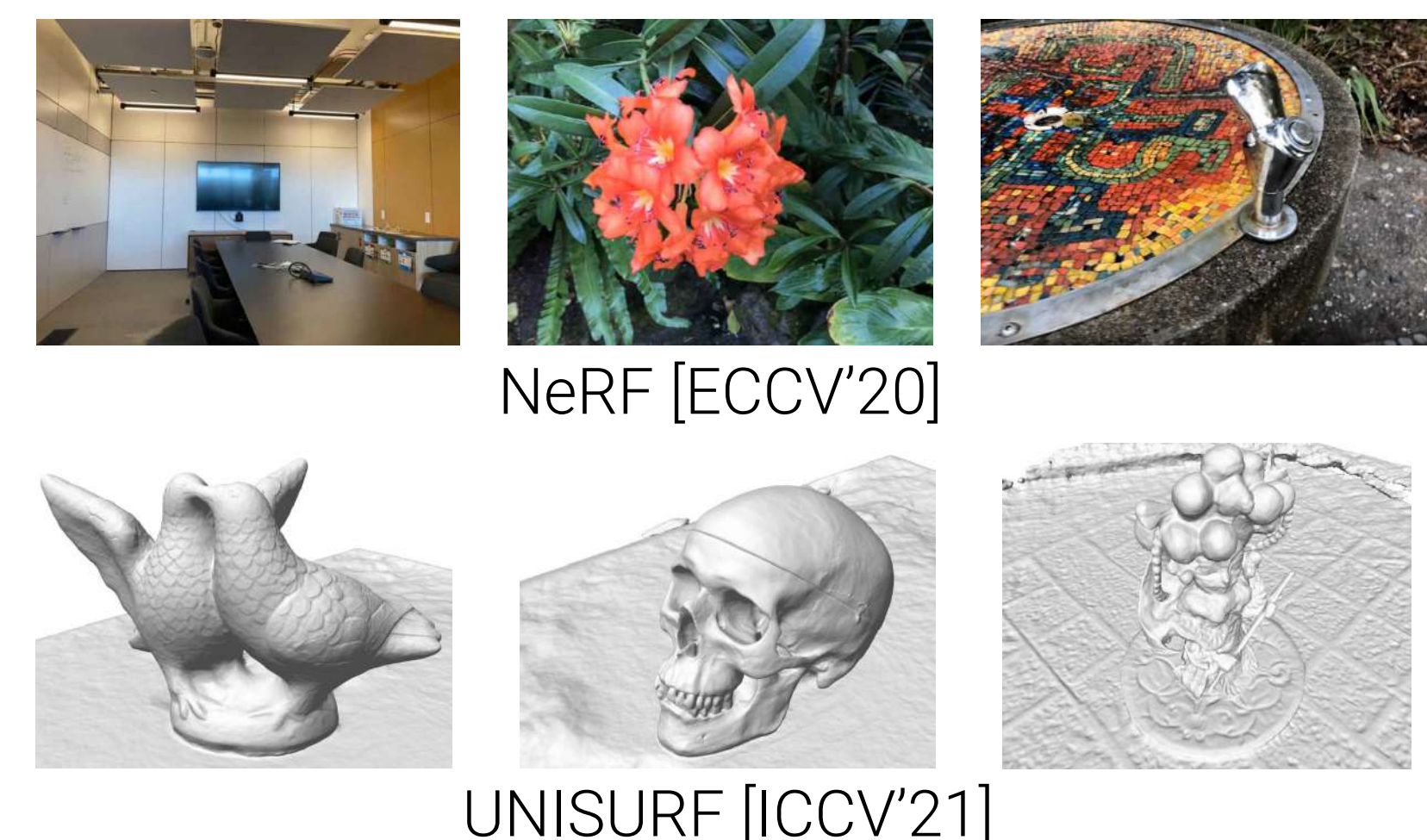
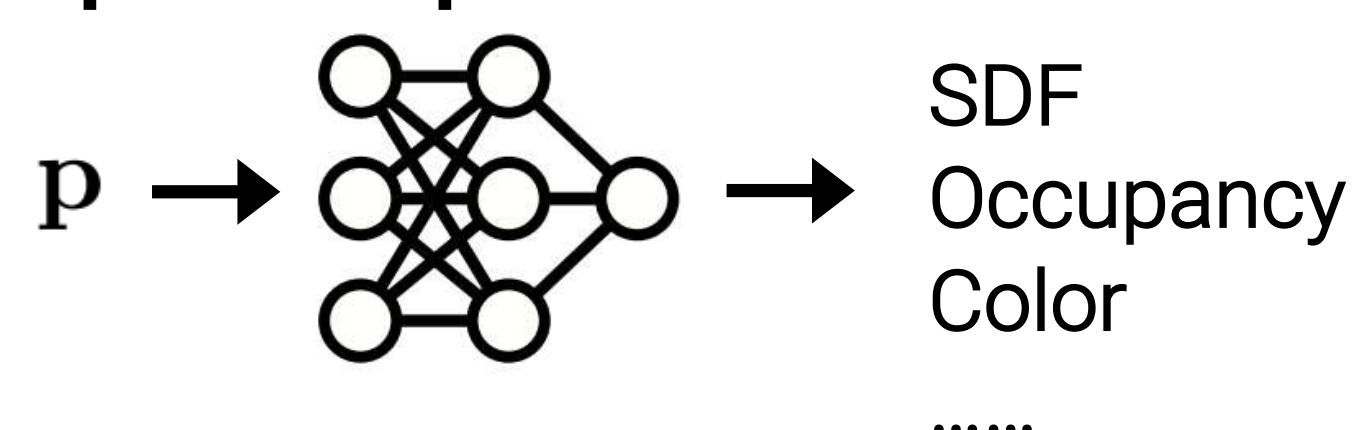




1. Motivation

Neural implicit representations are awesome!



- + Simple
- + HQ shape & color
- Offline
- Camera poses needed

2. Goal

Input: an RGB-D sequence for large-scale indoor scenes



Output: On-the-fly learn neural implicit surfaces & camera poses

iMAP [ICCV'21]

A single MLP

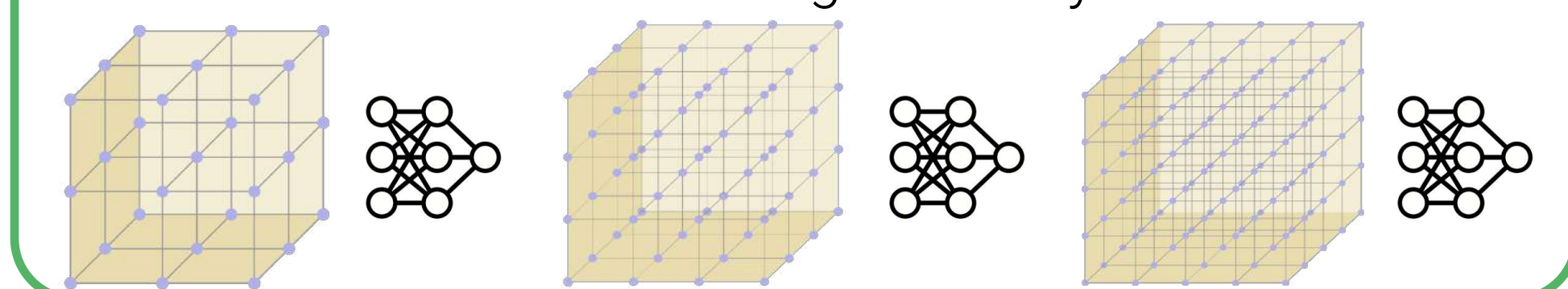


Predicted Poses
 GT Poses

- Fail to scale up to larger scenes
- Catastrophic forgetting
- Slow convergence

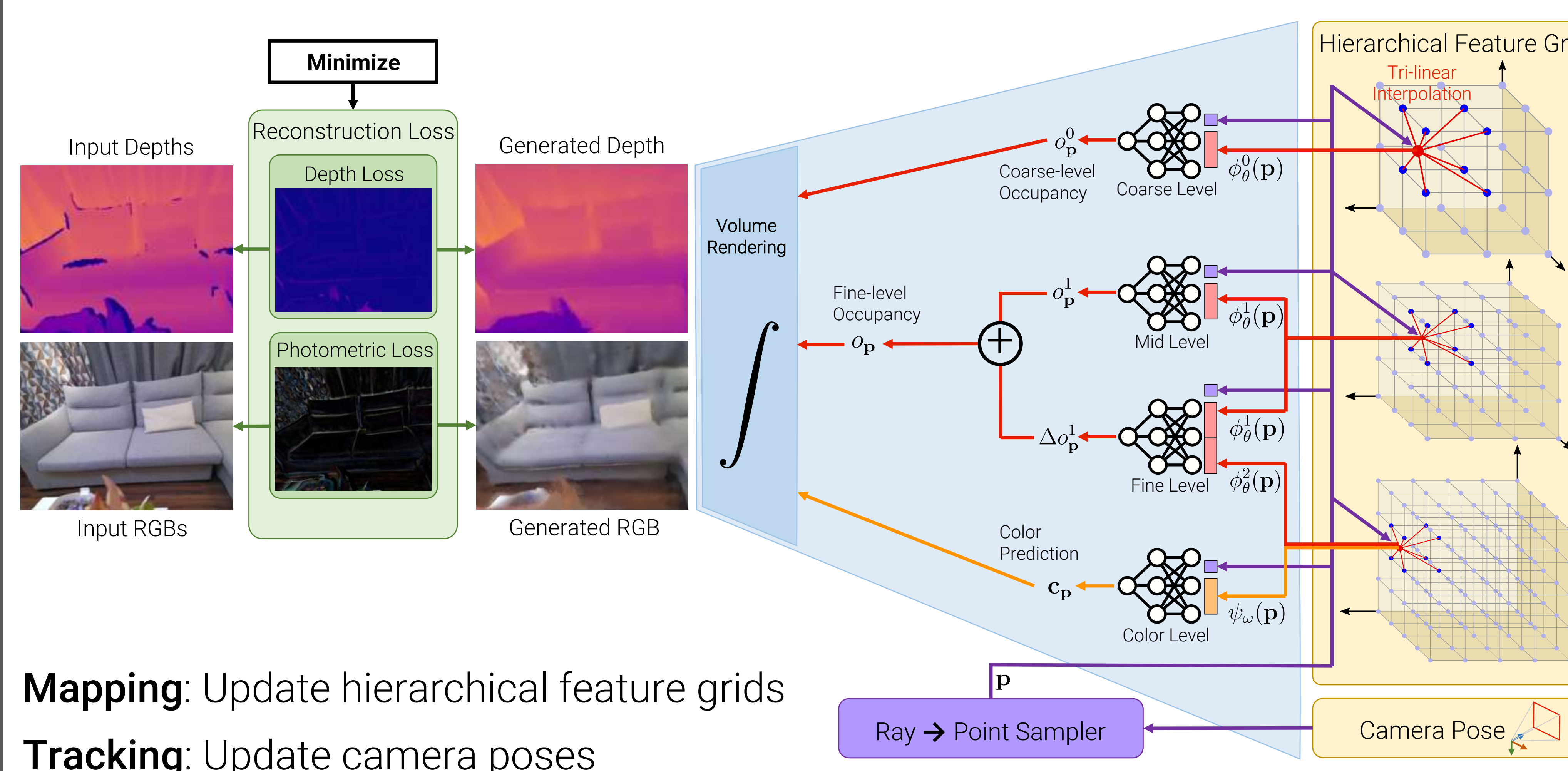
NICE-SLAM

Multi-res feature grids + tiny MLPs



- Scalable to large-scale scenes
- Local update → No forgetting
- Fast convergence

3. Pipeline

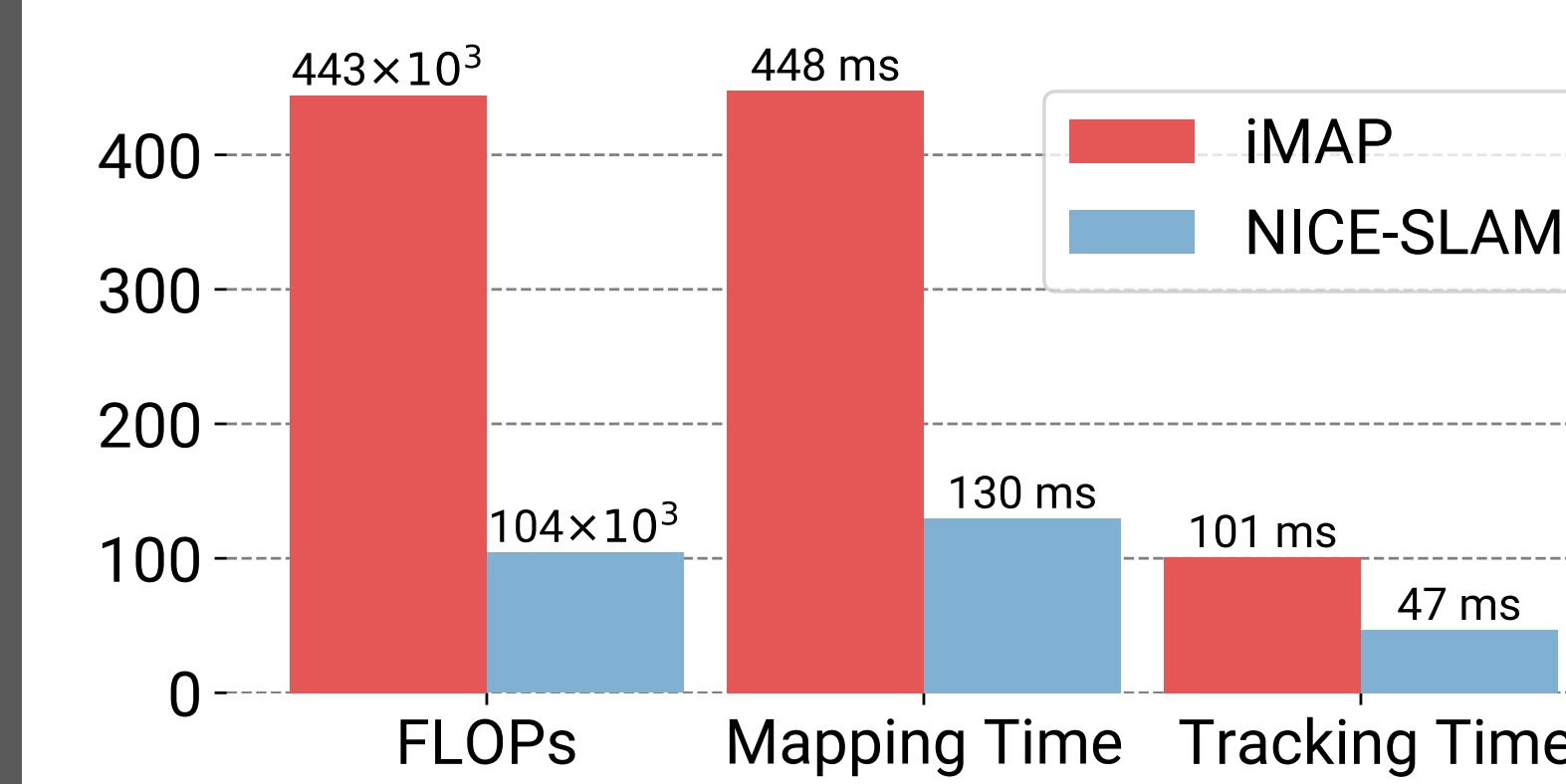


Mapping: Update hierarchical feature grids

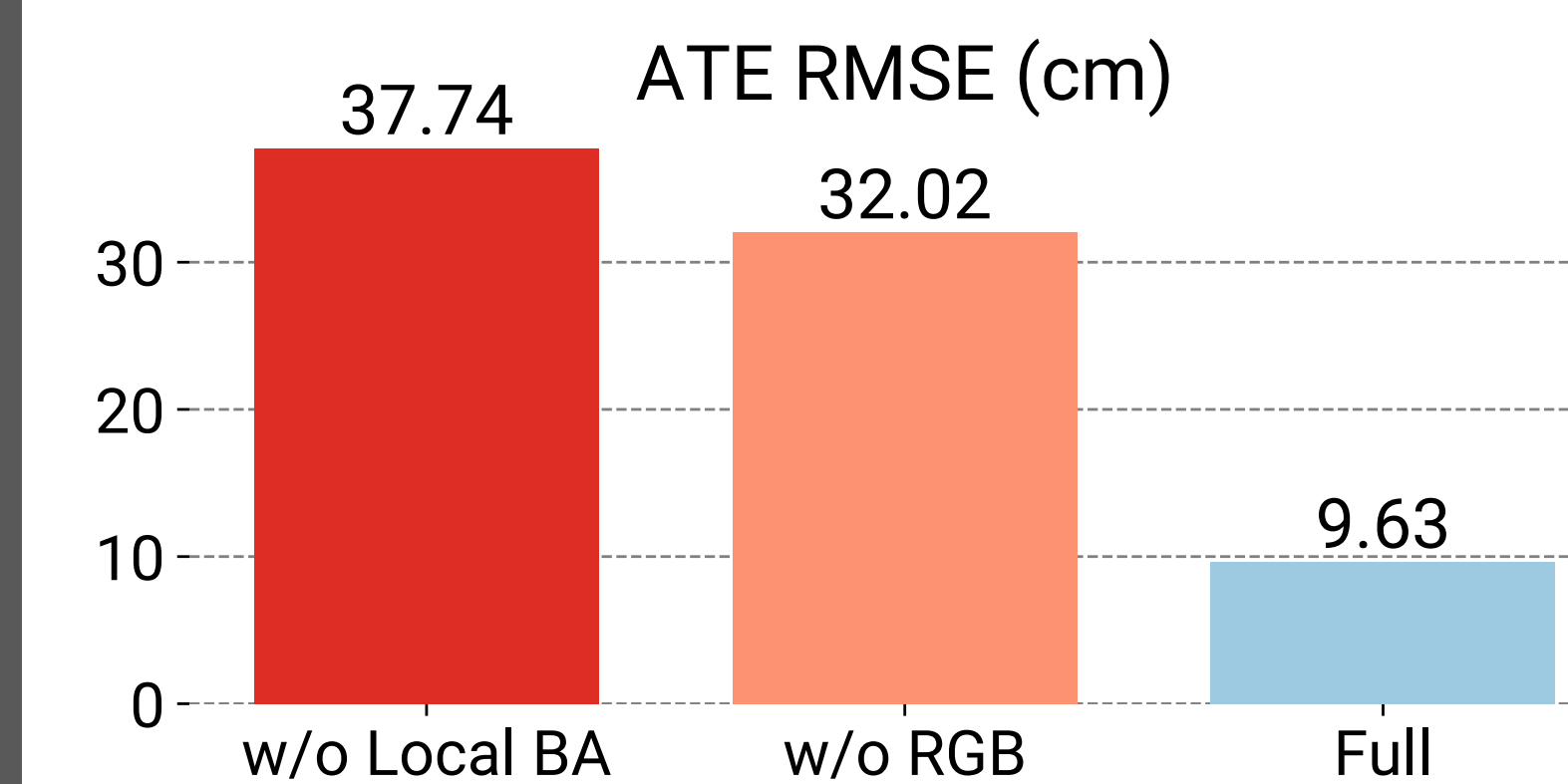
Tracking: Update camera poses

5. More Studies

Computation & Runtime



Ablation Study



4. Results

Replica

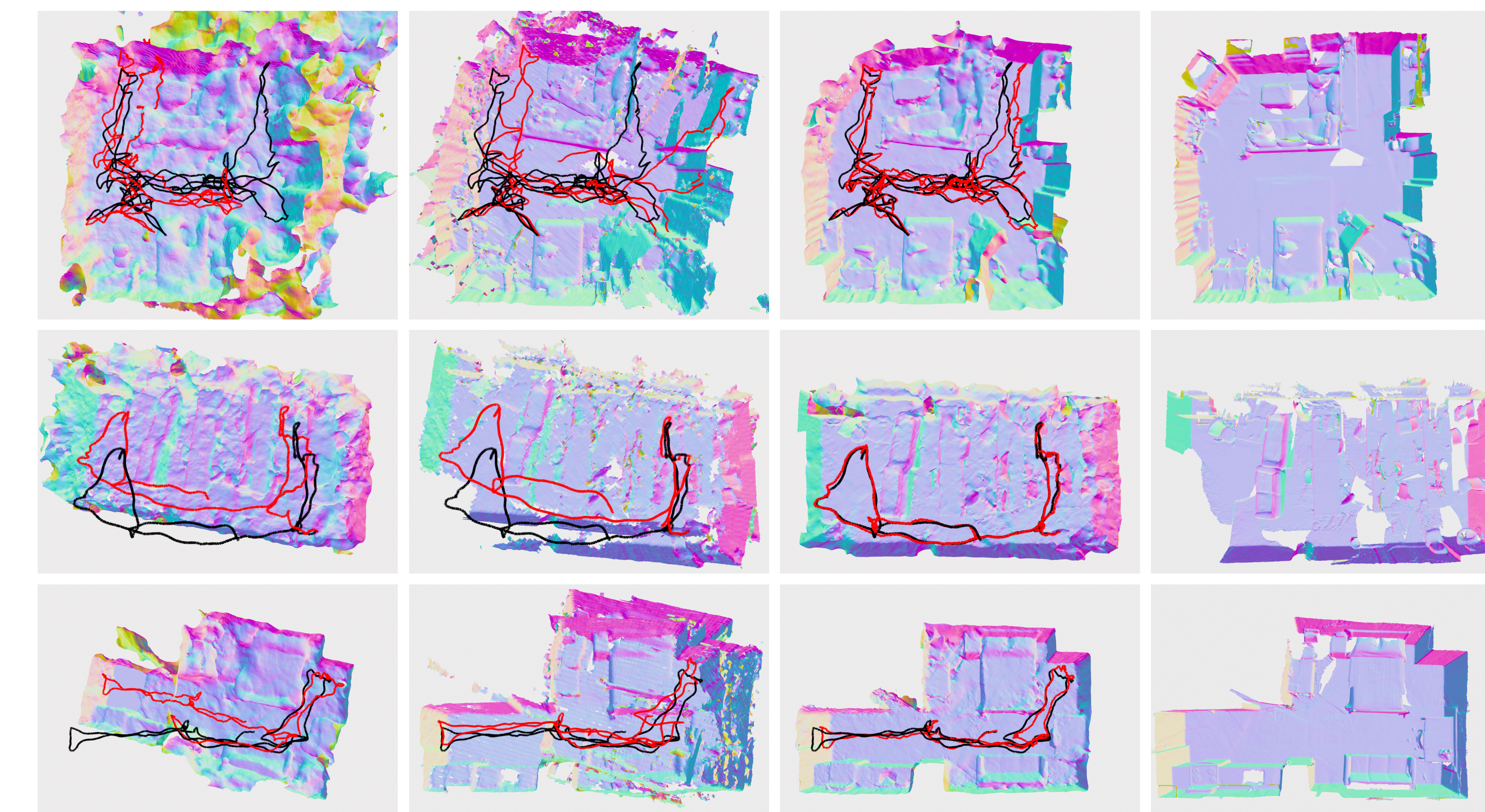


iMAP*
 (Our re-implementation)

NICE-SLAM

GT Mesh

ScanNet



iMAP*
 (Our re-implementation)

DI-Fusion

NICE-SLAM

GT Mesh