



SceneNN: A Scene Meshes Dataset with aNNotations

Binh-Son Hua¹

Quang-Hieu Pham¹
Lap-Fai Yu³

Duc Thanh Nguyen²
Sai-Kit Yeung¹

Minh-Khoi Tran¹

International Conference on 3D Vision 2016

¹ Singapore University of Technology and Design

² Deakin University

³ University of Massachusetts Boston

Motivation

Dataset	Quantity	Annotation	Format	Pose
NYU v2	1449 frames	All	Image	N
SUN RGB-D	10K frames	All	Image	N
RGB-D v2	17 scenes	All	Cloud	Y
TUM	47 scenes	N.A.	Image	Y
SUN3D	254 scenes	8 scenes	Cloud	Y
Ours	100 scenes	All	Mesh	Y

SceneNN dataset

<http://www.scenenn.net>

- **100+** RGBD scenes
- **Raw videos** from 2,000 to more than 10,000 frames
- Reconstructed **triangle meshes** in PLY format
- Per-frame camera poses
- **Per-vertex** and **per-pixel** labelling
- **Annotated bounding boxes**, object poses
- Categories: Workplace (27) Bedroom (19) Living room (8) Kitchen (11) Study space (9) Meeting space (3) Lounge (8) Other (10), Asus Xtion and Kinect v2 comparison (6)



Input

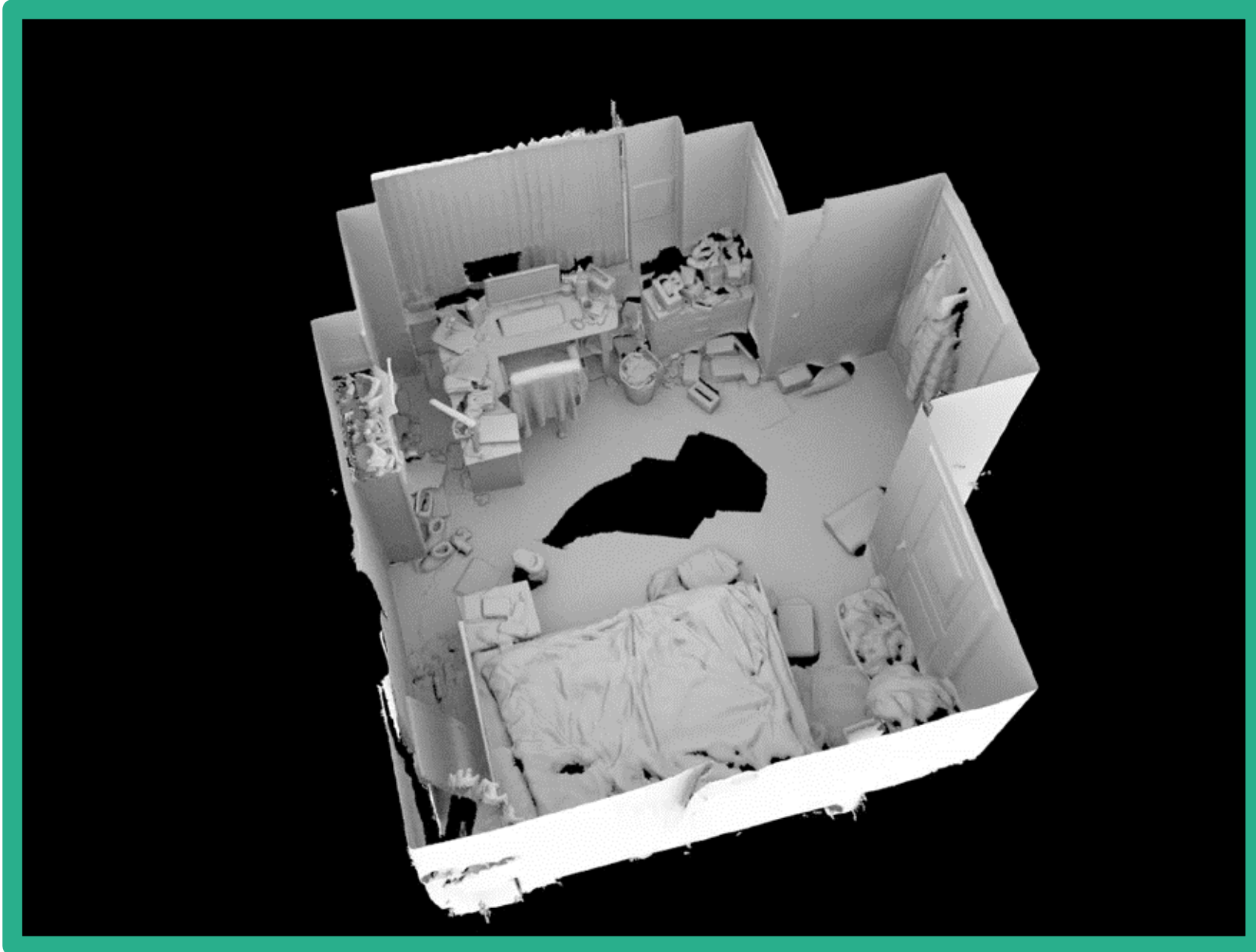


RGB

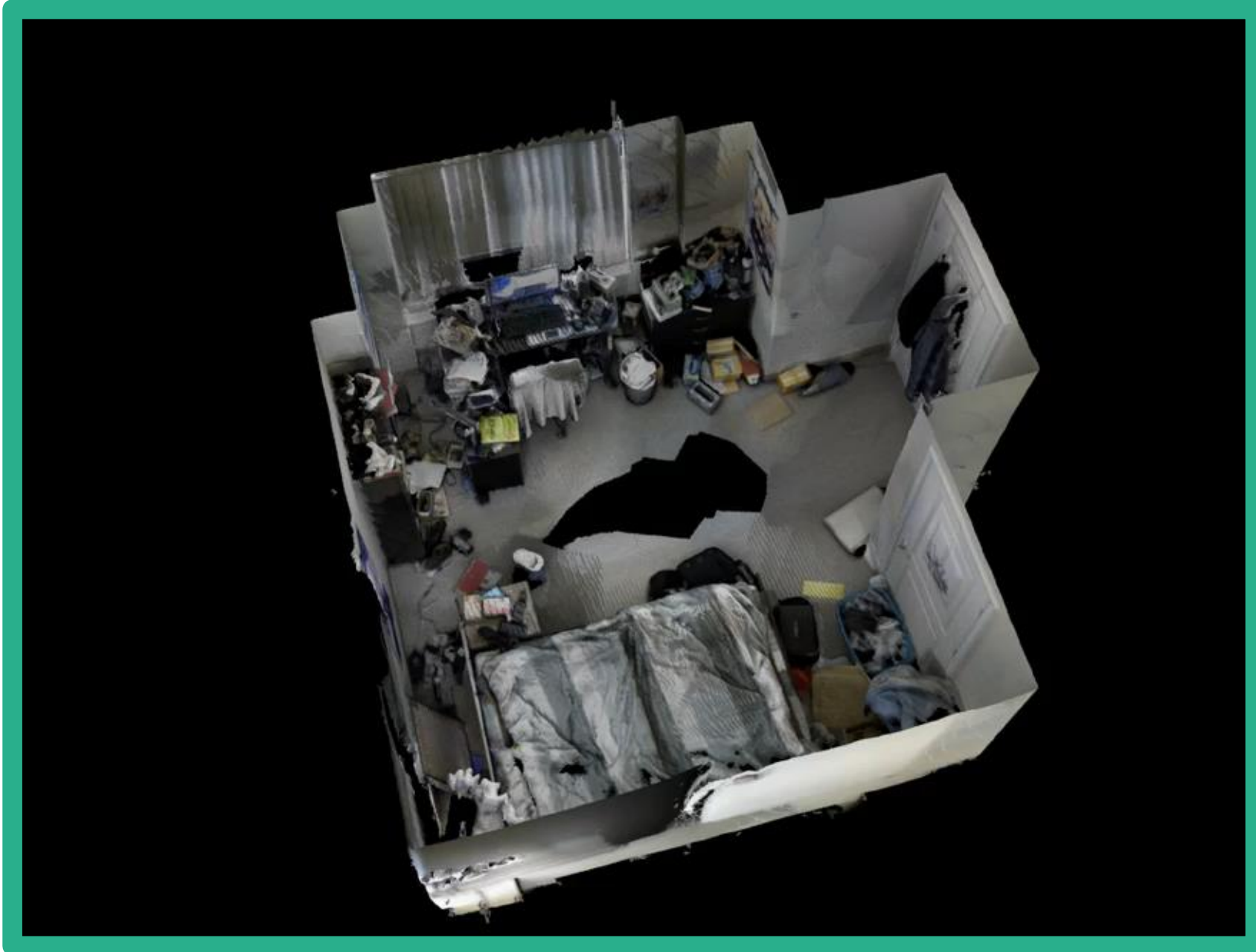


Depth

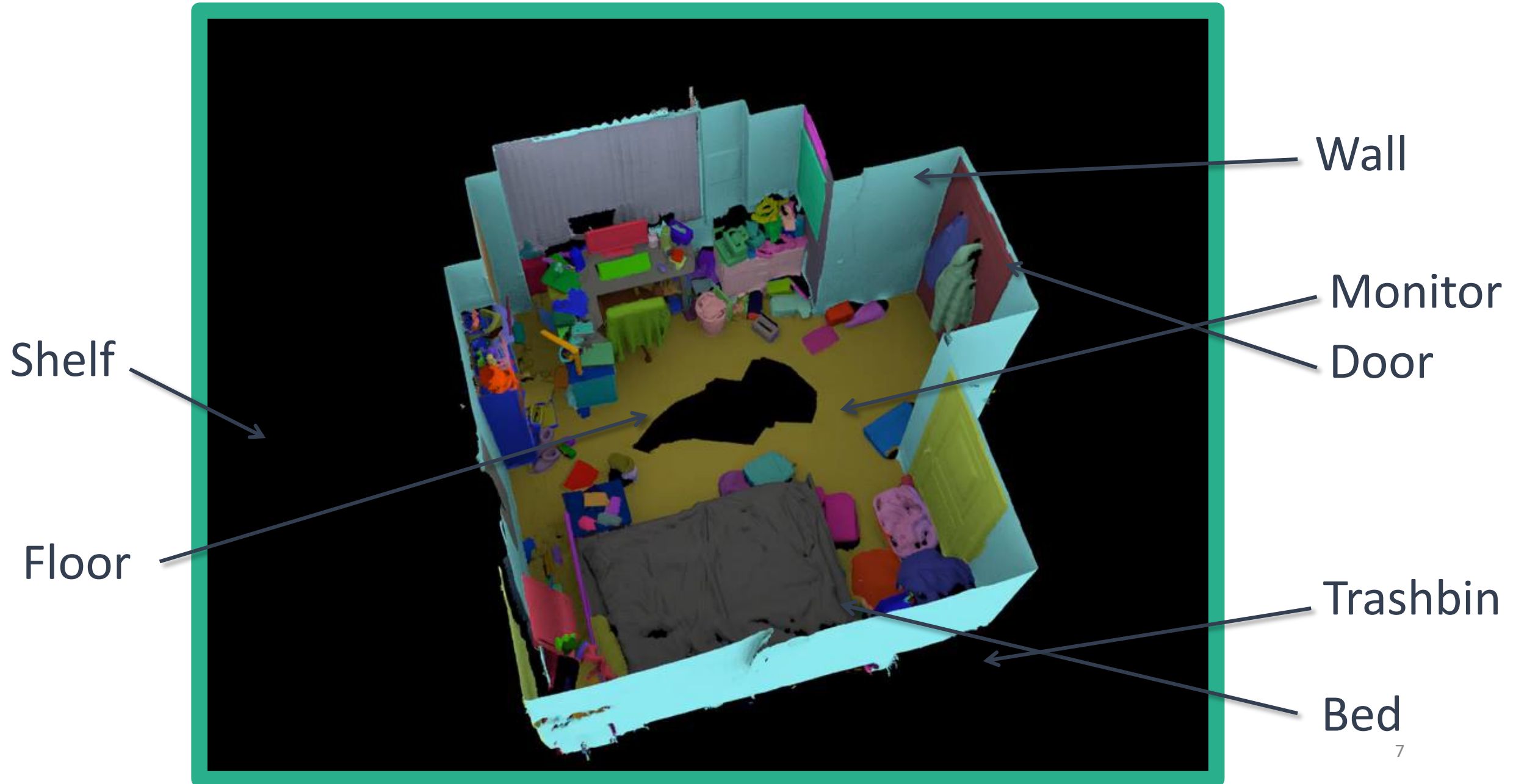
3D reconstruction



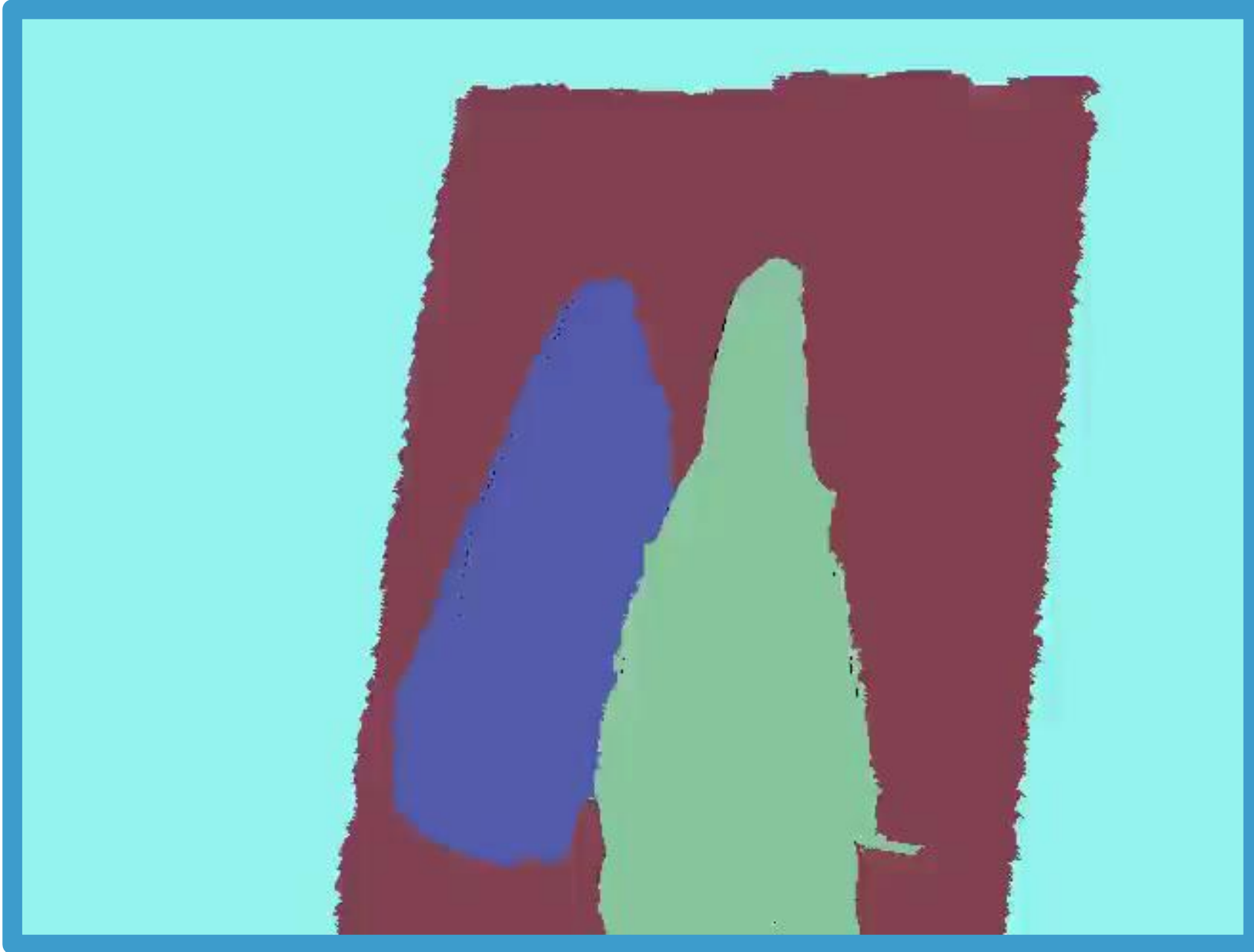
3D reconstruction



Output: 3D segmentation and annotation



Output: 2D segmentation and annotation



Workflow summary

Capture



Reconstruct



Annotate in 3D



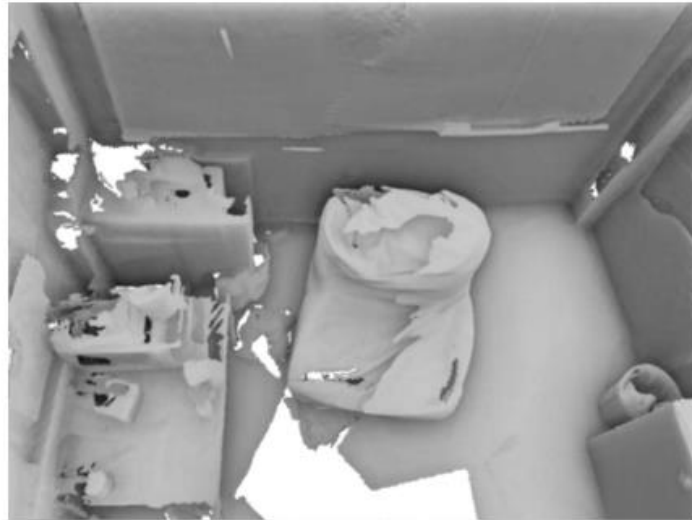
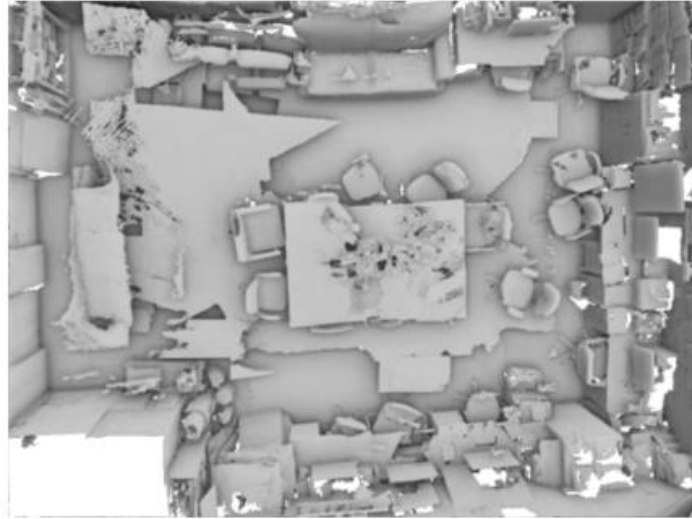
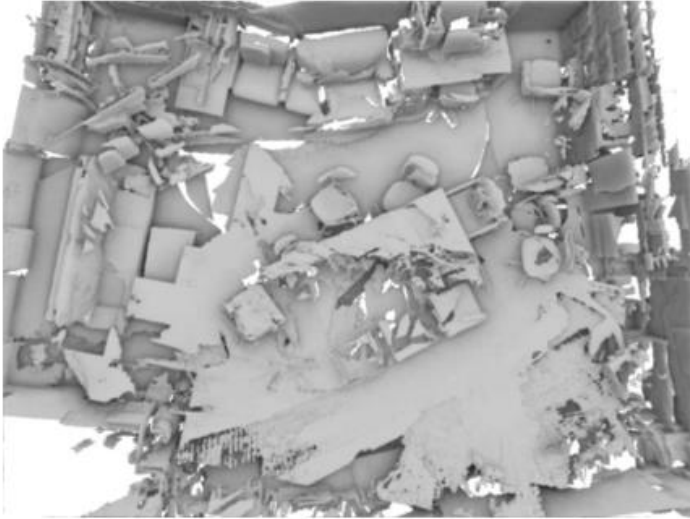
- Triangle mesh
- Camera poses



- monitor
- keyboard
- poster
- cabine
- bookshelves
- trash bin

- Per-vertex and per-pixel labels
- Bounding boxes, object poses

RGBD Reconstruction



DVO SLAM

[Kerl et al., IROS 2013]

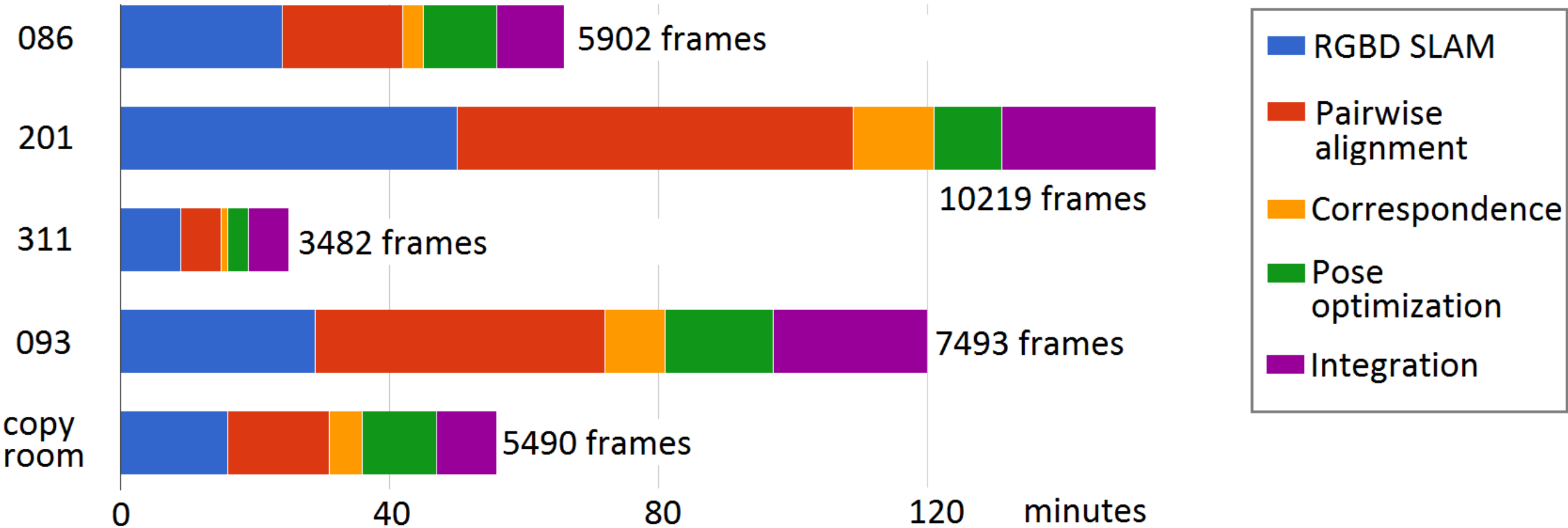
Elastic Fusion

[Whelan et al., RSS 2015]

Elastic Reconstruction

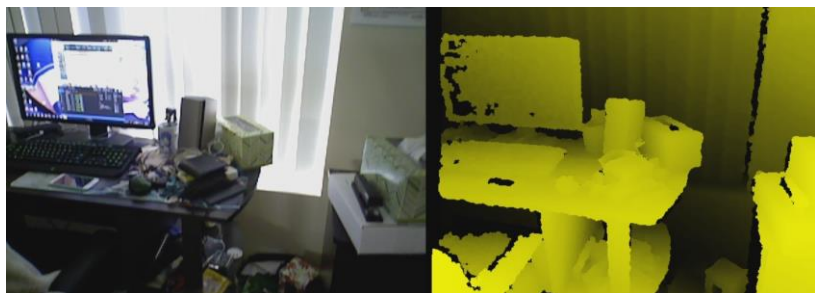
[Choi et al., CVPR 2015]

Reconstruction statistics



Annotation

Capture



Reconstruct



Annotate in 3D

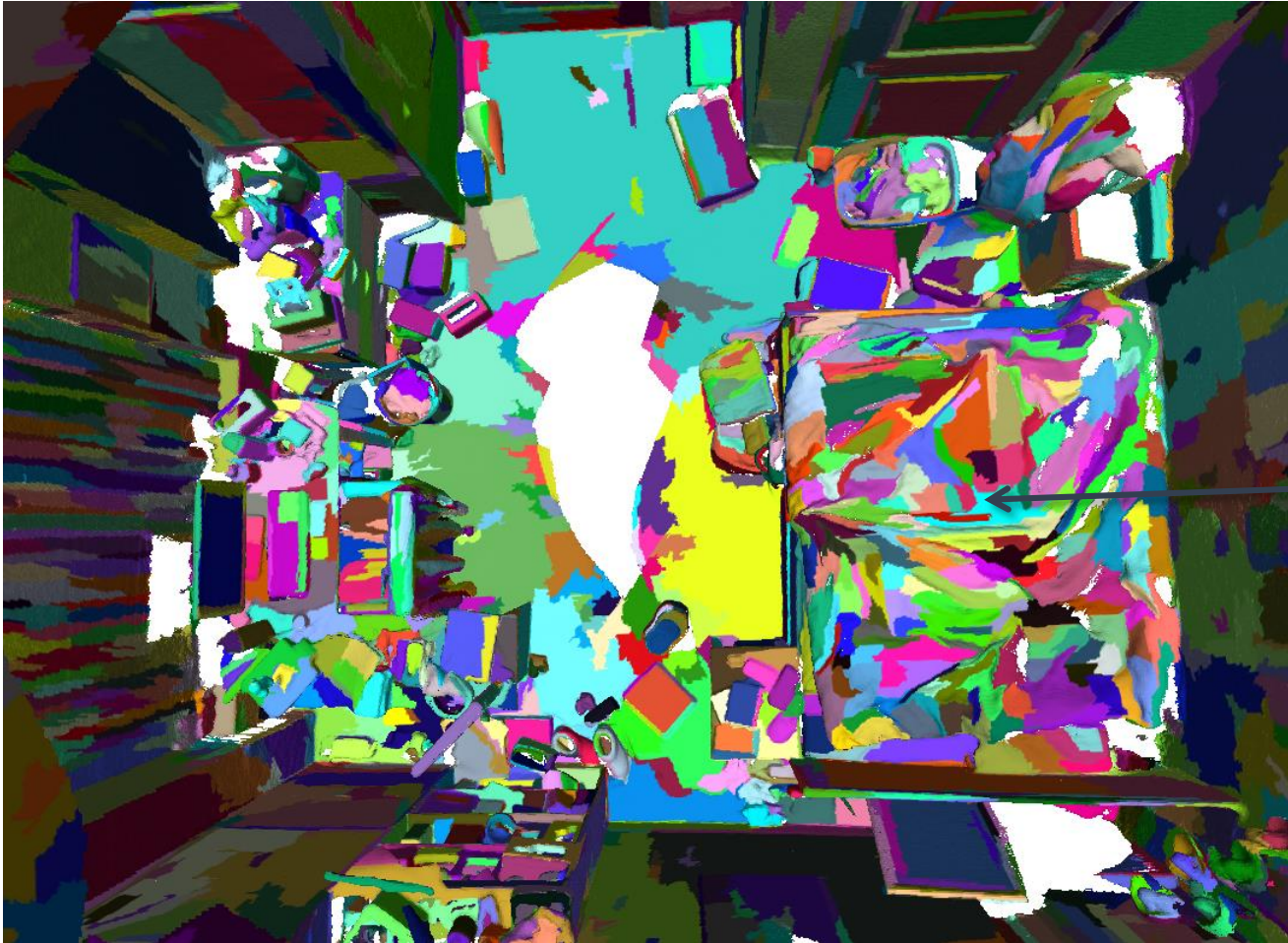


- Triangle mesh
- Camera poses



- Per-vertex and per-pixel labels
- Bounding boxes, object poses

Graph-based segmentation



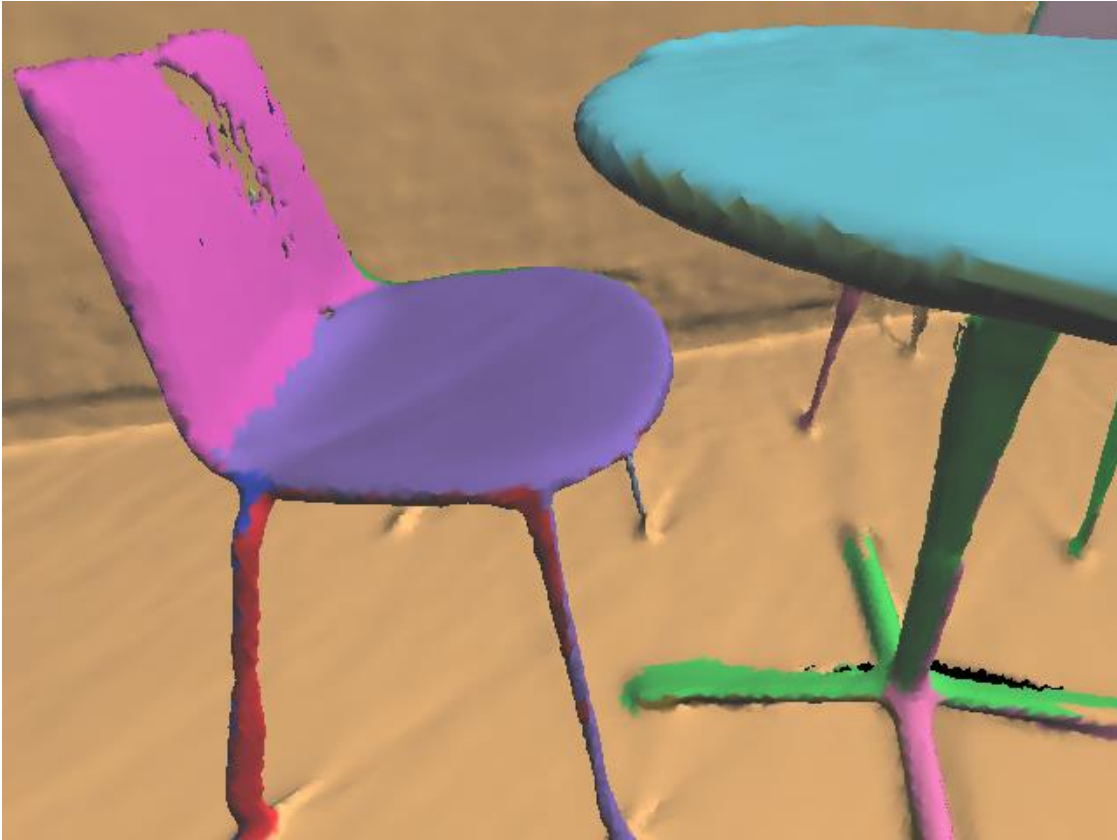
Supervertices

Markov random field

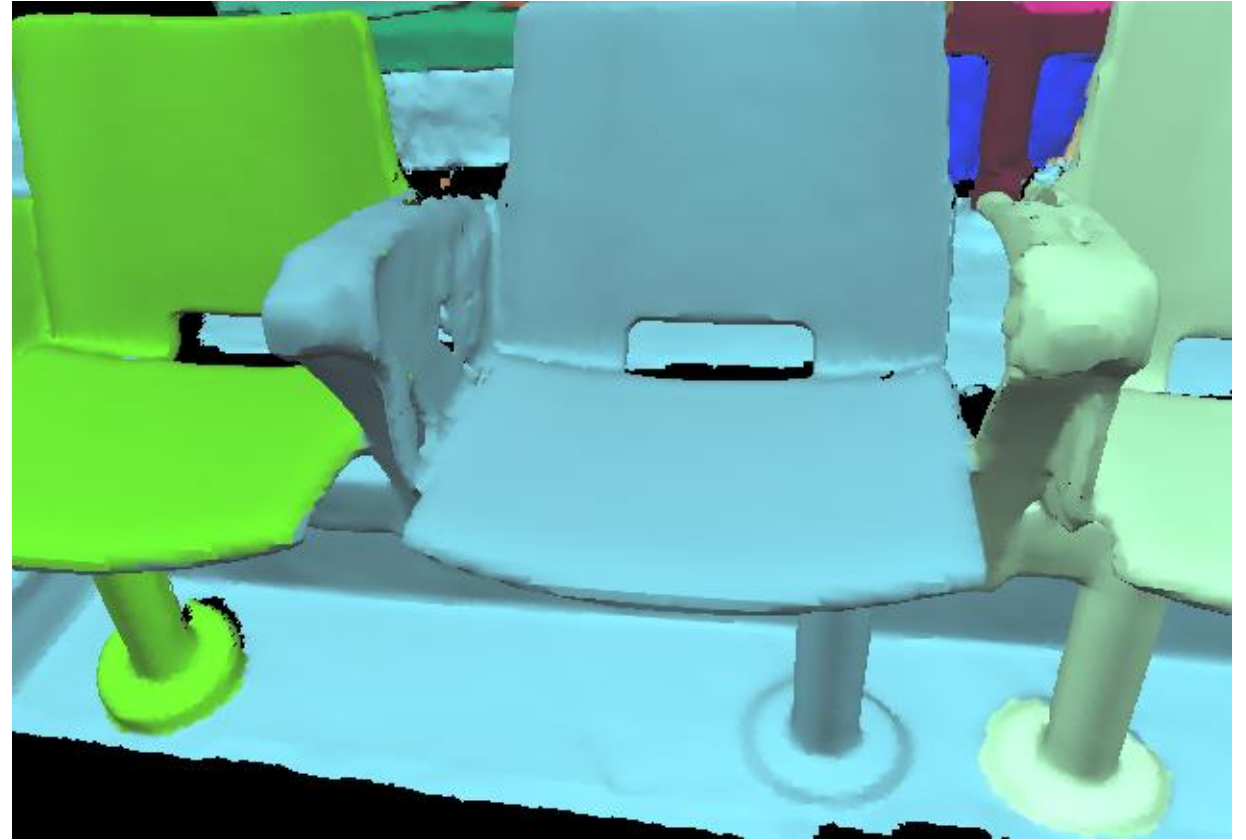


Regions

Imperfect segmentation



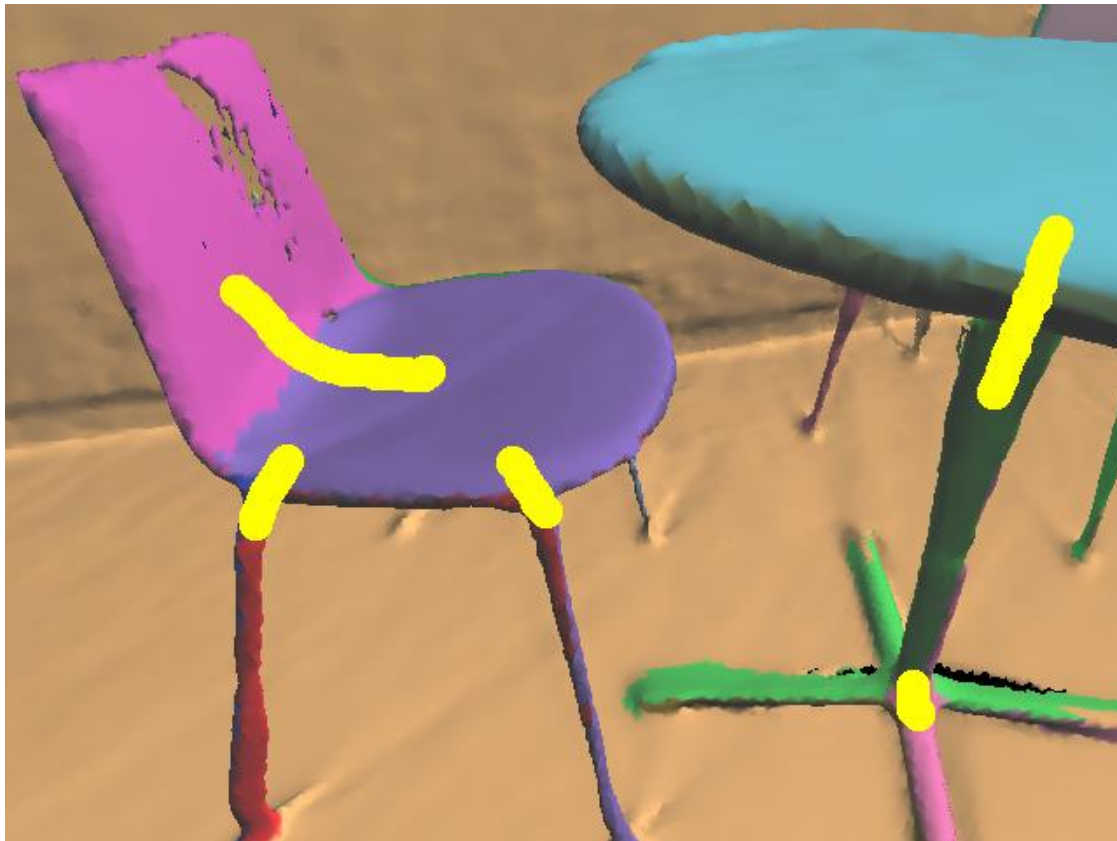
Over-segmentation



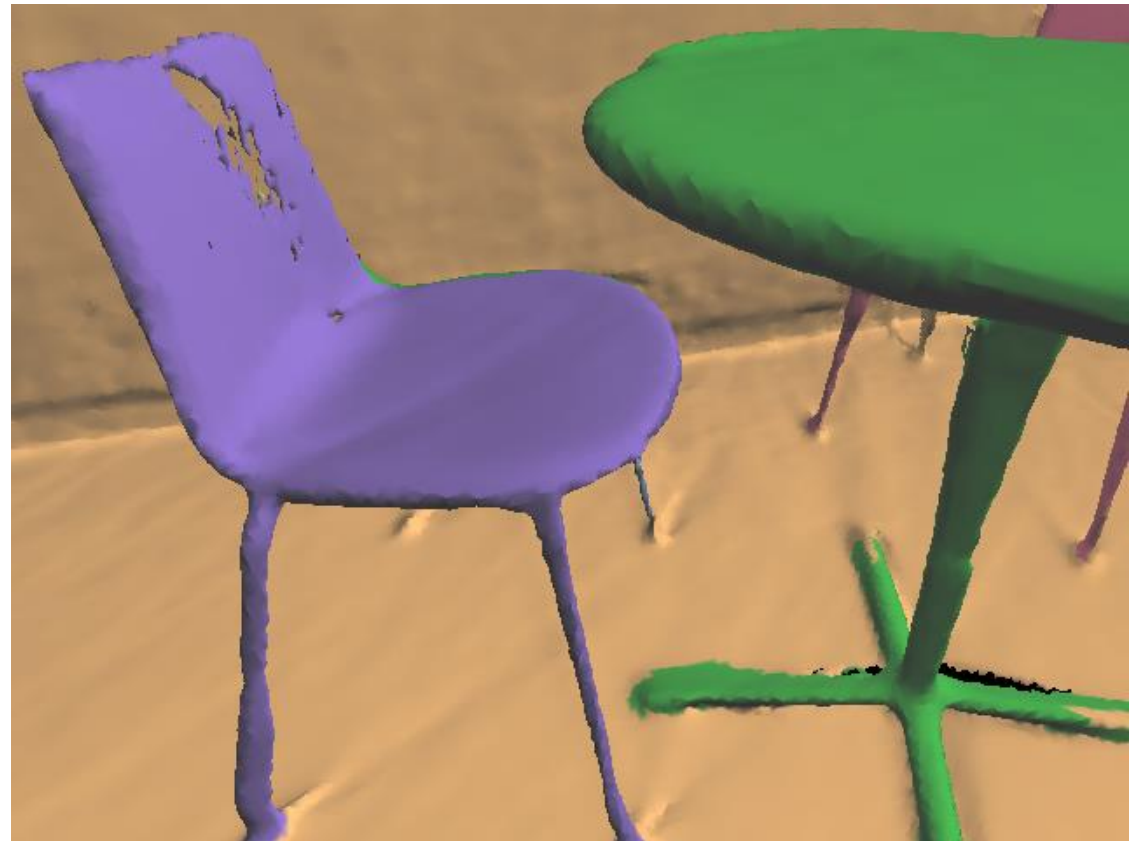
Under-segmentation

User interaction

Merge

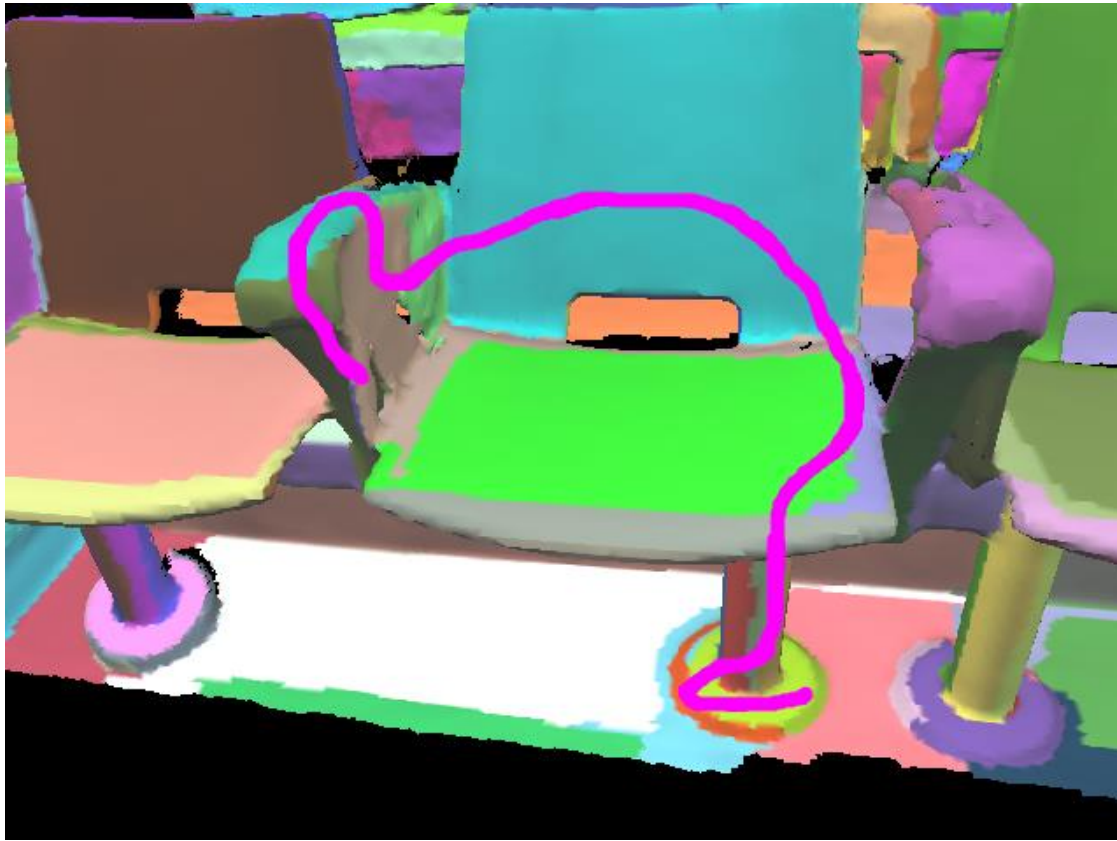


Before

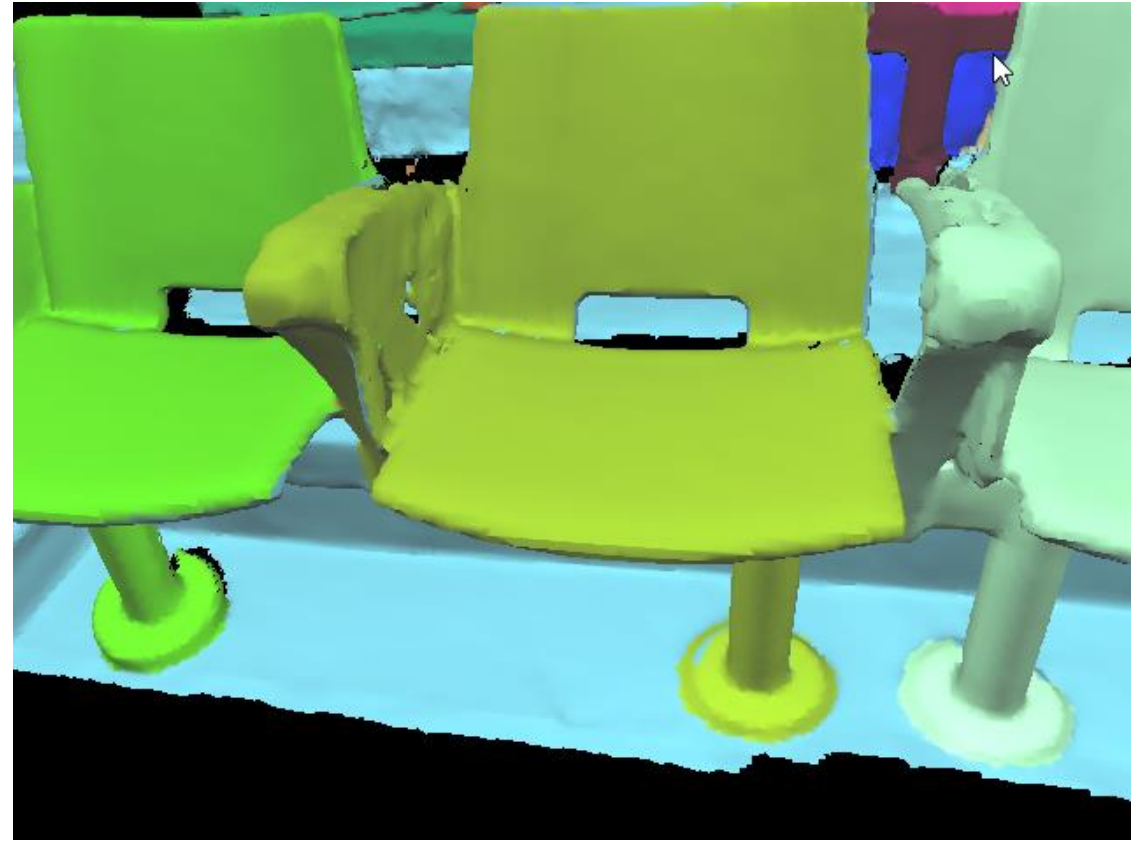


After

Extract

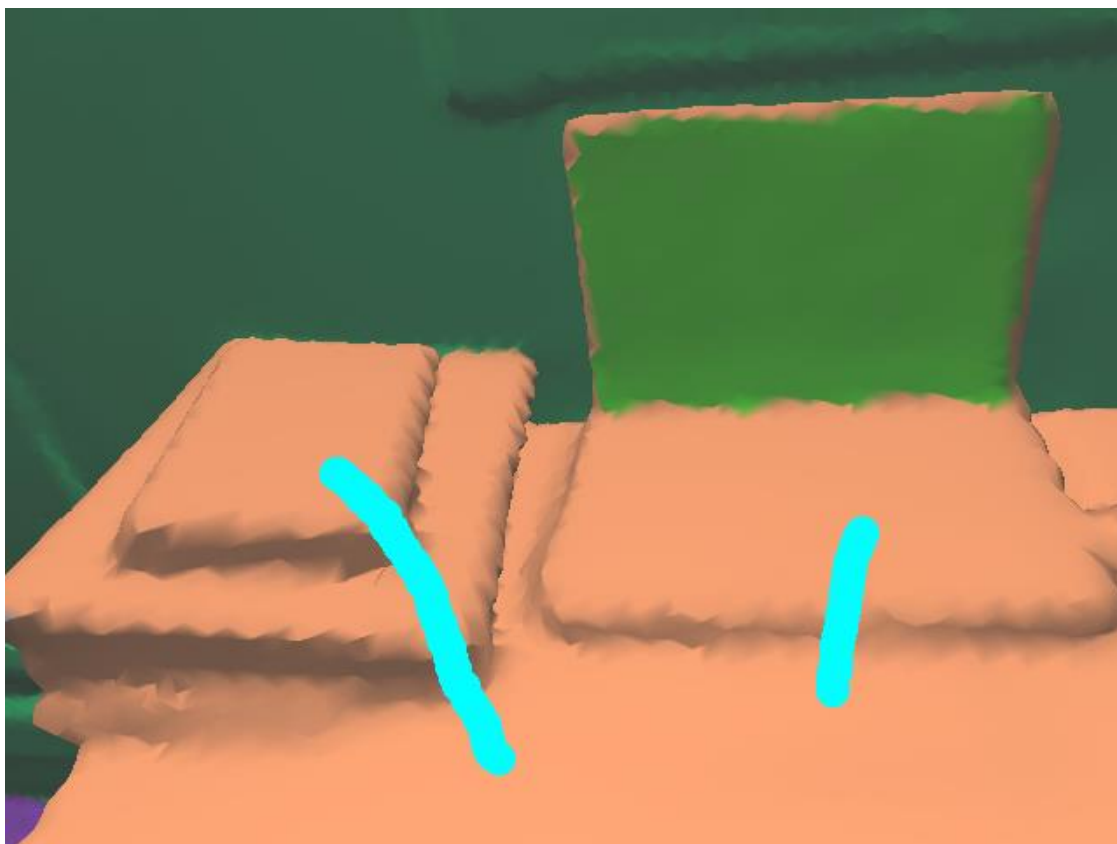


Before

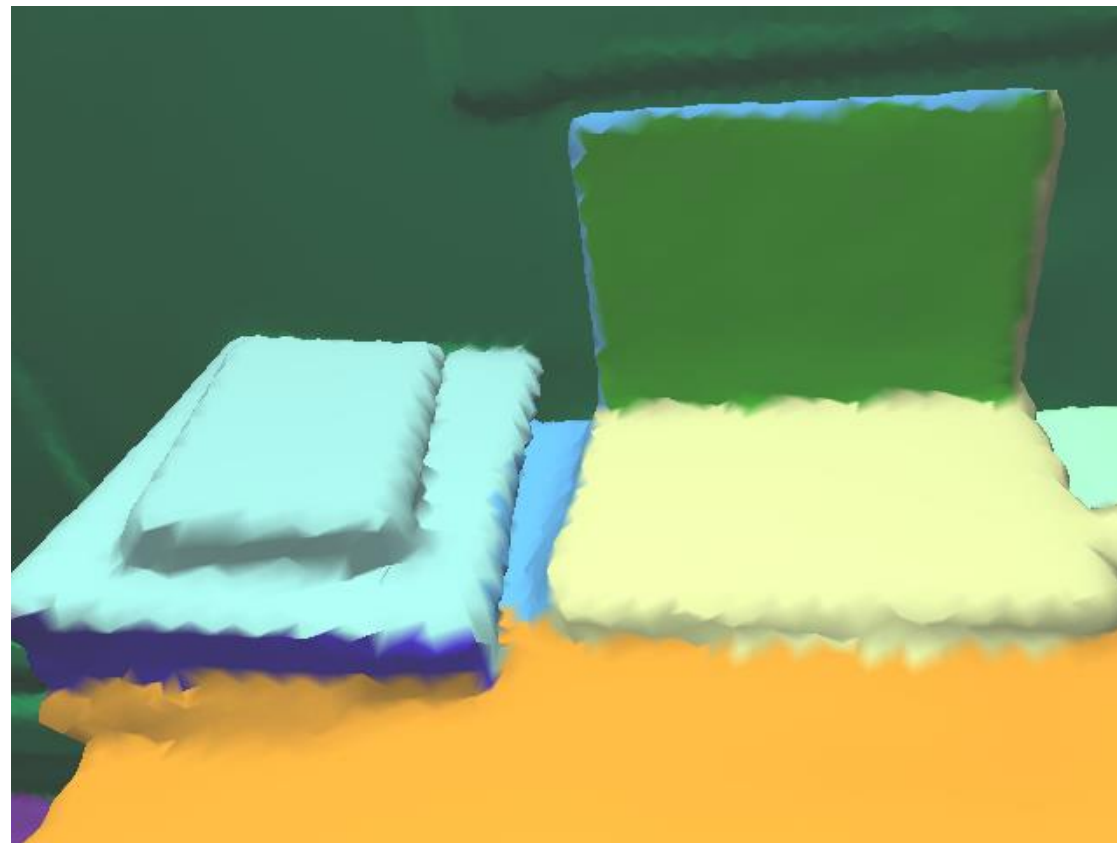


After

Split



Before



After

Example

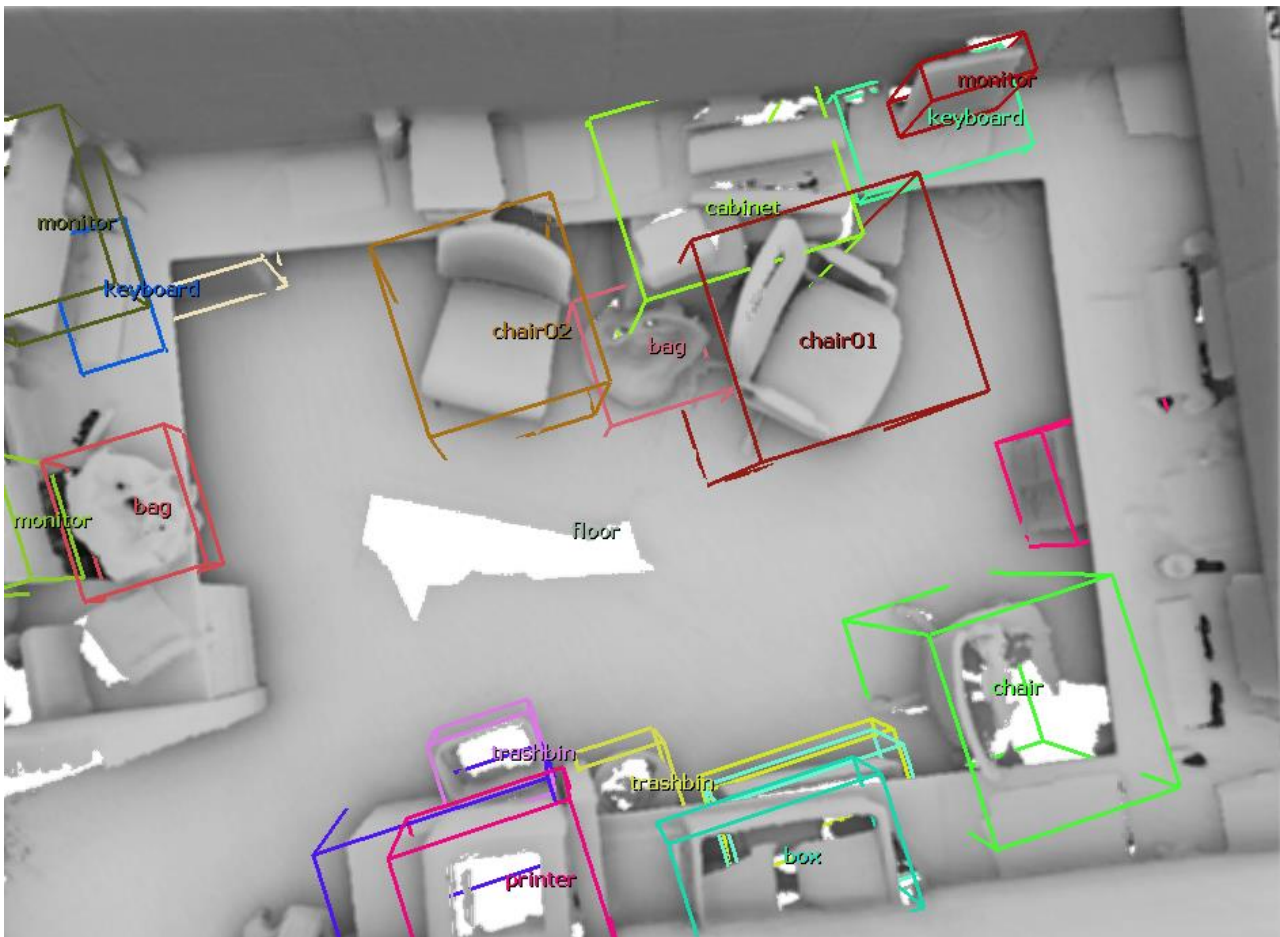


Final results



This image shows a 3D point cloud of a room interior, where different objects and surfaces are segmented into distinct, color-coded regions. The walls are a solid blue, the floor is a light green, and various furniture items like chairs, tables, and a sofa are rendered in a variety of colors including orange, red, pink, purple, and yellow. The segmentation is highly detailed, capturing the geometry of the objects.

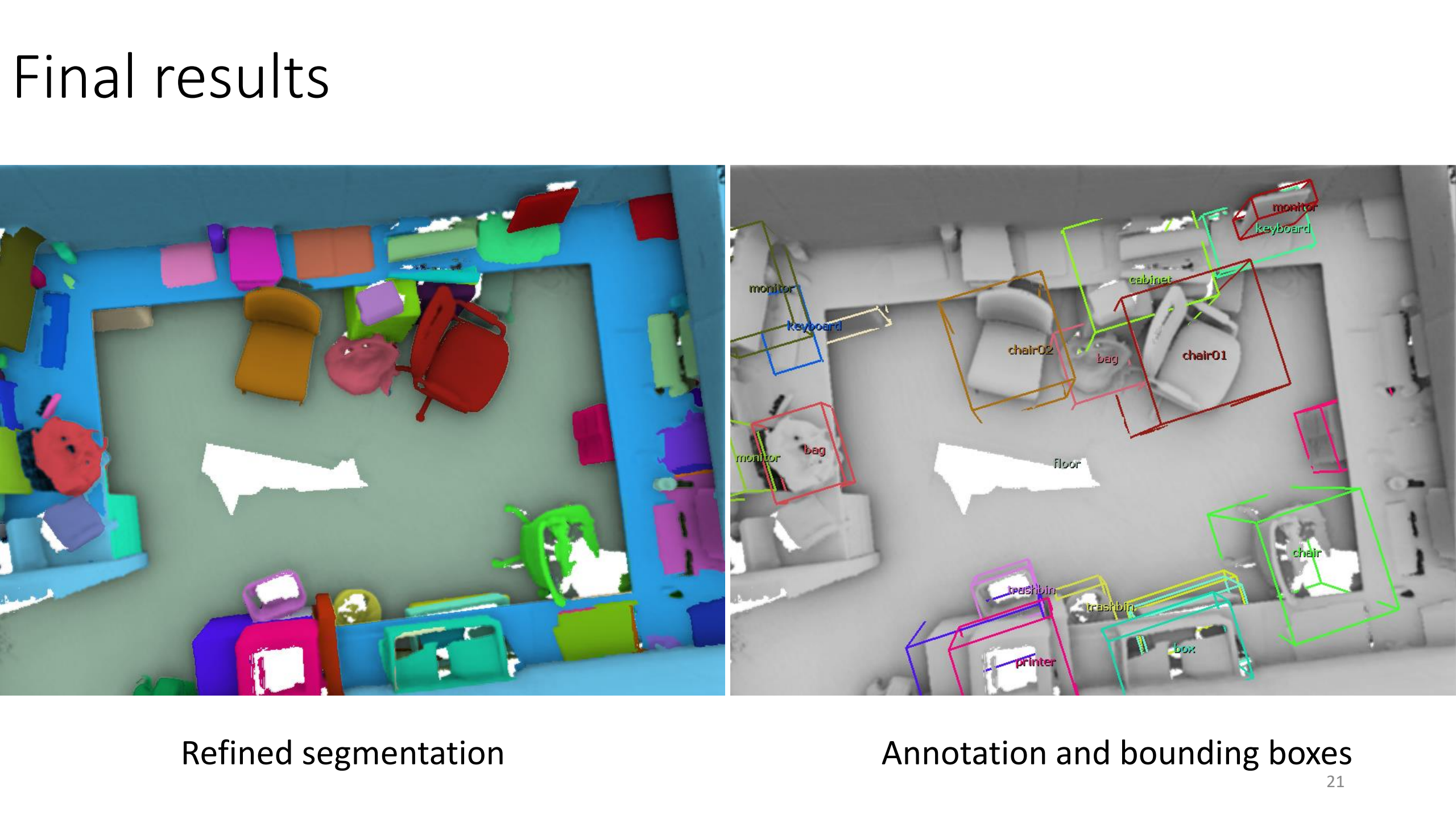
Refined segmentation



This image displays the same 3D point cloud of the room, but with colored bounding boxes overlaid on specific objects. Each box is accompanied by a text label identifying the object. The labels include 'monitor', 'keyboard', 'cabinet', 'chair01', 'chair02', 'bag', 'floor', 'trashbin', 'printer', and 'box'. The bounding boxes are colored in a way that matches the object they enclose, providing a clear visual link between the label and the object in the scene.

Annotation and bounding boxes

21

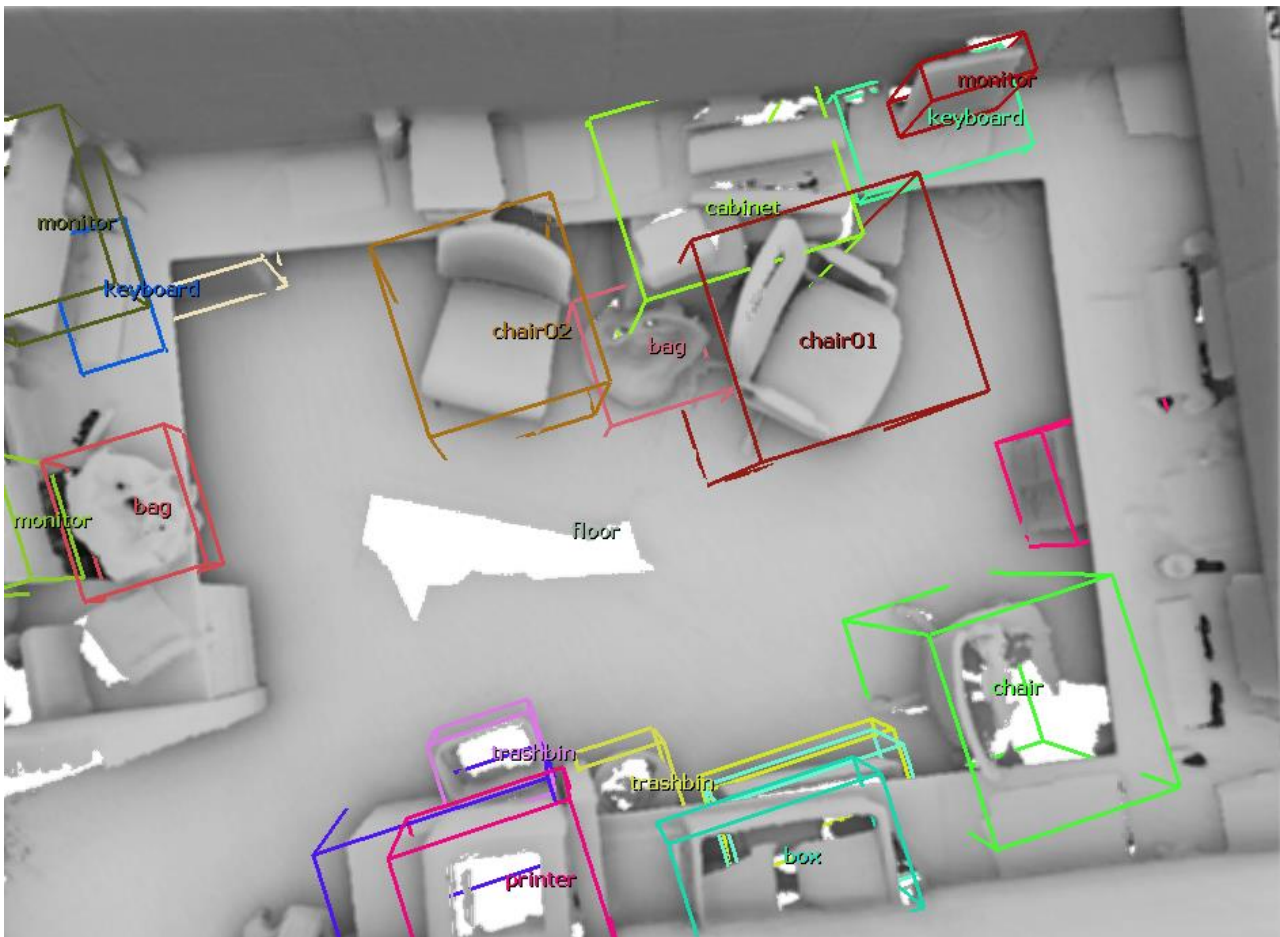


Final results



This image shows a 3D point cloud of a room interior, where different objects and surfaces are segmented into distinct colors. The walls are light blue, the floor is light green, and various furniture items like chairs, tables, and a sofa are rendered in a variety of colors including red, orange, yellow, green, blue, and purple. The segmentation is refined, showing clear boundaries between different object classes.

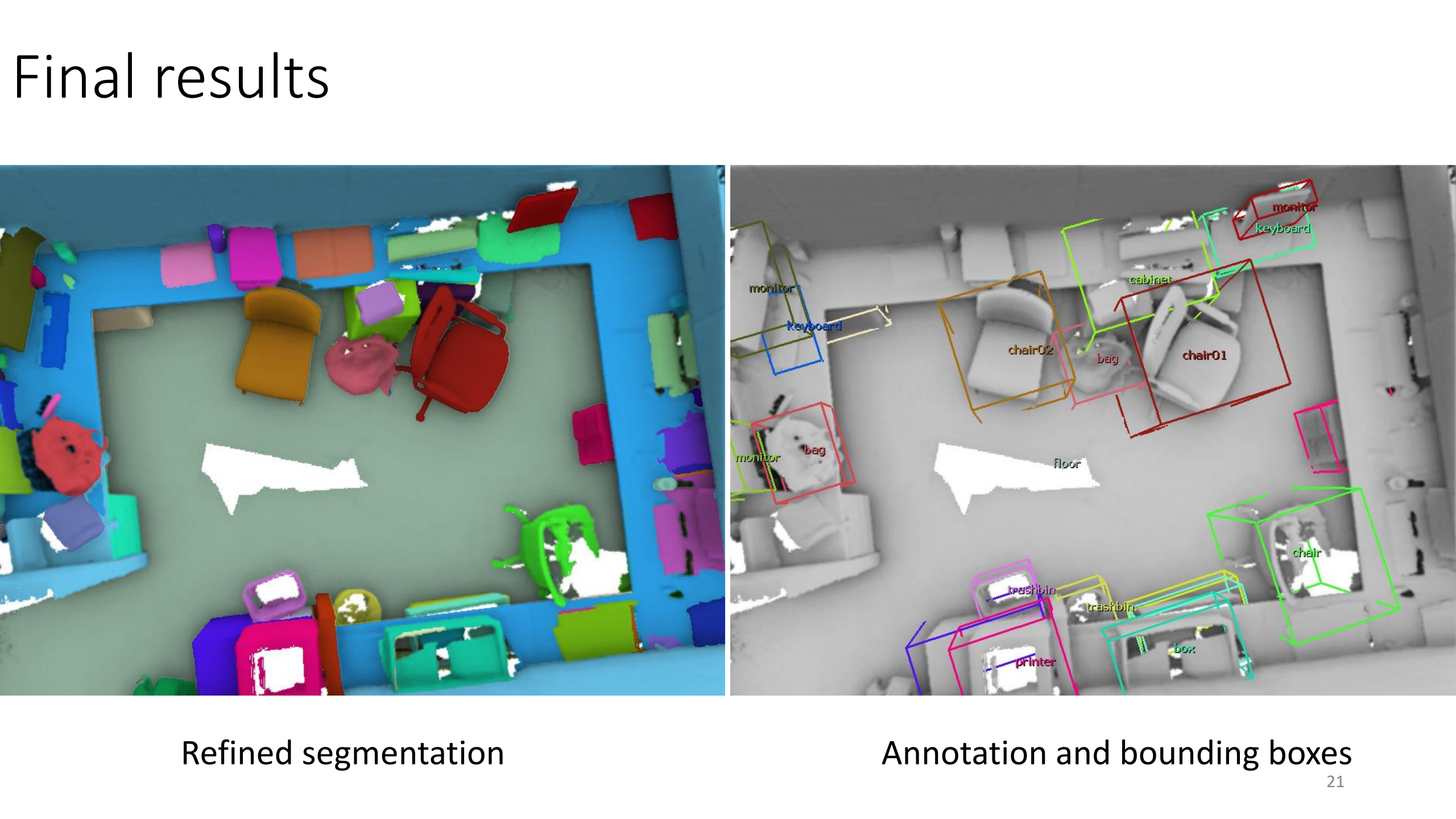
Refined segmentation



This image displays the same 3D point cloud of the room, but with colored bounding boxes and text labels identifying specific objects. The labels include 'monitor', 'keyboard', 'cabinet', 'chair01', 'chair02', 'bag', 'floor', 'trashbin', 'printer', and 'box'. The bounding boxes are colored to match the object they enclose, providing a clear visual link between the label and the object in the scene.

Annotation and bounding boxes

21

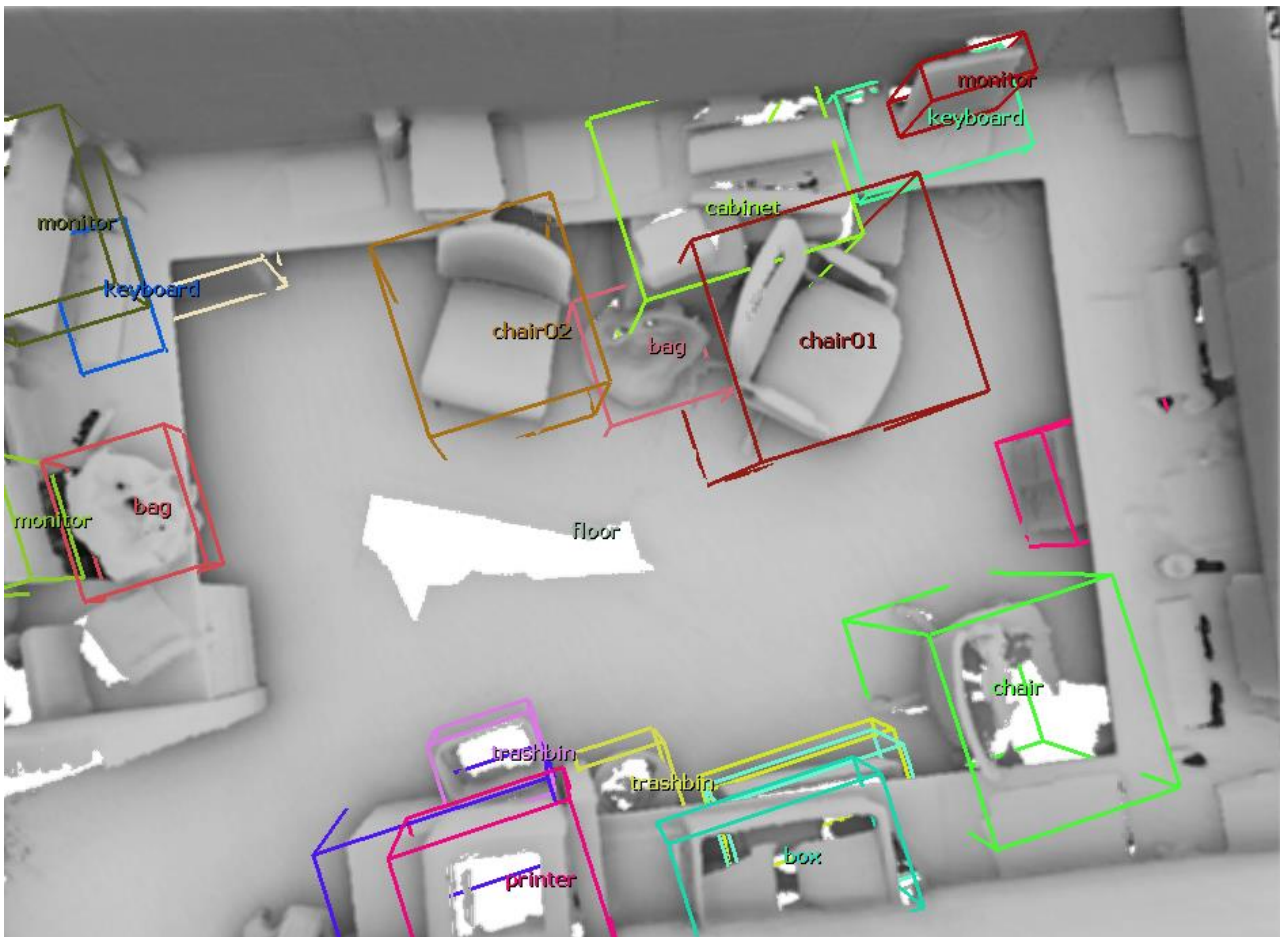


Final results



This image shows a 3D point cloud of a room interior, where different objects and surfaces are segmented into distinct, color-coded regions. The floor is a light blue, walls are a darker blue, and various furniture items like chairs, tables, and a sofa are rendered in different colors such as red, orange, green, and purple. The segmentation is refined, showing clear boundaries between the different parts of the scene.

Refined segmentation



This image shows the same 3D point cloud scene as the left image, but with colored bounding boxes drawn around specific objects. Each box is labeled with the name of the object it encloses. The labels include 'monitor', 'keyboard', 'cabinet', 'chair02', 'bag', 'chair01', 'floor', 'chair', 'trashbin', 'printer', and 'box'. The bounding boxes are colored in various colors like red, orange, green, and blue, matching the segmentation colors.

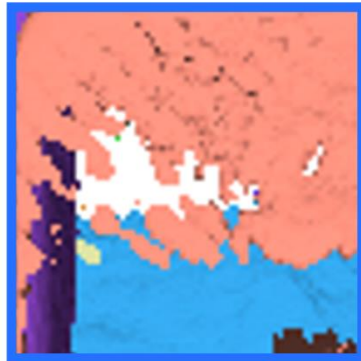
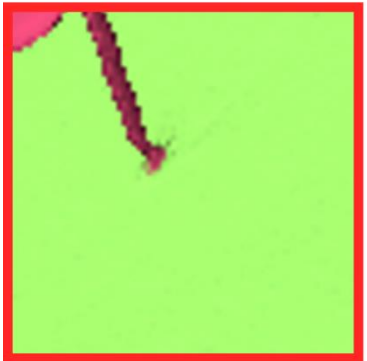
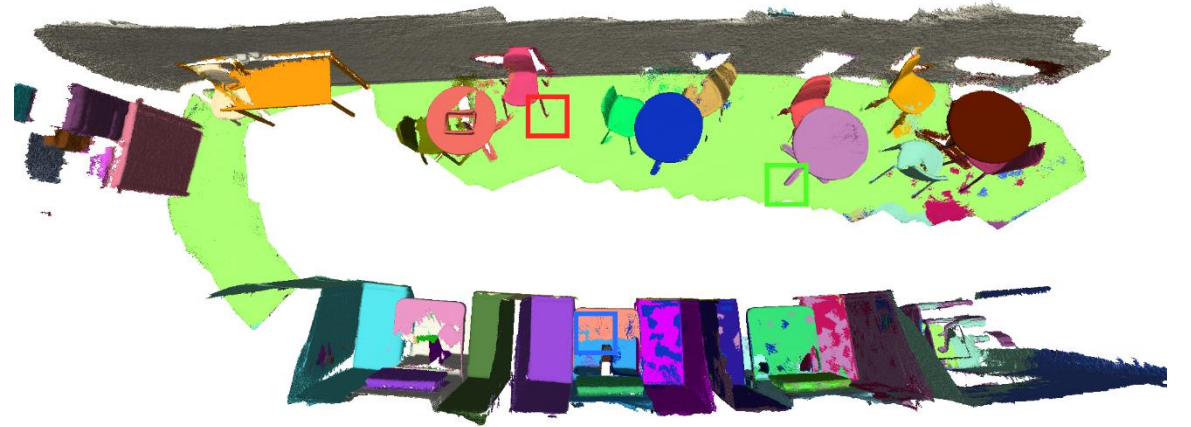
Annotation and bounding boxes

21

Proof-of-concept applications

- Annotation transfer
- Shape completion
- Scene relighting
- CAD scene synthesis
- Novel view synthesis

Annotation transfer

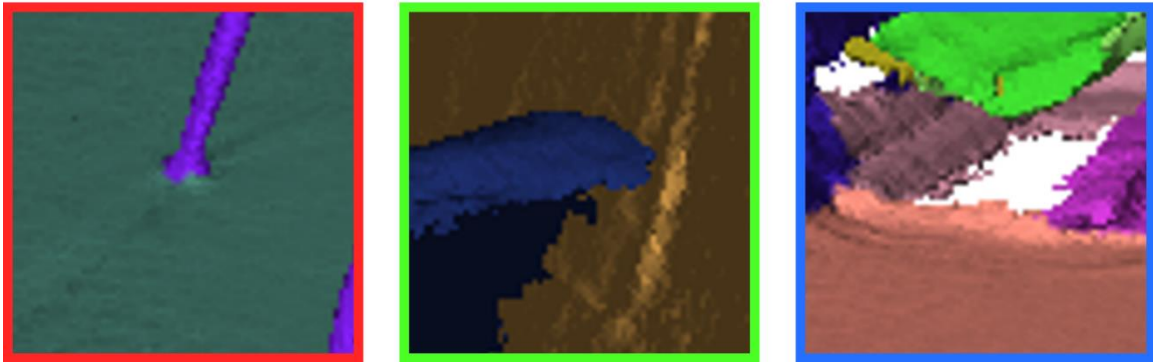
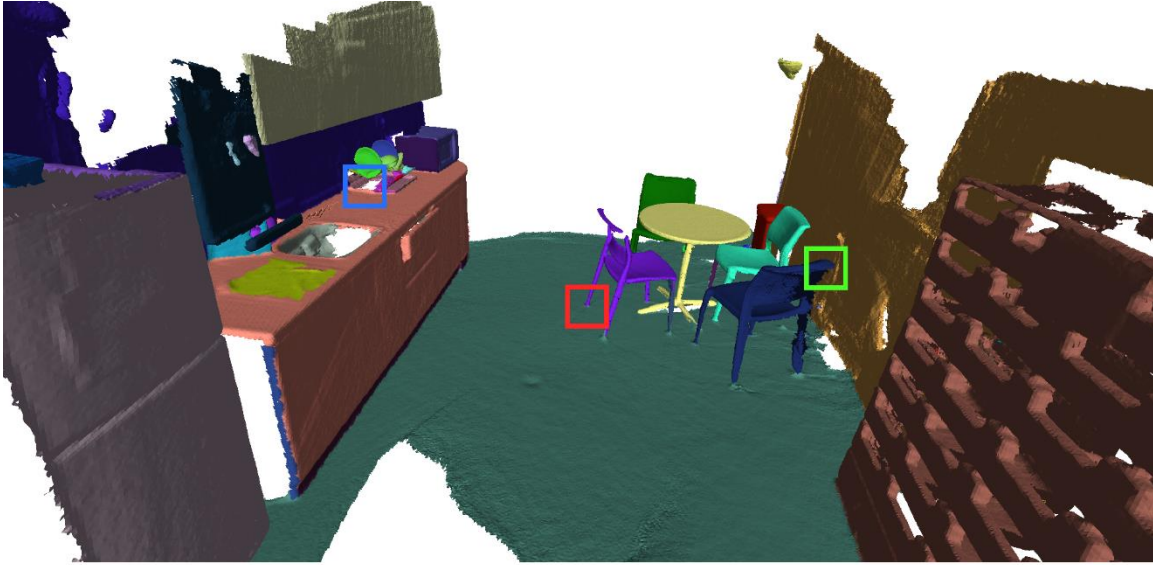


Input

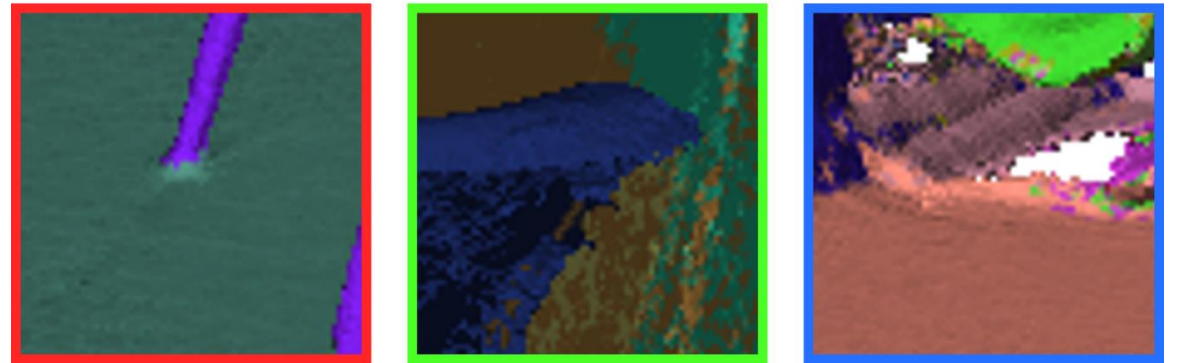
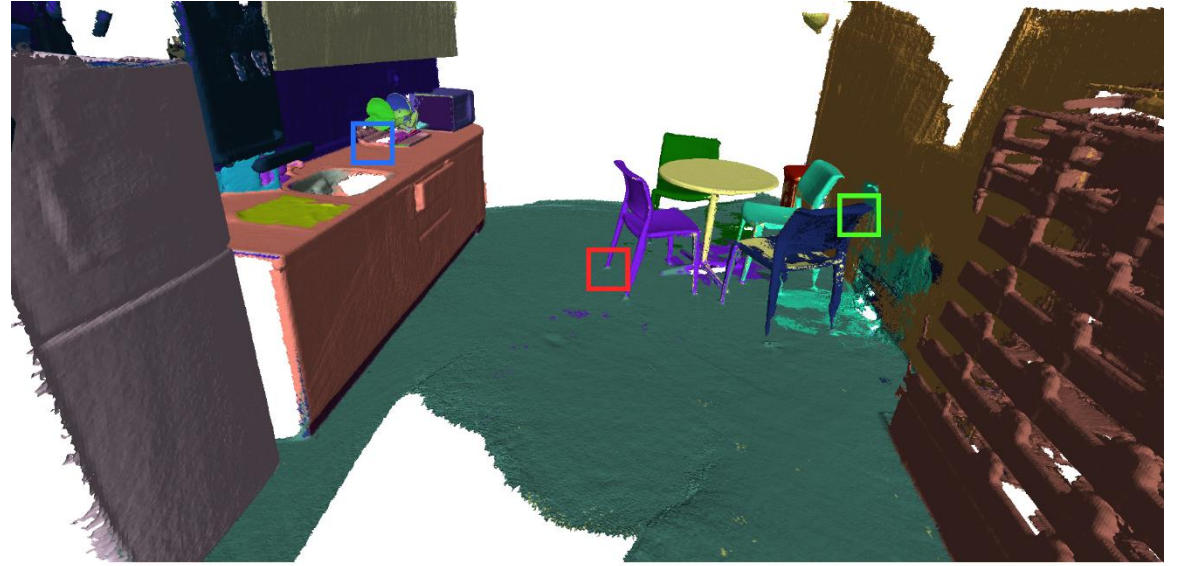


Nearest neighbour label propagated

Annotation transfer

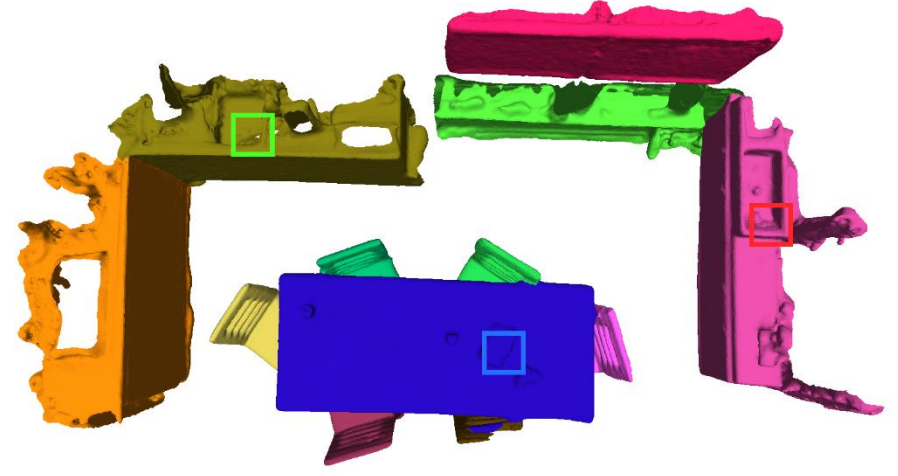
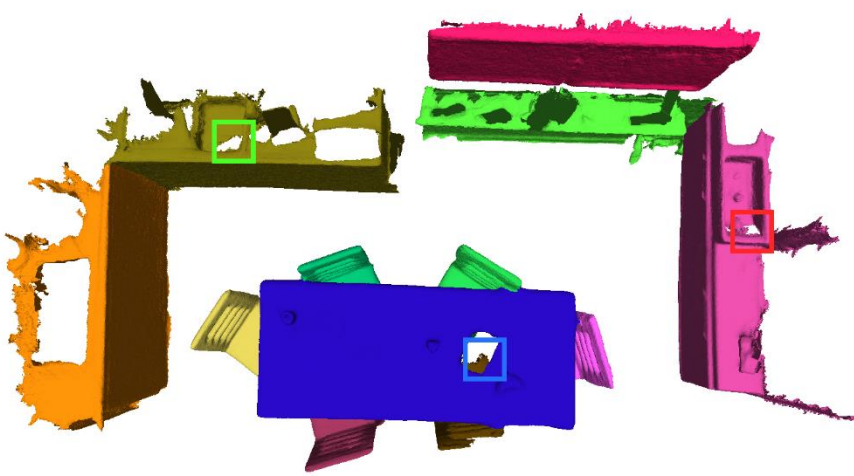


Input

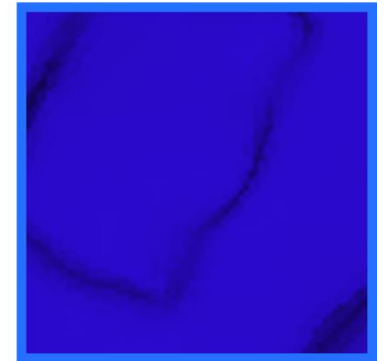


Nearest neighbour label propagated

Geometry-level shape completion

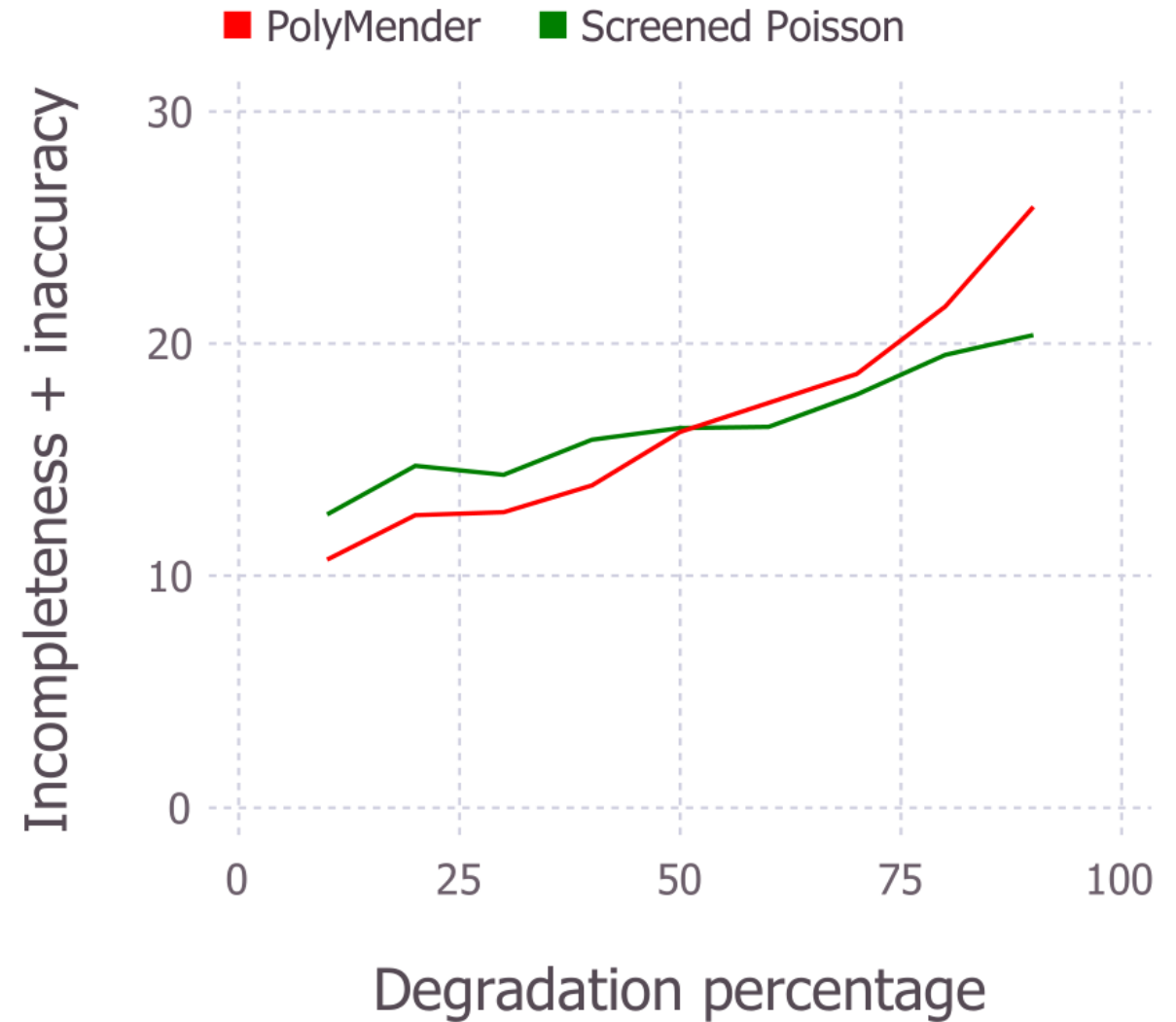
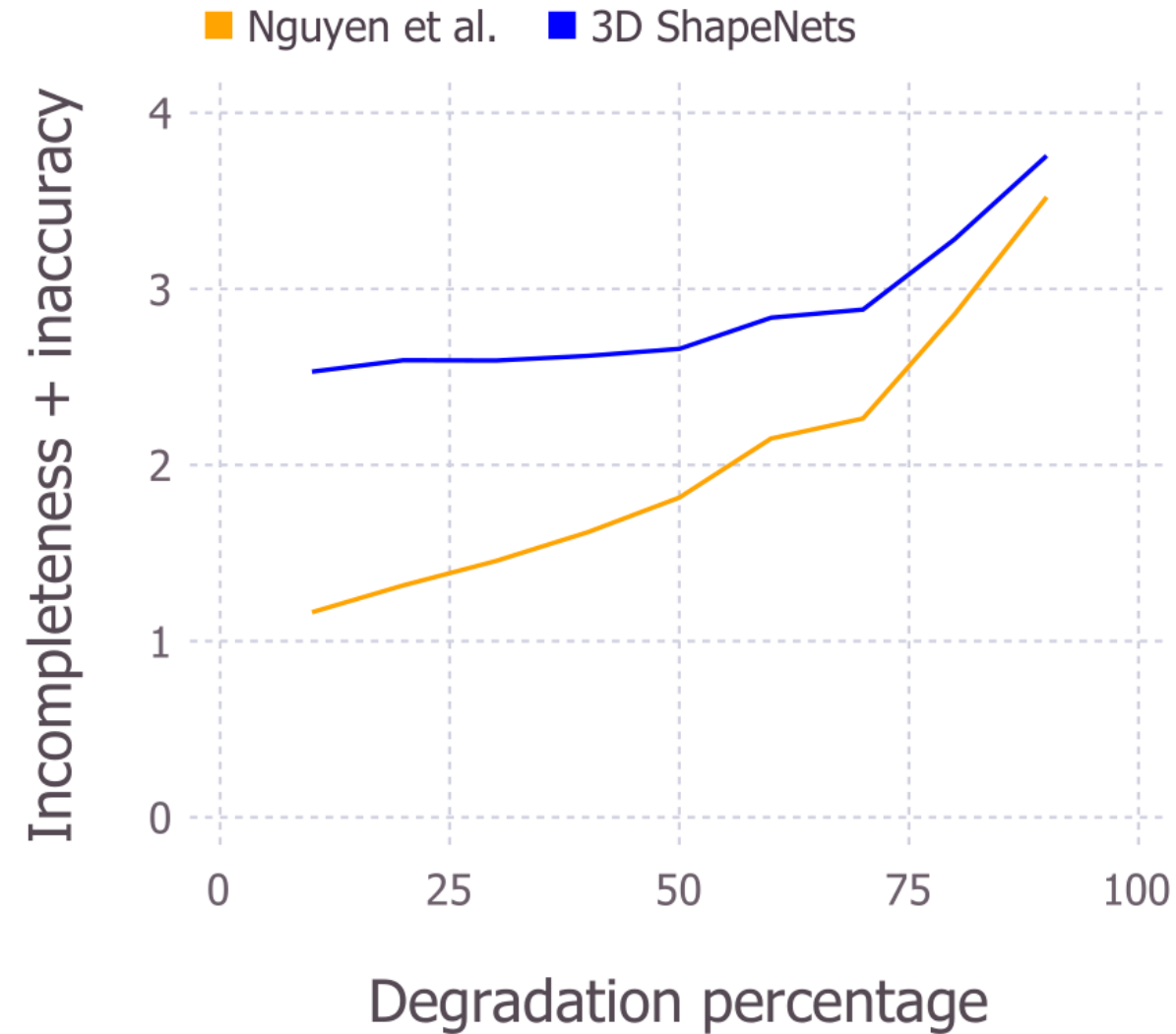


Input

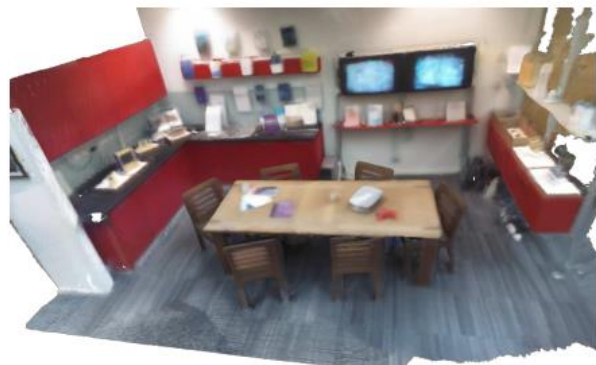
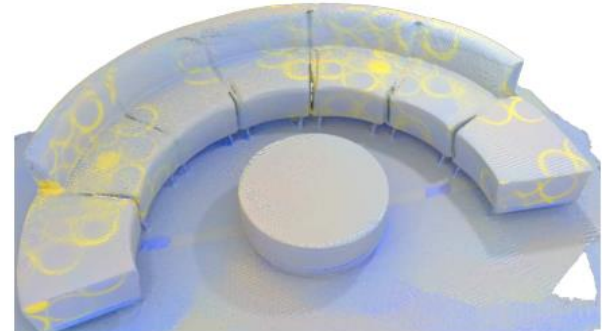
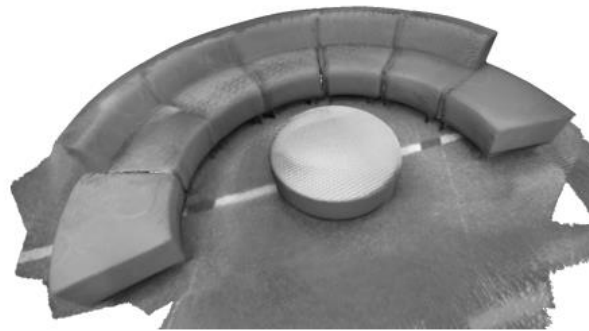
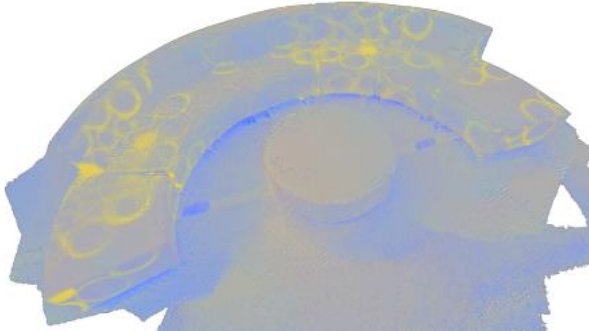
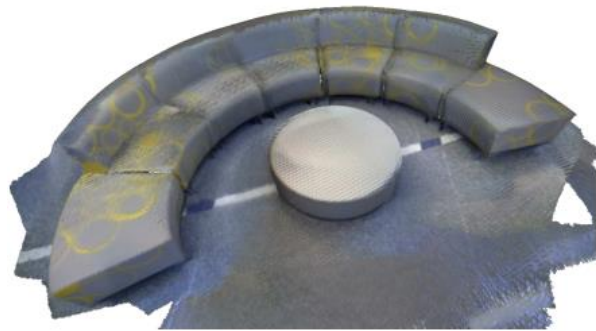
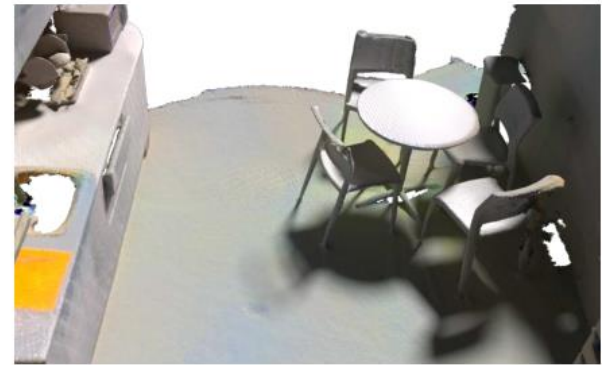
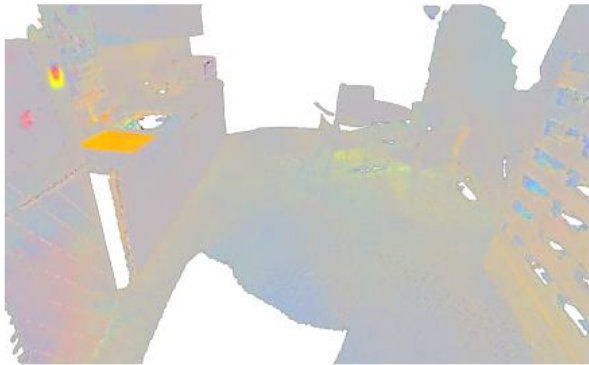


Screened Poisson reconstruction
[Kazhdan and Hoppe, SGP 2013]₂₅

Semantic-level object completion



Scene relighting



Input

Reflectance

Shading

Relighting²⁷

Scene synthesis

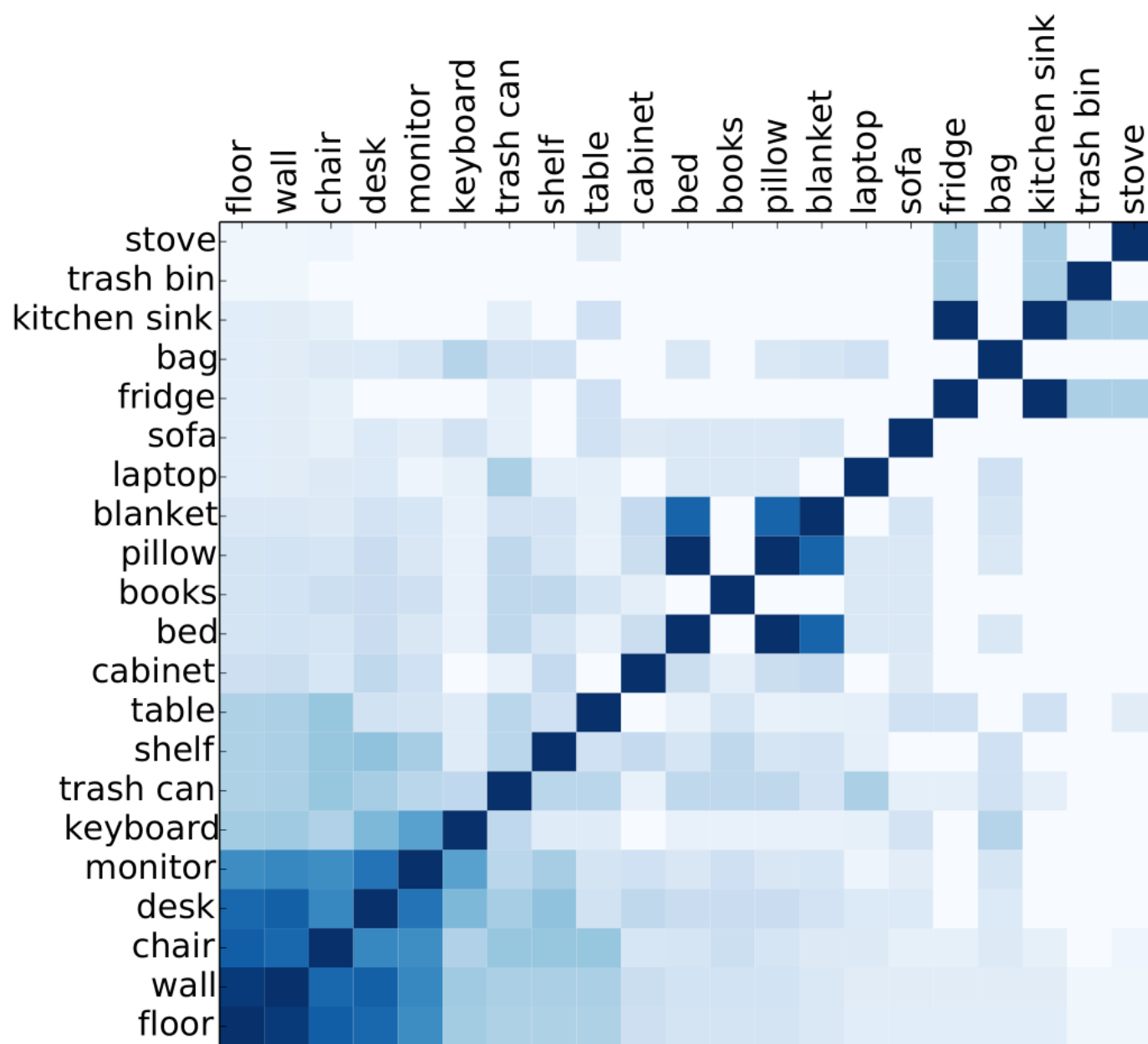


desk



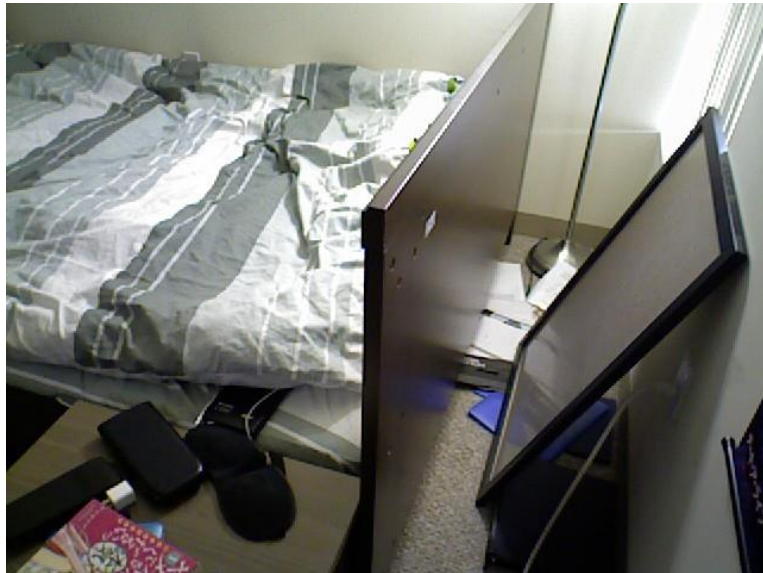
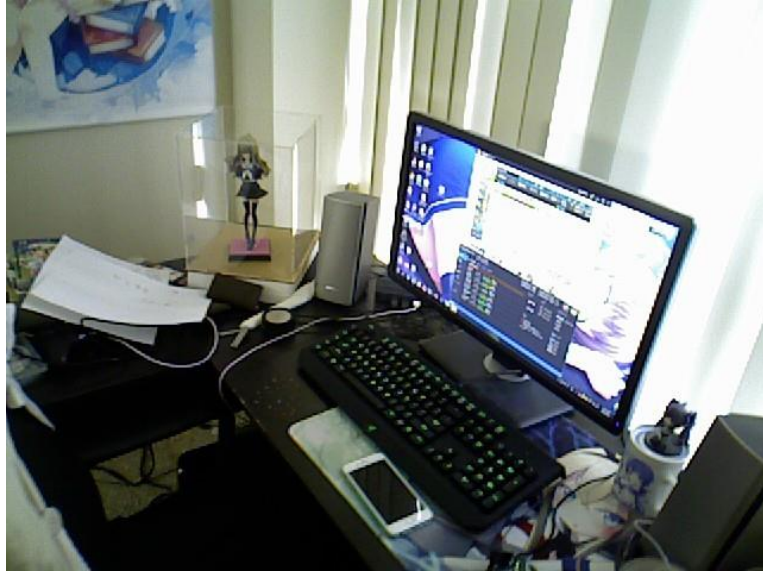
chair

Placement probabilities
[Chen et al., MM 2015]



Co-occurrence
[Xu et al., TOG 2013]

Novel view: widening FOV



Novel view: widening FOV / near plane clipping



- **100+** RGBD scenes with fine-grained annotation.
- Potential applications
 - Annotation transfer
 - Shape completion
 - Scene relighting
 - Scene analysis and synthesis
 - Novel view synthesis
- Future works
 - Online platform for benchmarking, e.g., object detection and recognition.
 - Extend and maintain the dataset with help from community.

Conclusions



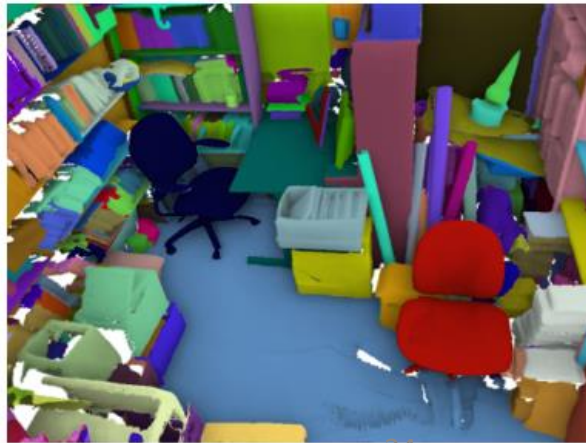
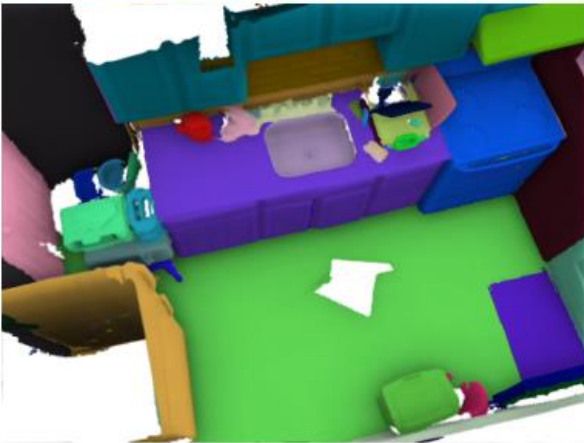
Challenges

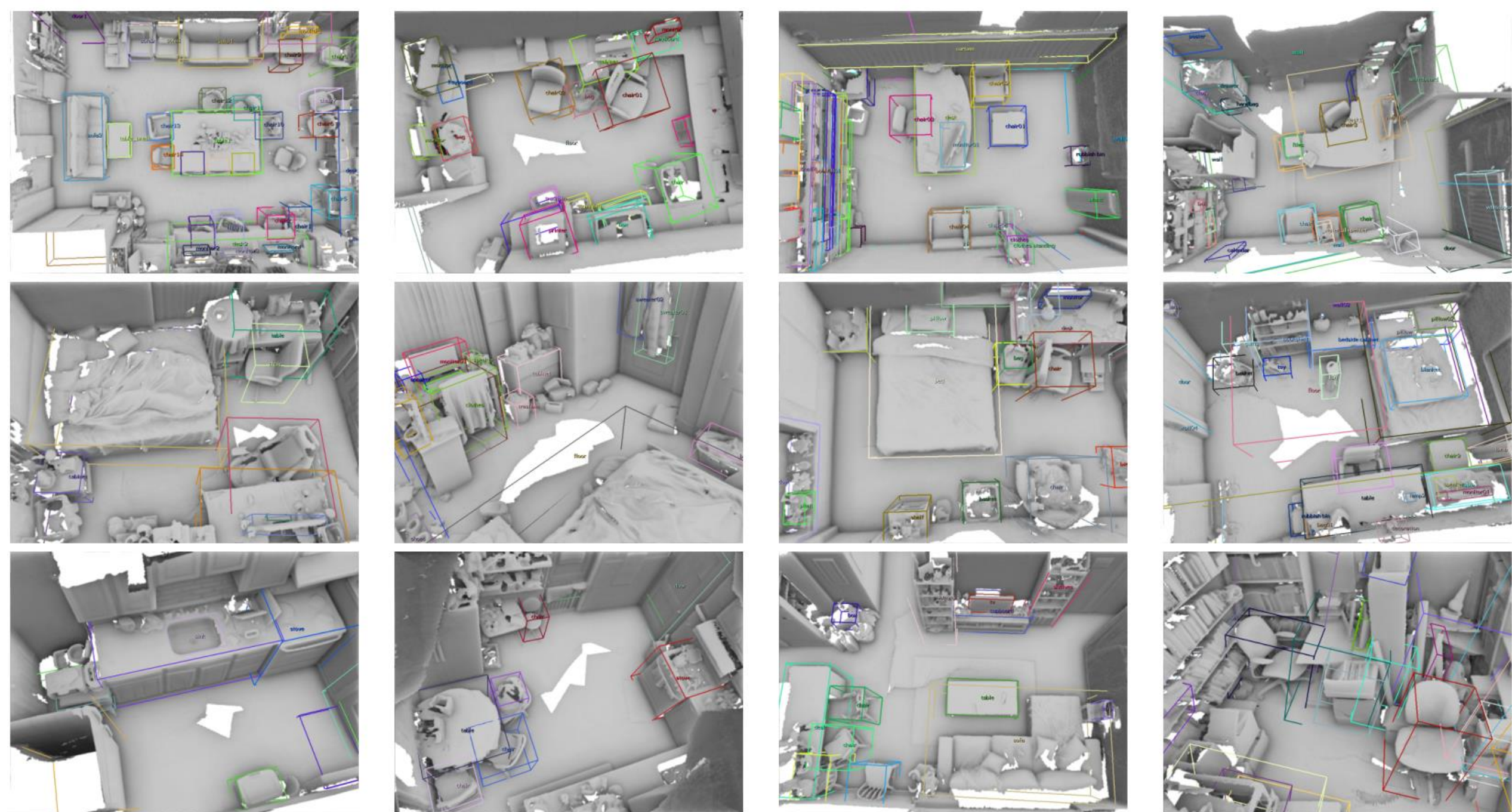
- Current reconstruction takes 1 – 2 hours to complete.
- Current annotation technique requires 30 – 45 minutes per scene for refining segmentation and annotation.
- Texture is generated using a baseline method.
- Resolution 640 x 480 is limited but no better consumer-grade RGBD camera is available.

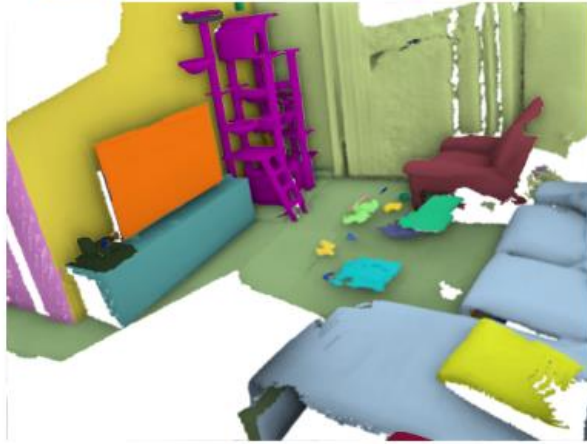
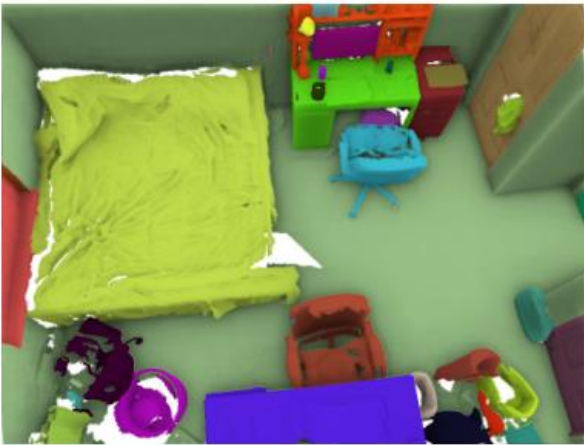
Before we end...

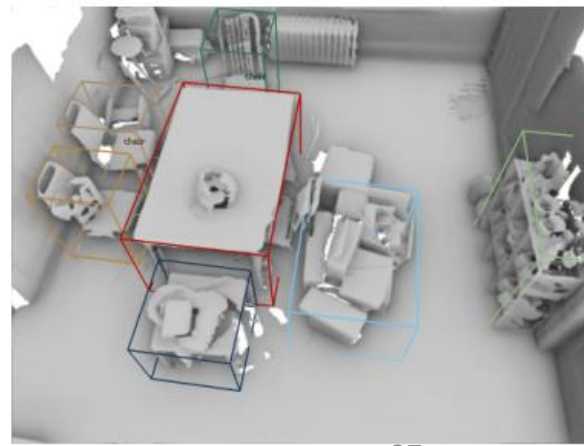
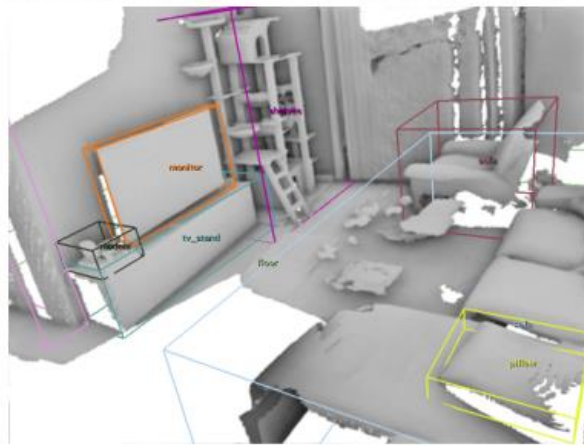
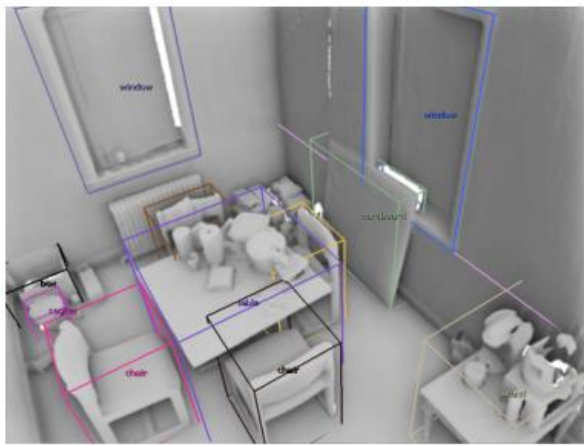
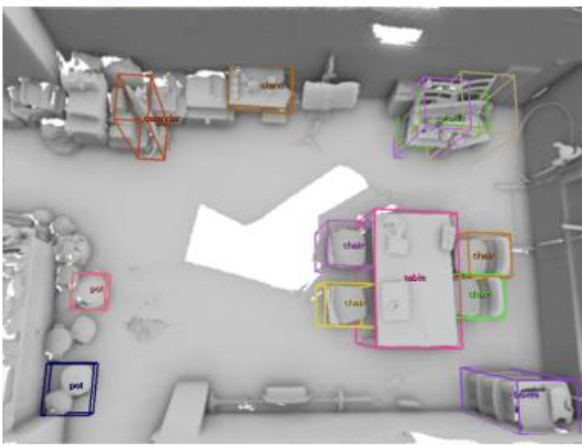
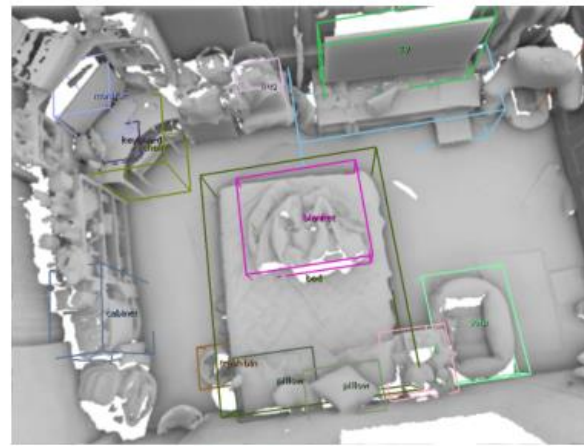
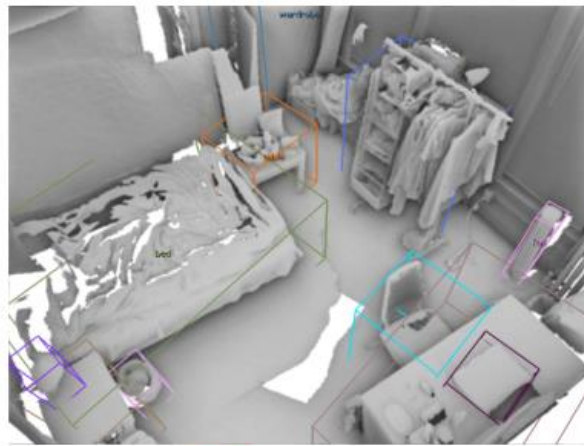
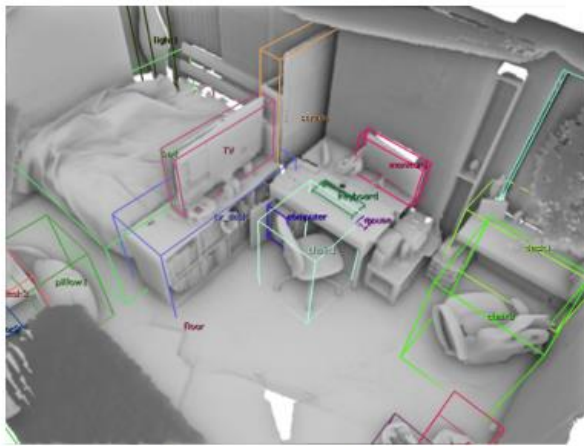
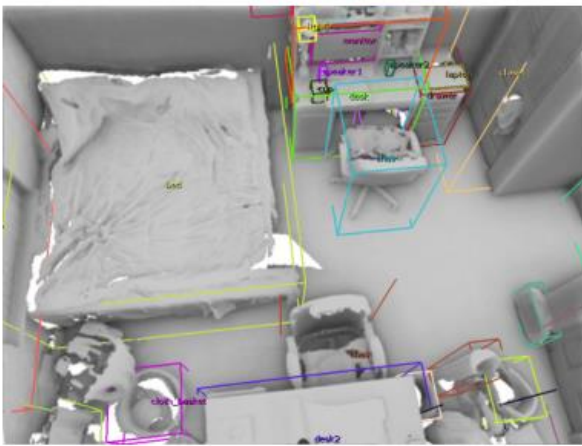
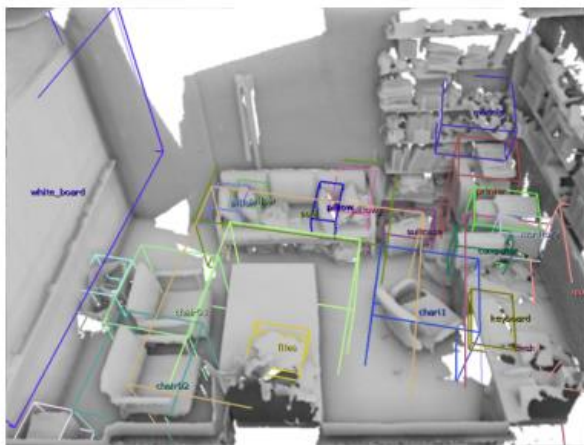
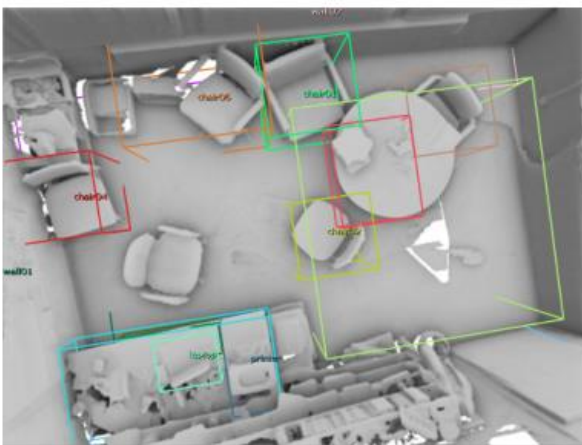
More segmentation and annotation from some popular categories:

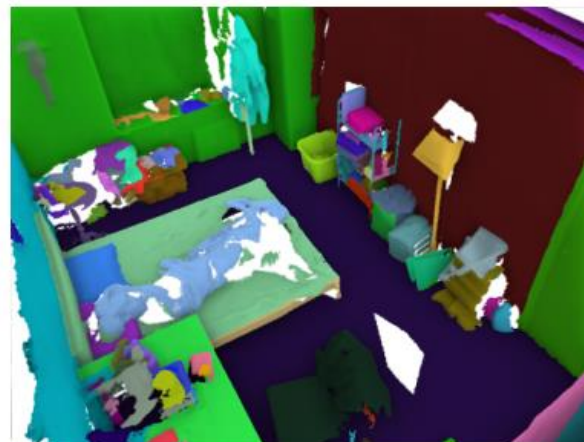
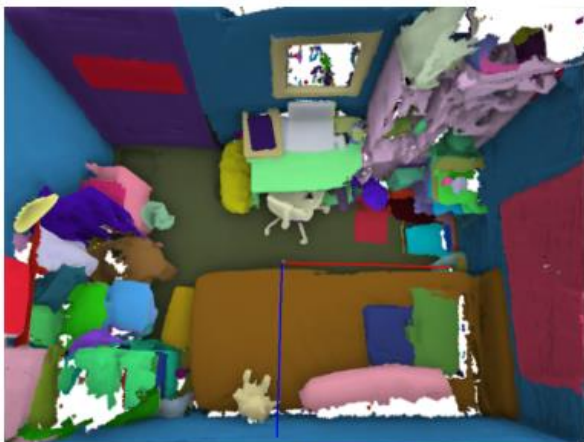
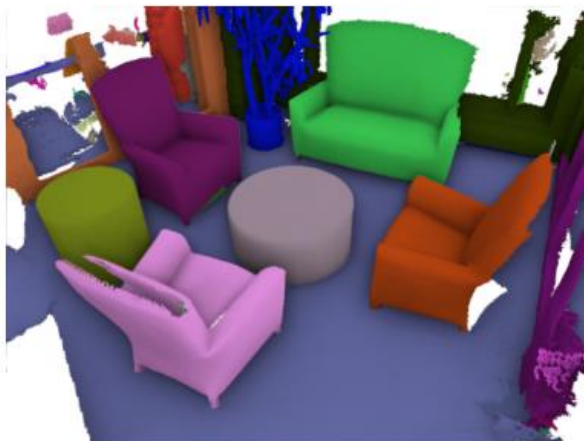
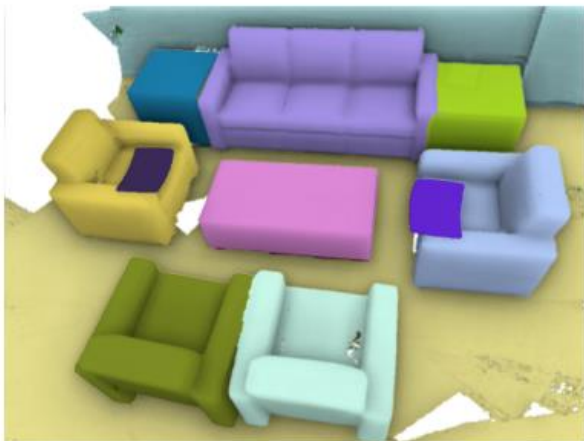
- Workplace
- Bedroom
- Livingroom
- Kitchen
- Lounge

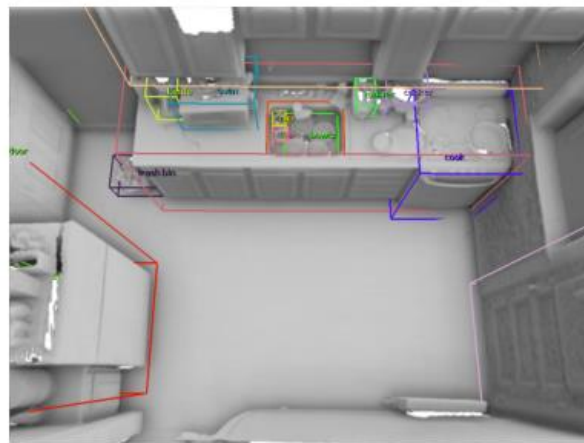
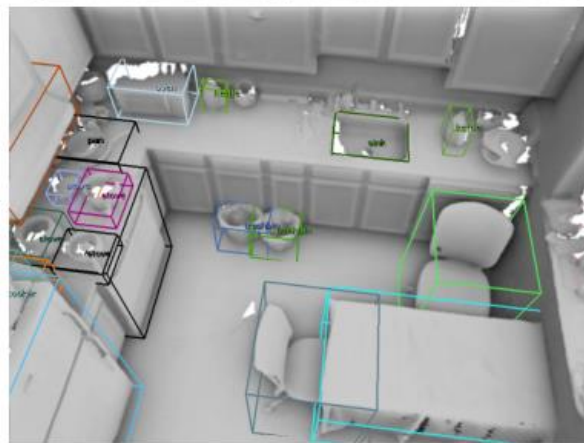
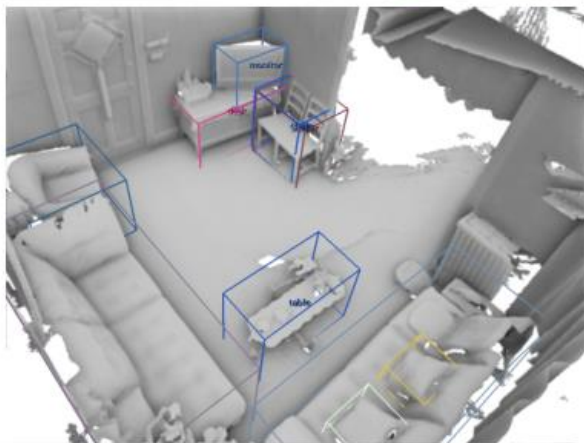
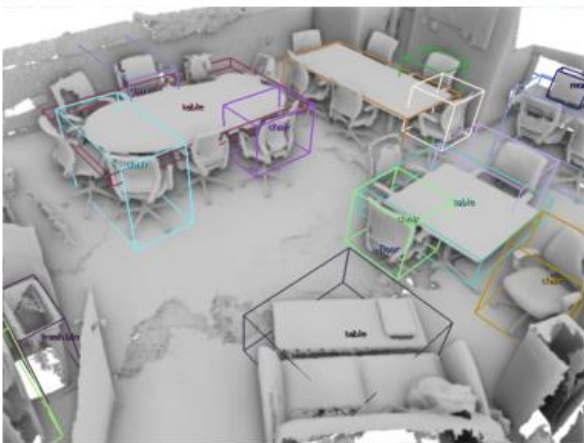
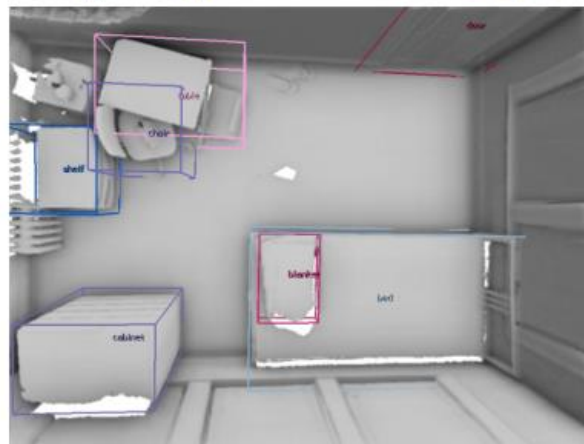
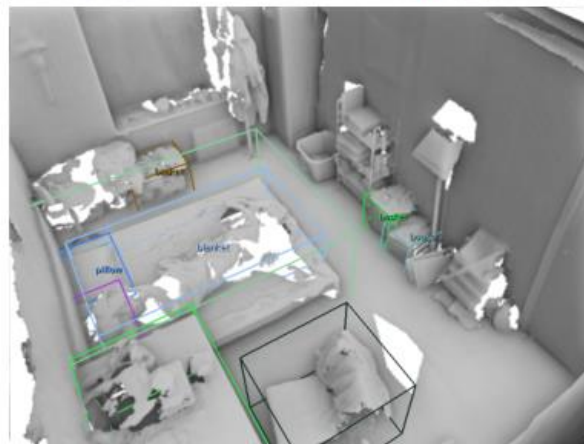
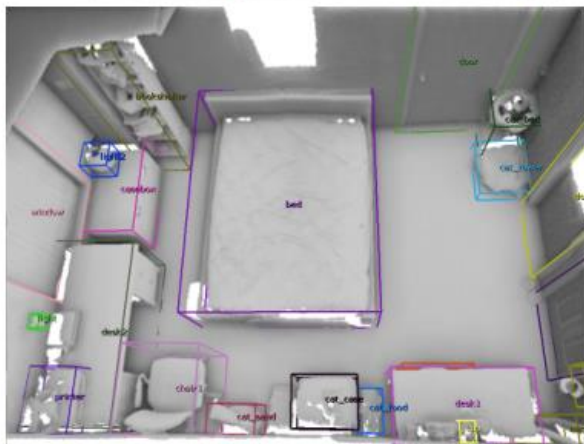
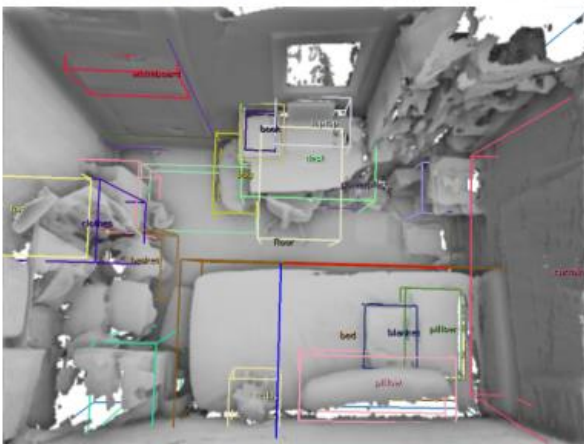
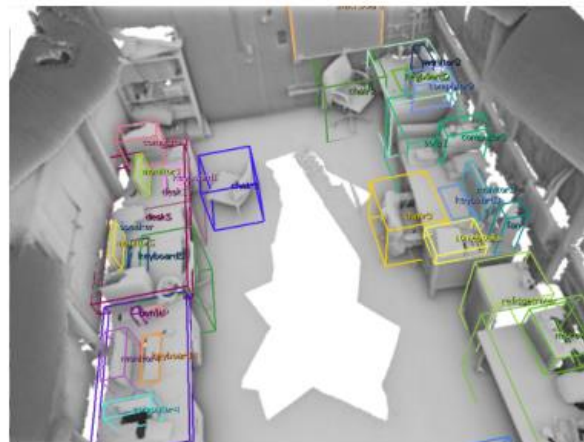
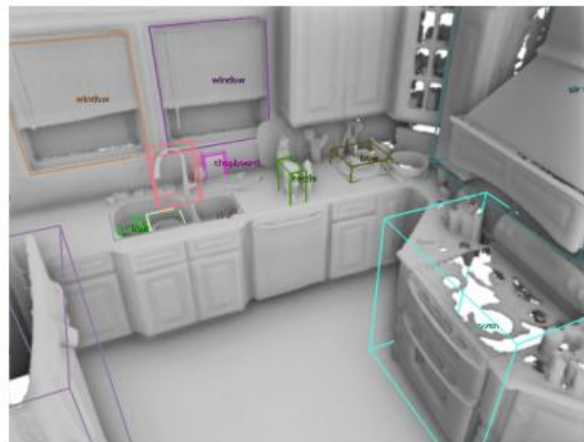
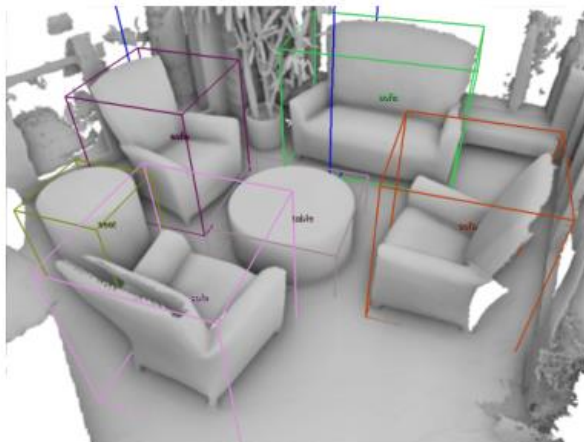
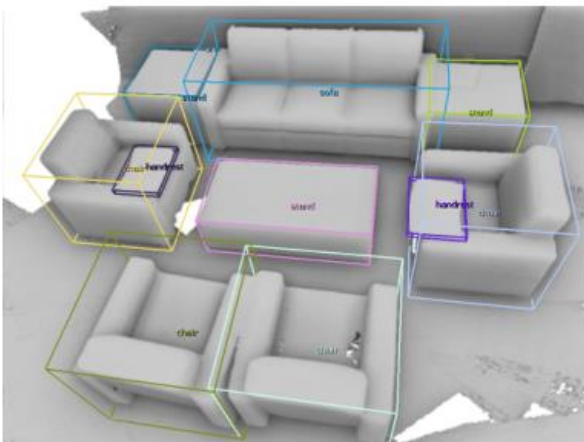












Acknowledgement

We would like to extend our sincere gratitude and appreciation to

- Fangyu Lin for his assistance with the data capture.
- Guoxuan Zhang for early development of the annotation tool.
- The anonymous reviewers for their constructive comments.
- Jamie Shotton (Microsoft Research) for the Kitchen scene data.
- NVIDIA for graphics card donation.
- Singapore National Research Foundation (NRF) and the U.S. National Science Foundation (NSF) for funding support.

The team



Thank you!

<http://www.scenenn.net>

- Dataset download script (Python)
- Annotation tool (Windows and Linux)
Code available upon request.
- Supplemental and technical reports
- Contact: scenenn@gmail.com

