaussian SLAM



RTG-SLAM: Real-time 3D Reconstruction at Scale Using Gaussian Splatting

Zhexi Peng*
zhexipeng@zju.edu.cn
State Key Lab of CAD&CG
Zhejiang University
Hangzhou, China

Tianjia Shao*
tjshao@zju.edu.cn
State Key Lab of CAD&CG
Zhejiang University
Hangzhou, China

Yong Liu
Jingke Zhou
zilae@zju.edu.cn
zhoujk@zju.edu.cn
State Key Lab of CAD&CG
Zhejiang University
Hangzhou, China

Yin Yang
yin.yang@utah.edu
University of Utah
Salt Lake City, USA

Jingdong Wang
welleast@gmail.com
Baidu Research
Beijing, China

Kun Zhou†
kunzhou@acm.org
State Key Lab of CAD&CG
Zhejiang University
Hangzhou, China

Ours: 16.28 fps; 7.3GB memory | Co-SLAM: 8.77 fps; 17GB memory | Point-SLAM: 0.22 fps; 9.4GB memory

Figure 1: A hotel room (about $56.3m^2 \times 1.7m$) reconstructed by our system and the state-of-the-art NeRF-based RGBD SLAM techniques (Co-SLAM [Wang et al. 2023], Point-SLAM [Sandström et al. 2023]) without any post-processing. Compared with the state-of-the-art NeRF-based RGBD SLAM, our system achieves comparable high-quality reconstruction but with around twice the speed and half the memory cost, and shows higher realism in novel view synthesis.

ABSTRACT

We present Real-time Gaussian SLAM (RTG-SLAM), a real-time 3D reconstruction system with an RGBD camera for large-scale environments using Gaussian splatting. The system features a compact Gaussian representation and a highly efficient on-the-fly Gaussian optimization scheme. We force each Gaussian to be either opaque or nearly transparent, with the opaque ones fitting the surface and dominant colors, and transparent ones fitting residual colors. By rendering depth in a different way from color rendering, we let a single opaque Gaussian well fit a local surface region without the need of multiple overlapping Gaussians, hence largely reducing the memory and computation cost. For on-the-fly Gaussian optimization, we explicitly add Gaussians for three types of pixels per frame: newly observed, with large color errors, and with large depth errors. We also categorize all Gaussians into stable and unstable ones, where the stable Gaussians are expected to well fit previously observed RGBD images and otherwise unstable. We only optimize the unstable Gaussians and only render the pixels occupied by unstable Gaussians. In this way, both the number of

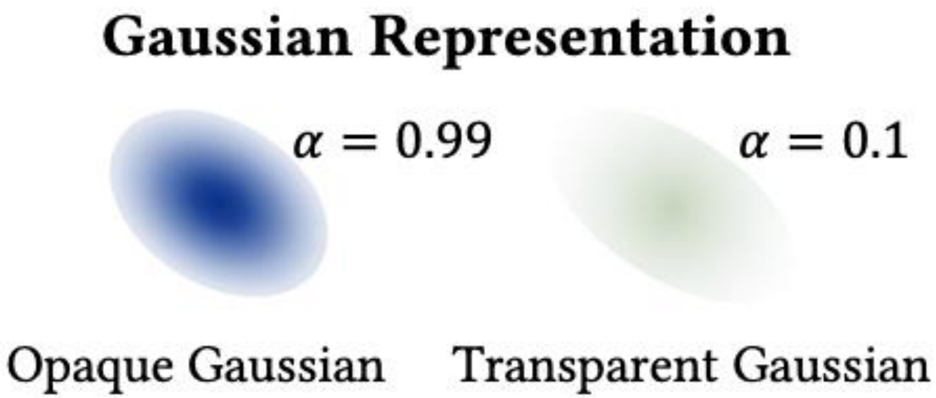
*Joint first authors
†Corresponding author

SIGGRAPH Conference Papers '24, July 27-August 1, 2024, Denver, CO, USA

ACM ISBN 979-8-4007-0525-0/24/07...$15.00
https://doi.org/10.1145/3641519.3657455

Gaussian Splatting for Robot Navigation

- RTG-SLAM
  - 2024 SIGGRAPH
  - Real-time Gaussian SLAM

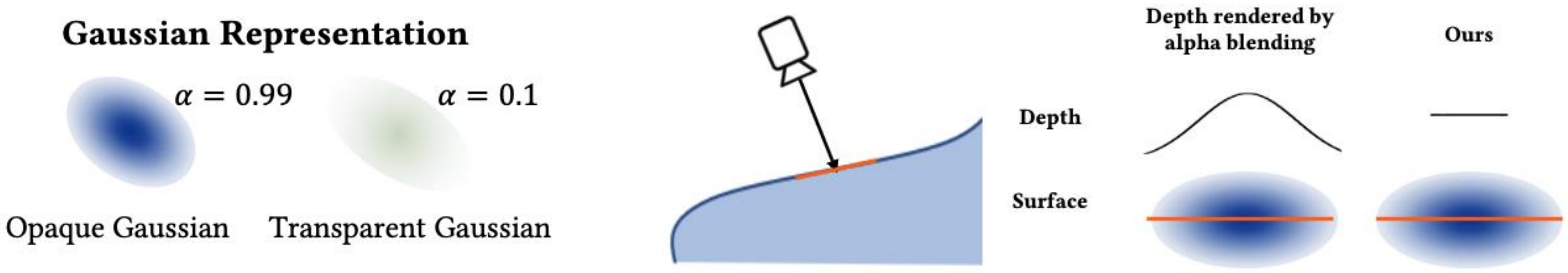


Type	α	Function	Depth Rendering	Effect
Opaque	0.99	Fit surface & dominant color	Ellipsoid disc intersection	Compact geometry
Transparent	0.1	Fit residual color only	Ignored in depth	Color refinement

Compact Gaussian Representation in RTG-SLAM

Gaussian Splatting for Robot Navigation

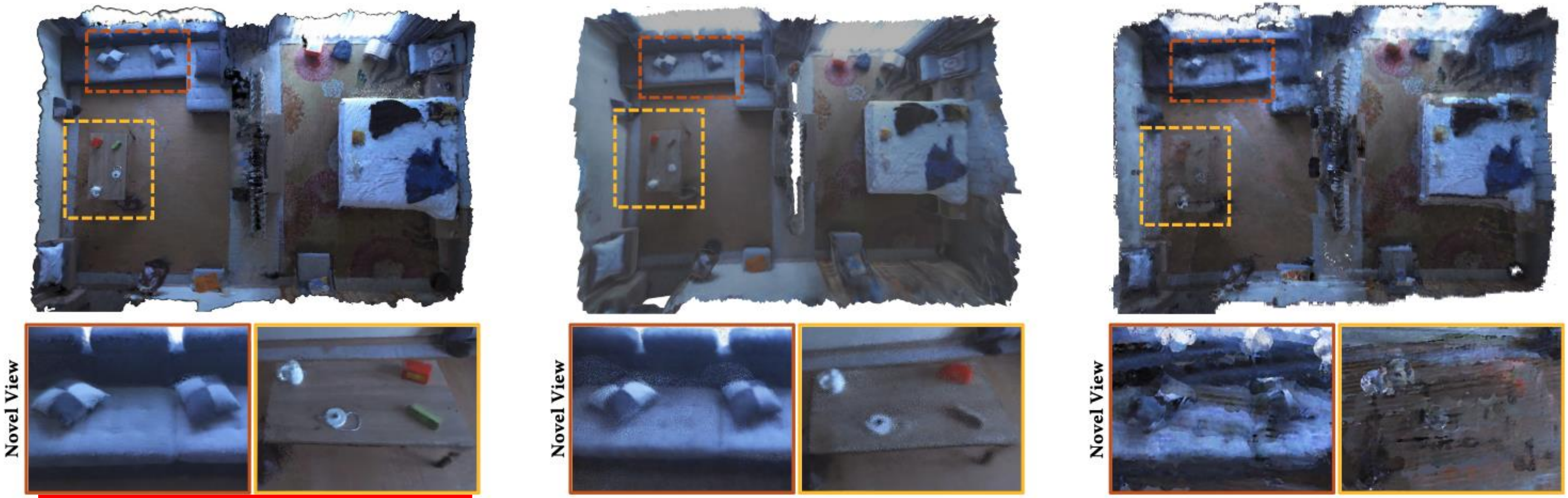
- RTG-SLAM
  - 2024 SIGGRAPH
  - Real-time Gaussian SLAM



Compact Gaussian Representation in RTG-SLAM

Gaussian Splatting for Robot Navigation

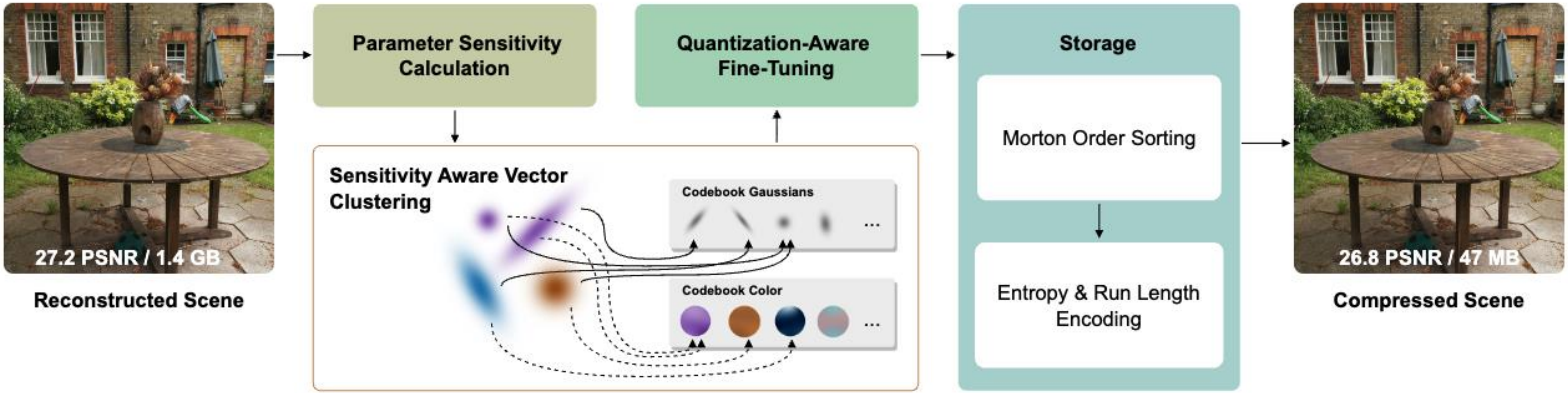
- RTG-SLAM
  - 2024 SIGGRAPH
  - Real-time Gaussian SLAM



Ours: 16.28 fps; 7.3GB memory   Co-SLAM: 8.77 fps; 17GB memory   Point-SLAM: 0.22 fps; 9.4GB memory

Gaussian Splatting for Robot Navigation

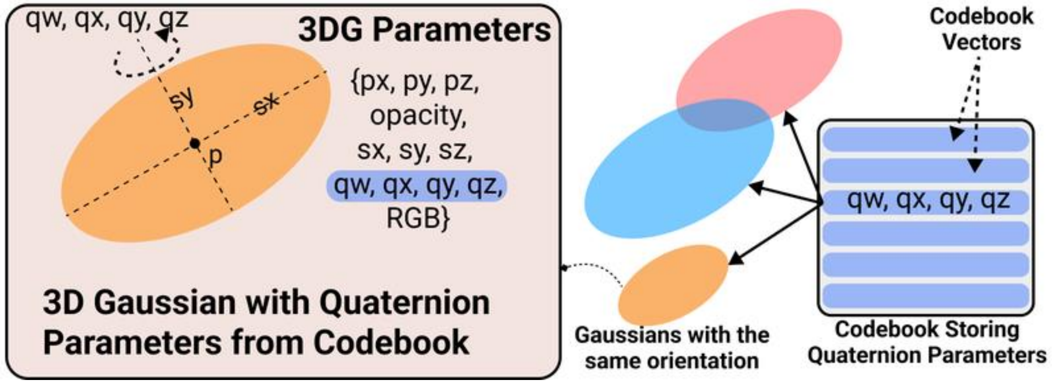
- Compressed 3D Gaussian Splatting for Accelerated Novel View Synthesis
  - 2024 CVPR
  - Sensitivity-aware Vector Clustering
  - Quantization-aware Fine Tuning



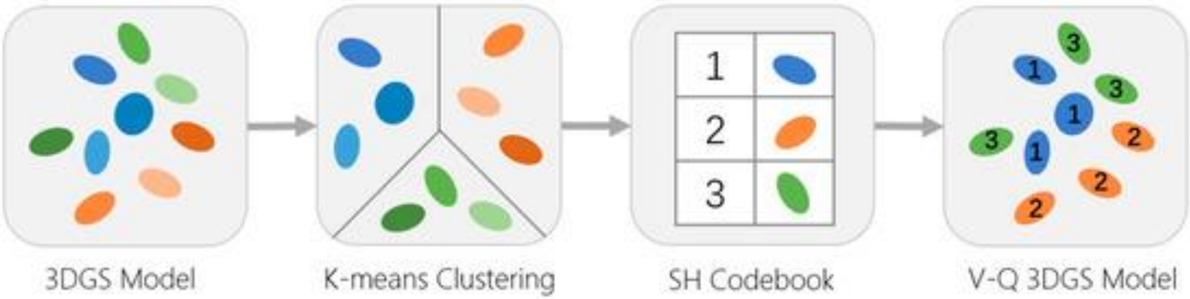
Proposed Compression Pipeline

Gaussian Splatting for Robot Navigation

- Compressed 3D Gaussian Splatting for Accelerated Novel View Synthesis
  - 2024 CVPR
  - Sensitivity-aware Vector Clustering
    - Remove redundant gaussians
    - compression without losing visual quality
  - Sensitivity measure
    - How much does changing a parameter affect image quality?
    - High-sensitivity parameters: Preserved (added directly to codebook)
    - Low-sensitivity parameters: **Clustered together** (compressed)



Codebook for GS[9]



GS Clustering[10]

KAIST

[9] ContraGS: Codebook-Condensed and Trainable Gaussian Splatting for Fast, Memory-Efficient Reconstruction
[10] High-Fold 3D Gaussian Splatting Model Pruning Method Assisted by Opacity

Summary

- **Graphics for Robot Navigation**
  - Actionable information for machines
  - Different goals require different representations
- **Navigation Requirements**
  - Geometric accuracy for collision-free planning
  - Semantic understanding for command-driven tasks
  - Real-time performance on mobile platforms
- **Two Key Challenges for GS in Navigation**
  - Visual→ Metric geometry conversion
  - Real-time onboard inference optimization
- To be continued in Team Project ....

Q&A

Quiz