PlaneMatch: Patch Coplanarity Prediction for Robust RGB-D Reconstruction

Yifei Shi, Kai Xu, Matthias Niessner, Szymon Rusinkiewicz, Thomas Funkhouser

Princeton University
National University of Defense Technology
Technical University of Munich
Google

RGB-D Reconstruction

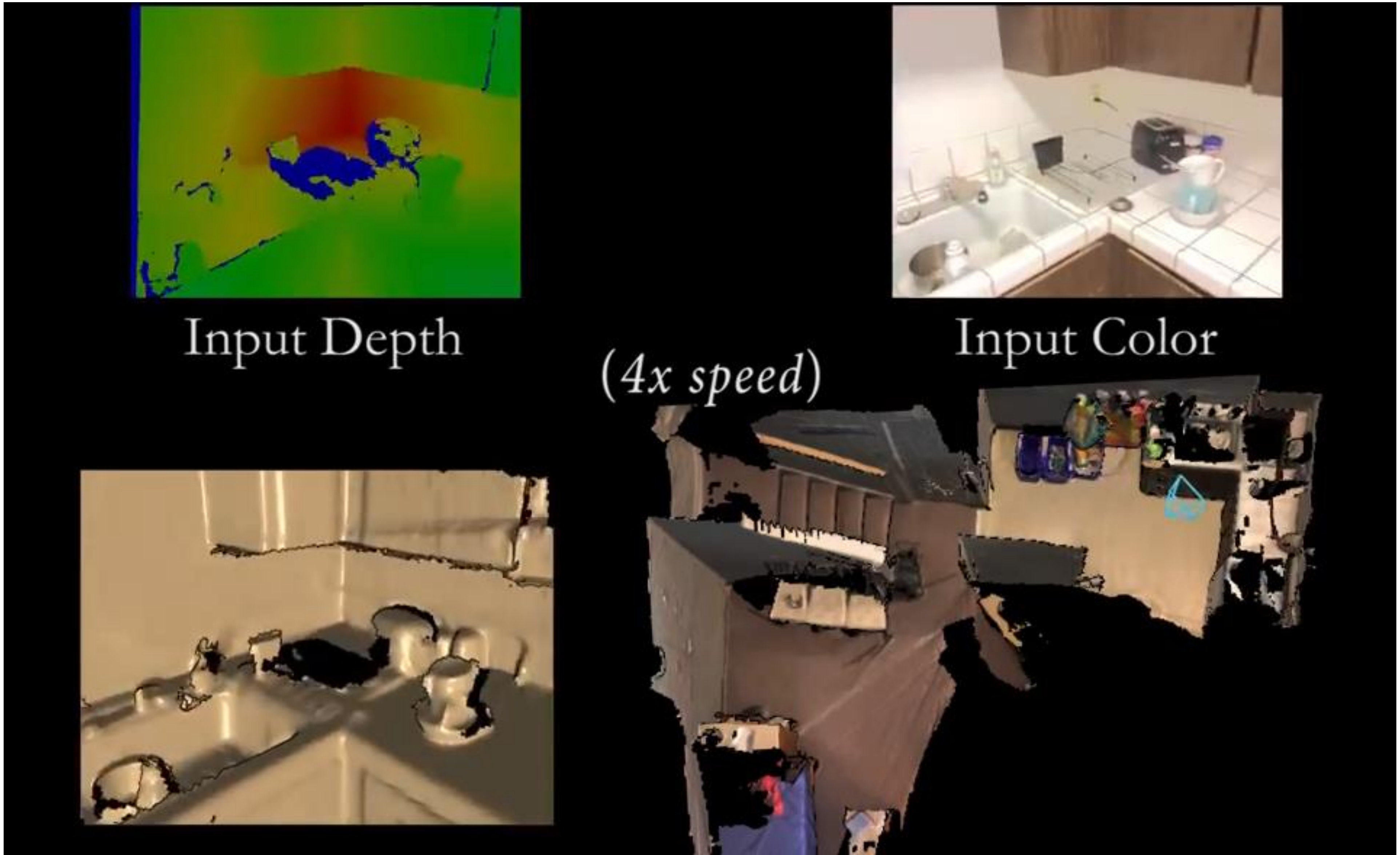
Microsoft Kinect

Structure Sensor

Xtion

RGB-D Reconstruction

Bundle Fusion [Dai et al. 17]



RGB-D Reconstruction

KinectFusion
[Newcombe/Izadi et al. 2011]

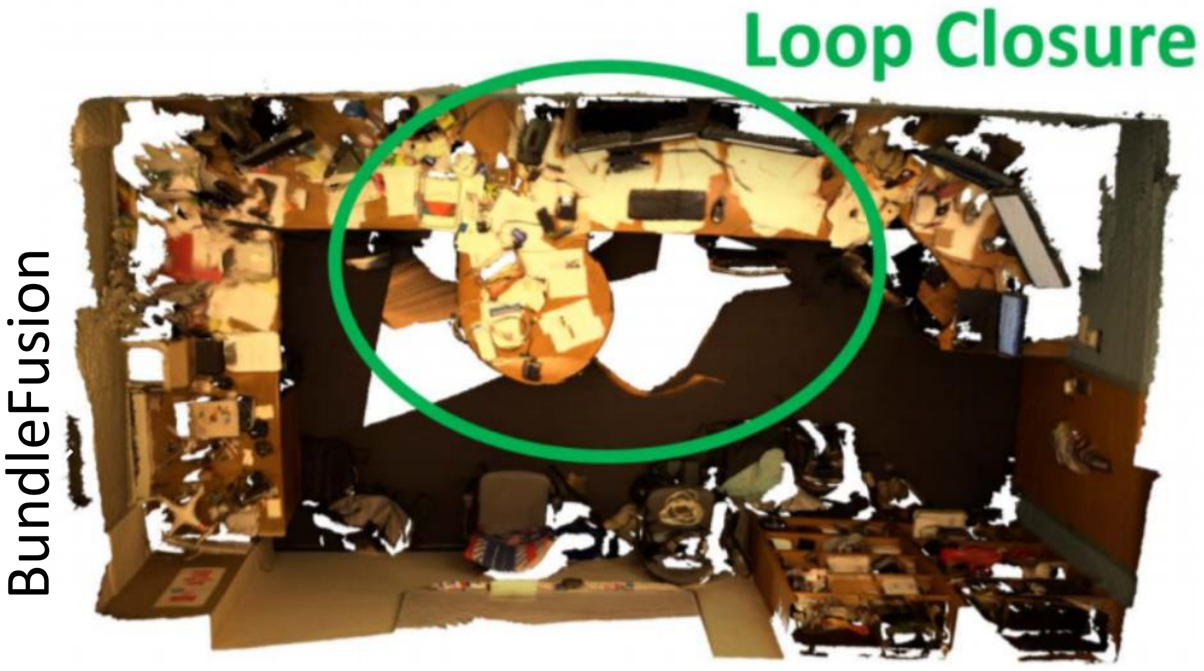
VoxelHashing
[Niessner et al. 2013]

BundleFusion
[Dai et al. 2017]

Robust Recon.
[Choi et al. 2015]

ElasticFusion
[Whelan et al. 2016]

Loop Closure



Loop Closure -> Feature Descriptor

RGB Features:

- SIFT, SURF, ORB, Freak, …
- LIFT, MatchNet, …

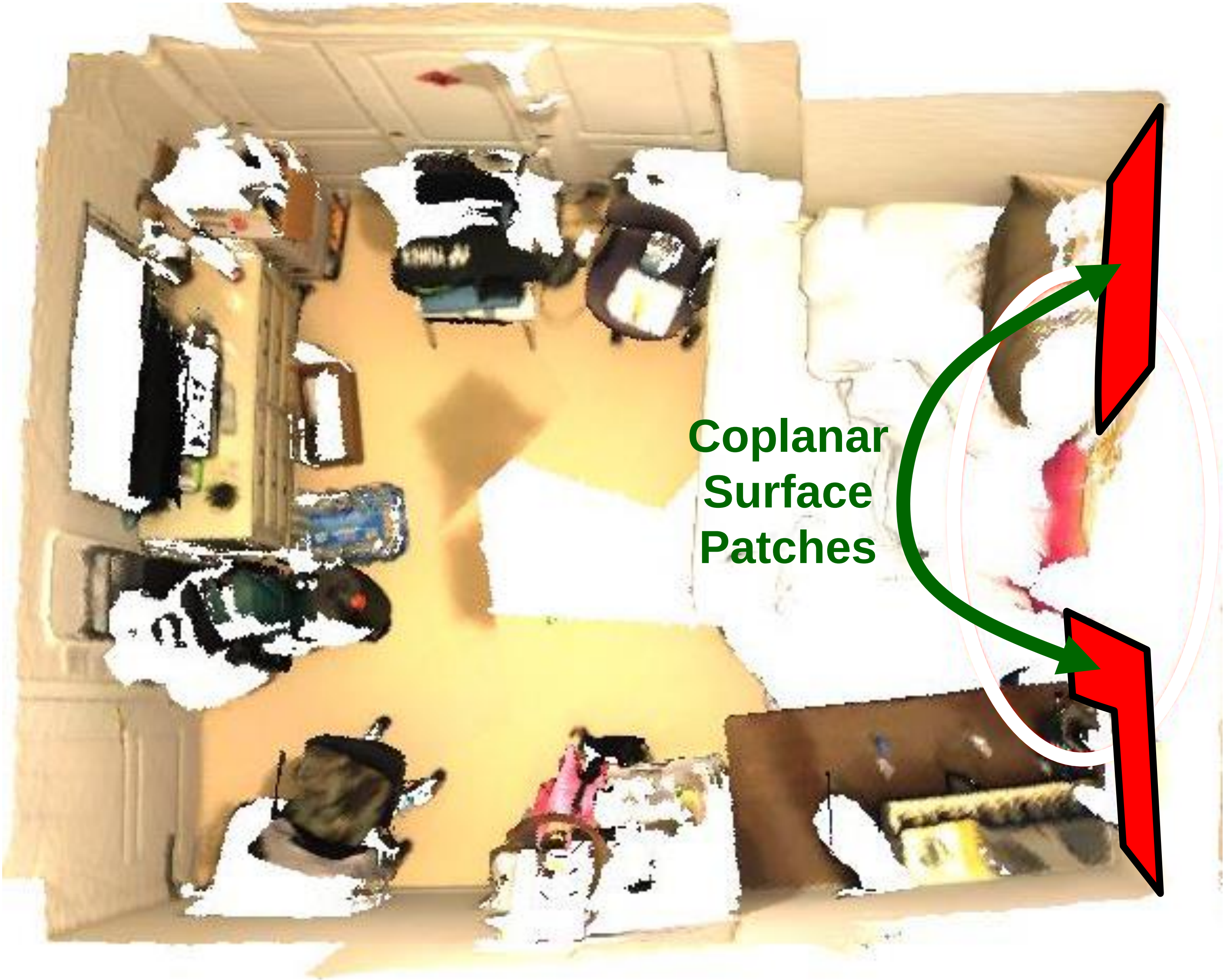
Geometric Features:

- SHOT, FPFH, SpinImages, …
- 3DMatch, …

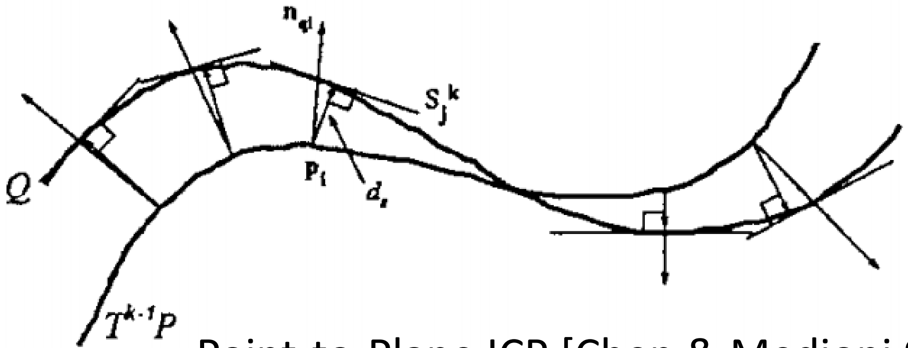
Keypoint-based

Are there additional primitives?

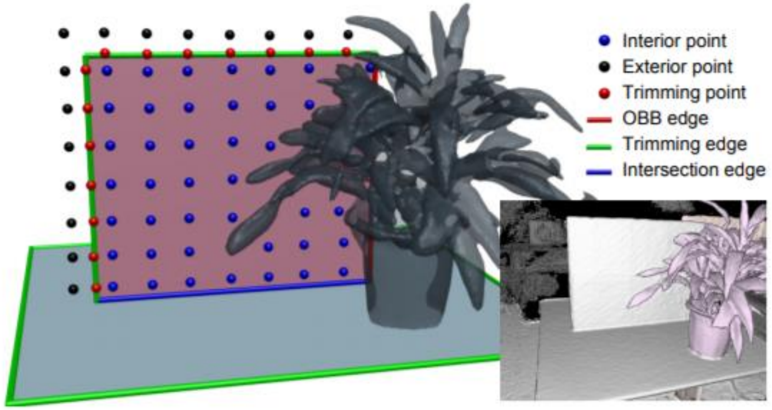
Our Idea: Planar Feature Descriptors



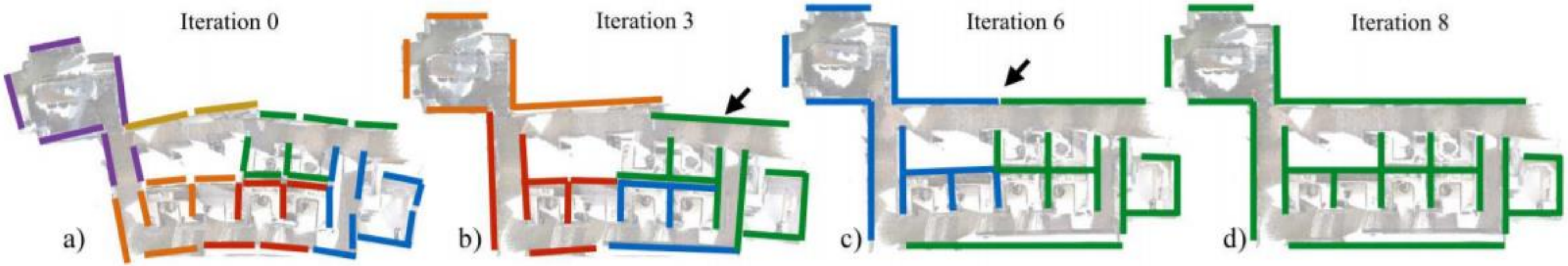
Existing Planar Matching is Local



Point-to-Plane ICP [Chen & Medioni 91]

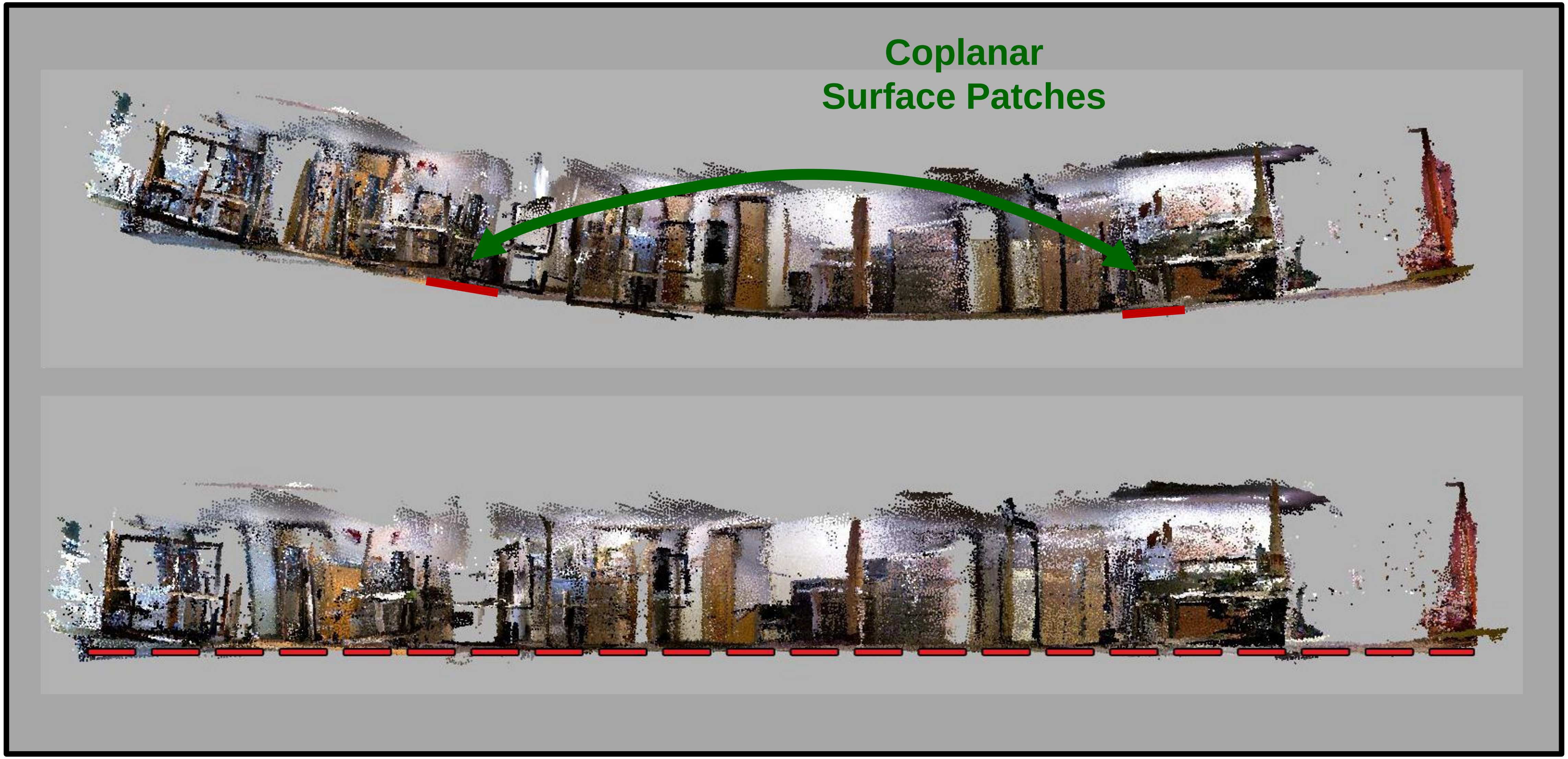


Online Structure Analysis [Zhang et al. 2015]

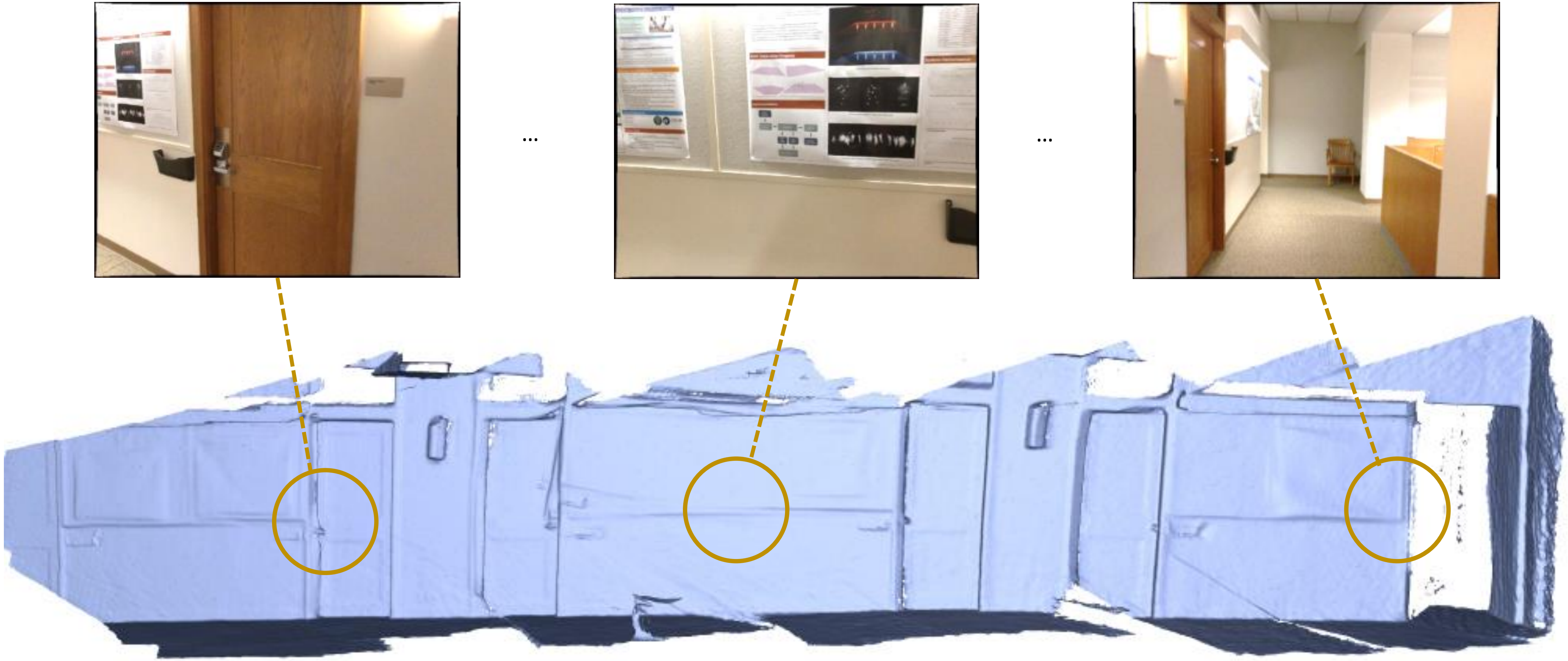


Fine-to-Coarse Registration [Halber and Funkhouser 2017]

Long-Range Constraints for SLAM



Task: Co-planarity Matching?



PlaneMatch: Learning Co-planarity Features

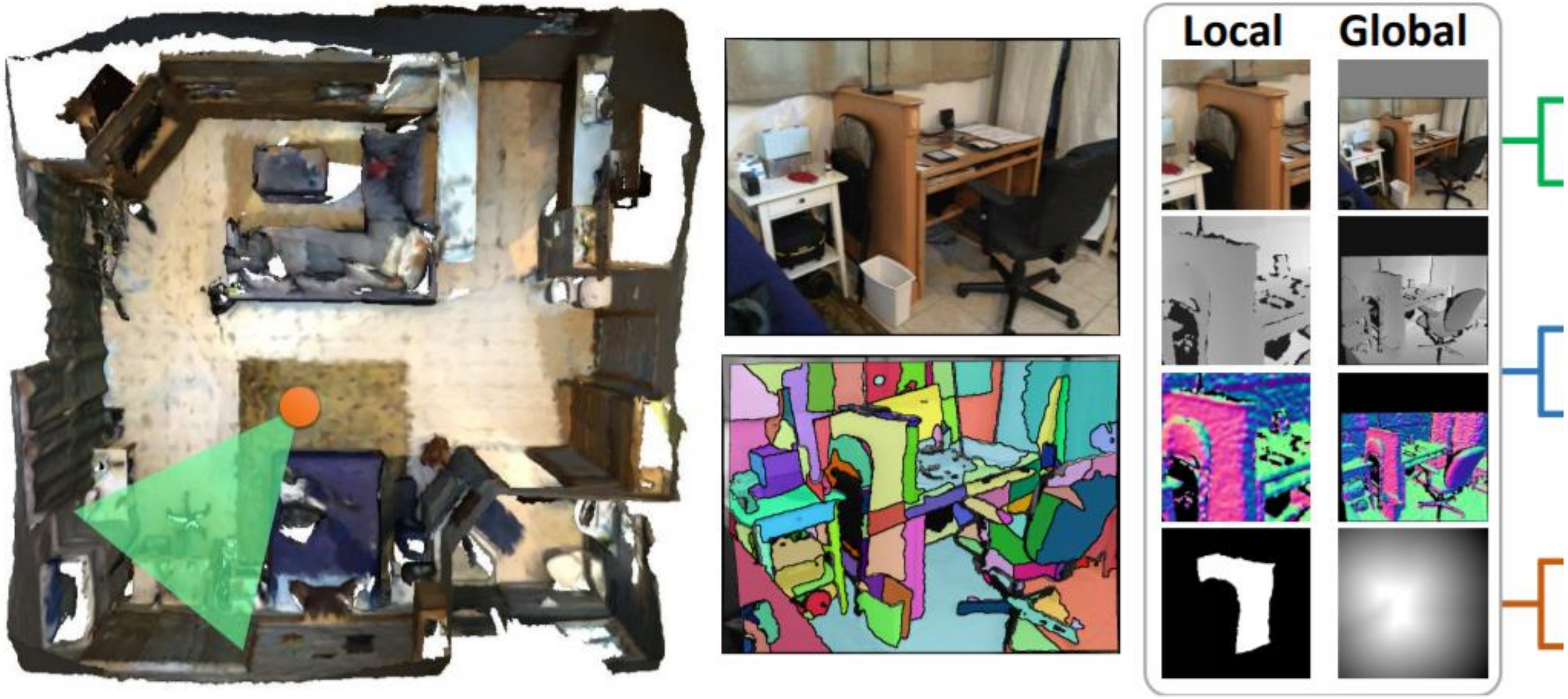
- Color
- Depth
- Normals
- Plane Segmentation (Mask)
- …

Learn from 3D data!

Siamese Network Architecture

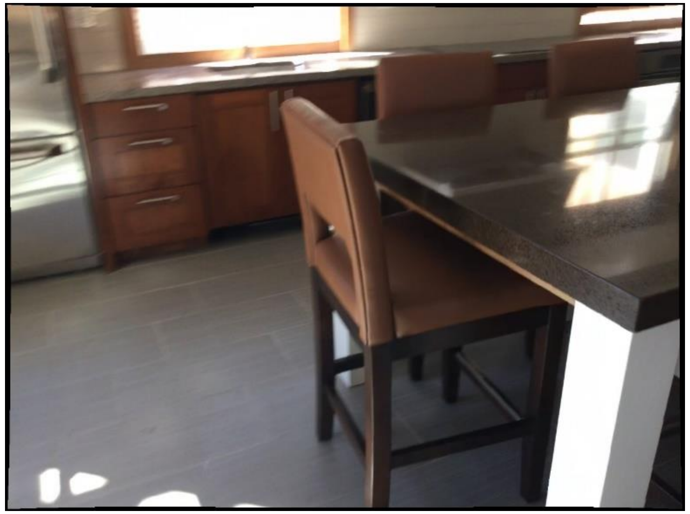
Siamese Network Architecture

Siamese Network Architecture

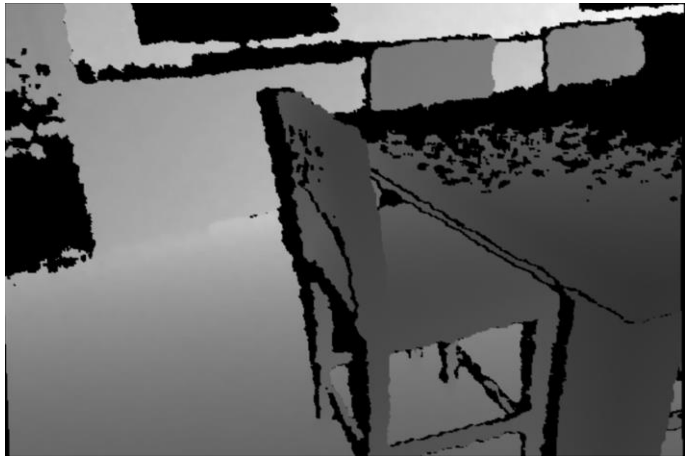


Siamese Network Architecture

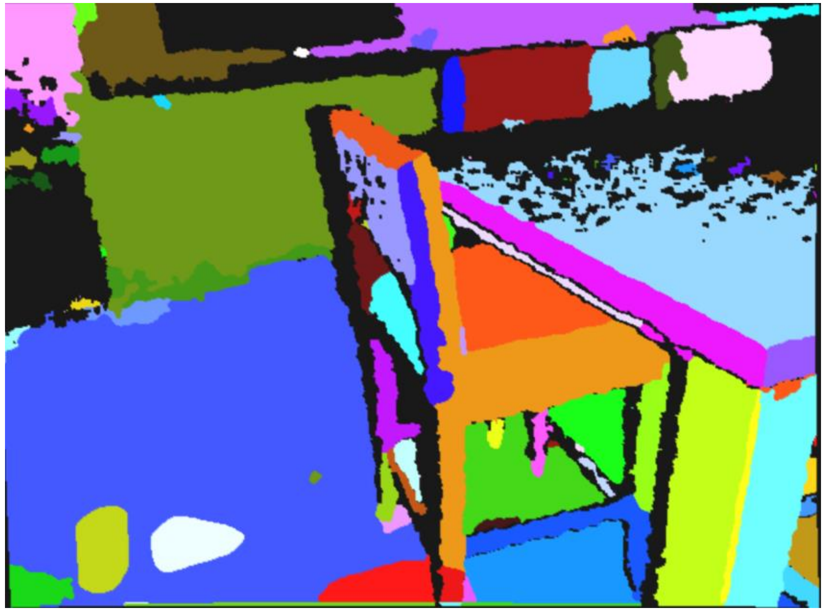
Step 1: Extract Planar Patches



RGB



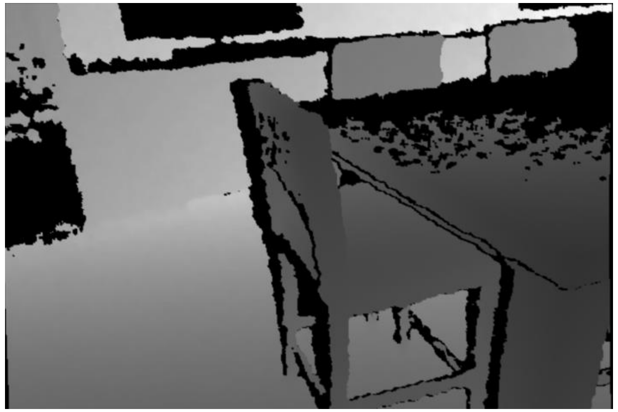
Depth



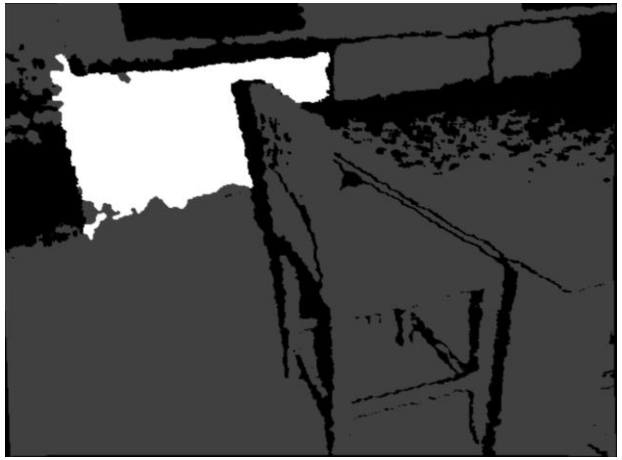
Planar Patches

Step 2: Extract Global Rep. / Patch

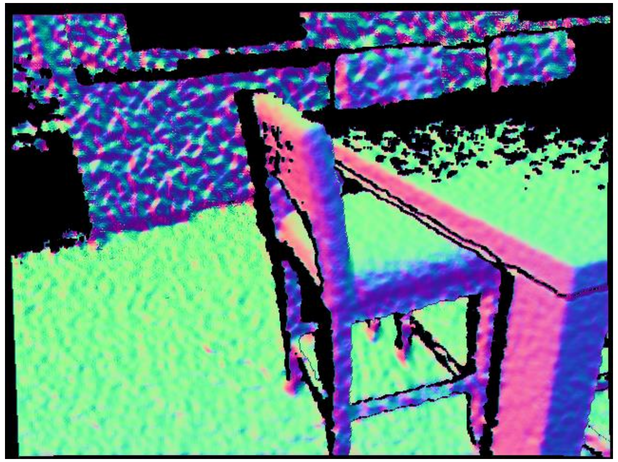
RGB



Depth

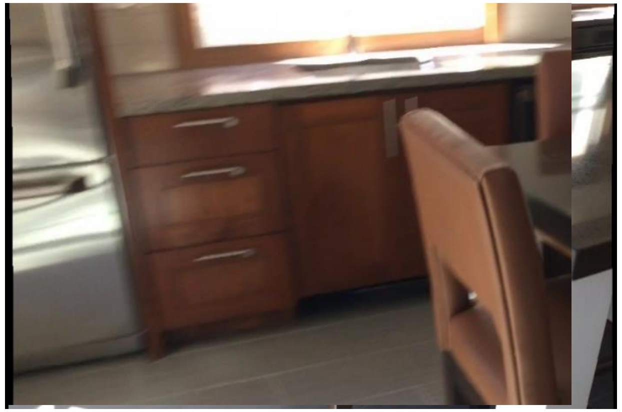


Patch Mask

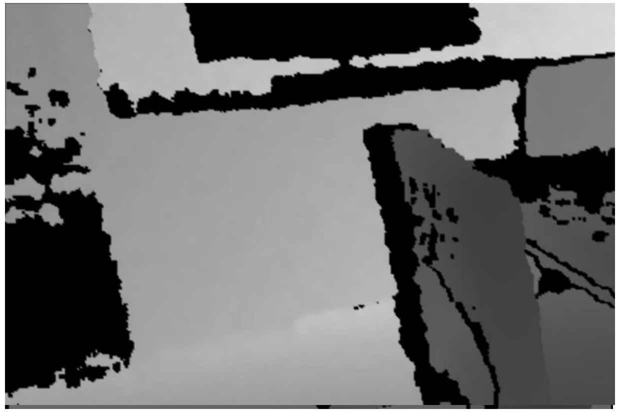


Normals

Step 3: Extract Local Rep. / Patch

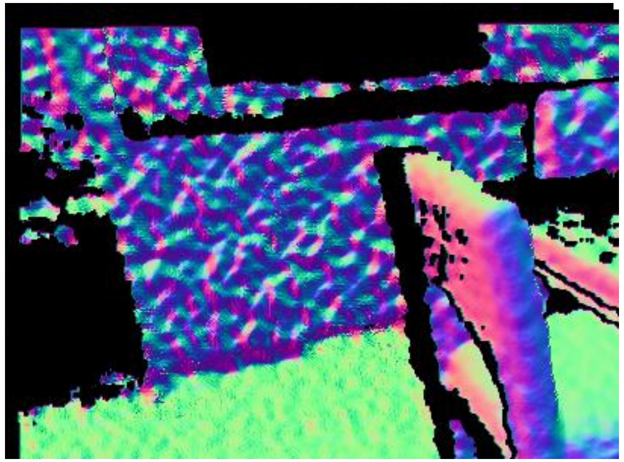


RGB



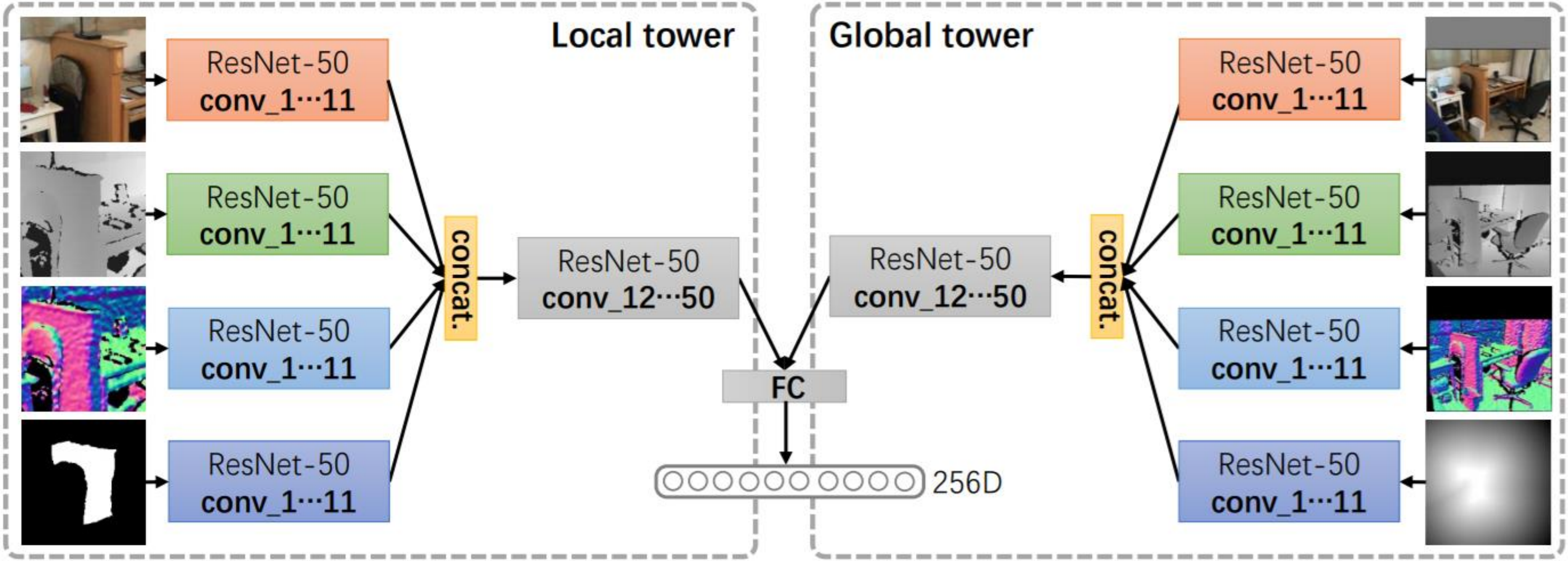
Depth

Patch Mask



Normals

Local / Global Representations



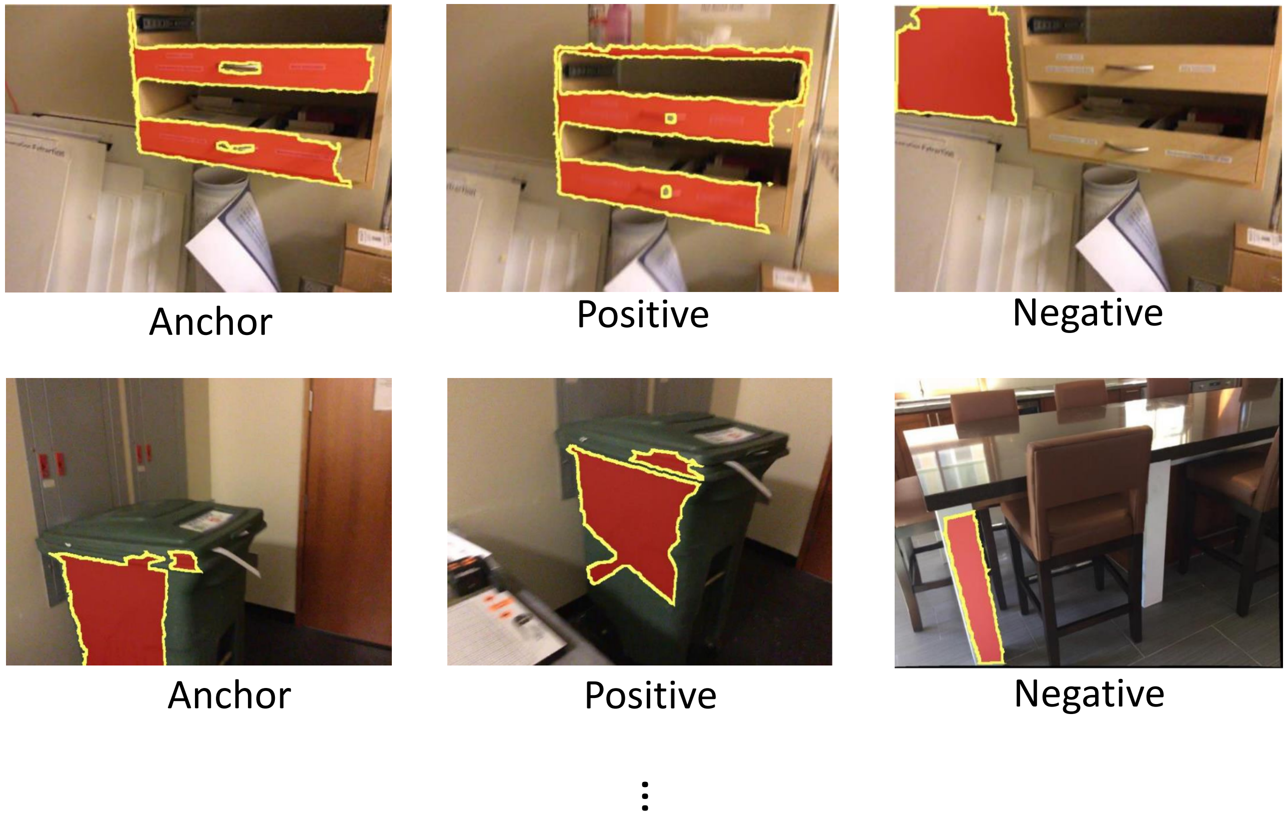
Local Representation

Global Representation

Siamese Network Architecture

Training: Self-Supervised Learning

ScanNet [Dai et al. 2017]

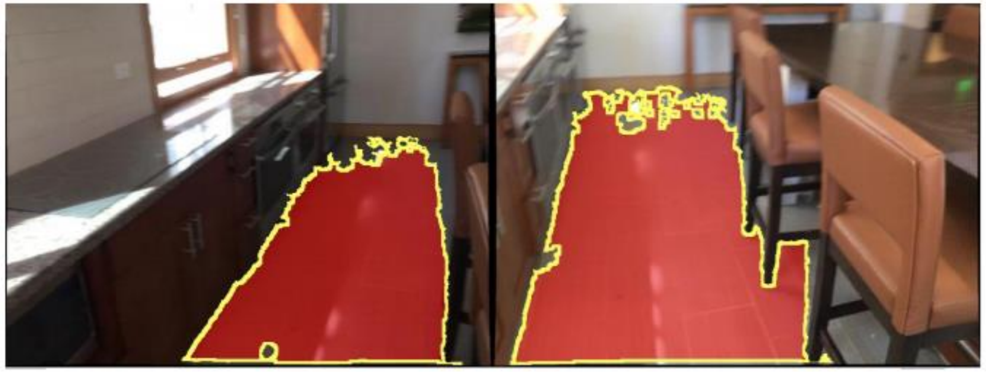


10 million triplets

Triplets for Training

Anchor Positive Negative

Benchmark for Task of Co-planarity Matching



Positive pair (6k)

Negative pair (6k)

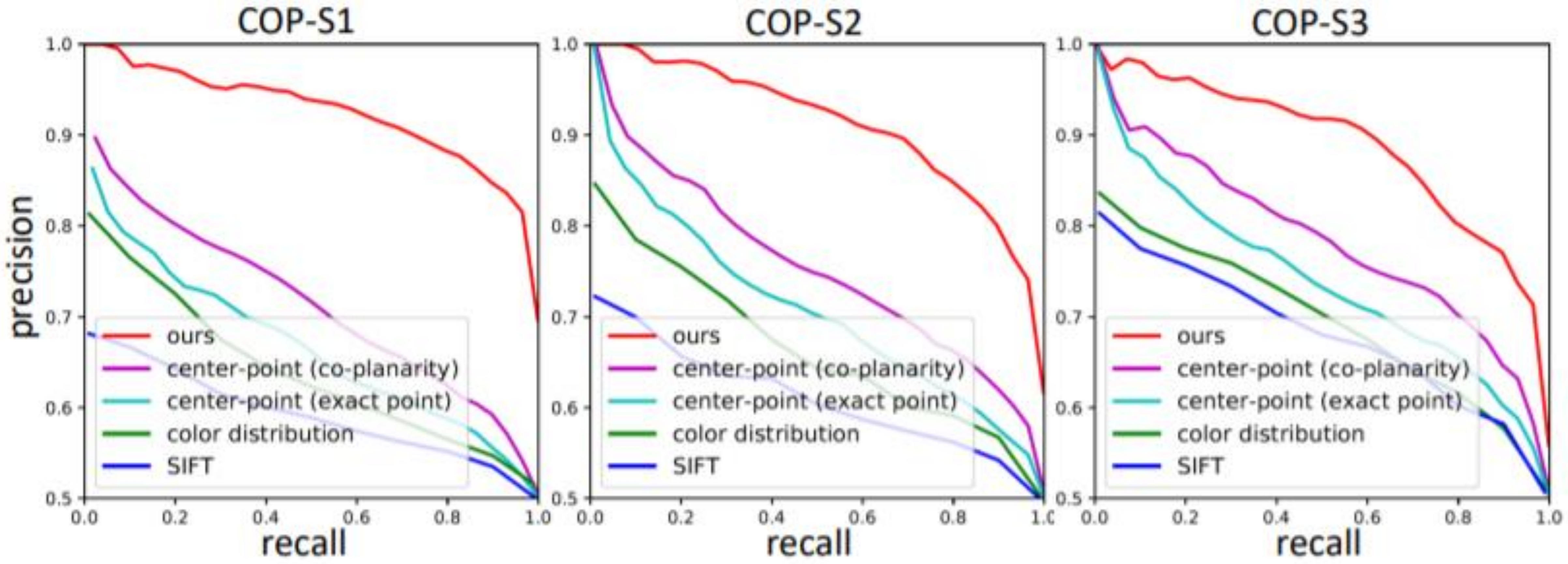
By patch size

COP-S	S1	S2	S3
patch size	0.25~10 m^2	0.05~0.25 m^2	0~0.05 m^2

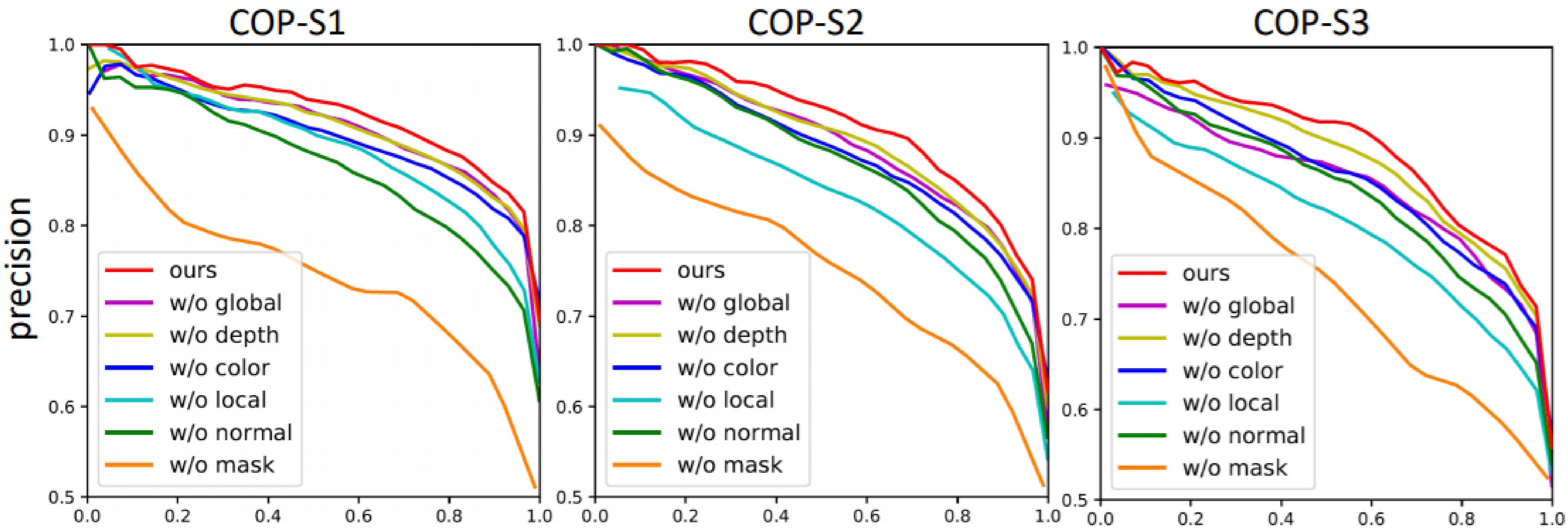
By pair distance

COP-D	D1	D2	D3
pair distance	0~0.3 m	0.3~1 m	1~5 m

PlaneMatch Evaluation



PlaneMatch Ablation Study



PlaneMatch Registration

$$E(T, s) = E_{\text{data-cop}}(T, s) + E_{\text{reg-cop}}(s) + E_{\text{data-kp}}(T, s) + E_{\text{reg-kp}}(s)$$

T : transformation matrix s : indicator variables ($\in [0,1]$)

PlaneMatch Registration

$$E(T,s) = \boxed{E_{\text{data-cop}}(T,s)} + E_{\text{reg-cop}}(s) + E_{\text{data-kp}}(T,s) + E_{\text{reg-kp}}(s)$$

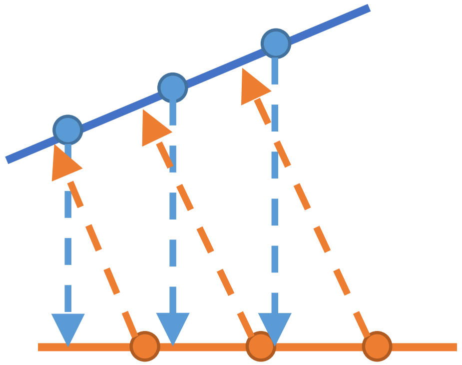
T : transformation matrix s : indicator variables ($\in [0,1]$)

$$E_{\text{data-cop}}(T,s) = \sum_{\pi \in \Pi_{\text{cop}}} w_\pi \, s_\pi \, d^2_{\text{cop}}(T,\pi)$$

Pairs predicted by coplanarity network

Π_{cop}: plane pair set d_{cop}: plane-to-plane distance

w_π: confidence weight



PlaneMatch Registration

$$E(T,s) = E_{\text{data-cop}}(T,s) + E_{\text{reg-cop}}(s) + \boxed{E_{\text{data-kp}}(T,s)} + E_{\text{reg-kp}}(s)$$

T : transformation matrix $\quad$ s : indicator variables ($\in [0,1]$)

$$\boxed{E_{\text{data-kp}}(T,s) = \sum_{\pi \in \Pi} w_\pi \, s_\pi \, d^2(T,\pi)}$$

Pairs predicted by SIFT keypoints

Π_{kp}: point pair set $\quad$ d_{kp}: point-to-point distance

w_π: confidence weight

PlaneMatch Registration

$$E(T,s) = E_{\text{data-cop}}(T,s) + \boxed{E_{\text{reg-cop}}(s)} + E_{\text{data-kp}}(T,s) + \boxed{E_{\text{reg-kp}}(s)}$$

T : transformation matrix

s : indicator variables ($\in [0,1]$)

$$E_{\text{reg-cop}}(s) = \sum_{\pi \in \Pi} \mu \left(\sqrt{s_\pi} - 1\right)^2$$

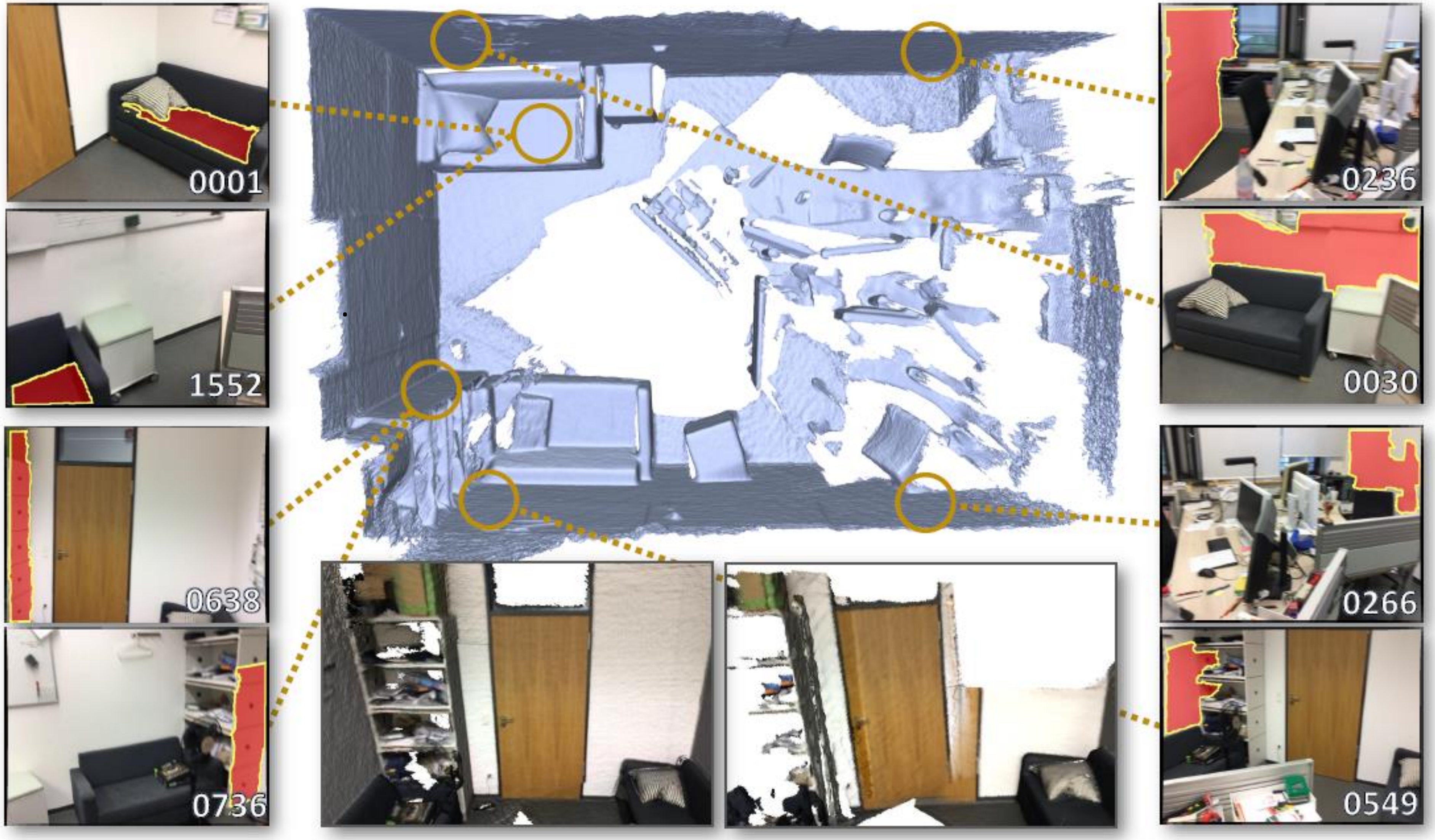
μ : threshold for error (0.01 m)

If $d^2 > \mu$, $s_\pi = 0$

If $d^2 < \mu$, $s_\pi = 1$

Robust optimization following [Choi et al. 15] / [Zollhoefer et al. 14] / [Zach et al. 14]

PlaneMatch Registration Results



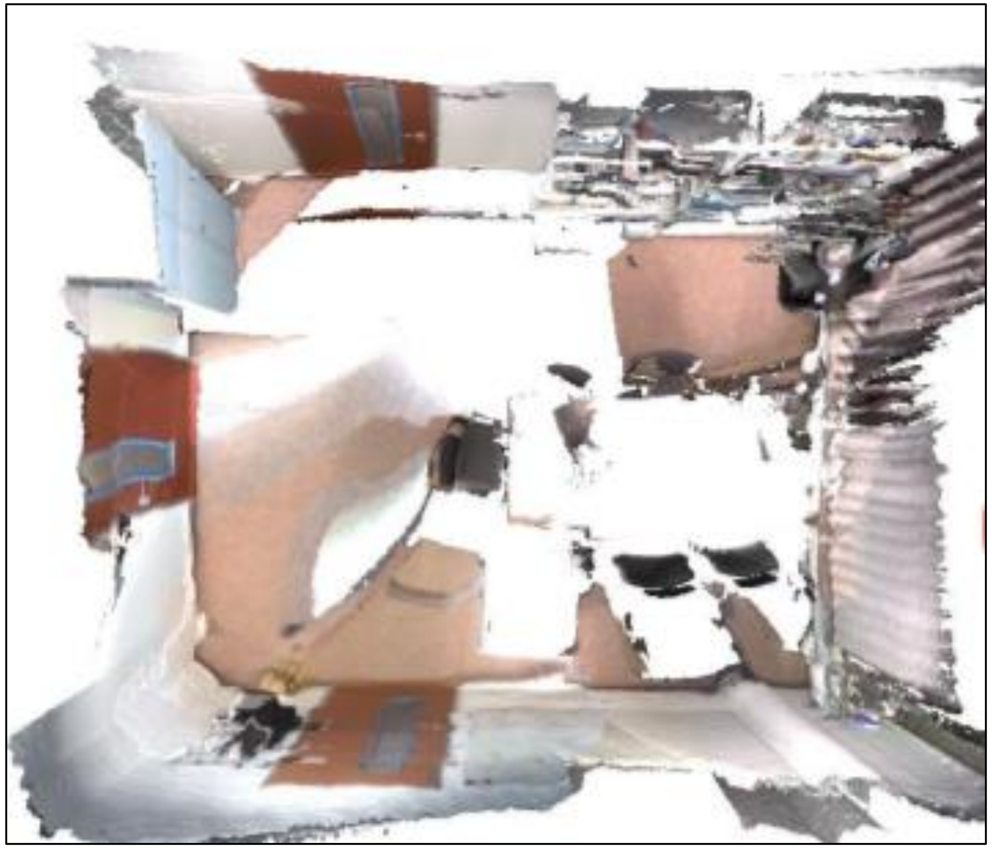
PlaneMatch Registration Results

BundleFusion [Dai et al.17]

PlaneMatch (Ours)

PlaneMatch Registration Results

BundleFusion [Dai et al.17]



PlaneMatch (Ours)

Evaluation on TUM-RGBD

Method	fr1/desk	fr2/xyz	fr3/office	fr3/nst
RGB-D SLAM	2.3	**0.8**	3.2	1.7
VoxelHashing	2.3	2.2	2.3	8.7
Elastic Fusion	2.0	**1.1**	1.7	1.6
Redwood	2.7	9.1	3.0	192.9
Fine-to-Coarse	5.0	3.0	3.9	3.0
BundleFusion	1.6	**1.1**	2.2	**1.2**
Ours	**1.4**	**1.1**	**1.6**	1.5
BundleFuison+Ours	**1.3**	**0.8**	**1.5**	**0.9**

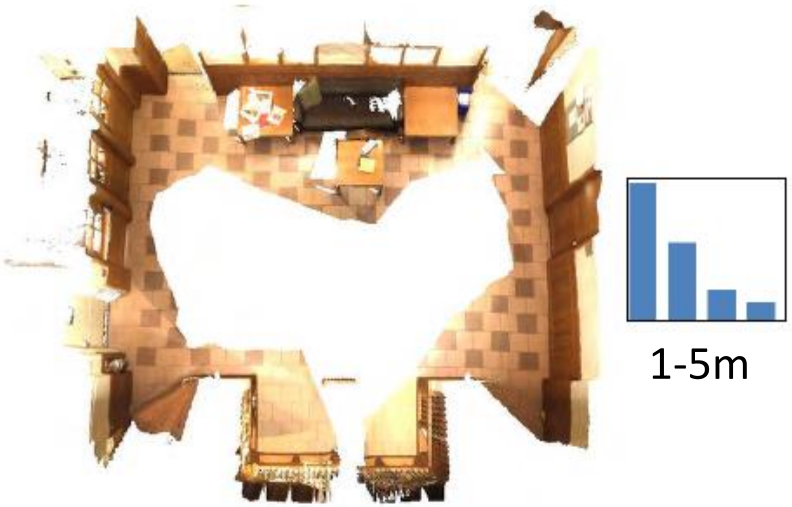
RMSE in cm (lower is better)

Ablation on TUM-RGBD

Method	fr1/desk	fr2/xyz	fr3/office	fr3/nst
Key-point Only	5.6	4.4	5.2	2.6
Coplanarity Only	2.5	2.1	3.7	–
Ours	**1.4**	**1.1**	**1.7**	**1.5**

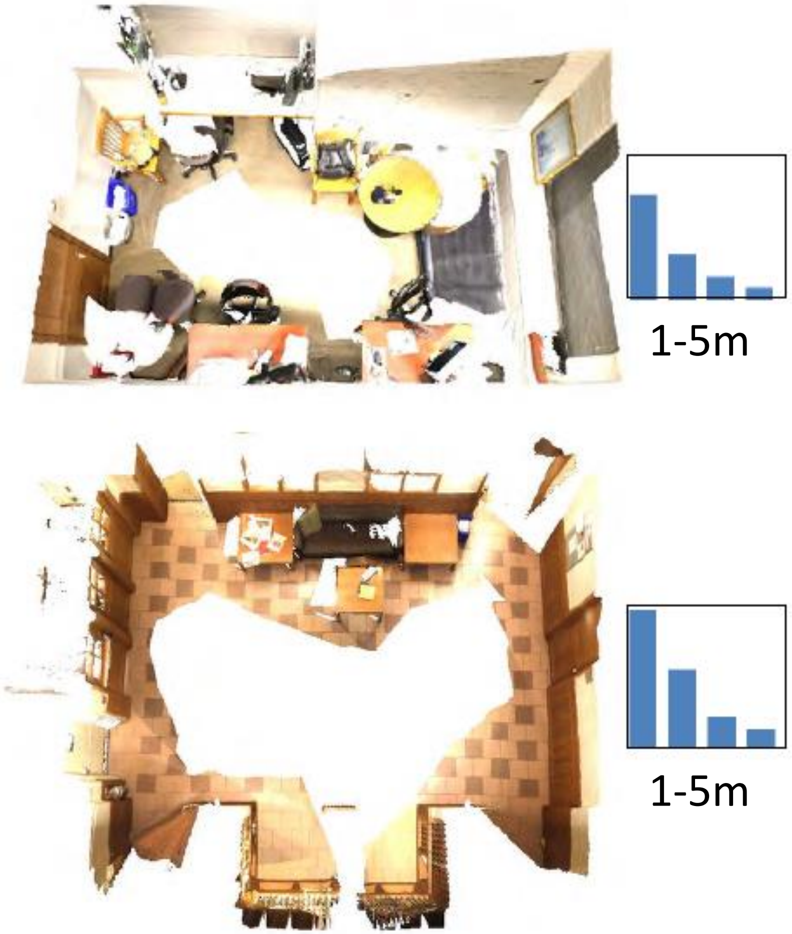
RMSE in cm (lower is better)

Effect of Long-range Co-planar Pairs

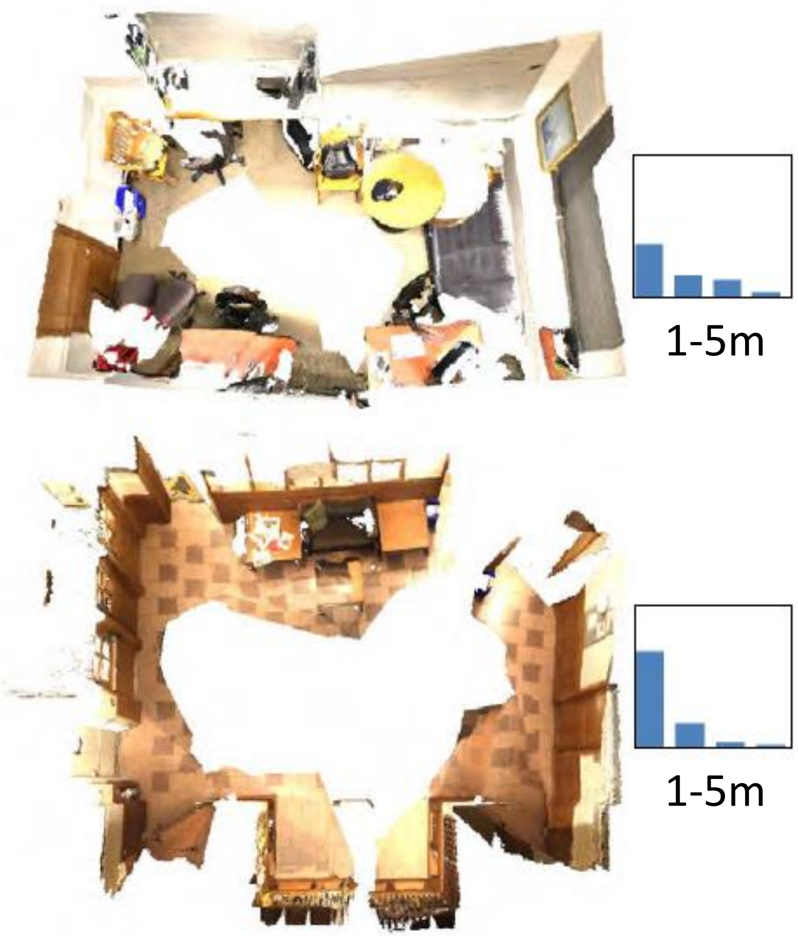


0% deduction

Effect of Long-range Co-planar Pairs

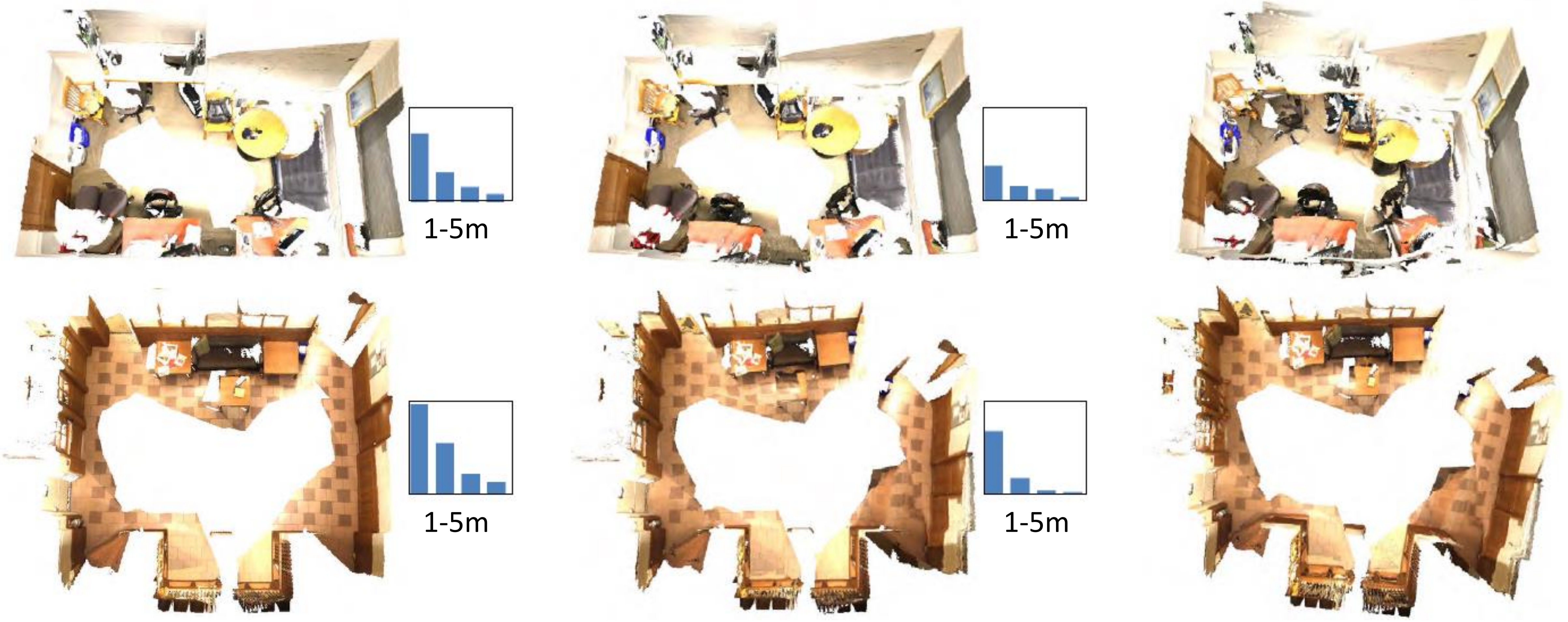


0% deduction



50% deduction

Effect of Long-range Co-planar Pairs



0% deduction | 50% deduction | 100% deduction

Conclusion

1. New task: co-planarity matching
2. Feature learning using self-supervision
3. Integration with robust optimization into SLAM

Thank You!

Yifei Shi

Kai Xu

Szymon Rusinkiewicz

Tom Funkhouser