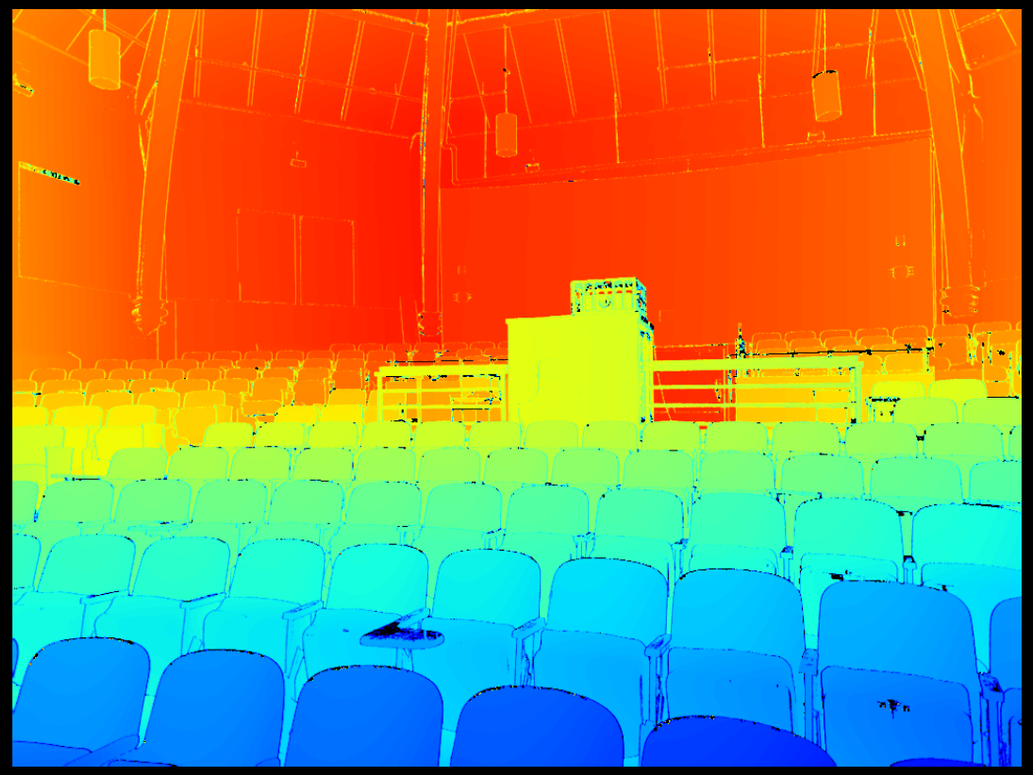
DIODE: A Diverse Indoor/ Outdoor DEpth Dataset

Igor Vasiljevic
TTI-Chicago

TOYOTA TECHNOLOGICAL INSTITUTE AT CHICAGO

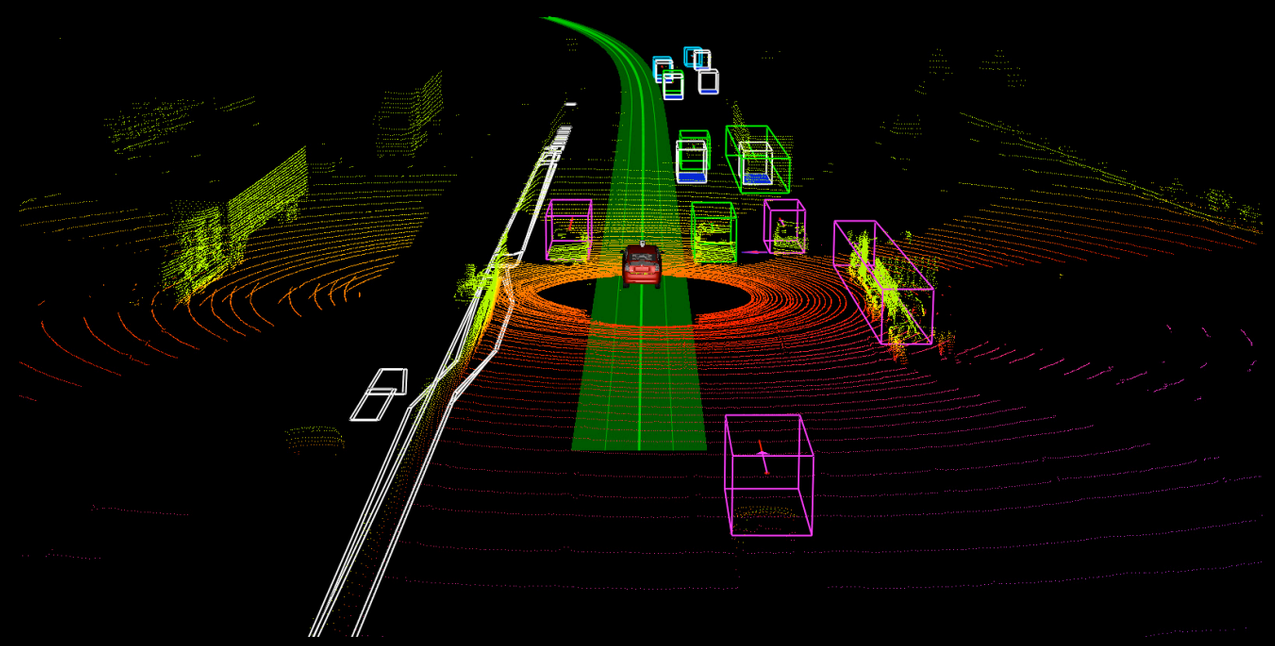
Depth Estimation

- **Many advances in 2D vision, but need 3D info to link to real world**
- **First step in 2.5D vision — depth sensors are expensive, noisy, and/or sparse, cameras are cheap**



RGB-D Data (Outdoor)

- **KITTI dataset — predominantly street scenes taken with Velodyne LiDAR sensor (tied to one task)**
- **Very sparse — 5-20% of pixels have depth values**



Geiger, Andreas, Philip Lenz, and Raquel Urtasun. "Are we ready for autonomous driving? the kitti vision benchmark suite." *CVPR 2012*

RGB-D Data (Indoor)

- **NYUv2 dataset — dense, but noisy (+ low angular resolution) depth from Kinect sensor, inpainted**
- **Predominantly office and apartment scenes**

Silberman, Nathan, et al. "Indoor segmentation and support inference from rgbd images." *ECCV 2012*

DIODE: Diverse Indoor/Outdoor DEpth Dataset

- **FARO Focus S50, an actuated survey-grade LiDAR**
- **Coaxially-mounted high-resolution color camera**
- **Large depth FOV (0.6-350m, 1mm resolution, 0.009° angular res.)**
- **Well suited for both indoors and outdoors**

DIODE: Diverse Indoor/Outdoor DEpth Dataset

- **Student offices, large residential buildings, hiking trails, lecture halls, parks, city streets, parking lots, etc.**
- **~5k indoor / ~5k outdoor RGB-D images**
- **Long scan duration: dynamic motion, curation**

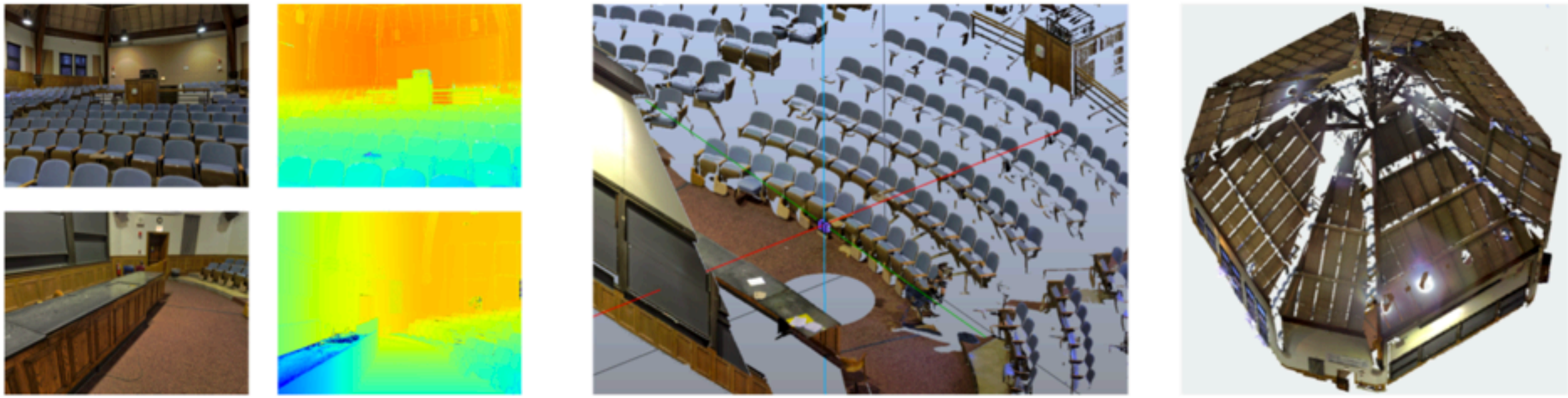
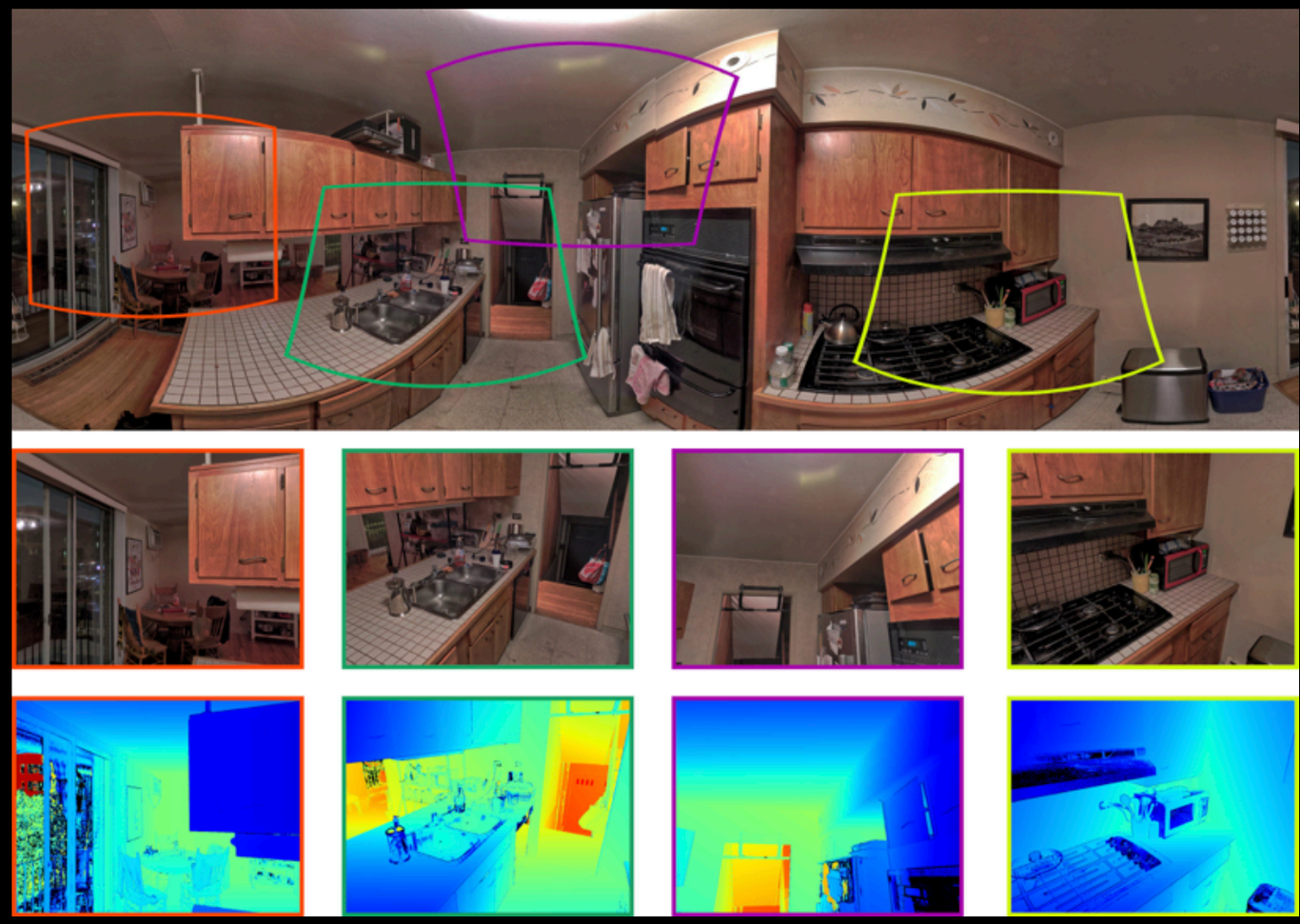
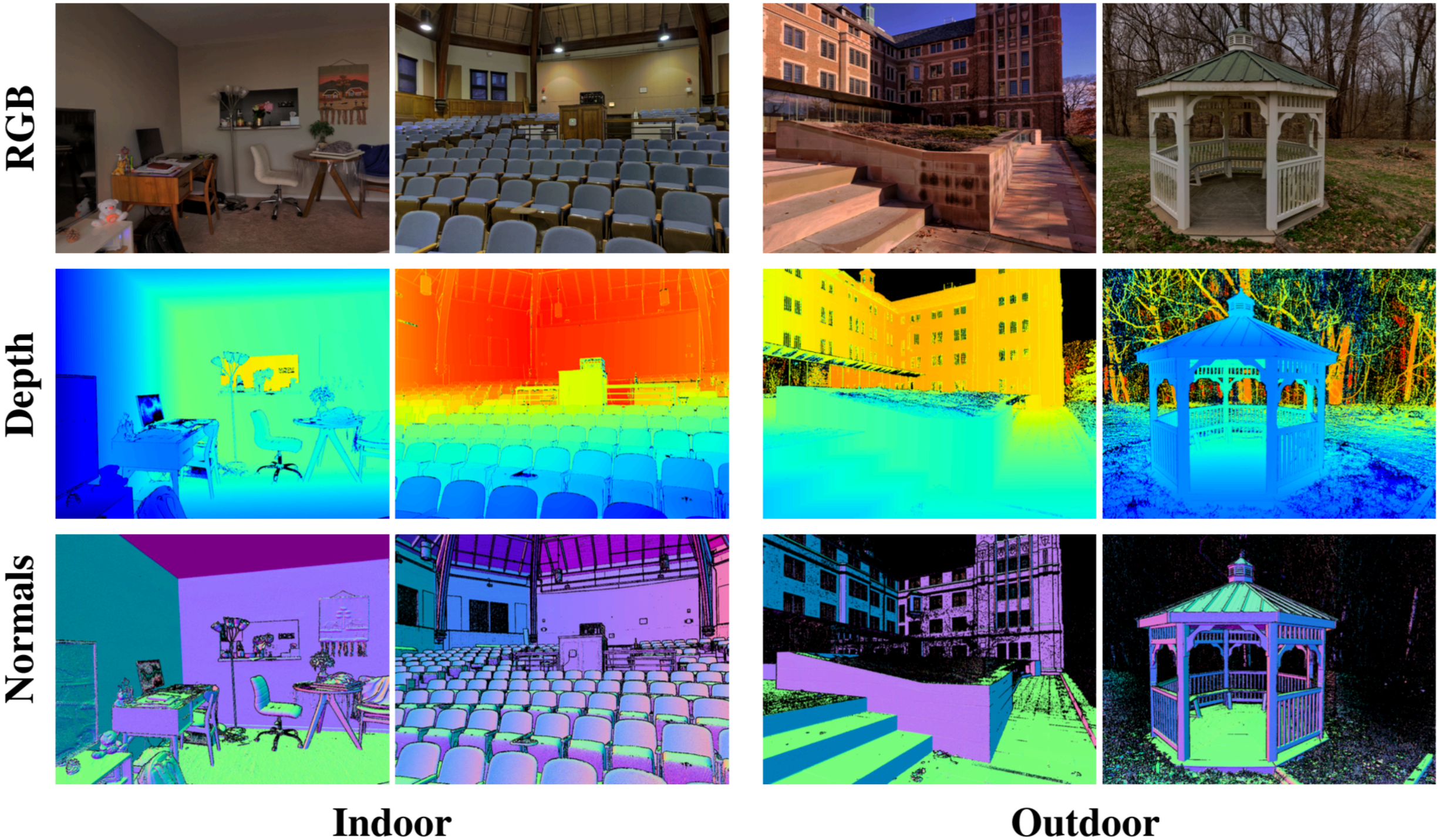


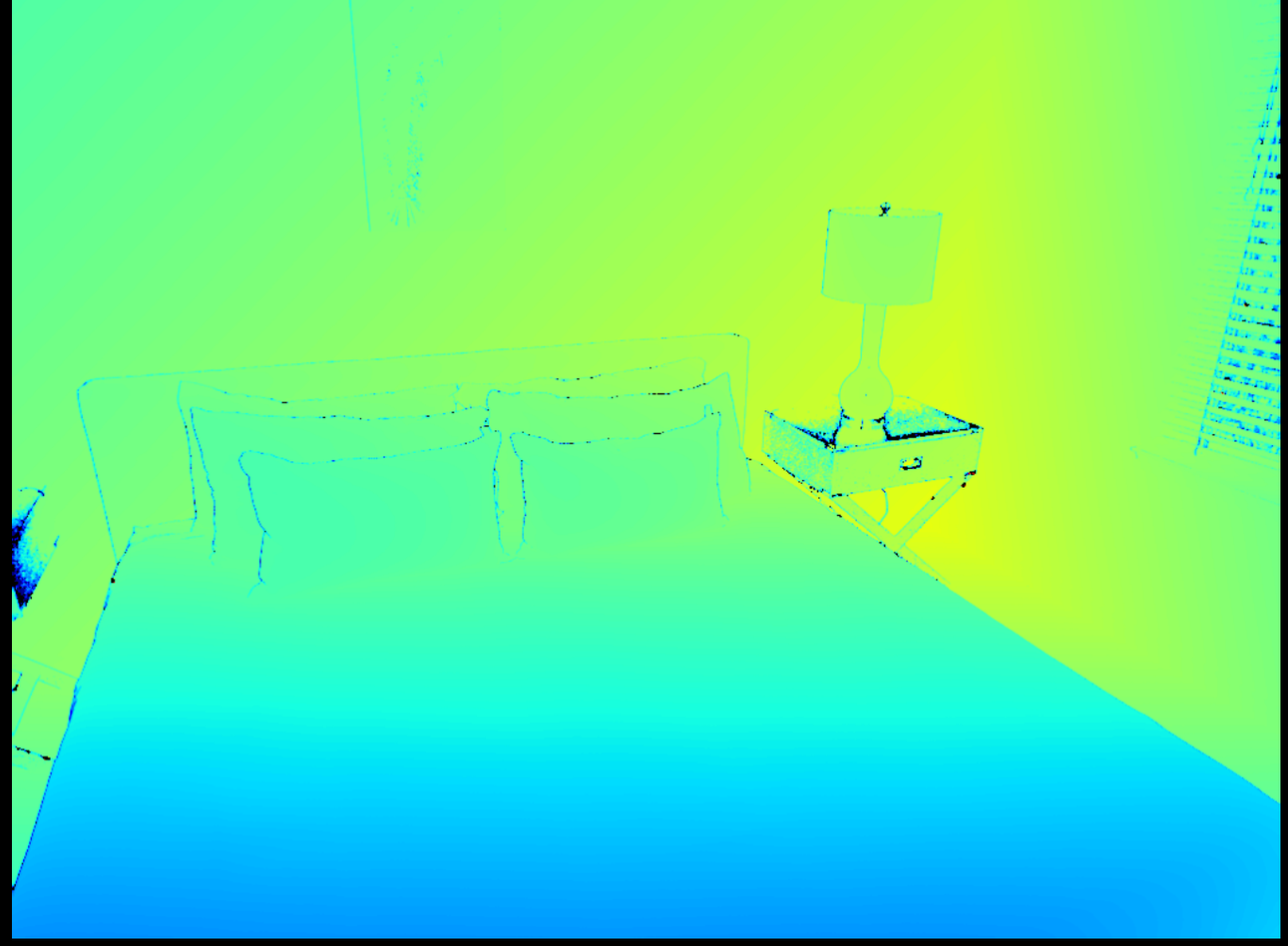
Figure 2. From left to right: RGB images and depth maps from single point cloud; side view of point cloud; top-down view of point cloud

Rectification

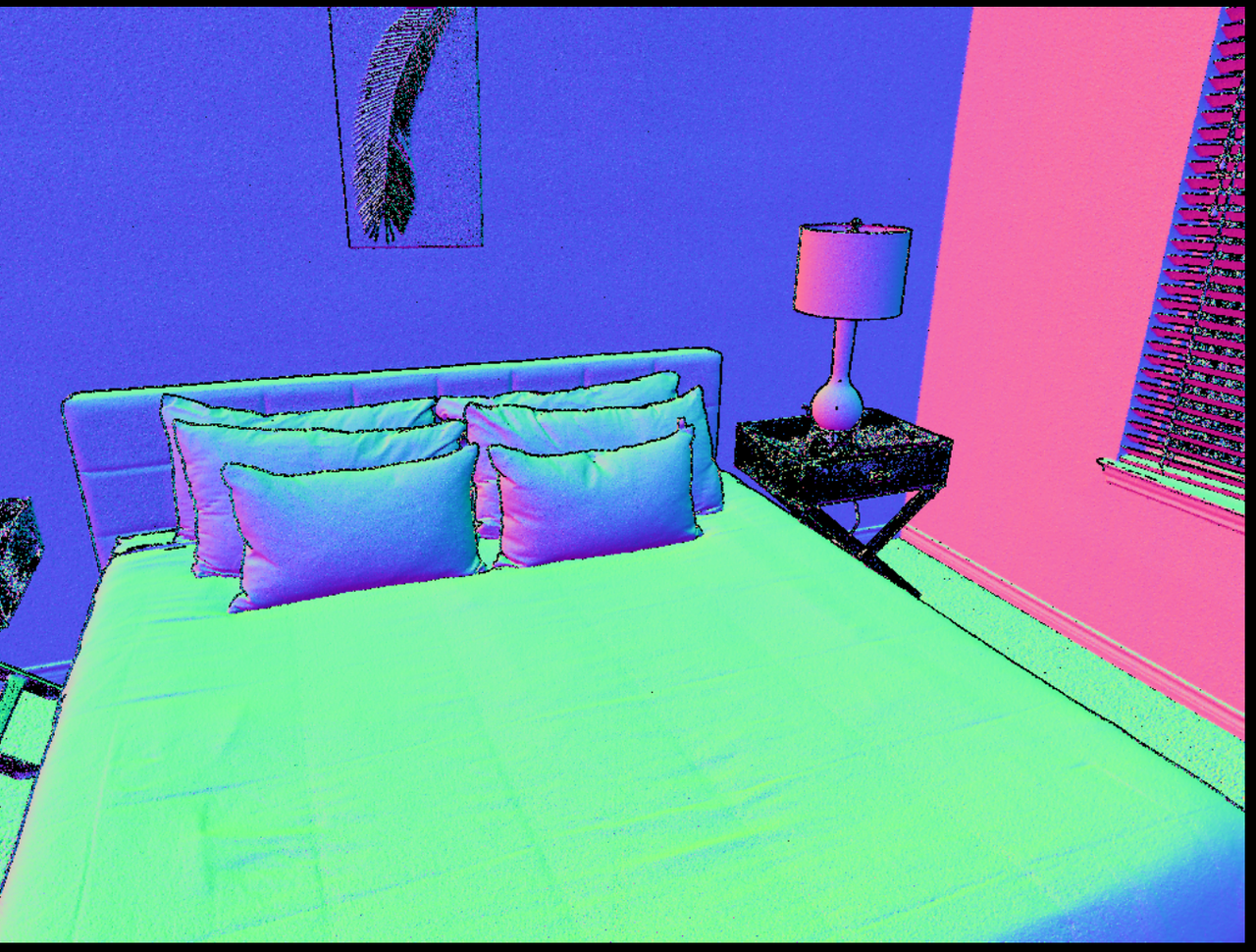
- Curved sections of the panorama need to be rectified
- Crop with respect to canonical angle, use 90° vertical and 67.5° horizontal
- Downsample 5× to obtain 768 x1024 cropped image



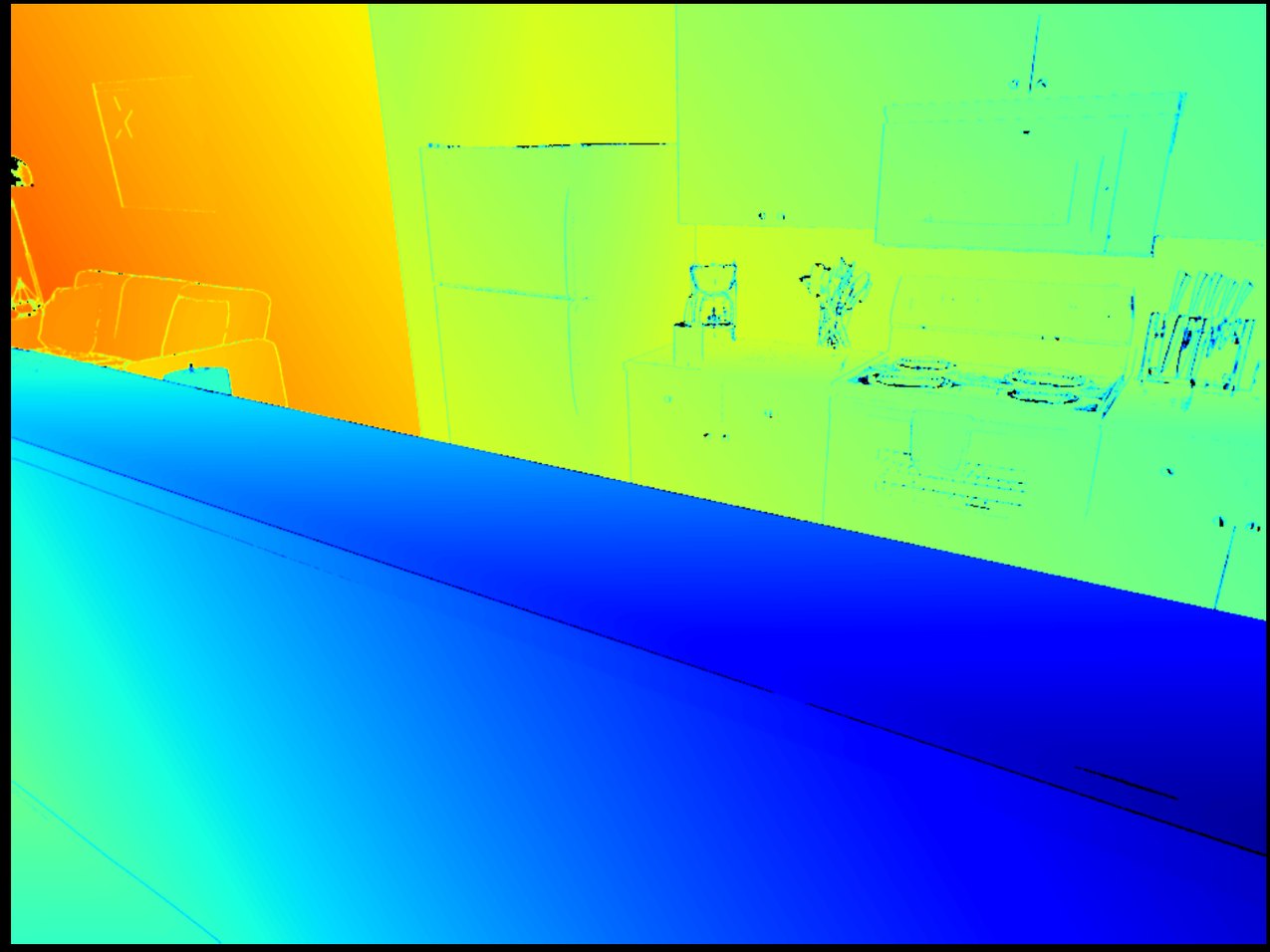

RGB
Depth
Normals
Indoor
Outdoor



Depth

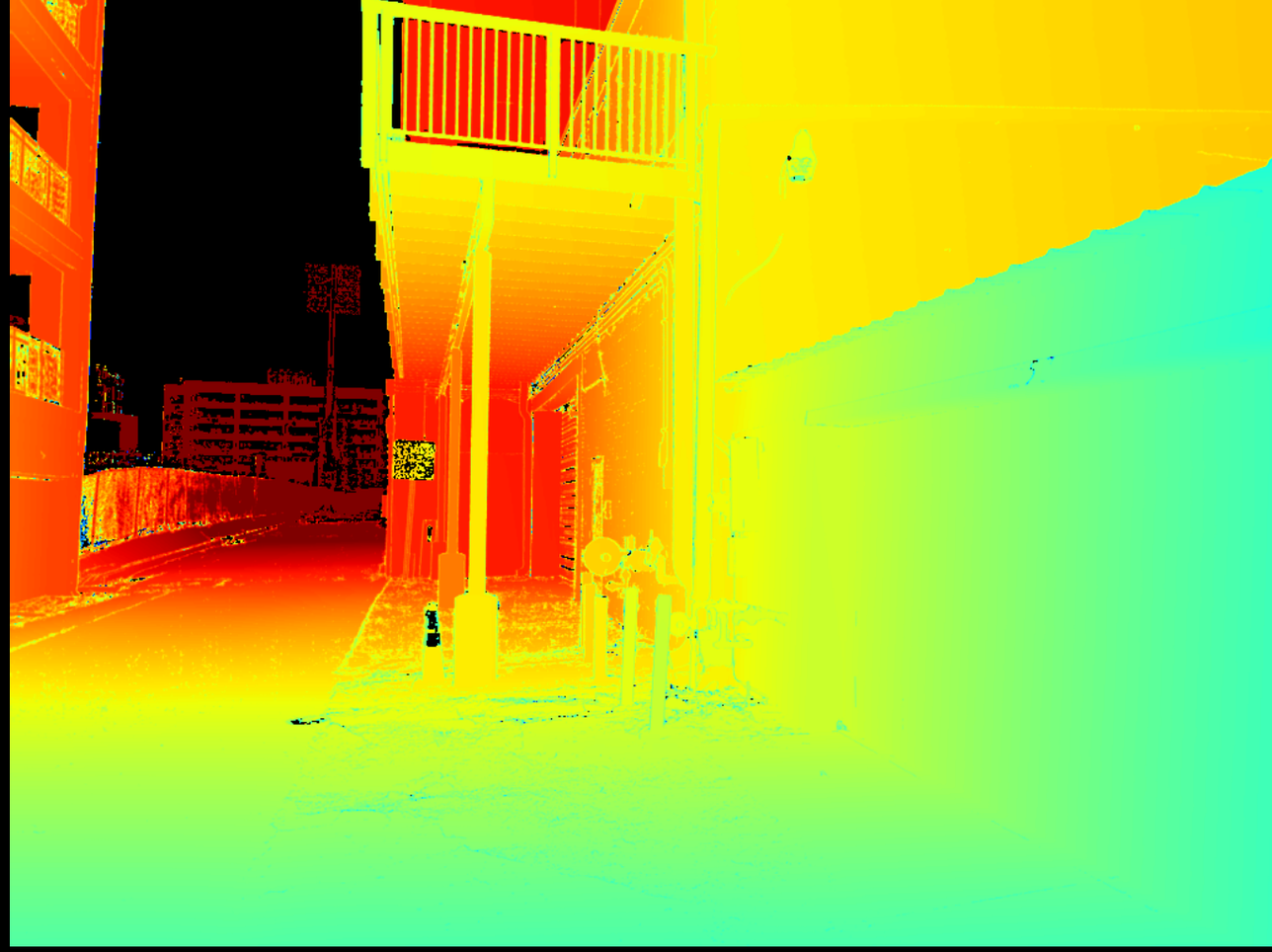


Normals

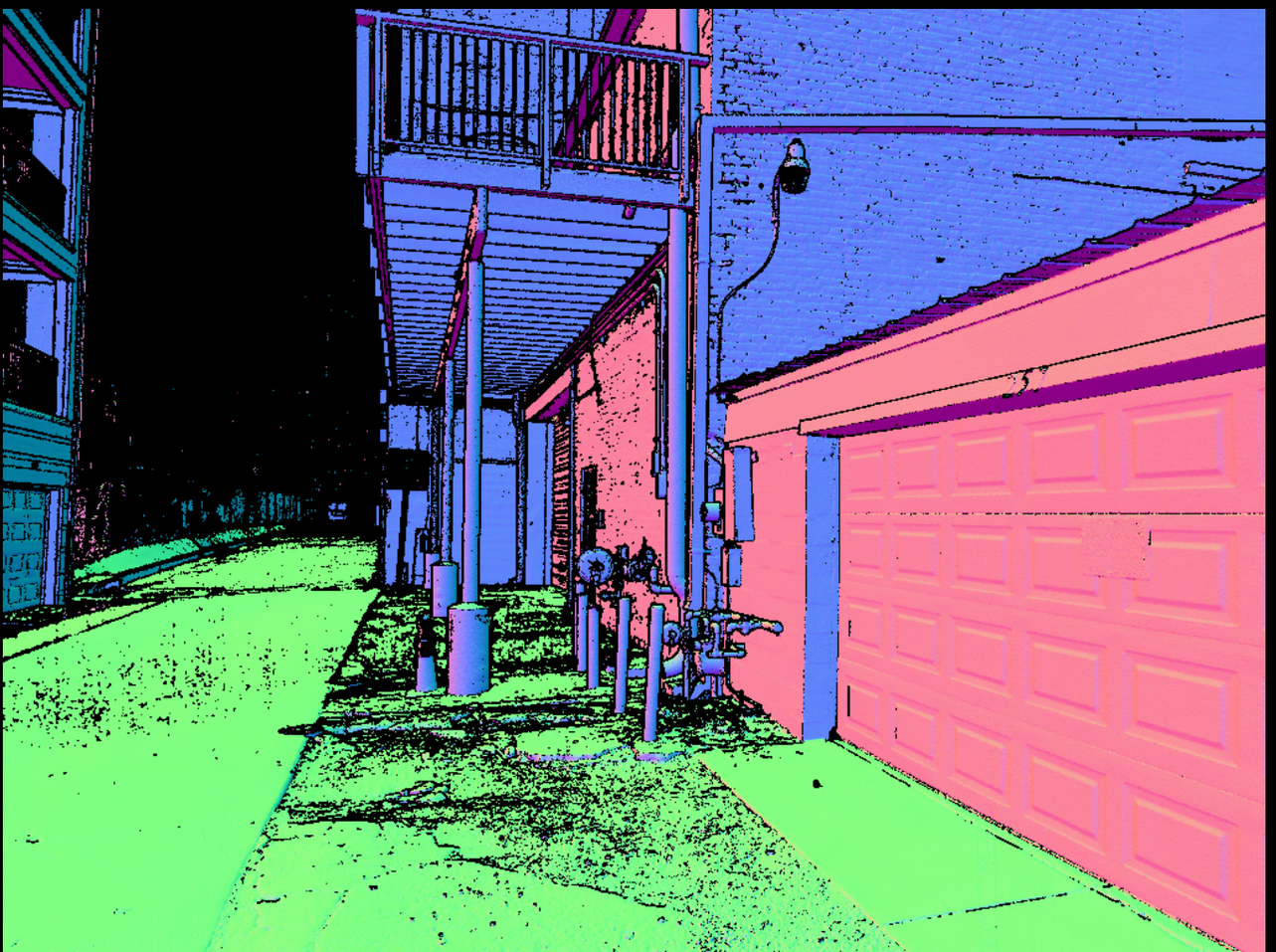


Depth

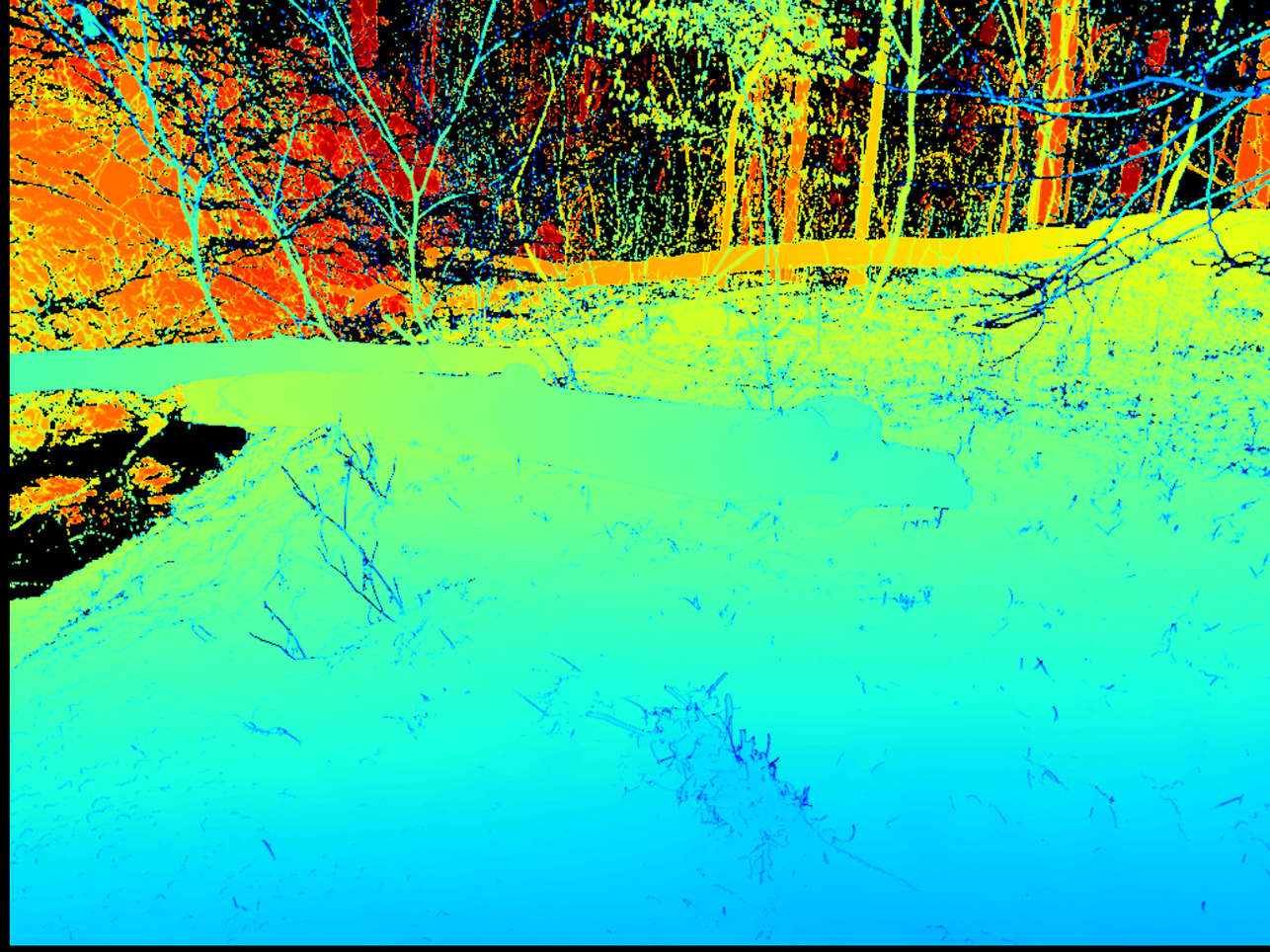
Normals



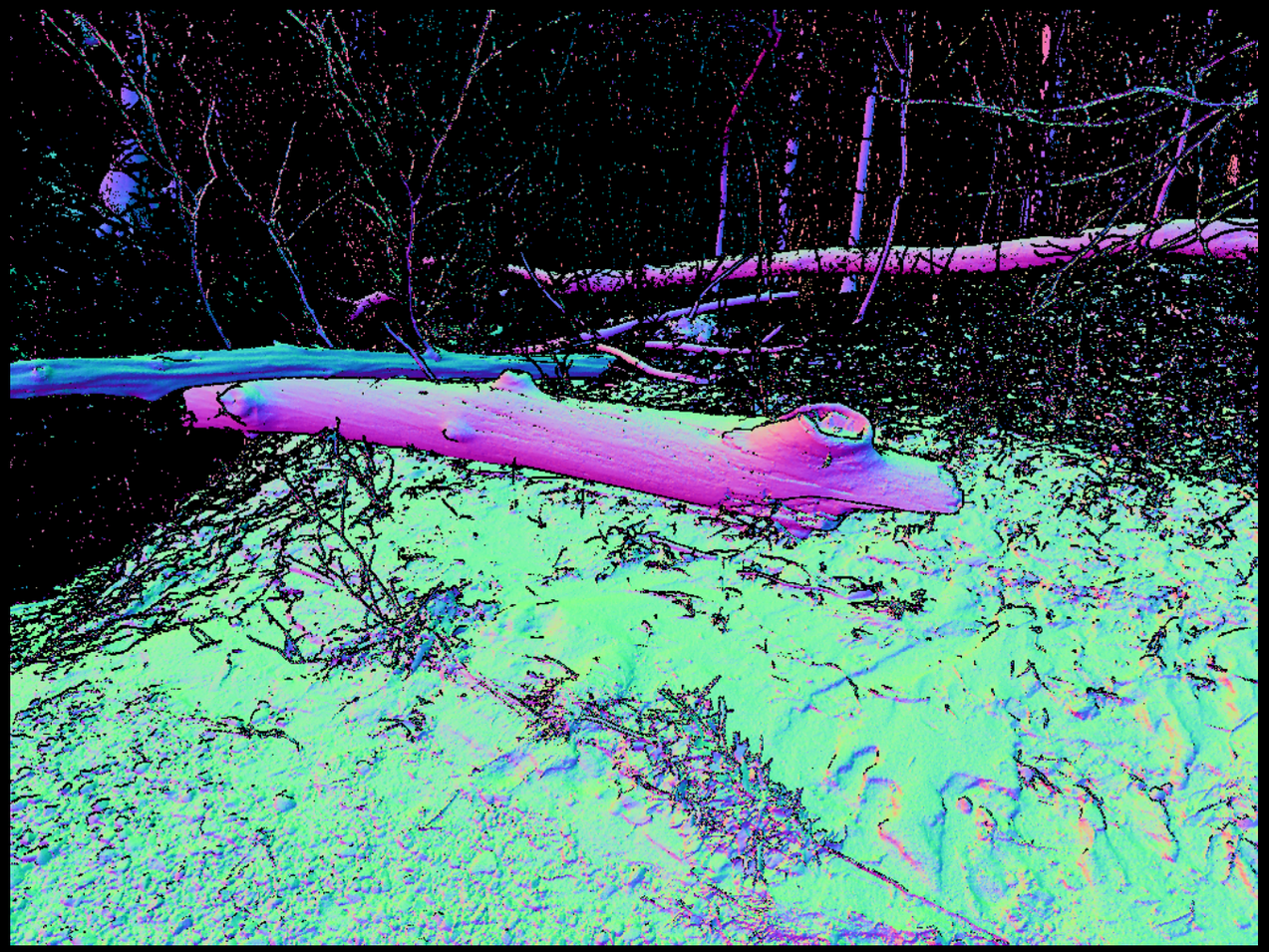
Depth



Normals



Depth

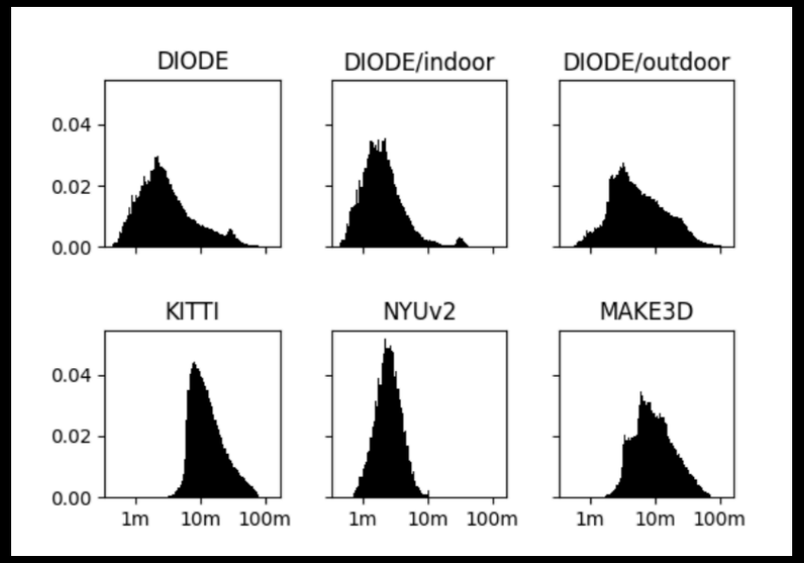


Normals

Dataset Comparison

Table 1. Statistics of DIODE compared to other popular RGBD datasets.

	DIODE	NYUv2	KITTI	Make3D
Return Density (Empirical)	82%	68%	16%	0.38%
# Images Indoor/Outdoor	4764/5554	1449/0	0/94000	0/534
Sensor Depth Precision	±1 mm	±1 cm	±2 cm	±3.5 cm
Sensor Angular Resolution	0.009°	0.09°	0.08°H, 0.4°V	0.25°
Sensor Max Range	350 m	5 m	120 m	80 m
Sensor Min Range	0.6 m	0.5 m	0.9 m	1 m



- **Sharper depth boundaries than available indoor data**
- **Higher density and vertical resolution than available outdoor data**

Conclusion

- High quality (and resolution) RGB-D and normals dataset
- Diversity of indoor and outdoor scenes
- Future directions/suggestions?
  - Higher resolution depth
  - Semantic annotations in 2D and 3D

Thank you

Collaborators

TOYOTA TECHNOLOGICAL INSTITUTE AT CHICAGO