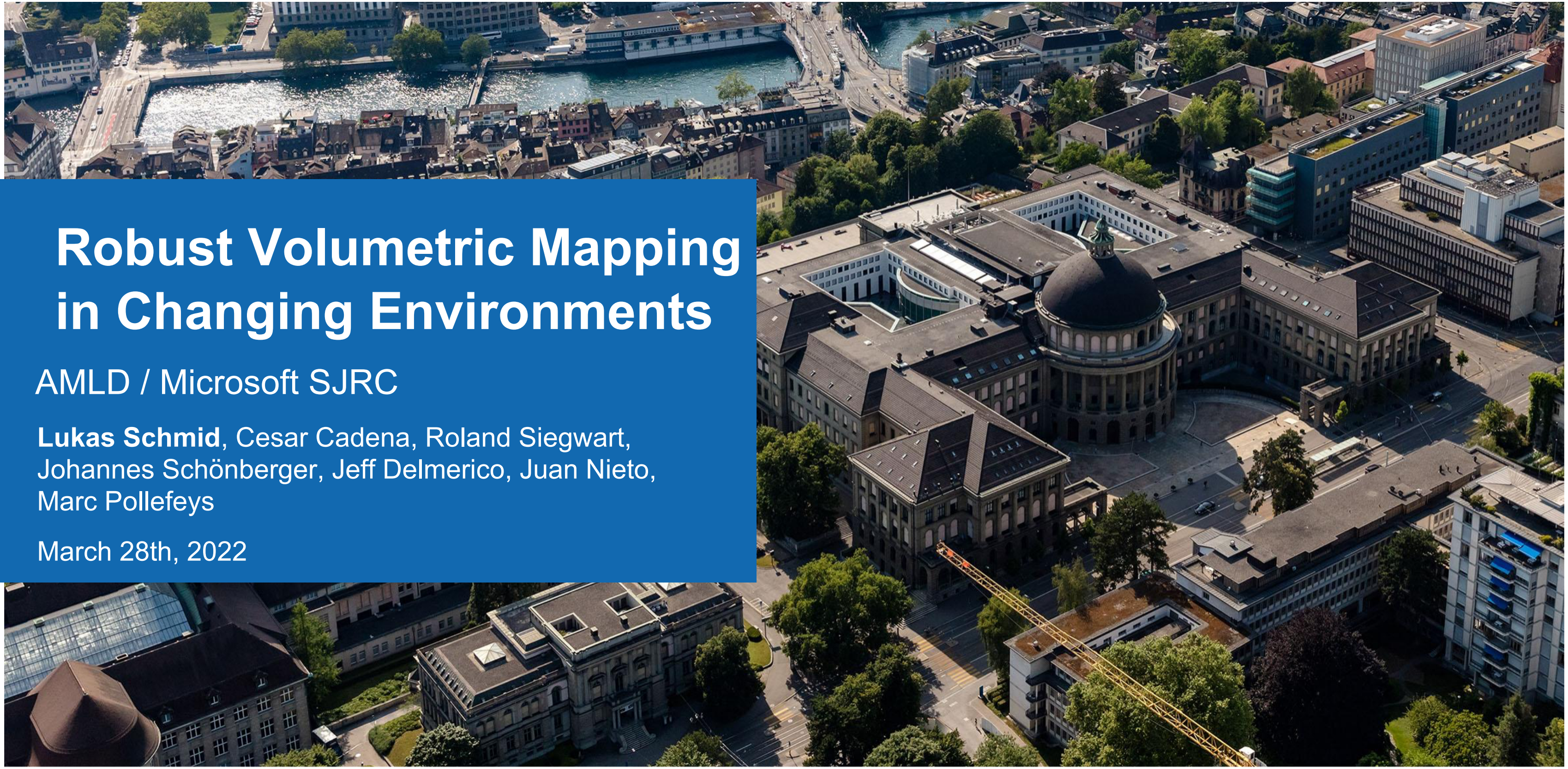


Robust Volumetric Mapping in Changing Environments

AMLD / Microsoft SJRC

Lukas Schmid, Cesar Cadena, Roland Siegwart, Johannes Schönberger, Jeff Delmerico, Juan Nieto, Marc Pollefeys

March 28th, 2022

Sponsored by

Robust Volumetric Mapping in Changing Environments

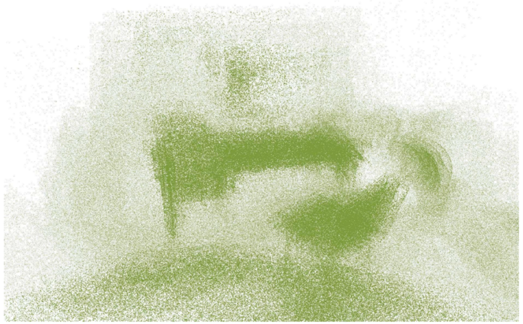
Goal:

Enable robots to act intelligently in environments that are shared with other agents over longer periods of time.

Robust Volumetric Mapping in Changing Environments

Challenges

Pose Errors



Accumulated noisy point clouds

TSDF reconstruction

Long-term Scene Changes

Day 1 Day 2 Changes

Monolithic Map

Robust Volumetric Mapping in Changing Environments

How can the map representation help?

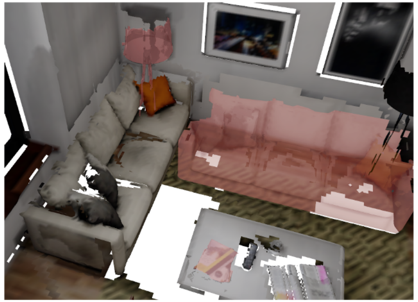
Pose Errors

TSDF reconstruction

[S. Lionar*, **L. Schmid***, C. Cadena, R. Siegwart, A. Cramariuc, "NeuralBlox: Real-Time Neural Representation Fusion for Robust Volumetric Mapping." 3DV 2021]

Long-term Scene Changes

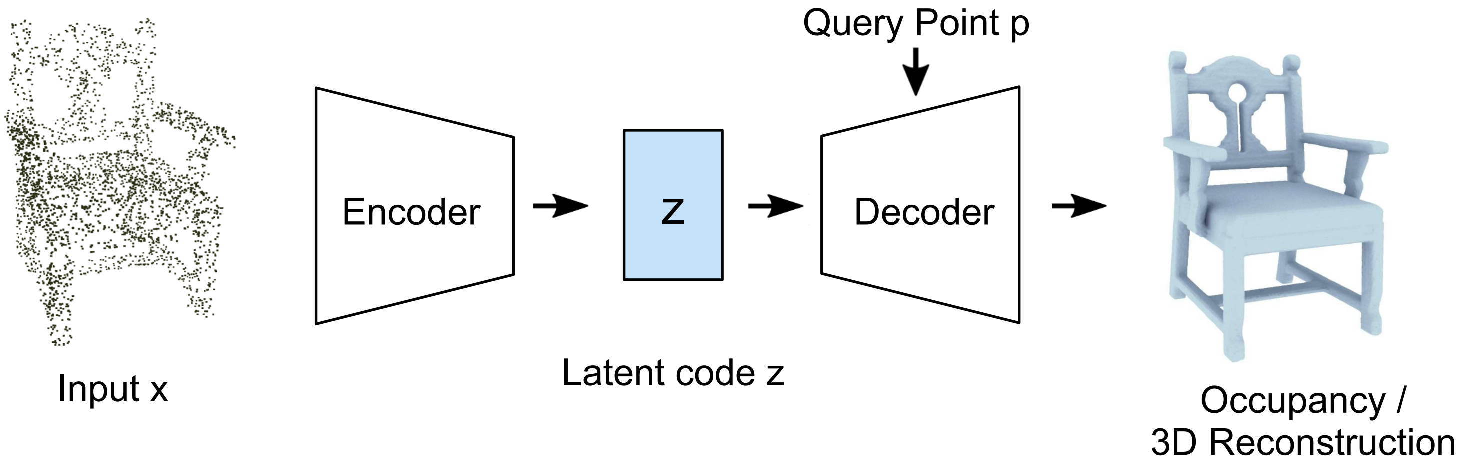
Monolithic Map



[**L. Schmid**, J. Delmerico, J. Schönberger, J. Nieto, M. Pollefeys, R. Siegwart, C. Cadena, "Panoptic Multi-TSDFs: a Flexible Representation for Online Multi-resolution Volumetric Mapping and Long-term Dynamic Scene Consistency." ICRA 2022]

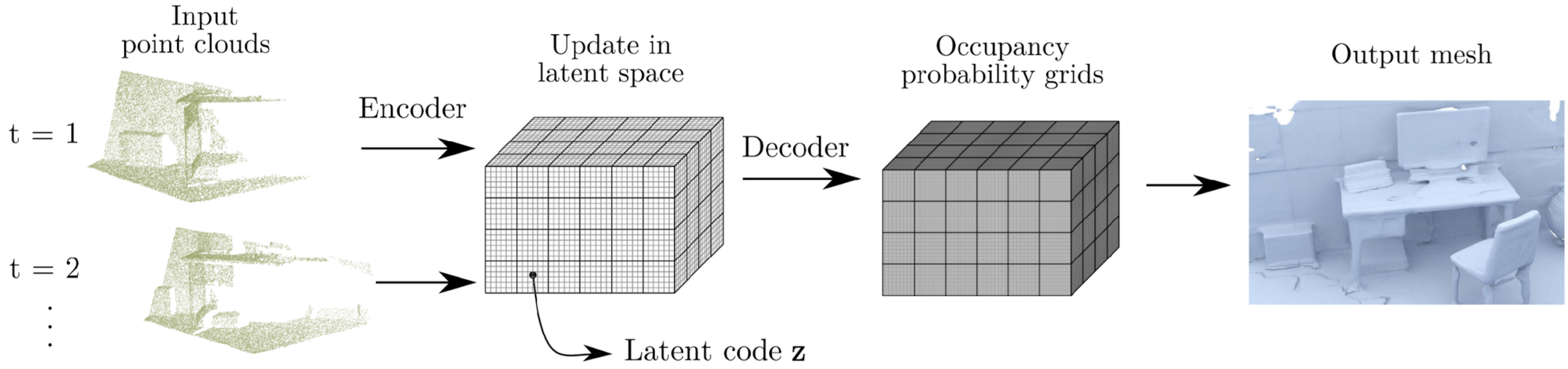
NeuralBlox: Real-Time Neural Representation Fusion

- Recent progress in neural shape representation.
- Can we incrementally build a neural implicit map?
- Can we leverage geometric context for more robust fusion?



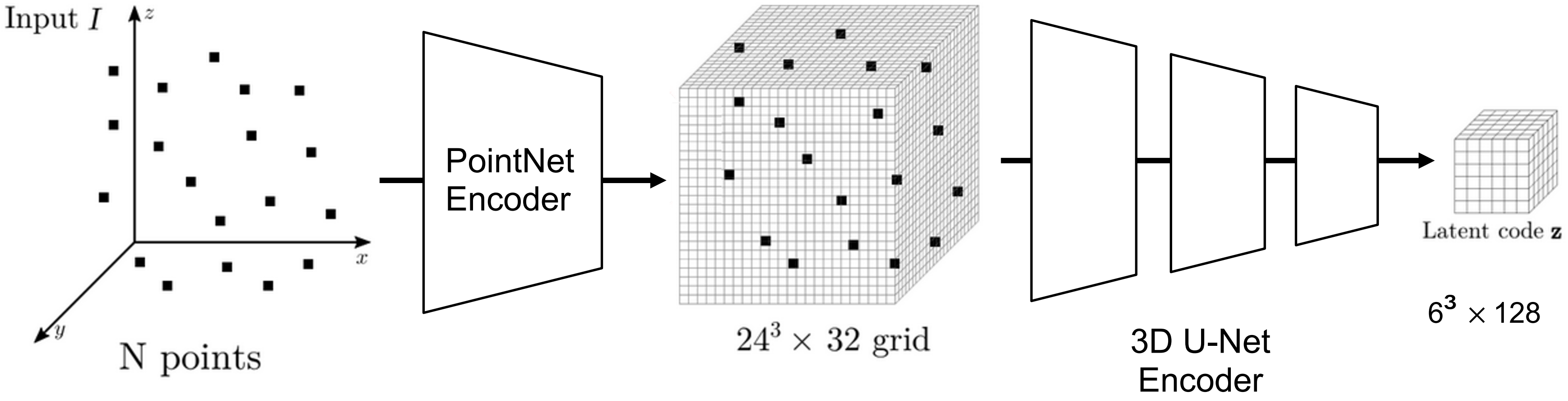
NeuralBlox: Real-Time Neural Representation Fusion

Pipeline Overview



NeuralBlox: Real-Time Neural Representation Fusion

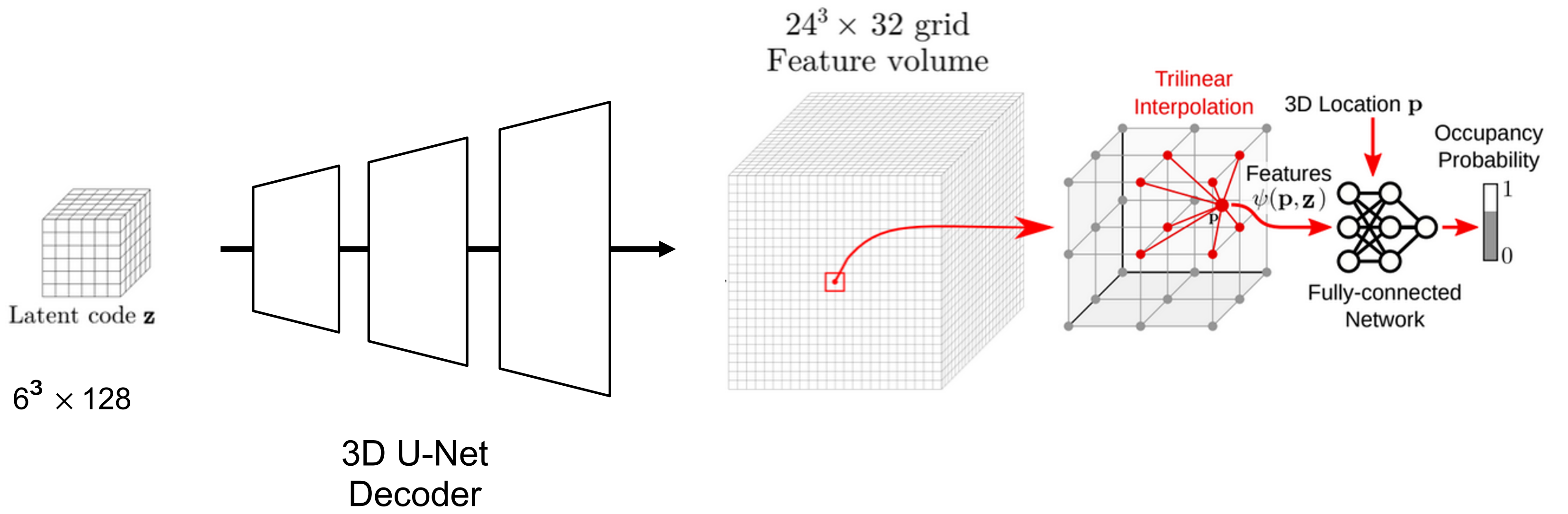
- Decouple problem, train shape representation independently on ShapeNet.



[S. Peng et al. "Convolutional Occupancy Networks." ECCV 2020]

NeuralBlox: Real-Time Neural Representation Fusion

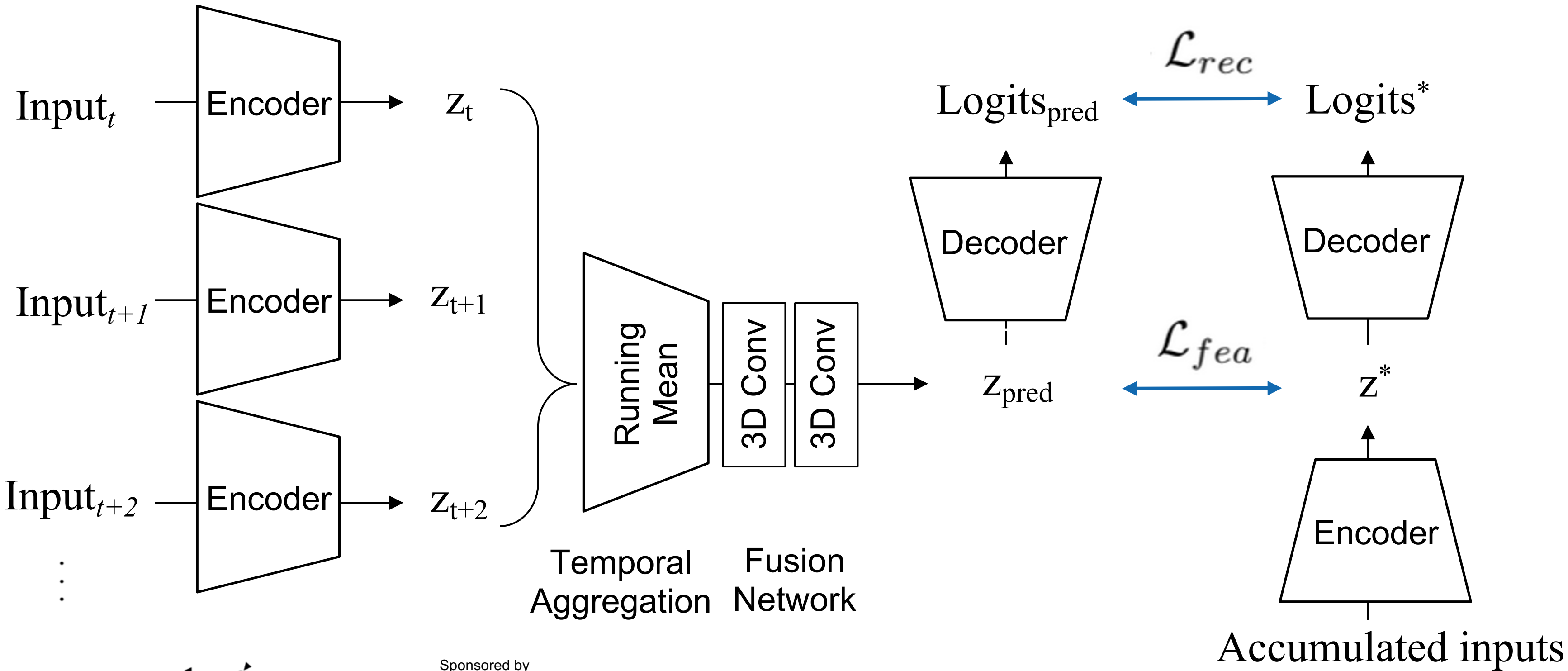
- Decouple problem, train shape representation independently on ShapeNet.



[S. Peng et al. "Convolutional Occupancy Networks." ECCV 2020]

NeuralBlox: Real-Time Neural Representation Fusion

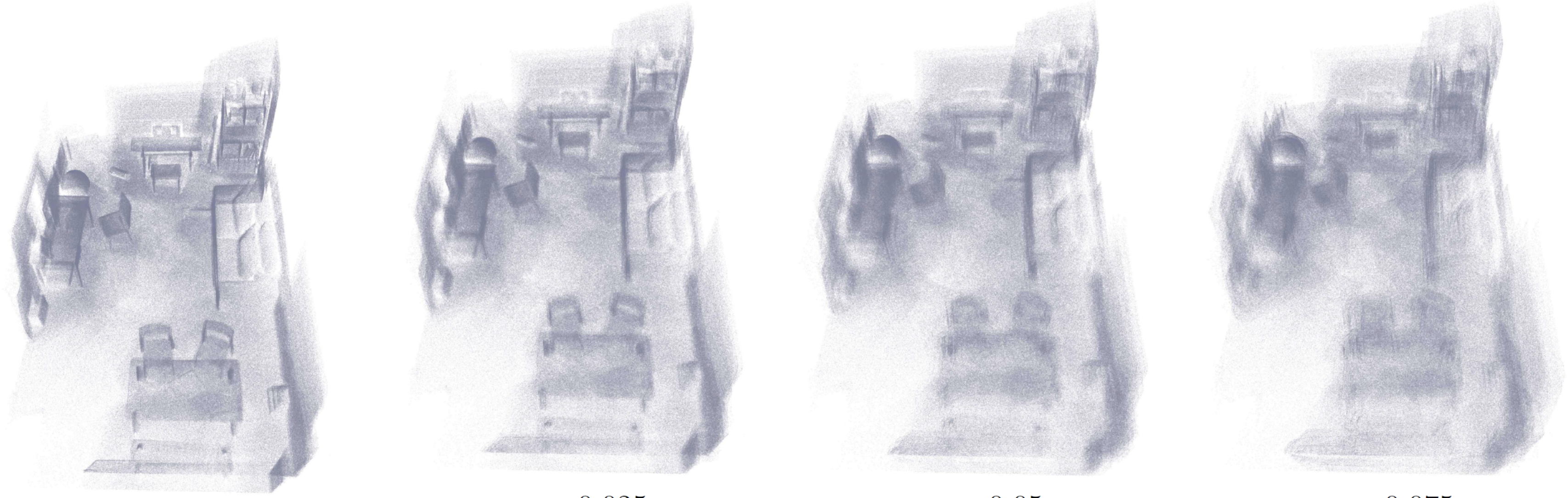
- Train incremental fusion of pre-trained neural implicit representations.



NeuralBlox: Real-Time Neural Representation Fusion

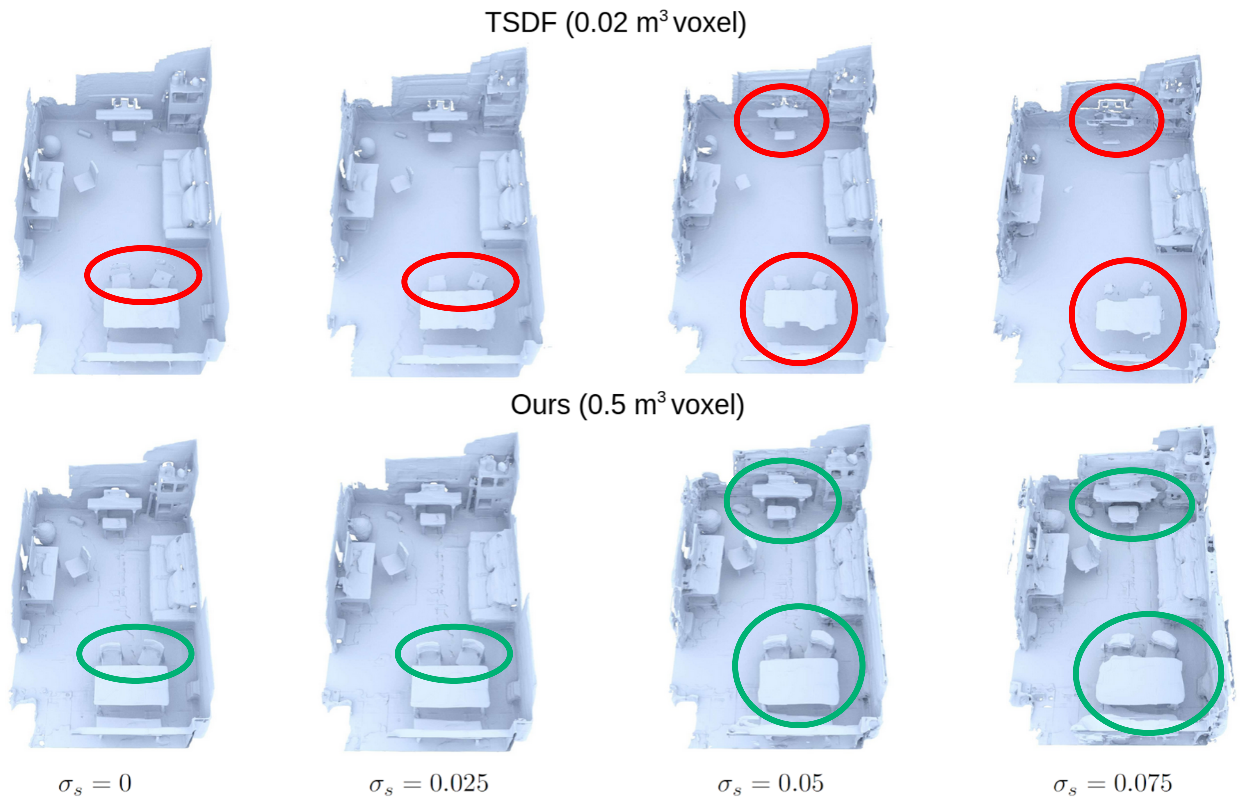
- Test on real scenes of Redwood dataset.

- Sensor pose uncertainty.

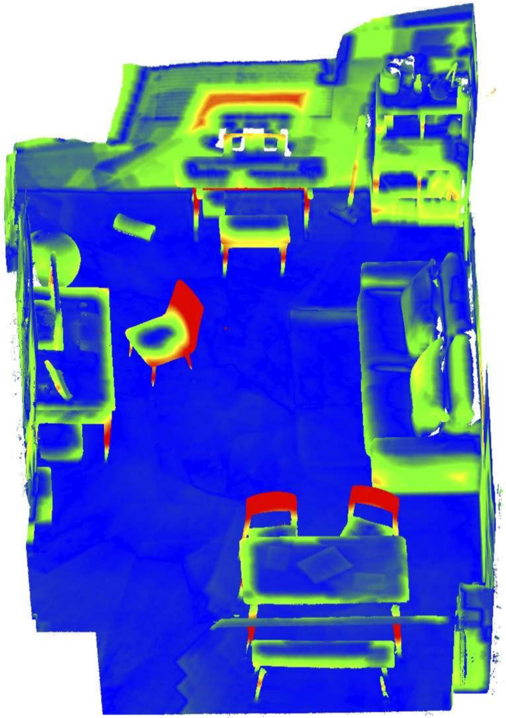


$\sigma_s = 0.025$ $\sigma_s = 0.05$ $\sigma_s = 0.075$

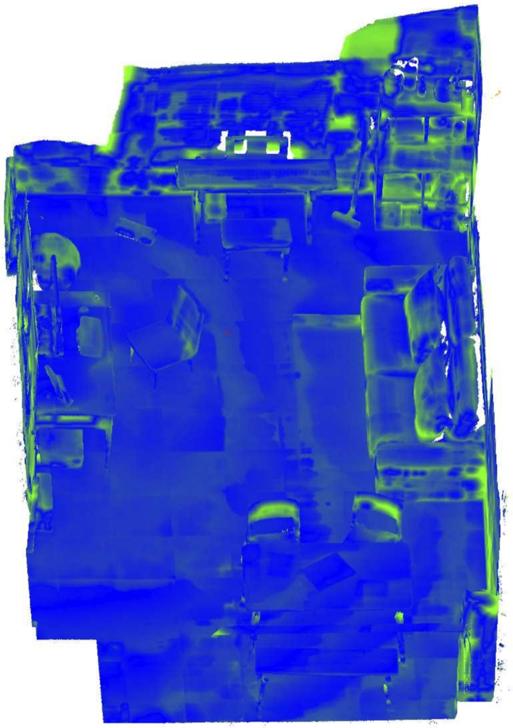
NeuralBlox: Real-Time Neural Representation Fusion



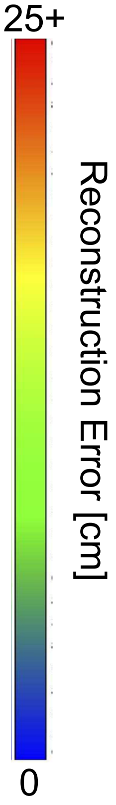
NeuralBlox: Real-Time Neural Representation Fusion



TSDF (0.02 m voxel)



Ours (0.5 m voxel)



NeuralBlox: Real-Time Neural Representation Fusion

σ_s (m)	***Left:*** *TSDF* (0.02 m^3) [3]. ***Right:*** Ours (0.5 m^3).							
	Accuracy		*Completeness*		*Recall*		*Recall**	
0	**0.0175**	0.0225	0.0137	**0.0135**	0.988	**0.990**	0.984	**0.986**
0.025	**0.0213**	0.0253	0.0202	**0.0155**	0.926	**0.990**	0.888	**0.986**
0.050	0.0377	**0.0350**	0.0352	**0.0202**	0.756	**0.952**	0.625	**0.928**
0.075	0.0535	**0.0526**	0.0486	**0.0241**	0.635	**0.886**	0.438	**0.827**

*the ground is excluded

NeuralBlox: Real-Time Neural Representation Fusion

- Robotic volumetric mapping based on implicit neural representations.
- Incremental fusion of sensor data directly in latent space.
- Can robustly capture surfaces in the presence of pose errors, sensing uncertainty, and sparse sensor inputs.
- Runs at interactive rates on a CPU only!

Encoding	Decoding	
4.7 FPS	15.6 voxel/s	2.1 FPS[3]

Panoptic Multi-TSDFs: Online multi-resolution volumetric mapping with long-term dynamic scene consistency

- World typically does not change at random, but in a semantically consistent way.

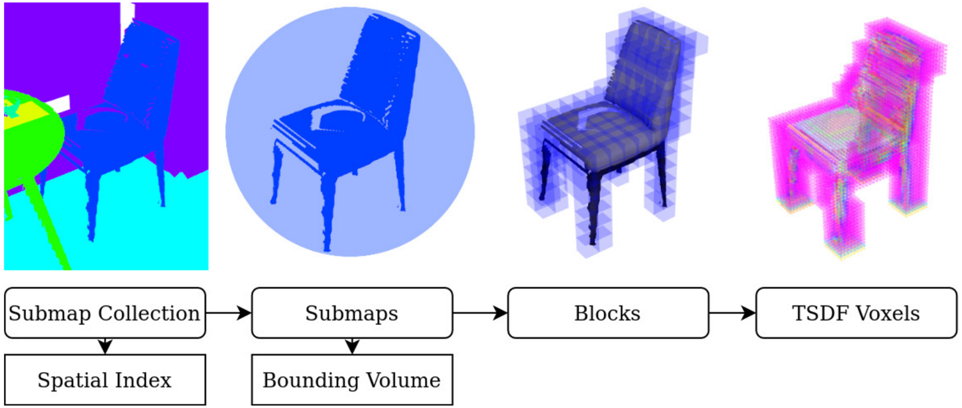
- Can we leverage high-level semantic information to build temporally consistent maps online?


Day 1 Day 2

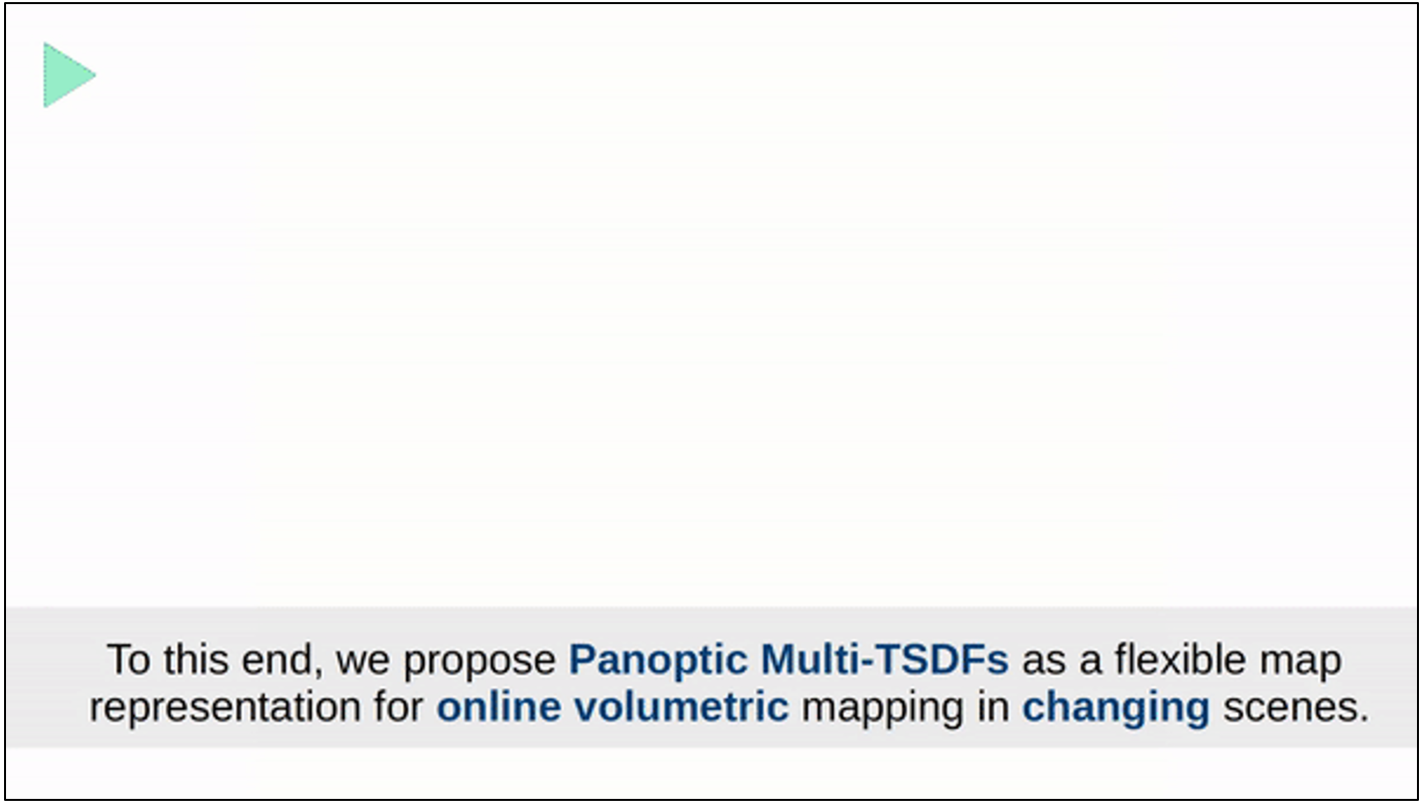

Single TSDF map

Panoptic Multi-TSDFs: Online multi-resolution volumetric mapping with long-term dynamic scene consistency

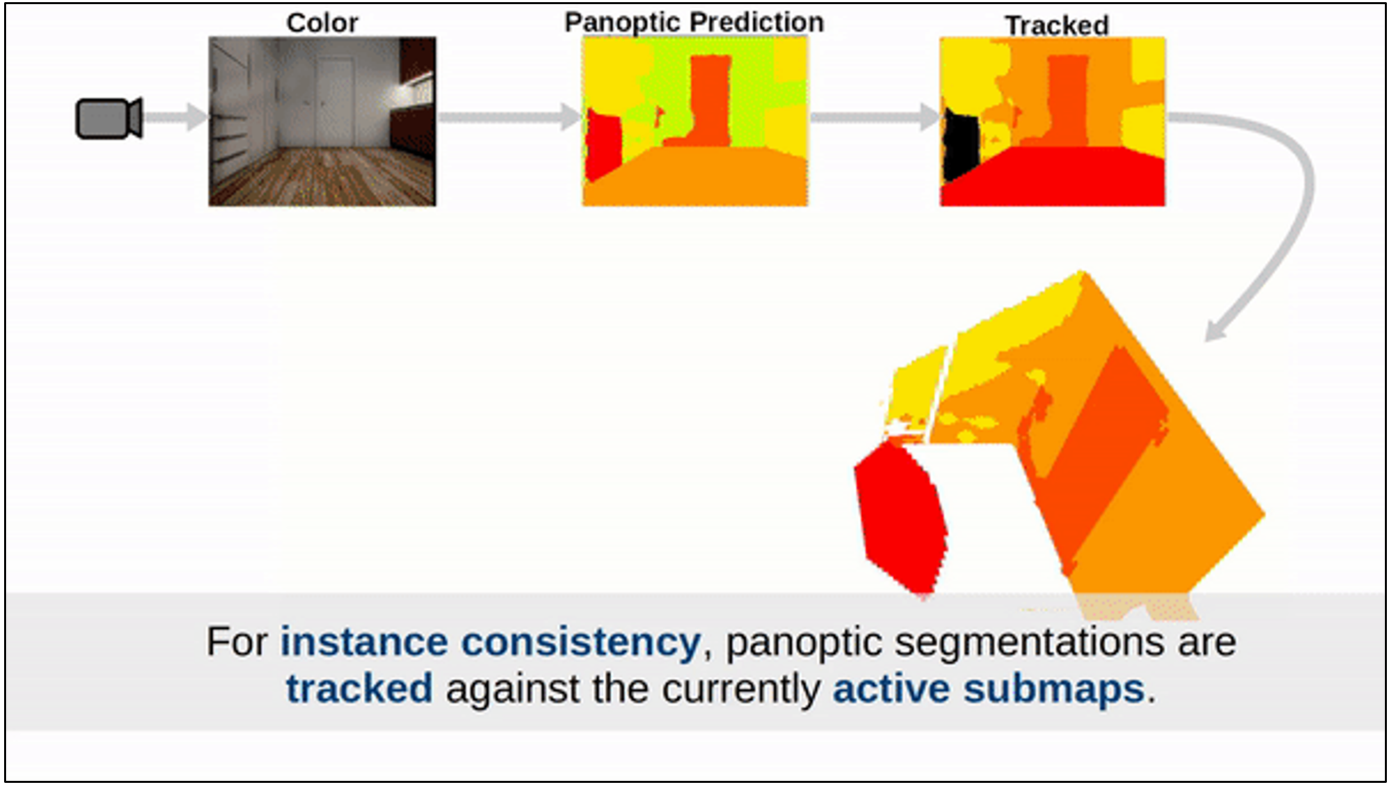
- Hierarchical map representation based on semantically and temporally consistent submaps.
- Only integrate submaps that are successfully tracked over time.
- Reason about change or persistence on the set of submaps.



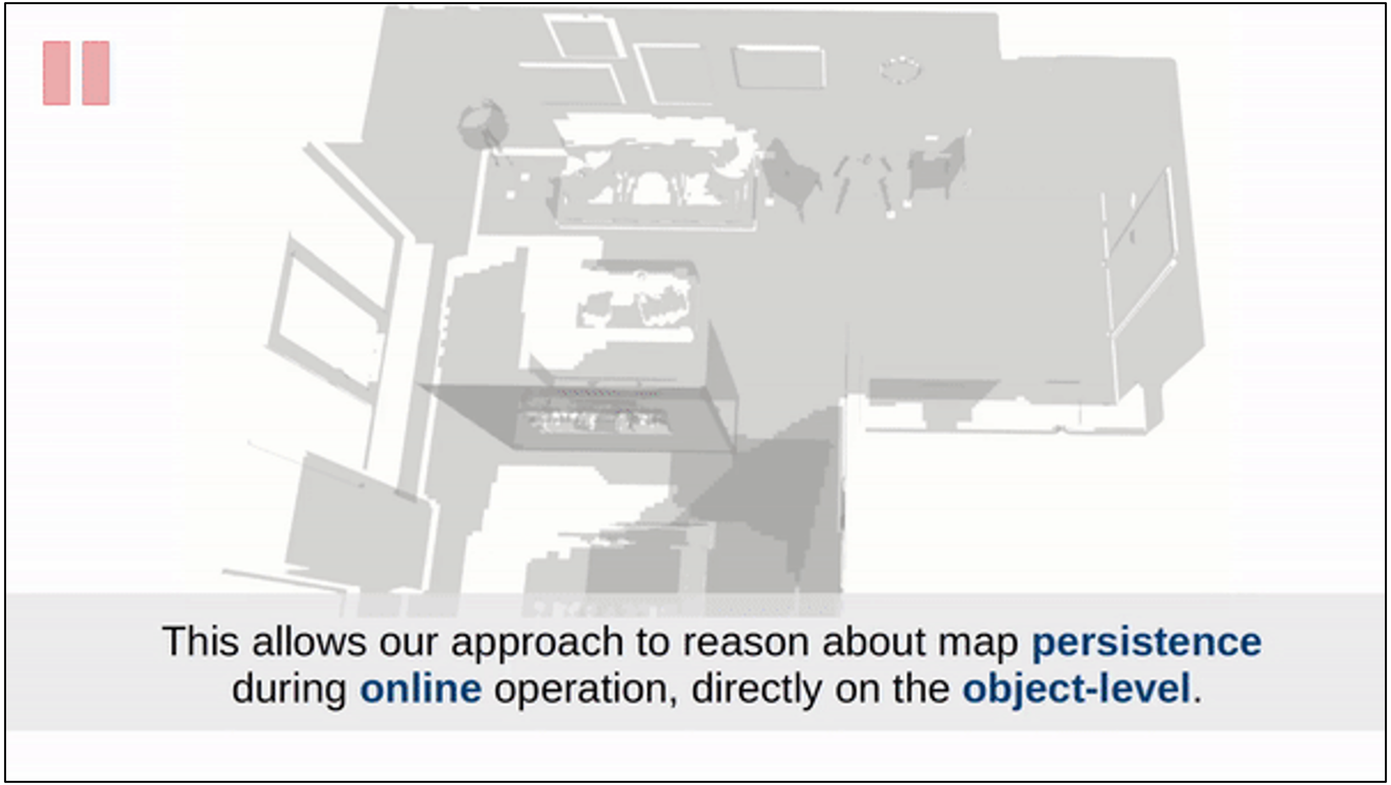
Panoptic Multi-TSDFs: Online multi-resolution volumetric mapping with long-term dynamic scene consistency



Panoptic Multi-TSDFs: Online multi-resolution volumetric mapping with long-term dynamic scene consistency

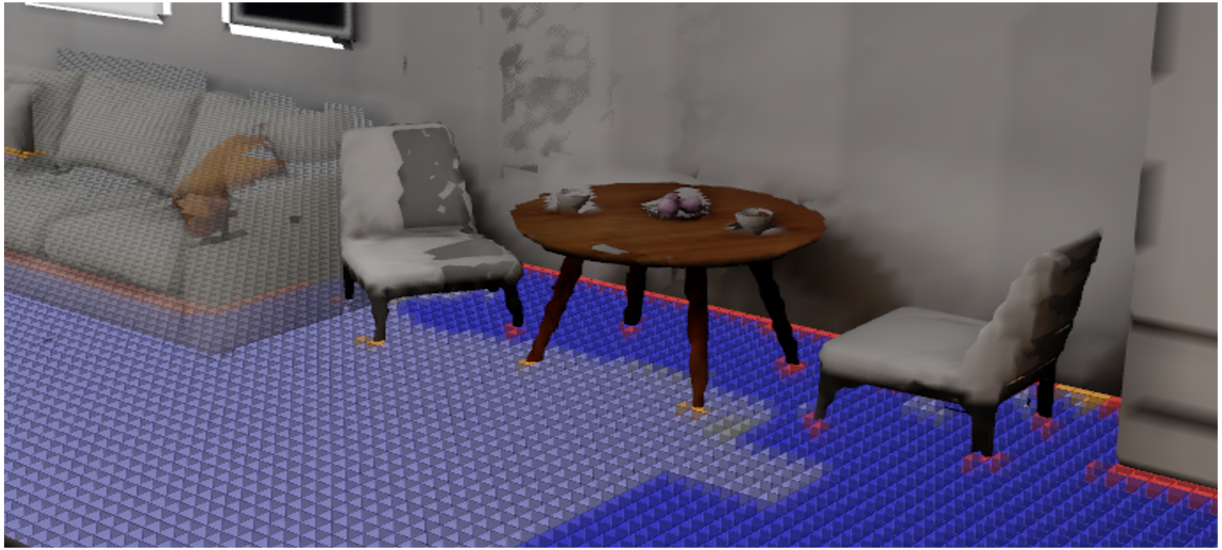


Panoptic Multi-TSDFs: Online multi-resolution volumetric mapping with long-term dynamic scene consistency



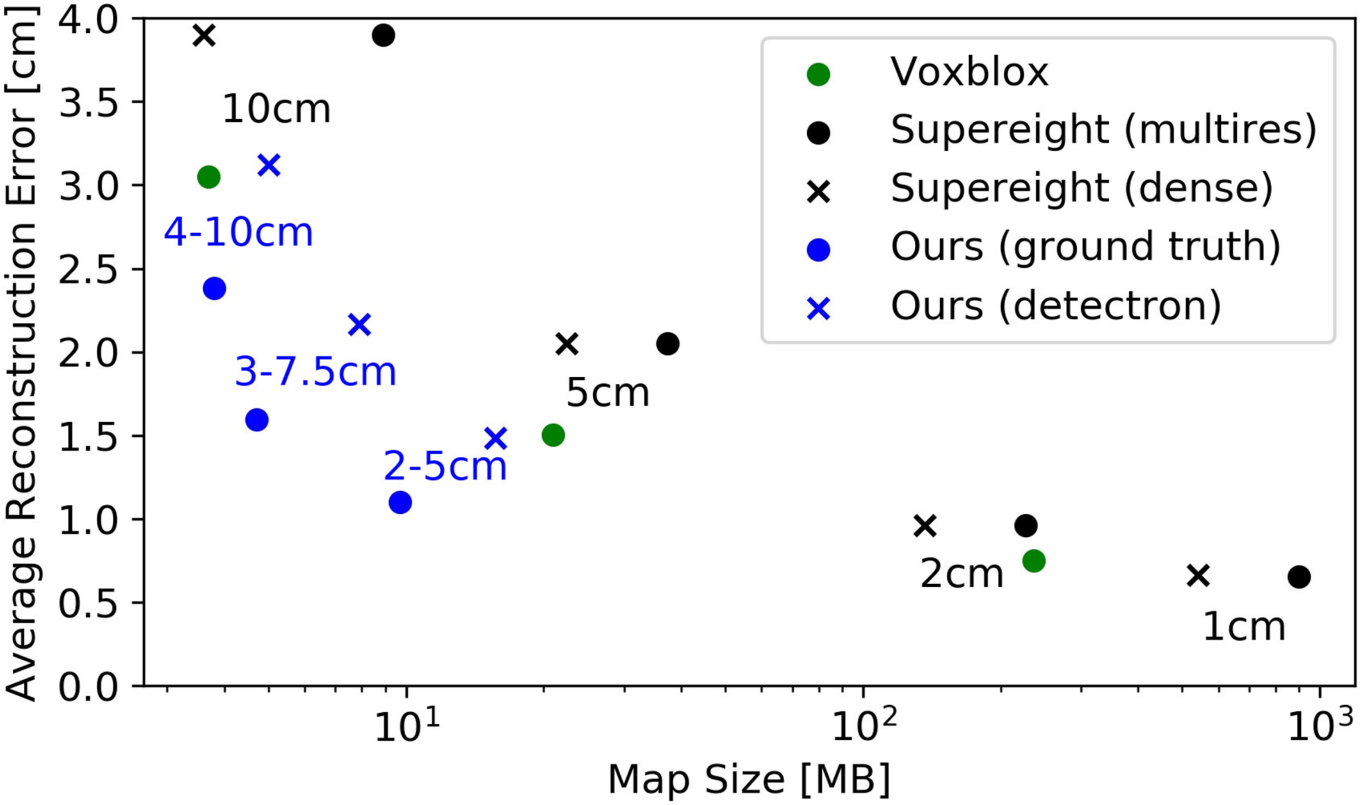
Panoptic Multi-TSDFs: Online multi-resolution volumetric mapping with long-term dynamic scene consistency

- Spatio-temporal map queries.

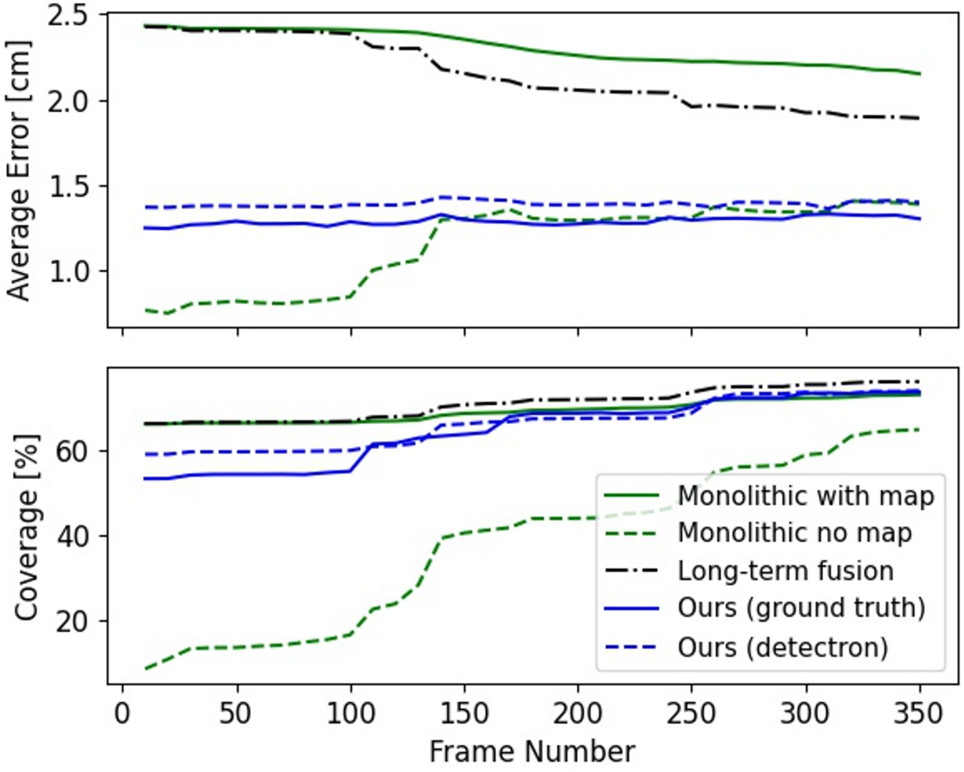


- Observed Free
- Observed Occupied
- Expected Free
- Expected Occupied
- Occupied due to semantic consistency
- Unknown

Panoptic Multi-TSDFs: Online multi-resolution volumetric mapping with long-term dynamic scene consistency



Panoptic Multi-TSDFs: Online multi-resolution volumetric mapping with long-term dynamic scene consistency



Panoptic Multi-TSDFs: Online multi-resolution volumetric mapping with long-term dynamic scene consistency

Setting	Resolution	Tracking	Integration	Management	FPS*
Flat, ground truth	2-5 cm	70.2 ± 8.4	104.3 ± 14.4	199.1 ± 54.1	5.1
	4-10 cm	63.9 ± 4.4	89.2 ± 7.8	182.1 ± 44.3	5.8
Flat, detectron	2-5 cm	57.8 ± 5.5	91.1 ± 11.1	192.1 ± 54.3	5.9
	4-10 cm	54.7 ± 5.1	80.3 ± 7.4	183.5 ± 49.2	6.5
RIO, detectron	2-5 cm	21.8 ± 4.6	21.8 ± 5.9	33.2 ± 23.8	21.3
	4-10 cm	16.8 ± 3.4	13.3 ± 3.5	9.5 ± 4.5	32.2

* Final frame rate is computed performing change detection every 10 frames.

Panoptic Multi-TSDFs: Online multi-resolution volumetric mapping with long-term dynamic scene consistency

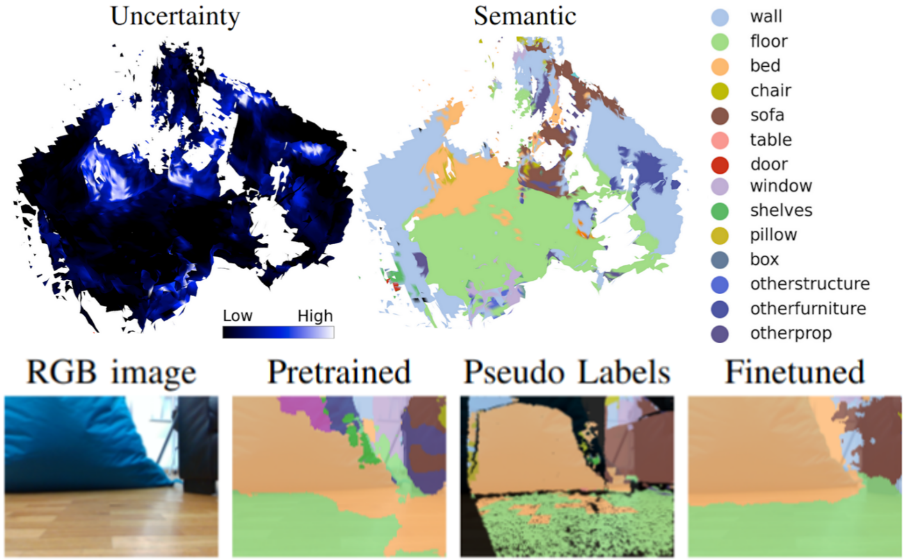
- Hierarchical map representation based on semantically and temporally consistent submaps.
- Reason about change or persistence on the set of submaps that together constitute a full volumetric map.
- Can be incrementally built, maintained, and queried on compute constrained robots.

Conclusions

- Learning-based methods show high potential for surface representation and scene understanding.
- Common map representations not yet there to effectively capture and reason about complex, changing, shared environments.
- Combine low-level geometry and high-level abstractions to reason at the adequate scale.

Outlook

- Spatio-temporal volumetric mapping.
- Active Perception to improve scene understanding and map building.



[**L. Schmid**, C. Ni, Y. Zhong, R. Siegwart, and O. Andersson, "Fast and Compute-efficient Sampling-based Local Exploration Planning via Distribution Learning." ArXiv 2022]

[R. Zurbrugg, H. Blum, C. Cadena, R. Siegwart, and **L. Schmid,** "Embodied Active Domain Adaptation for Semantic Segmentation via Informative Path Planning." ArXiv 2022]

NeuralBlox

Panoptic Mapping

Questions?

https://github.com/ethz-asl/neuralblox

https://github.com/ethz-asl/panoptic_mapping