SIGGRAPH 2022
VANCOUVER+ 8-11 AUG

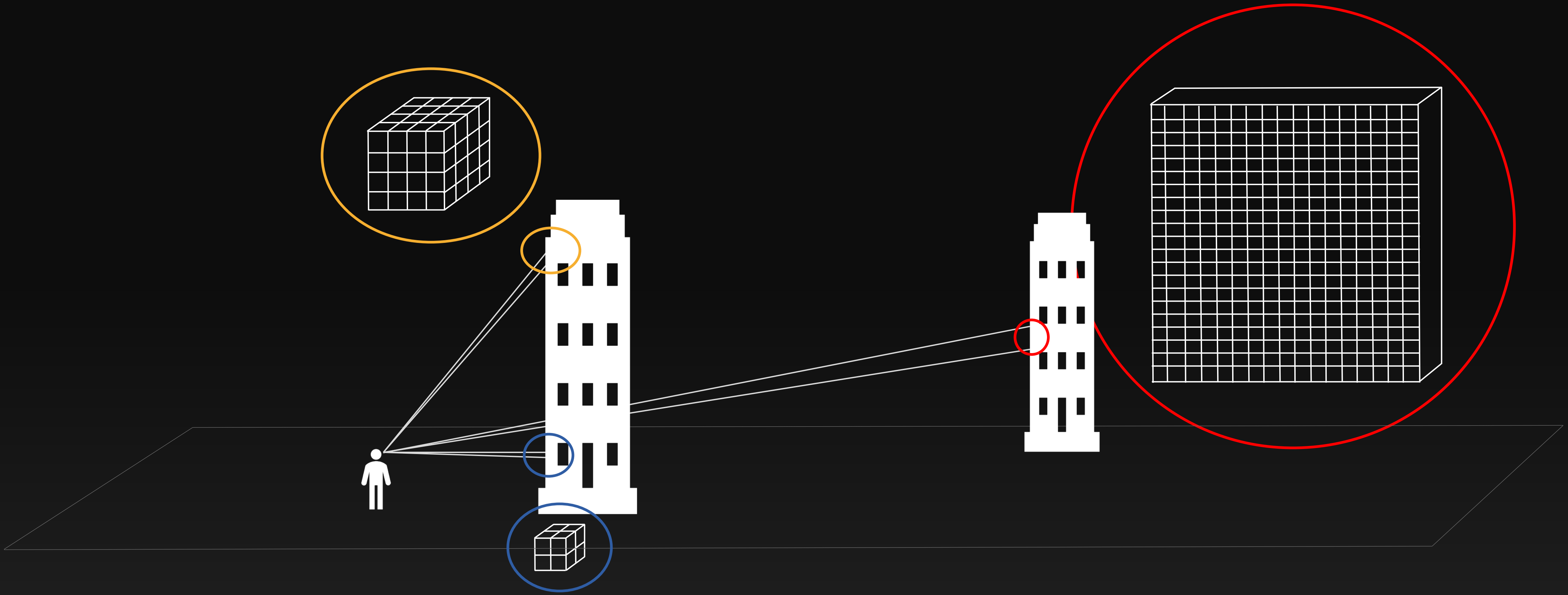
THE PREMIER CONFERENCE & EXHIBITION ON
COMPUTER GRAPHICS & INTERACTIVE TECHNIQUES

EGOCENTRIC SCENE RECONSTRUCTION FROM AN OMNIDIRECTIONAL VIDEO

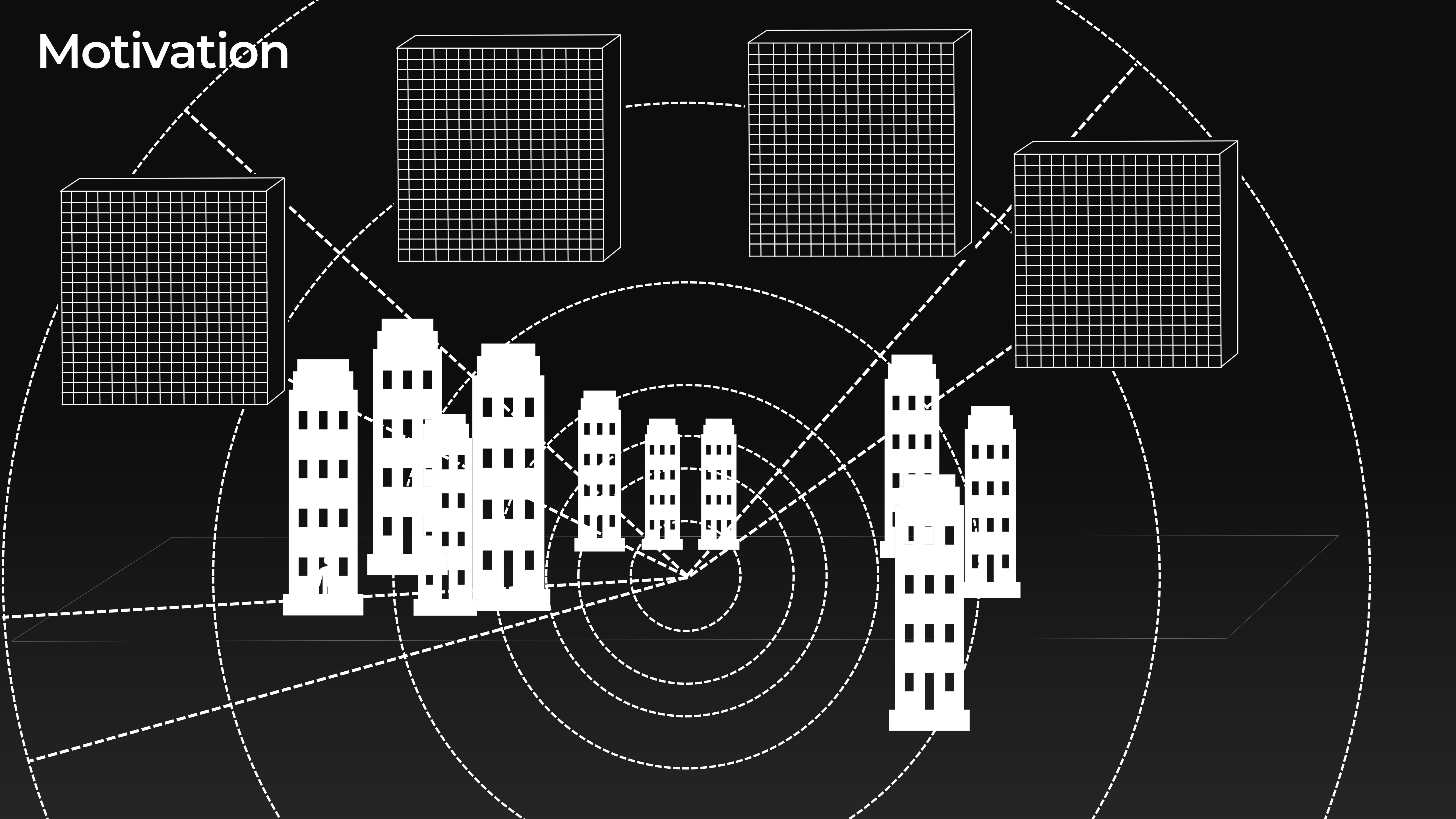
HYEONJOONG JANG ANDRÉAS MEULEMAN DAHYUN KANG
DONGGUN KIM CHRISTIAN RICHARDT MIN H. KIM


FARMACIA

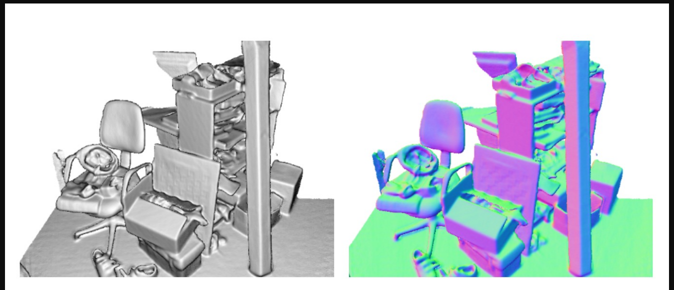
Motivation



Motivation



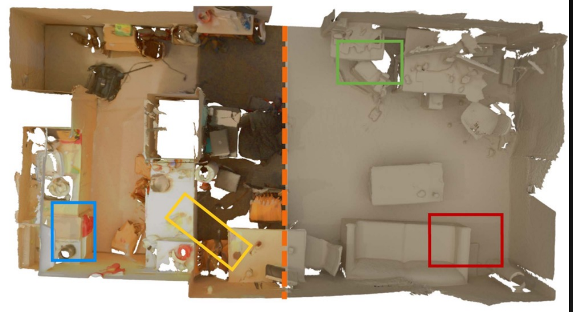
Related work



Newcombe et al. (2011)

Whelan et al. (2016)

Nießner et al. (2013)



Dai et al. (2017)

Active depth sensors have a short-range

Related work

Anderson et al. (2016)

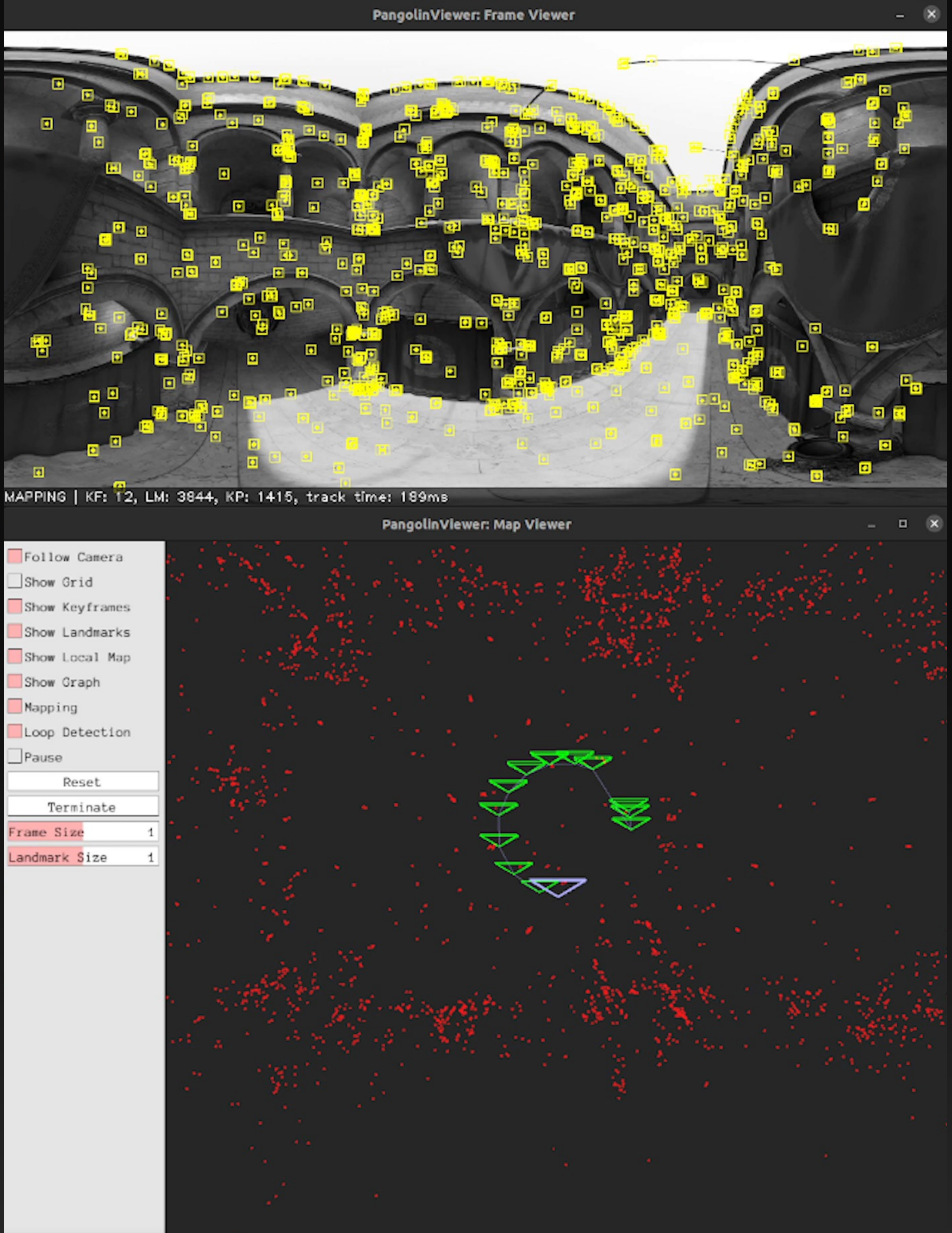
Parra Pozo et al. (2019)

Broxton et al. (2020)

High cost and not portable

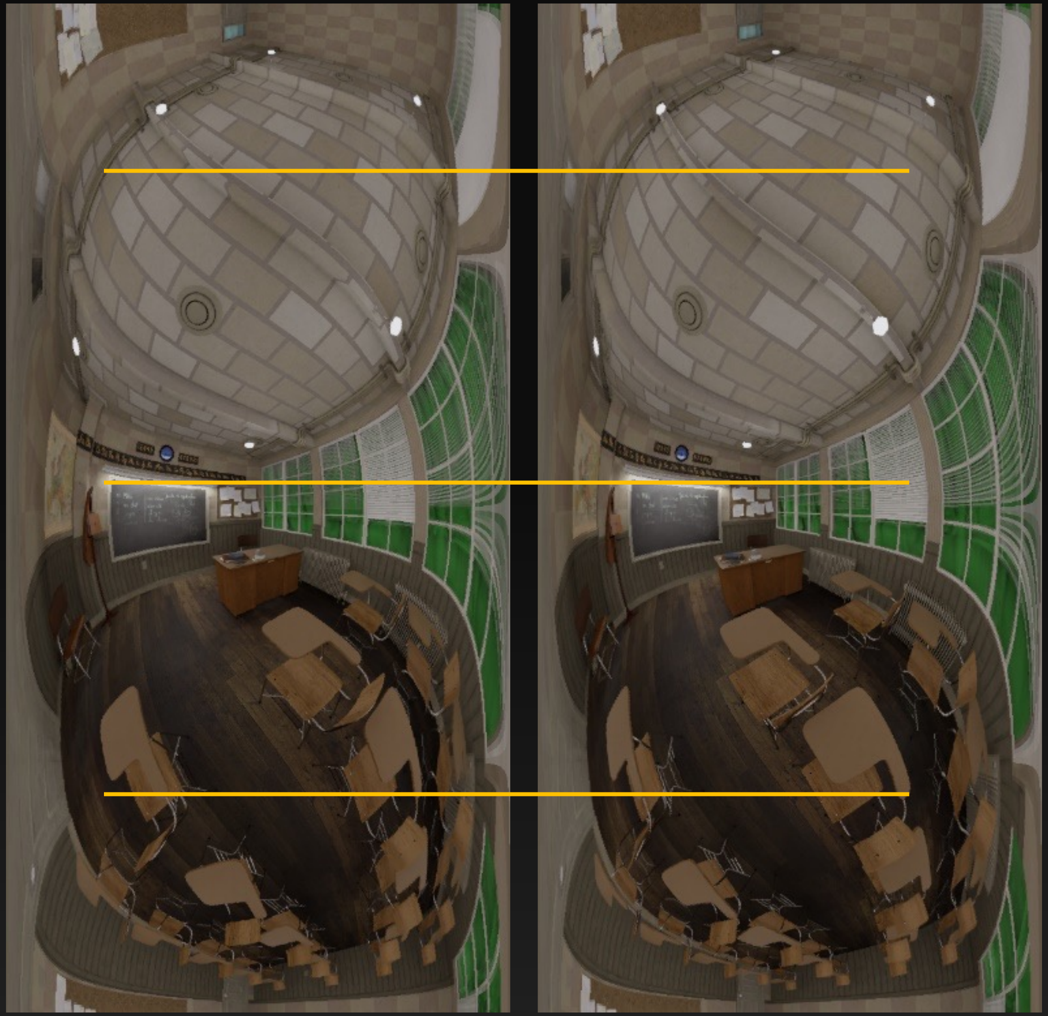
Pose estimation

- 2-pass OpenVSLAM
  1. Reconstruct 3D Map
  2. Estimate camera poses



Depth estimation

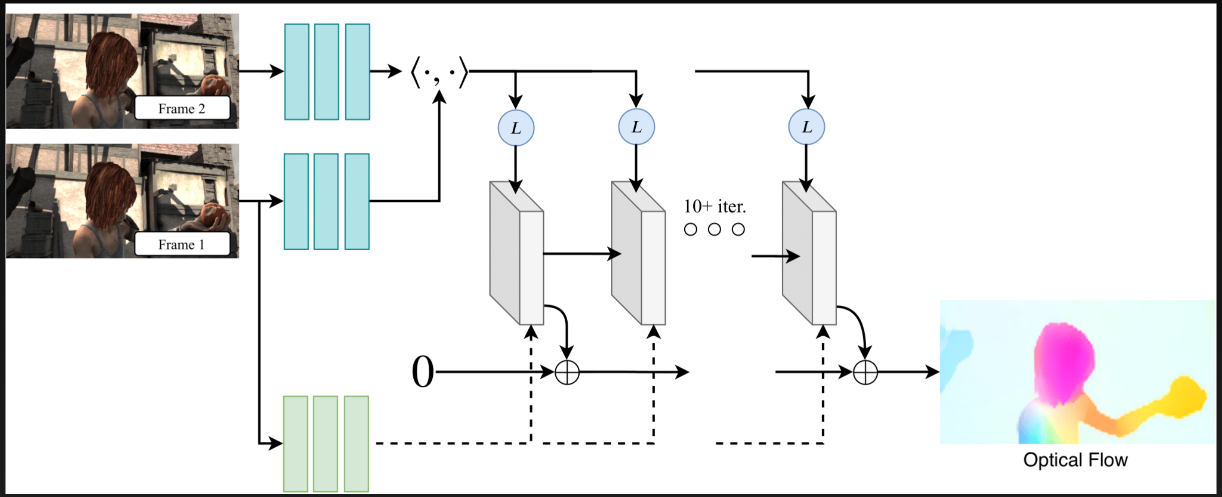
Given 360° image pair



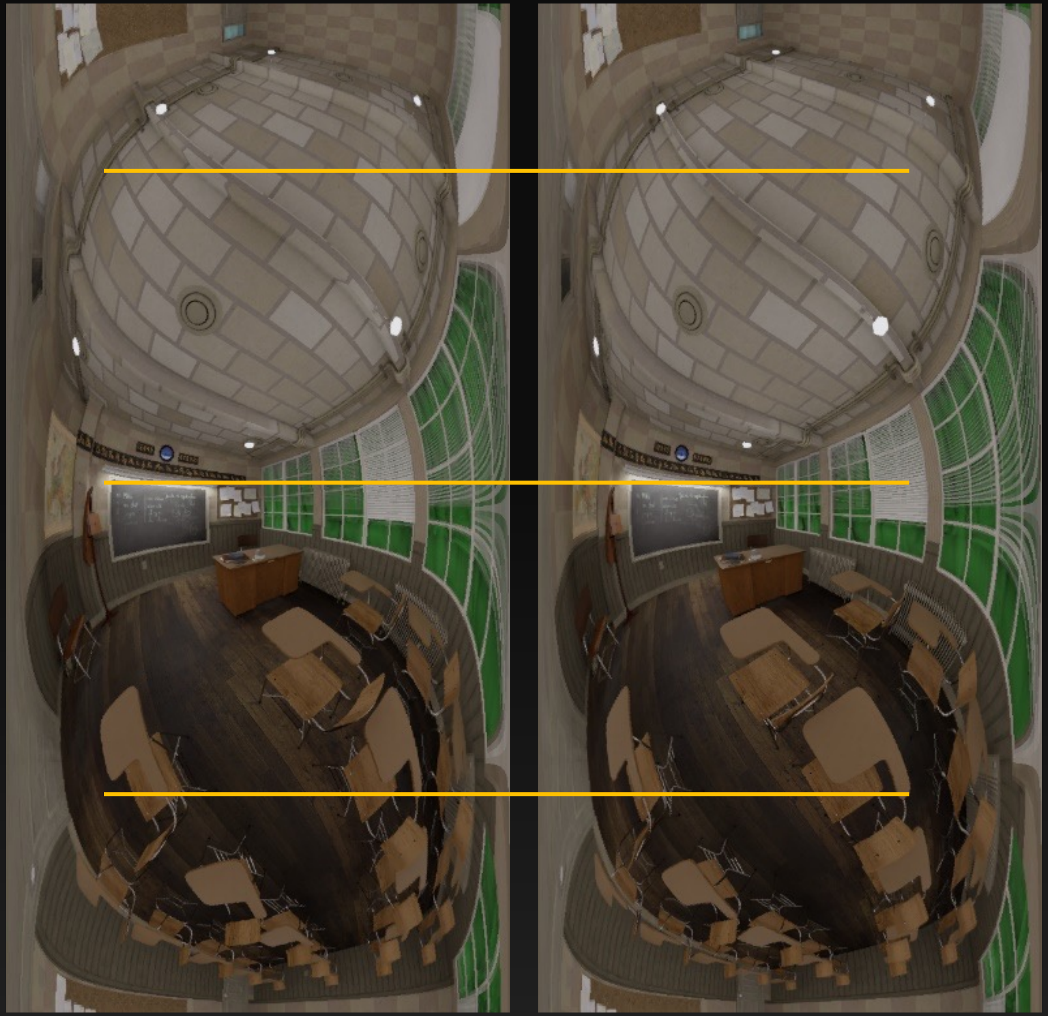
Rectified spherical stereo pair

Depth estimation

- 1D line search problem
  - Disparity estimation method
  - Depth estimation method
  - Optical flow estimation method

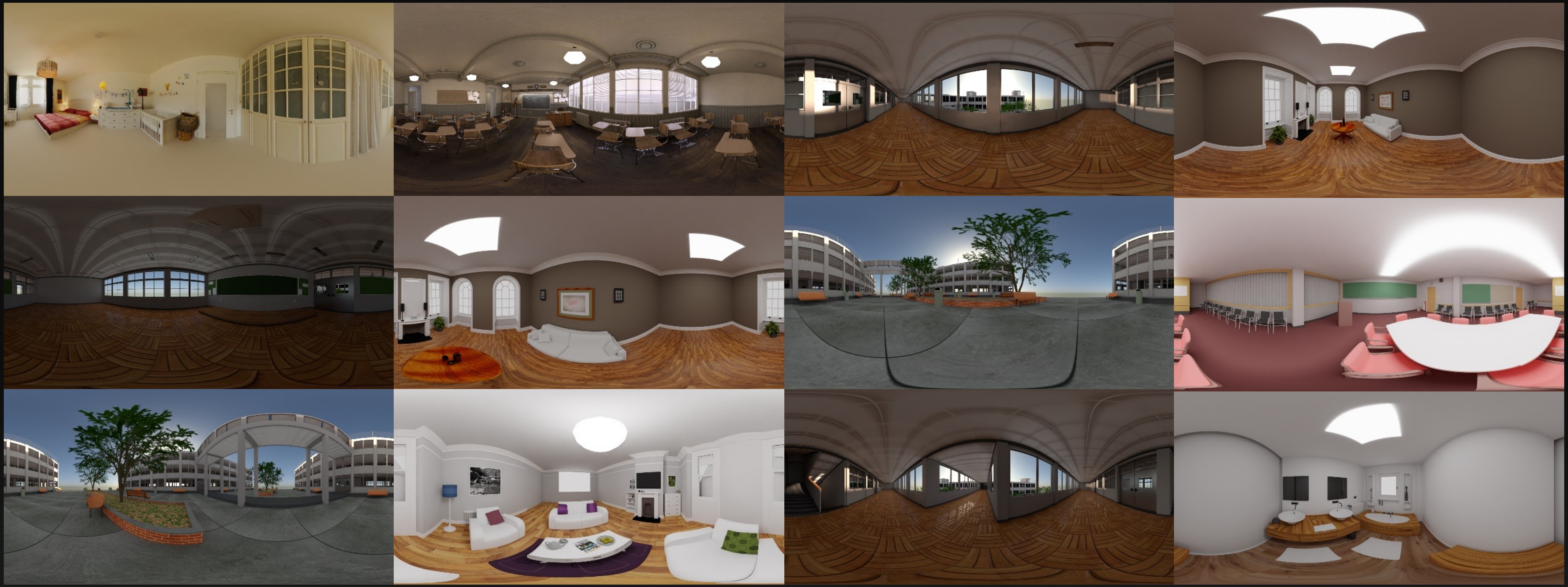


RAFT
Teed and Deng (2020)



Rectified spherical stereo pair

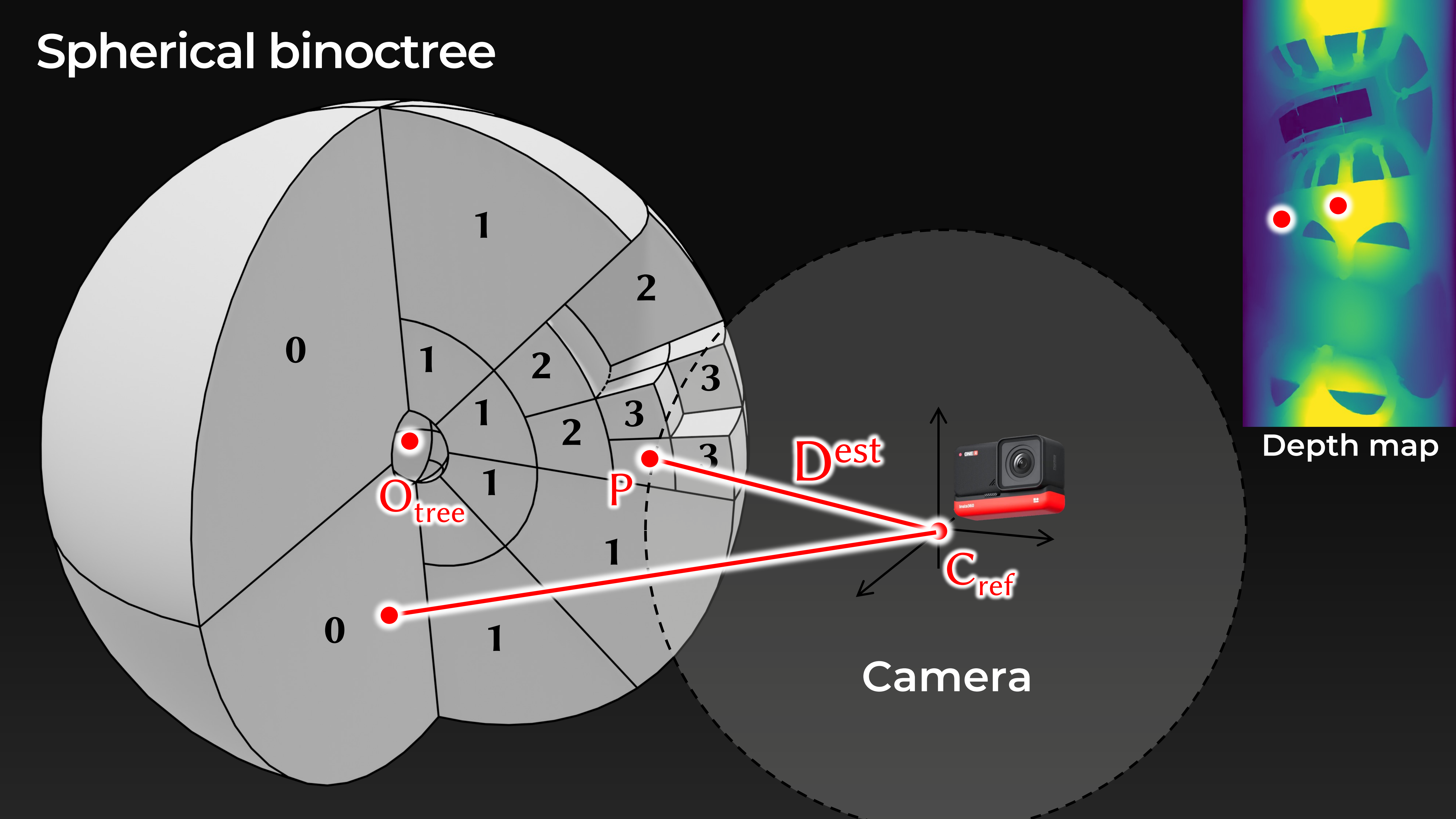
360° RGBD video training dataset




FARMACIA
Input video

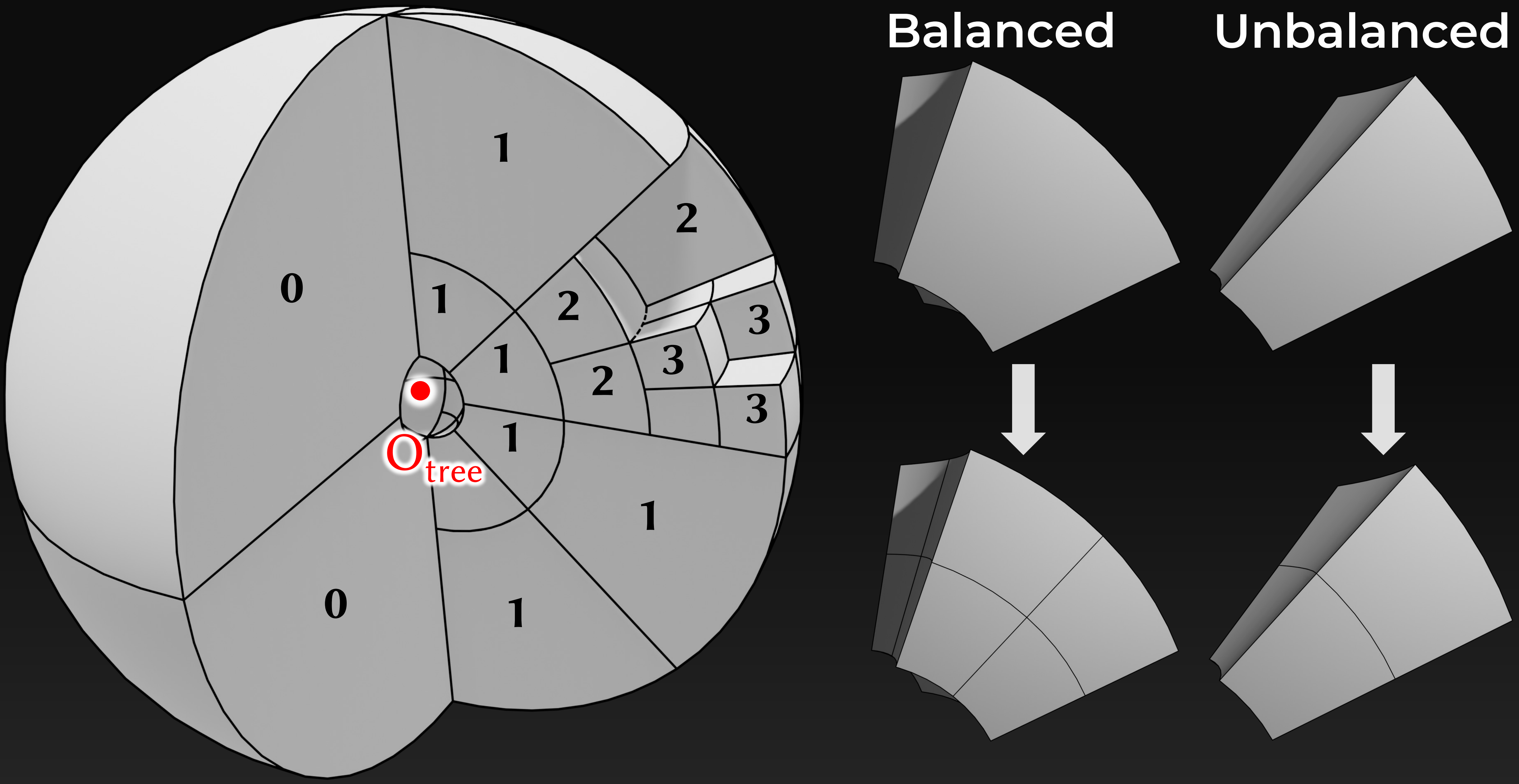

GIN
BAR
Input video


Input video

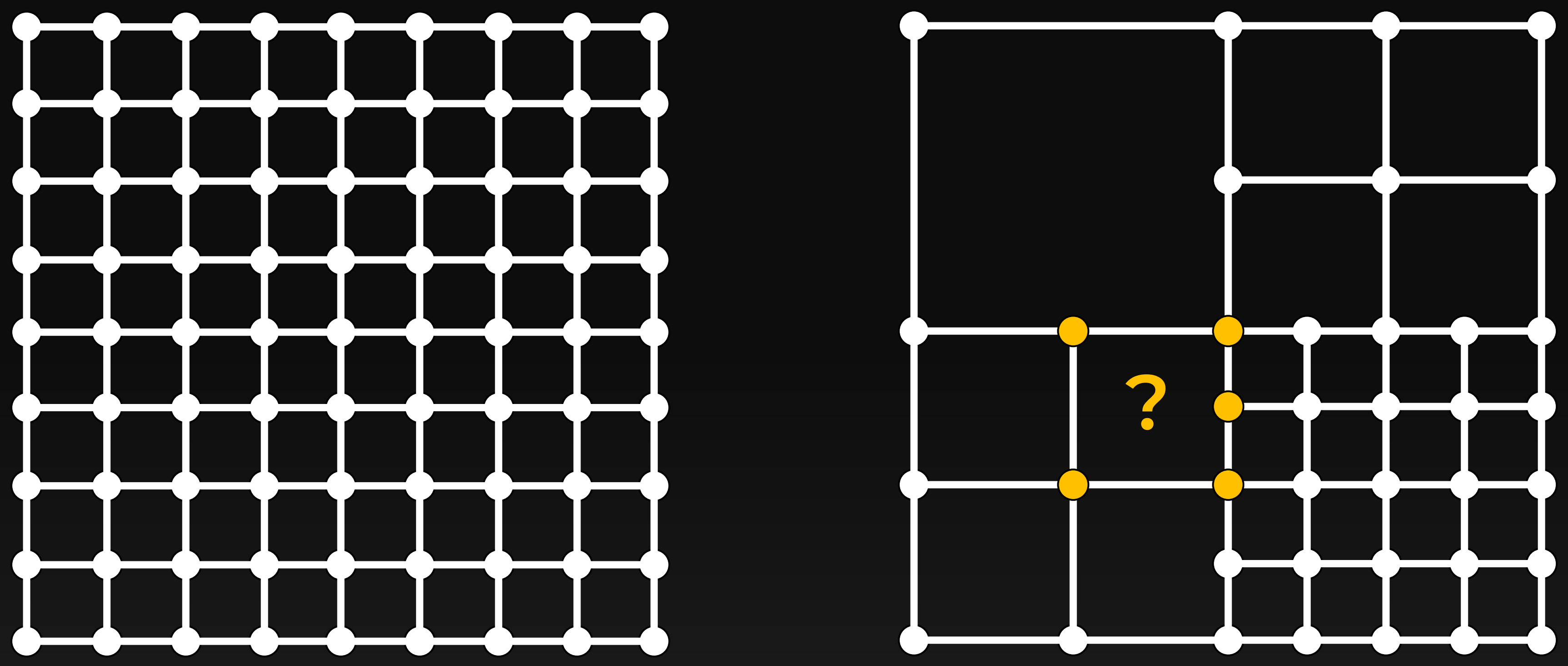


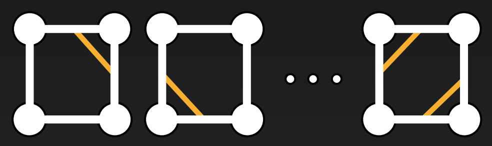
Spherical binoctree
0
1
2
3
0
1
1
1
1
2
2
3
3
3
1
1
O_{tree}
P
Dest
C_{ref}
Camera
Depth map

Spherical binoctree

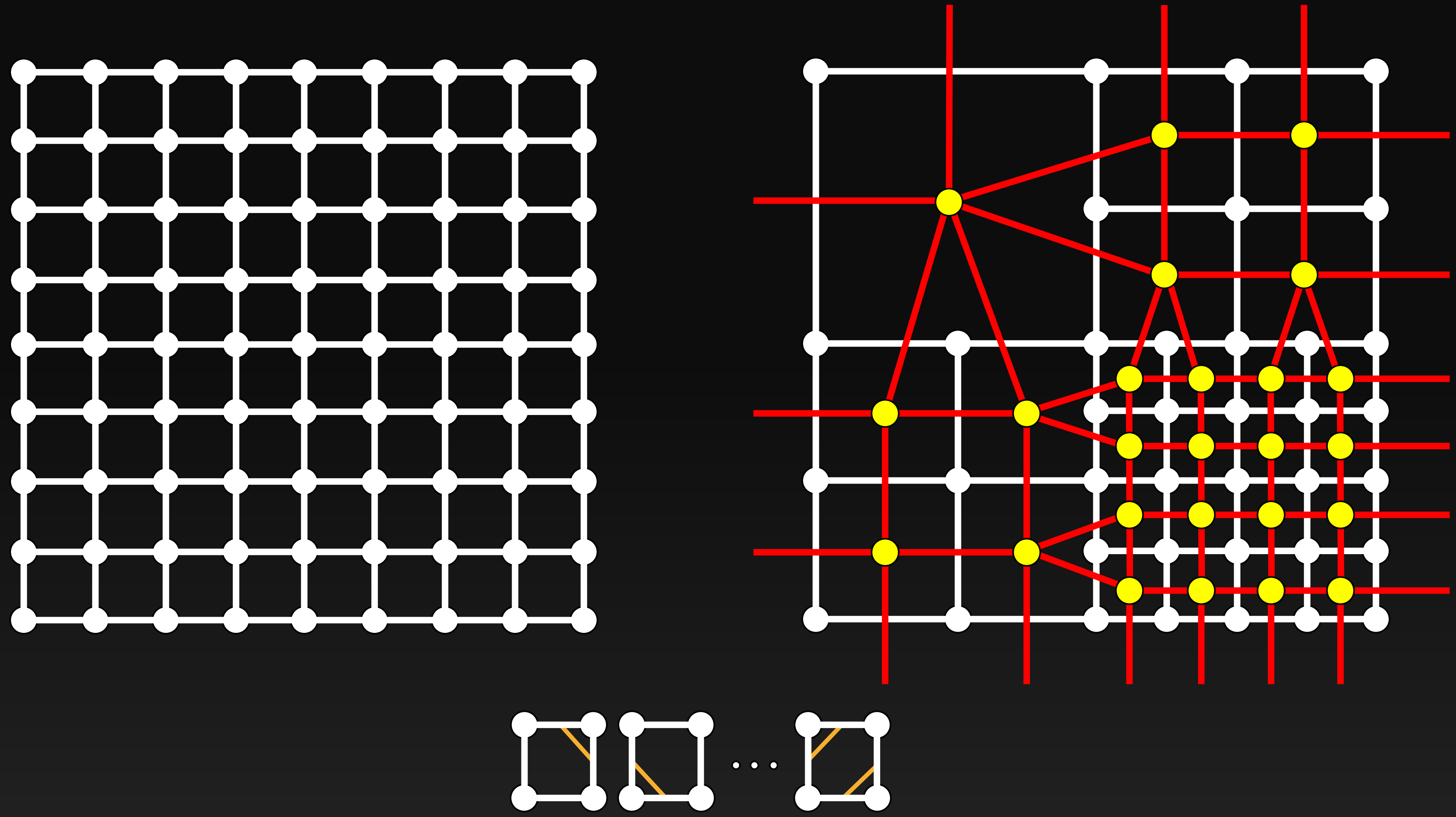


Mesh generation



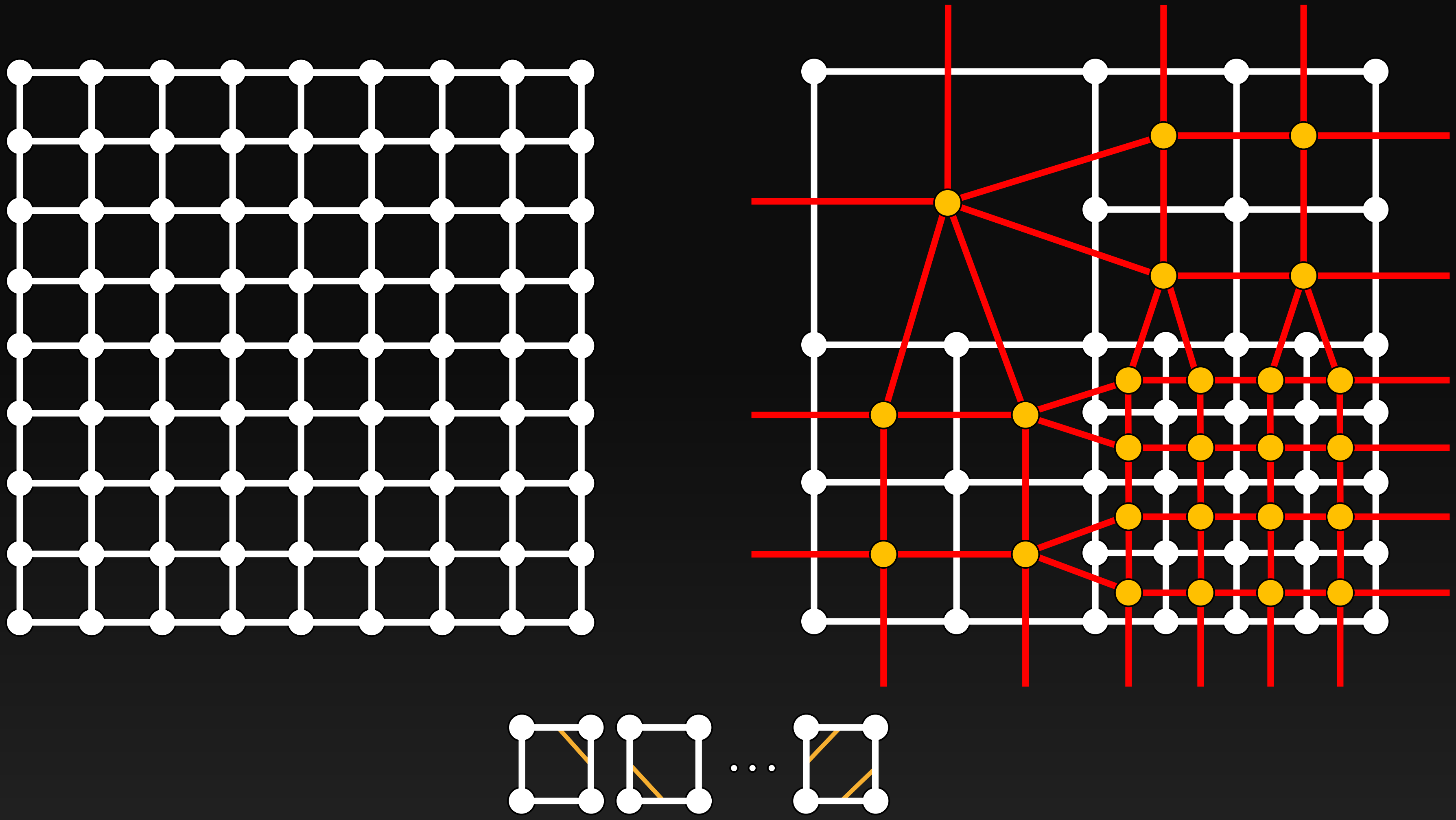


Dual marching cubes



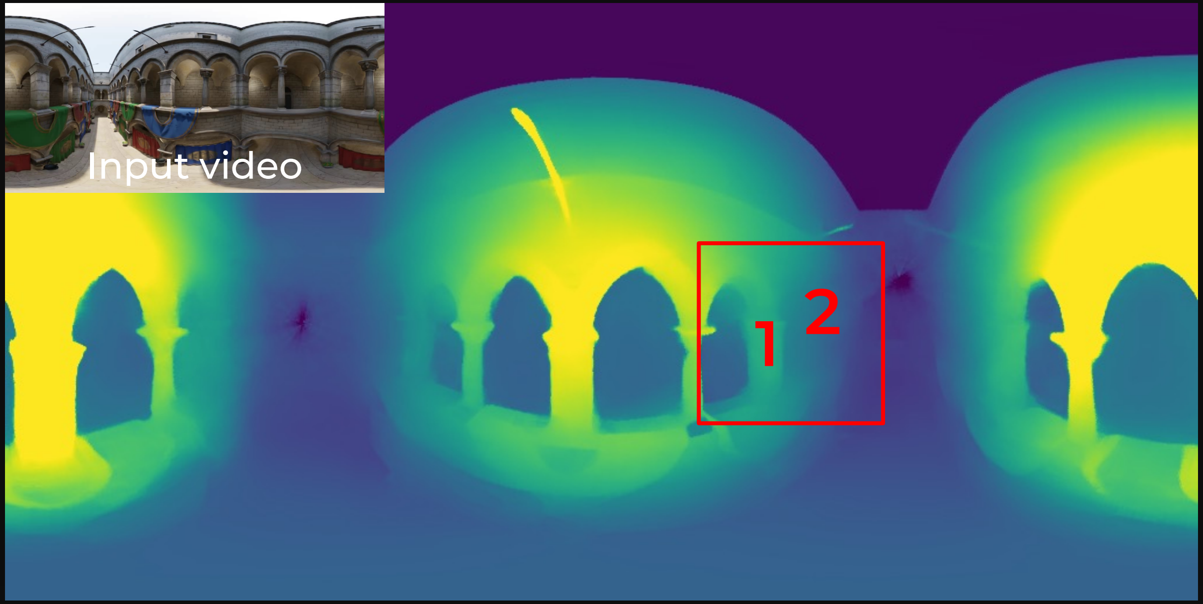
Dual Marching Cubes, Schaefer and Warren (2005)

Dual marching cubes



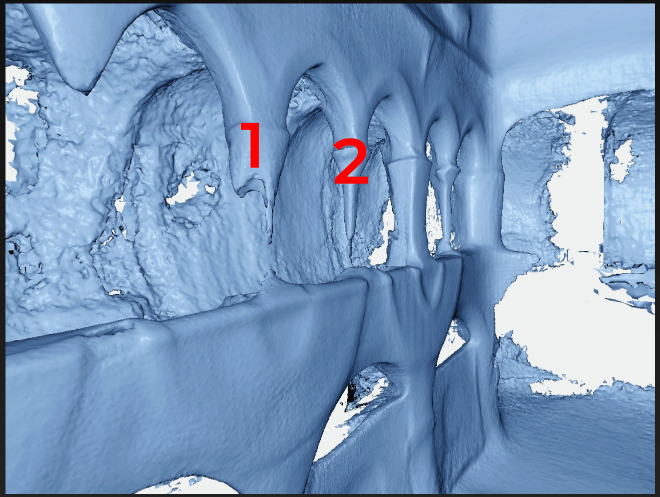
Dual Marching Cubes, Schaefer and Warren (2005)

TSDF integration



Estimated depth

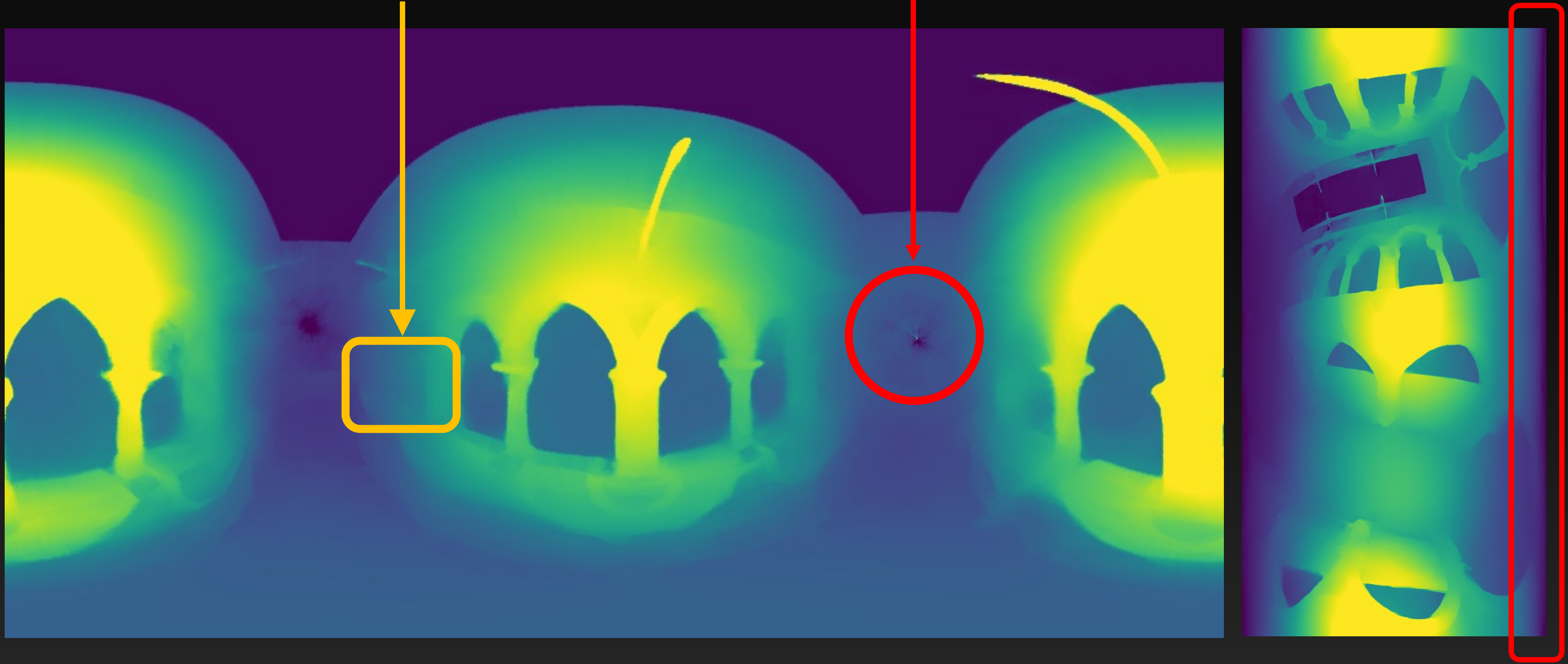
Closeup



Result mesh

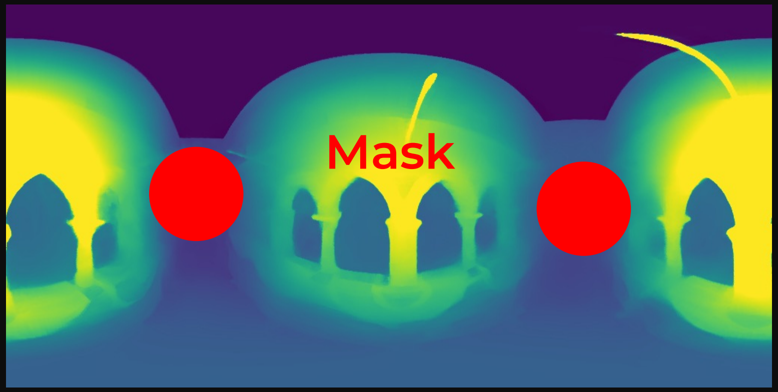
TSDF integration

- Low depth accuracy along the baseline axis
- Low depth accuracy for distant points



TSDF integration

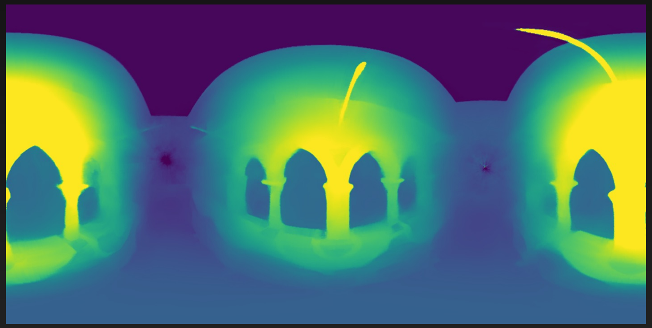
- Low depth accuracy along the baseline axis
- Low depth accuracy for distant points
- Lack of checking depth and color consistency



$$w_{\text{update}}(i,p) = w_{\text{p}}(i,p) \cdot \sum_{j \in N(i)} w_{\text{d}}(i,j,p) \cdot w_{\text{c}}(i,j,p)$$

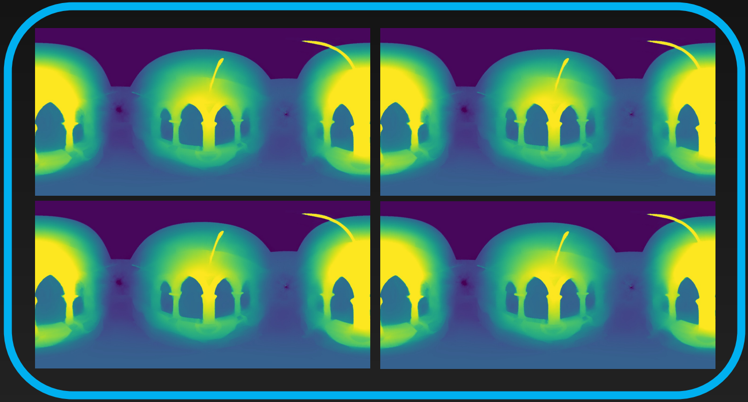
Small weight for distant points

Small weight for large difference of depth and color



i

reproject



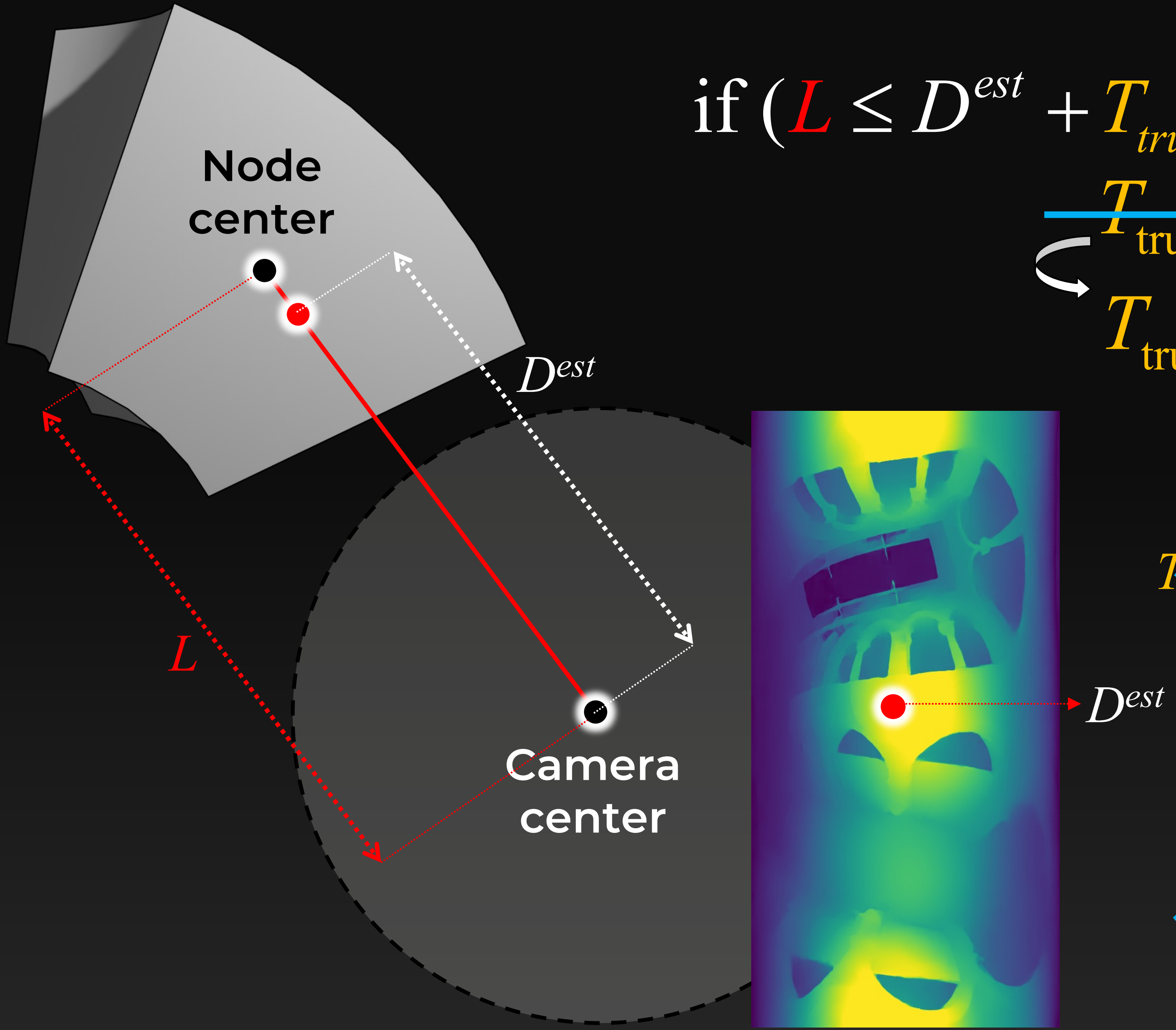
$N(i)$

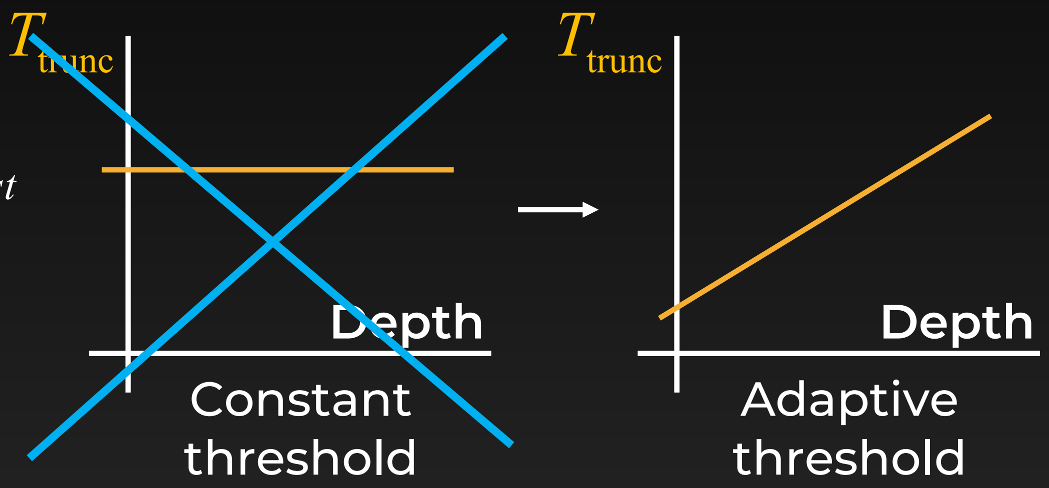
TSDF integration

$$\text{if } (L \leq D^{est} + T_{trunc}): W_{TSDF} = W_{TSDF} + w_{update}$$

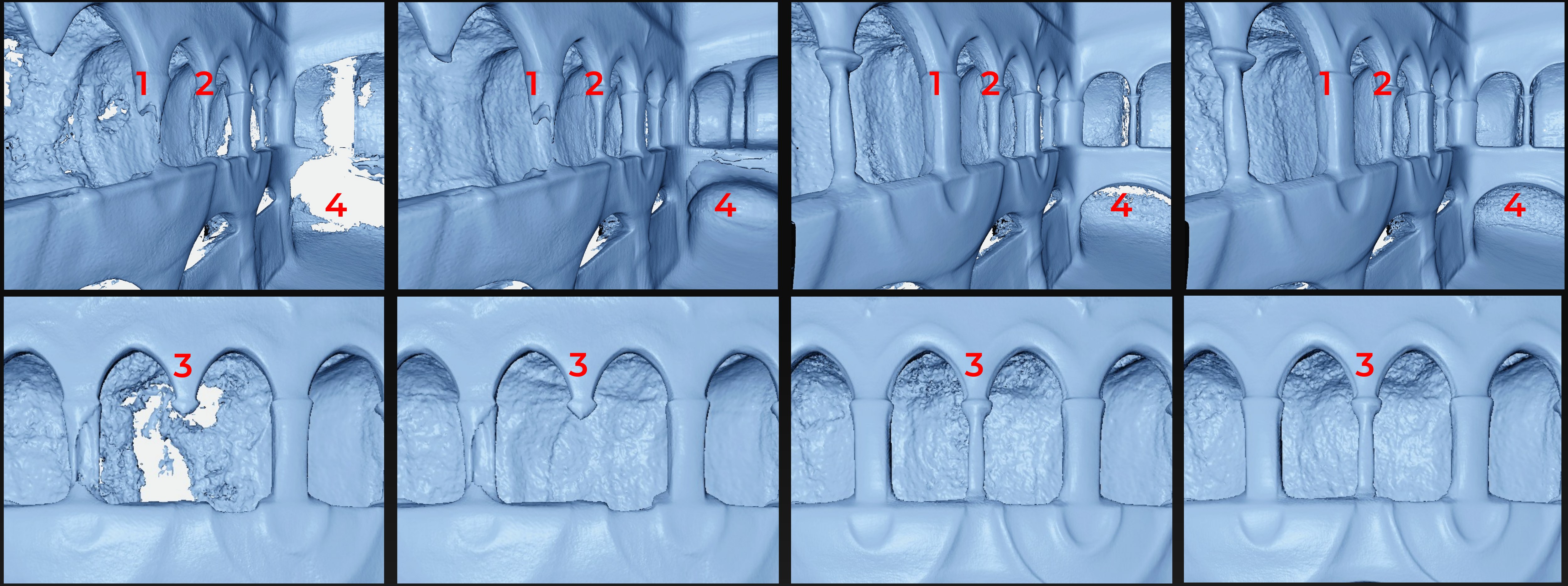
$$T_{\text{trunc}} = b$$

$$T_{\text{trunc}}(D^{est}) = aD^{est} + b$$





TSDF integration



Adaptive truncation(X) / Confidence weight (X)

Adaptive truncation(O) / Confidence weight (X)

Adaptive truncation(X) / Confidence weight (O)

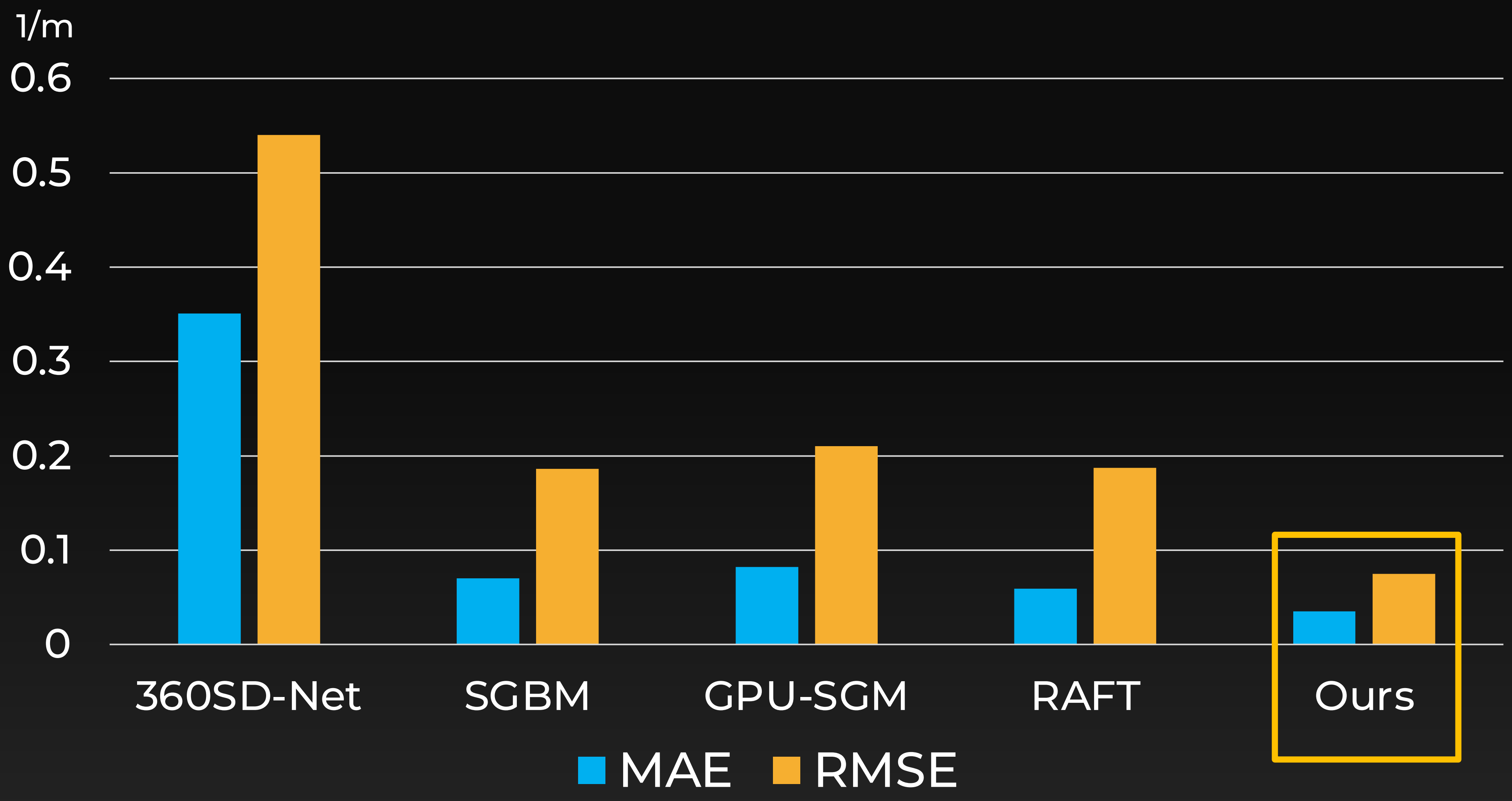
Adaptive truncation(O) / Confidence weight (O)

Reconstruction comparison


Input video
(small circular camera trajectory)

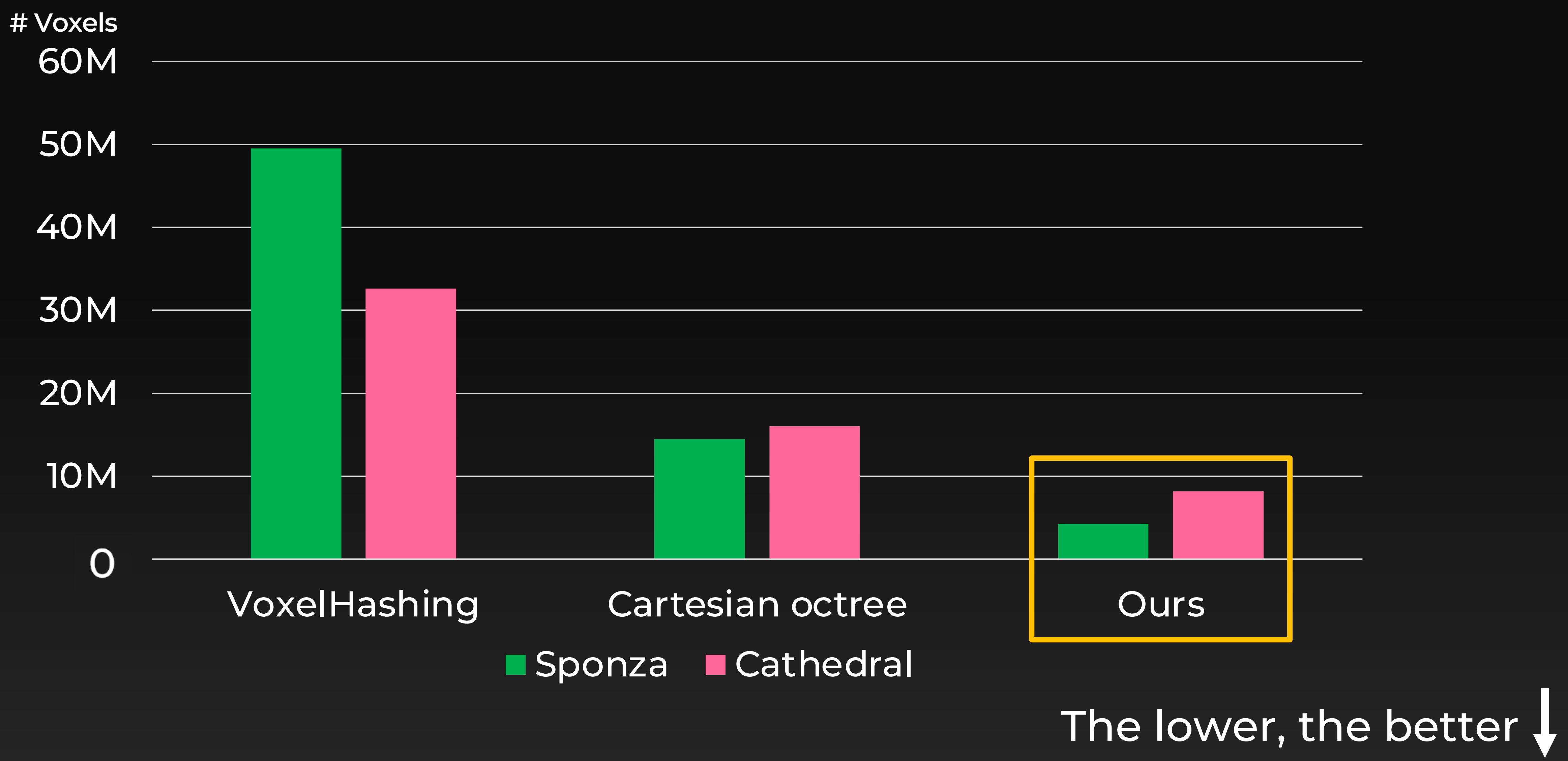

GT
COLMAP
Schonberger et al. (2016)
OmniSLAM
Won et al. (2020)
Ours

Depth accuracy comparison

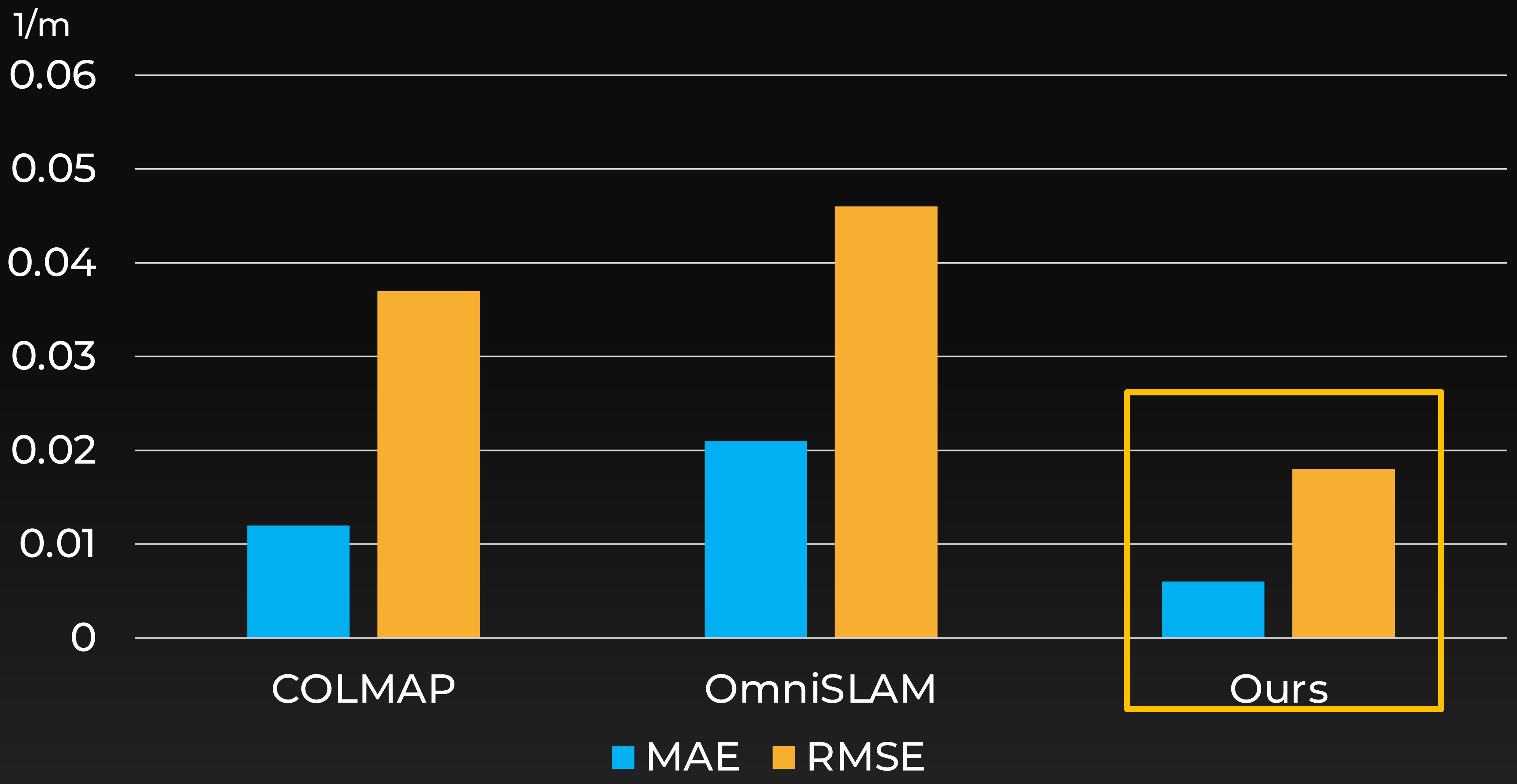


The lower, the better

Memory efficiency comparison



Mesh accuracy comparison



The lower, the better

COLMAP Ours


Vintage 'n' Rare

Limitations

Thin objects　Specular reflections　Dynamic objects

Conclusion

Scene scale 3D reconstruction from an omnidirectional video

- Accurate 360° depth estimation:

  ➔ 360° RGBD video dataset

- Efficient voxel allocation:

  ➔ Spherical binoctree data structure

- Full mesh from a short camera trajectory:

  ➔ adaptive truncation threshold

Project page: vclab.kaist.ac.kr/siggraph2022p2/

Thank you

Hyeonjoong Jang Andréas Meuleman Dahyun Kang
Donggun Kim Christian Richardt Min H. Kim

KAIST UNIVERSITY OF BATH