

# Derivation of Large-scale Geometrically and Semantically Annotated UAV Datasets from 3D City Models

Sidi Wu, Lukas Liebel, Marco Körner  
Chair of Remote Sensing Technology  
Technical University of Munich

# Major tasks of UAV applications

Image Understanding

Multi-view 3D  
Reconstruction

Geo-localization

☹ Limited volume

☹ Limited ground-truth semantic / geometric annotations

← Limitations of current open-source UAV datasets →

# Pipeline

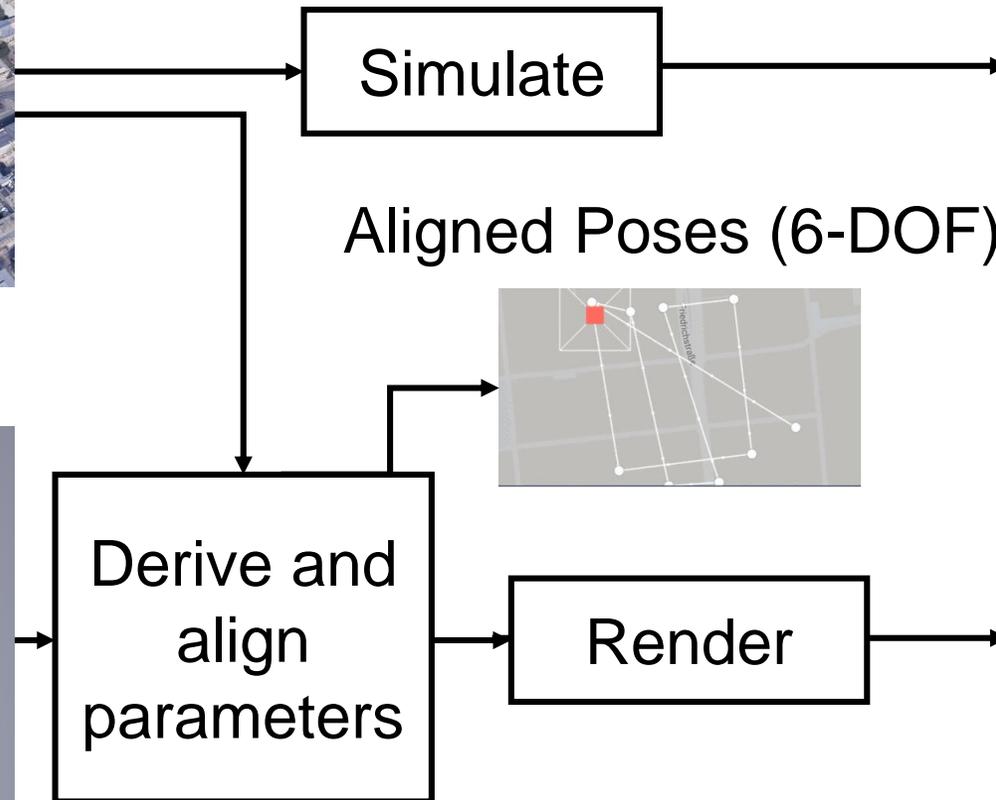
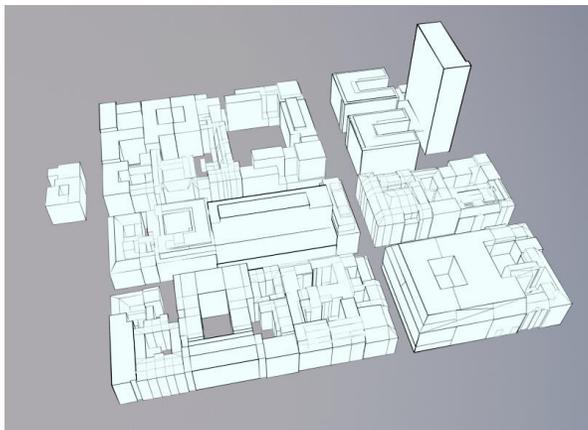
Google Earth Studio



UAV Image Sequences



3D City Models



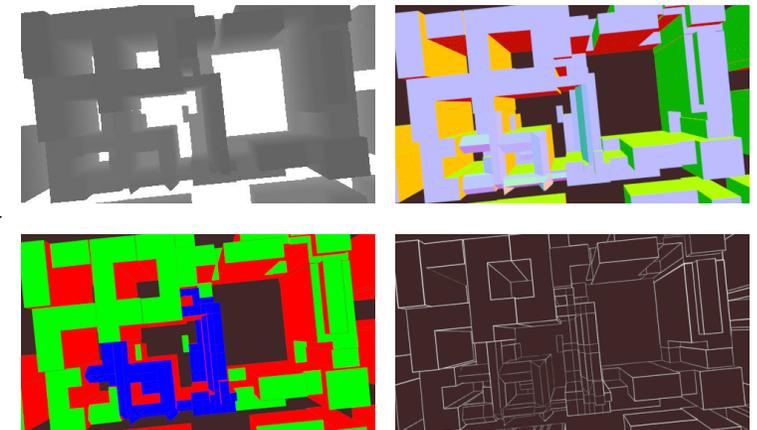
Simulate

Aligned Poses (6-DOF)



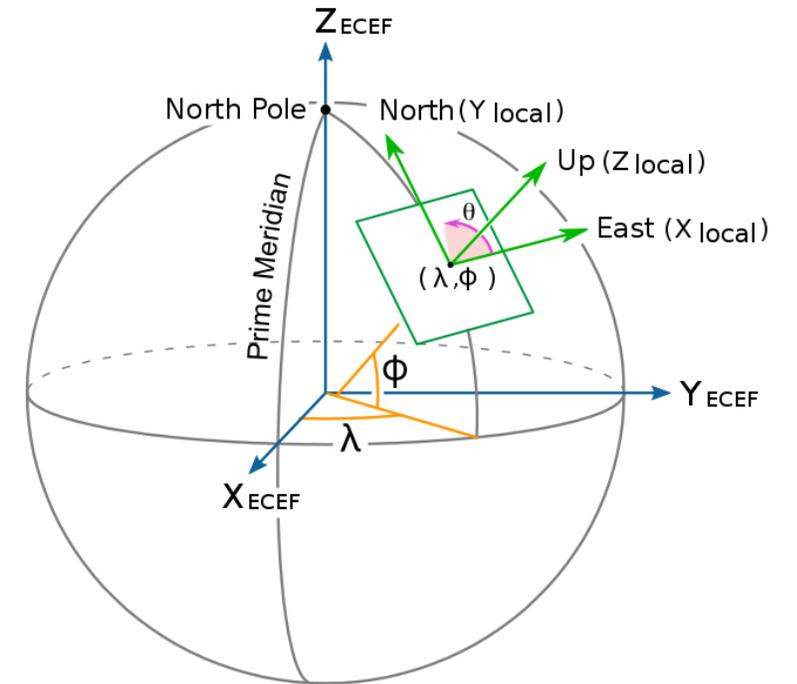
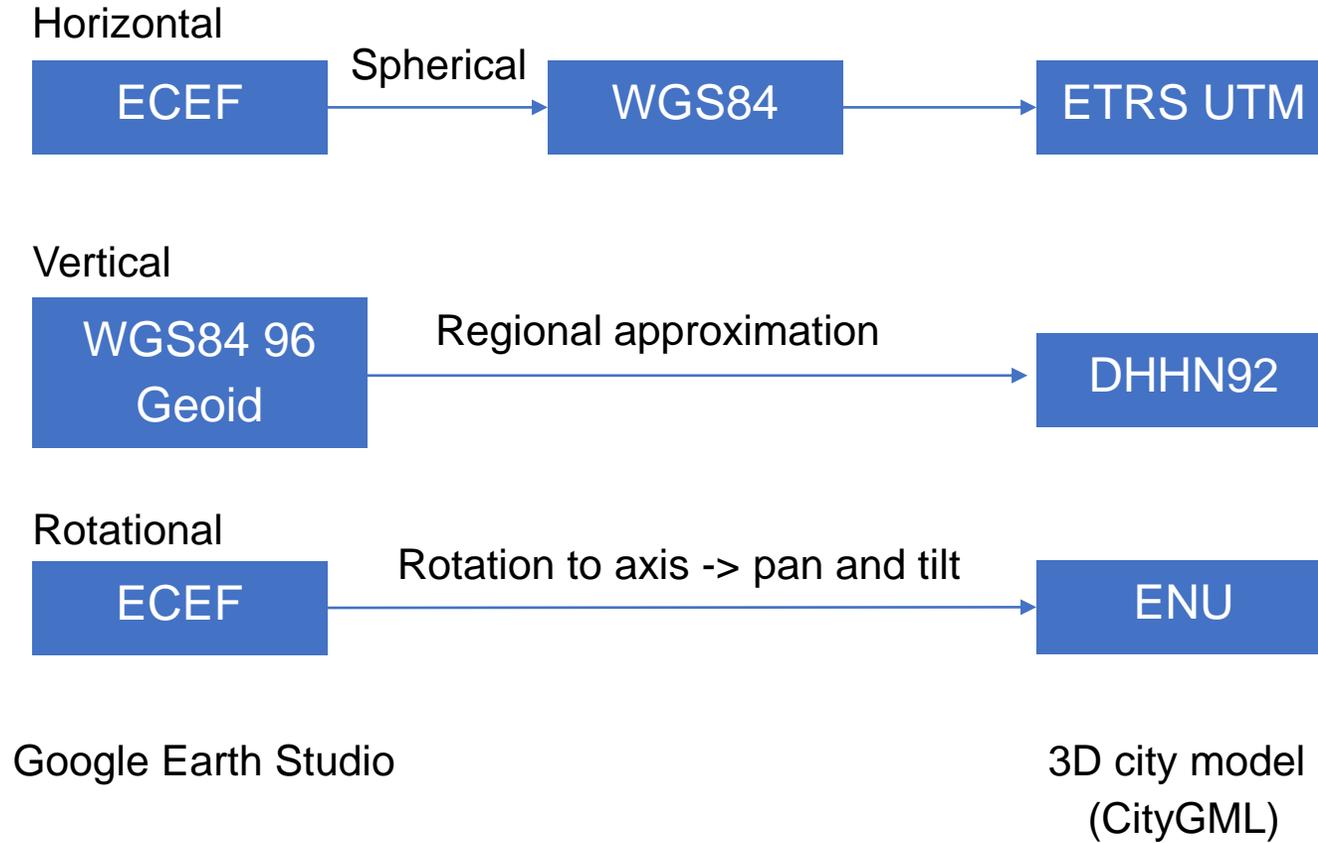
Render

Annotations



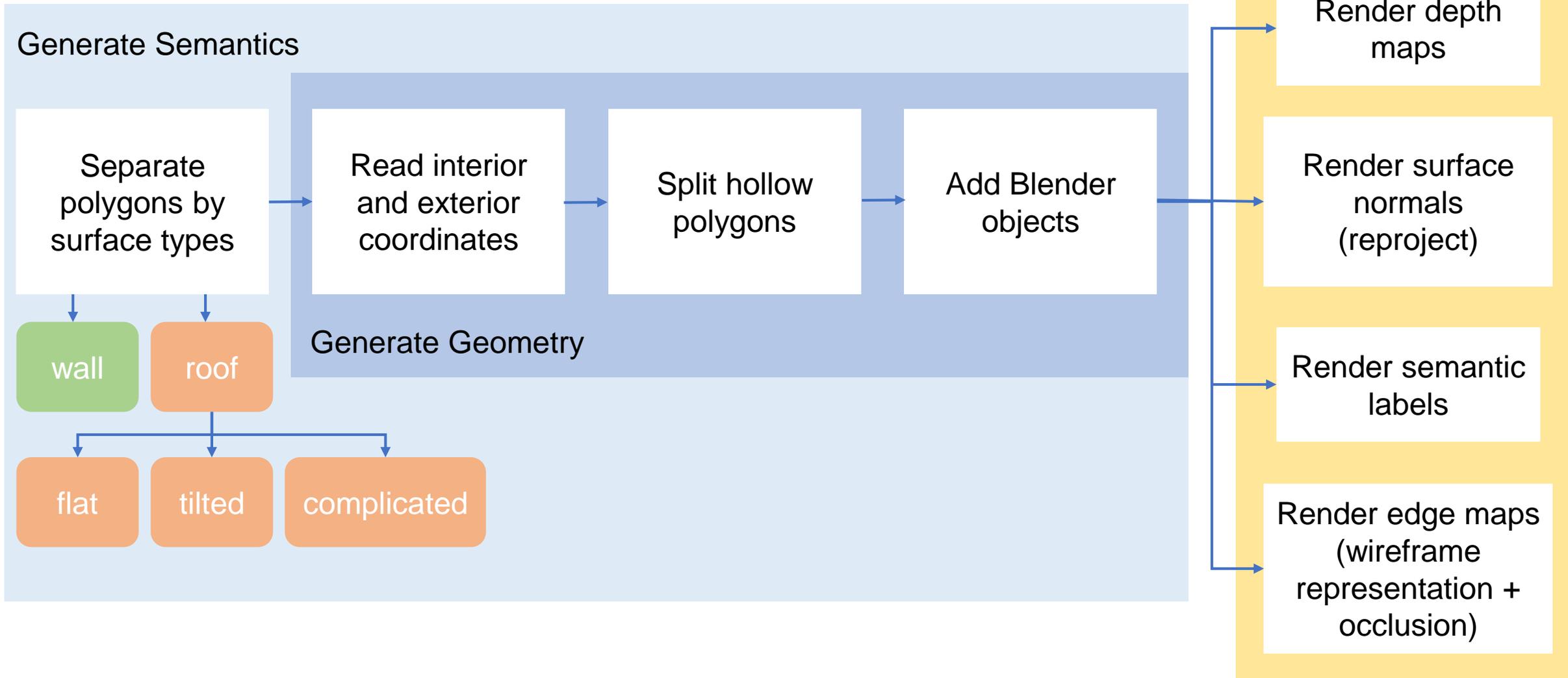
# Approaches

## Align extrinsic parameters



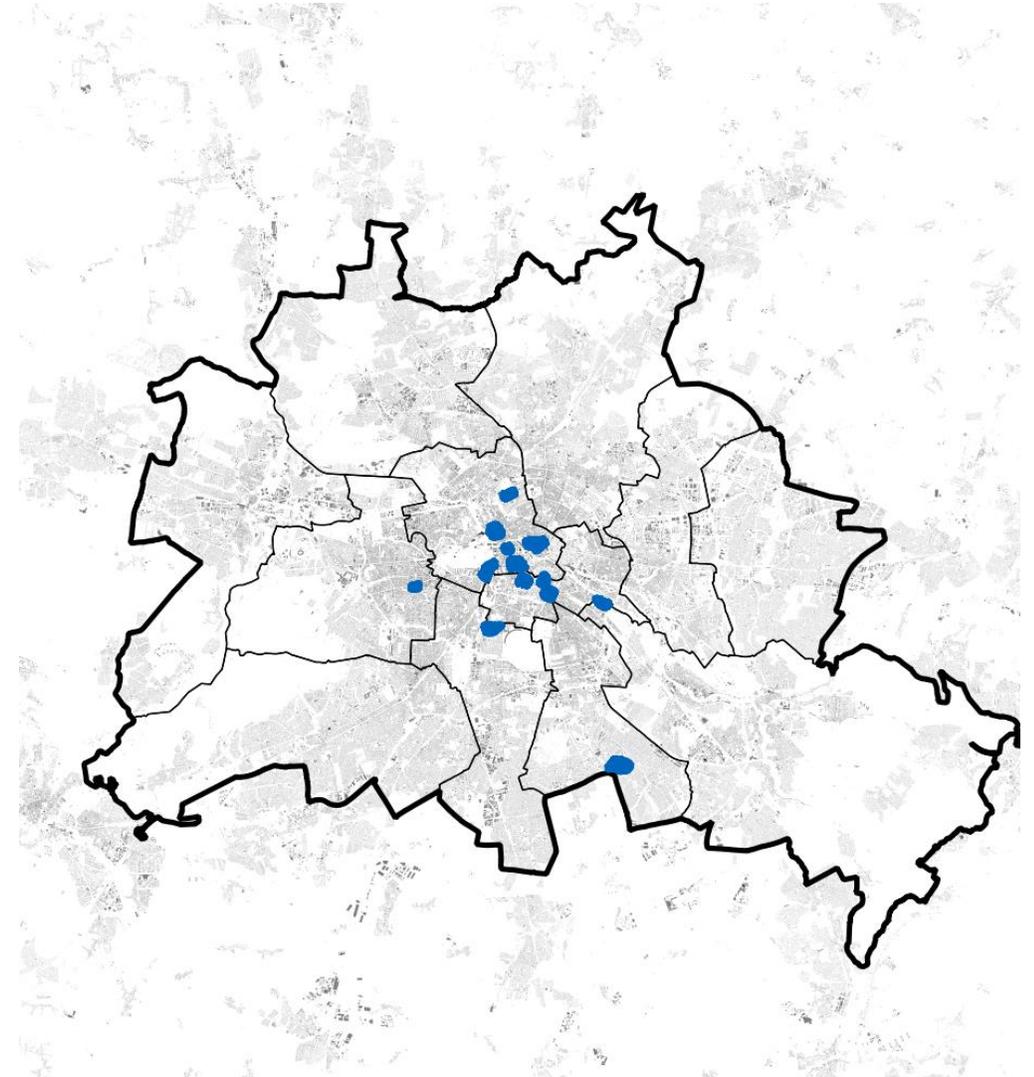
# Approaches

Parsing the 3D city model and generate annotations



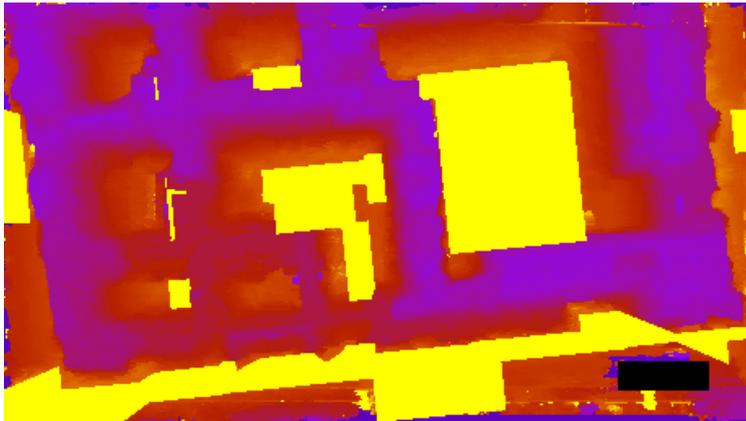
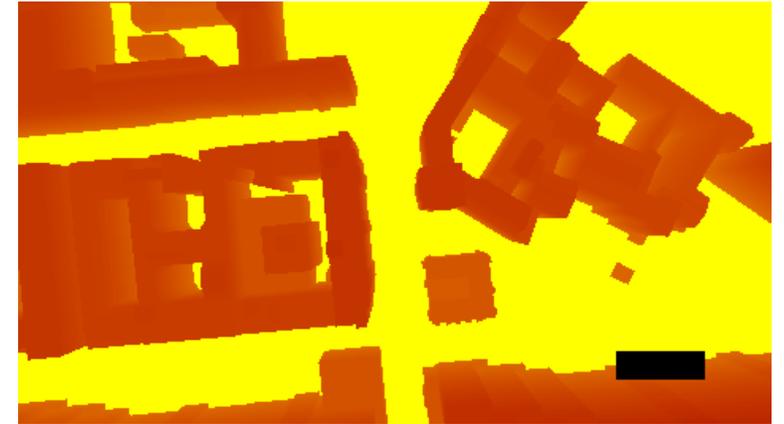
# Dataset

- 15 areas in Berlin, 144000 images
- Cruising height from 116m to 200m above sea level
- Field of View (FoV) - 70°, cruising speed - 13.4m/s, frame rate - 30fps
- <https://github.com/sian1995/largescaleuavdataset>



# Evaluation

- Data completeness: 81.8% usable buildings
- Geometric accuracy



Left: SfM-created depth

Right: generated ground-truth depth

# Evaluation

- Usability:

Dataset	Error metric			
	Abs Rel	Sq Rel	RMSE	RMSE (log)
KITTI [9]	0.208	1.768	6.856	0.283
Cityscapes [7]	0.267	2.686	7.580	0.334
Make3D [16]	0.383	5.321	10.470	0.478
Ours	0.092	1.646	10.930	0.116

[Zhou et al (2017), Unsupervised Learning of Depth and Ego-Motion from Video]

- Semantic accuracy: 95.73%