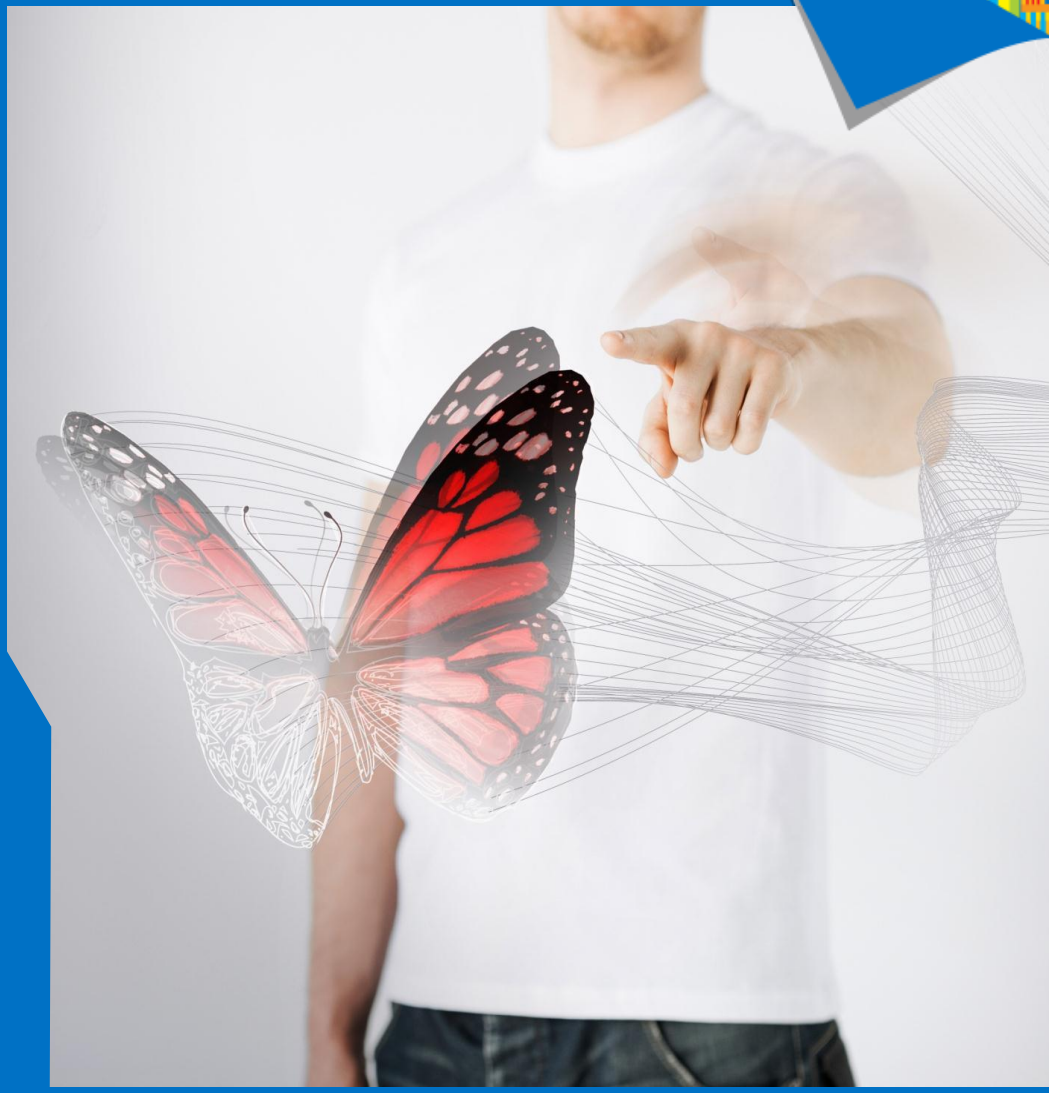
Towards Visual Understanding…



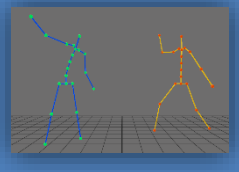
Real Sense is …

Middleware Algorithms

3D Capture

Object Recognition

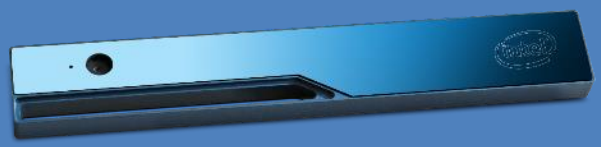
Human Tracking



Hand/Gesture

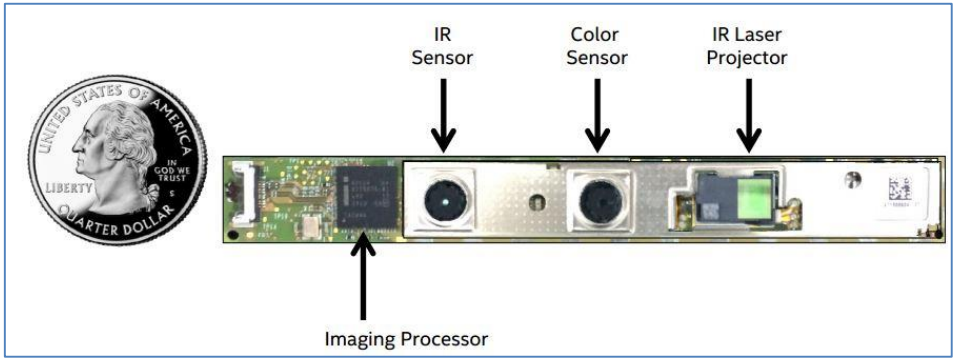
Face

3D Cameras



intel REALSENSE™ TECHNOLOGY

OEM Designs



HP Sprout*

Asus ROG G771JM*

Dell Inspiron* 15 5548

HP Envy* 15t Touch

Asus N551JQ*

NEC LaVie Hybrid* Advance HA850/AAS

Lenovo ThinkPad* E550

Asus X751LD*

Lenovo B5030

Lenovo Z41 & Z51

Fujitsu Esprimo* WH77/S

NEC LaVie Desk* AIO DA970/AAB

Acer Aspire* V 17 Nitro

Lenovo ThinkPad* Yoga 15

Asus Zen AIO Z240IC

MediaMarkt PEAQ

NEC Note Standard* NS850/AAB

Dell Inspiron* 23" 7000

PerC SW Research

Shahar Fleishman

Mark Kliger

Shai Aharon

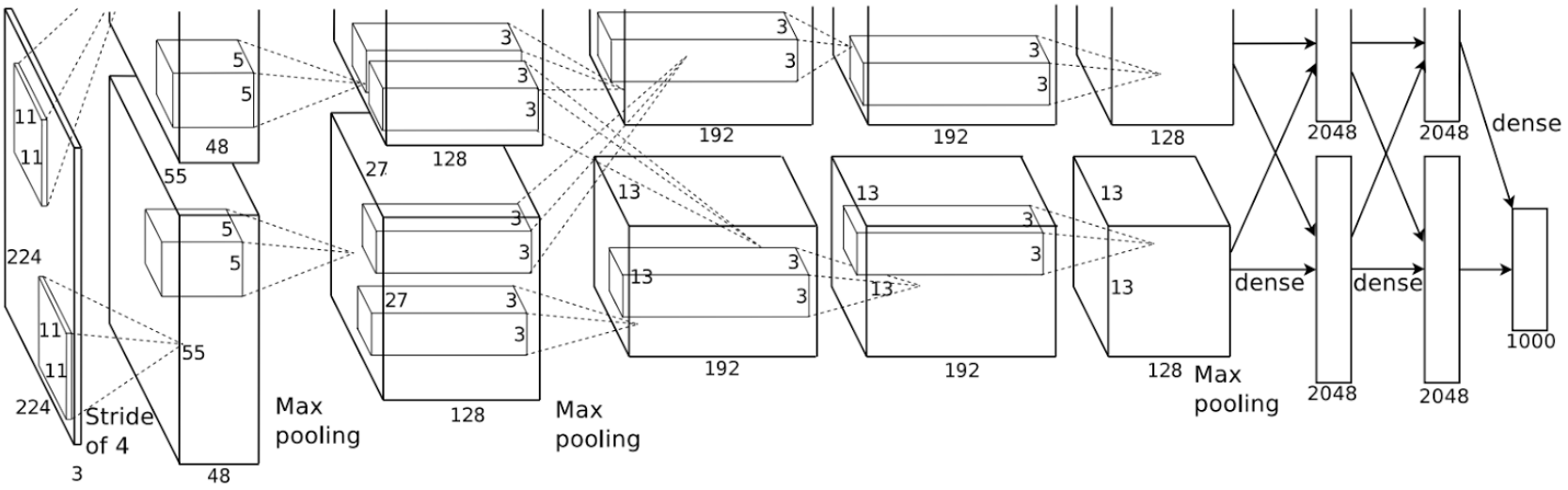
Jackie Assa

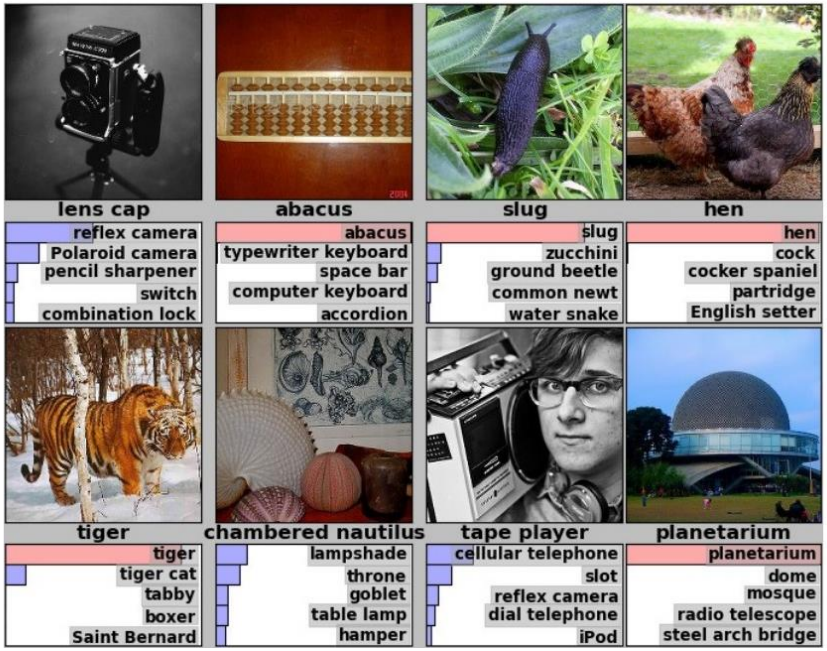
Zachi Karni

Dana Shavit

The Future…

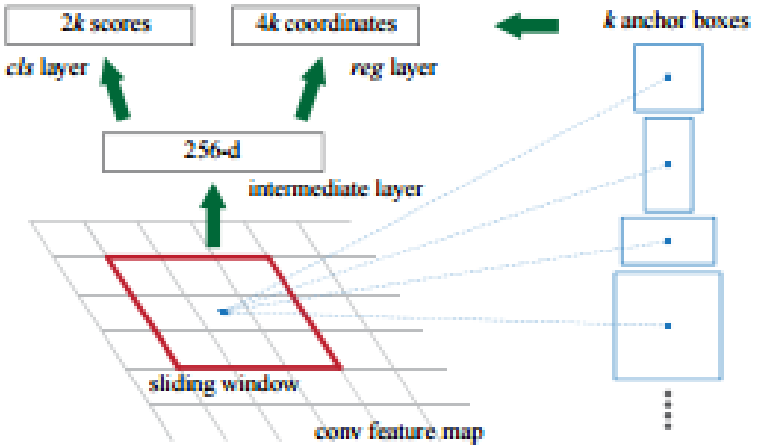
Deep Learning & CNN's ("Alexnet")



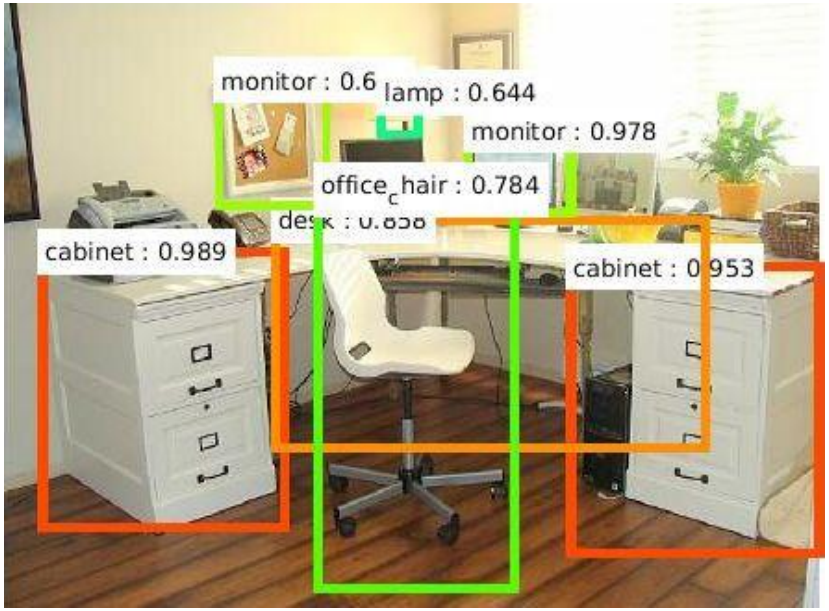


Krizhevsky, et al, 2012

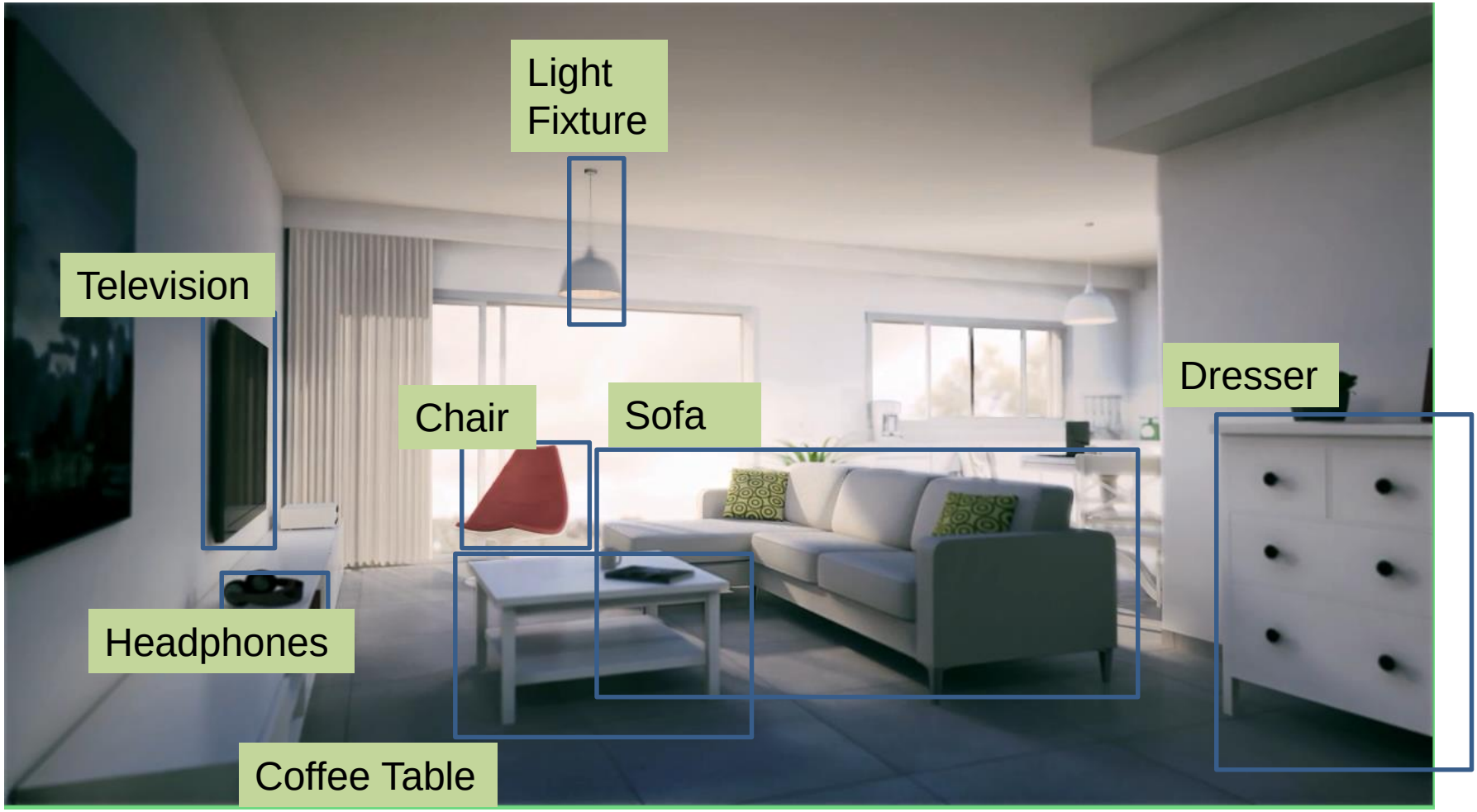
Object Detection (Faster R-CNN)



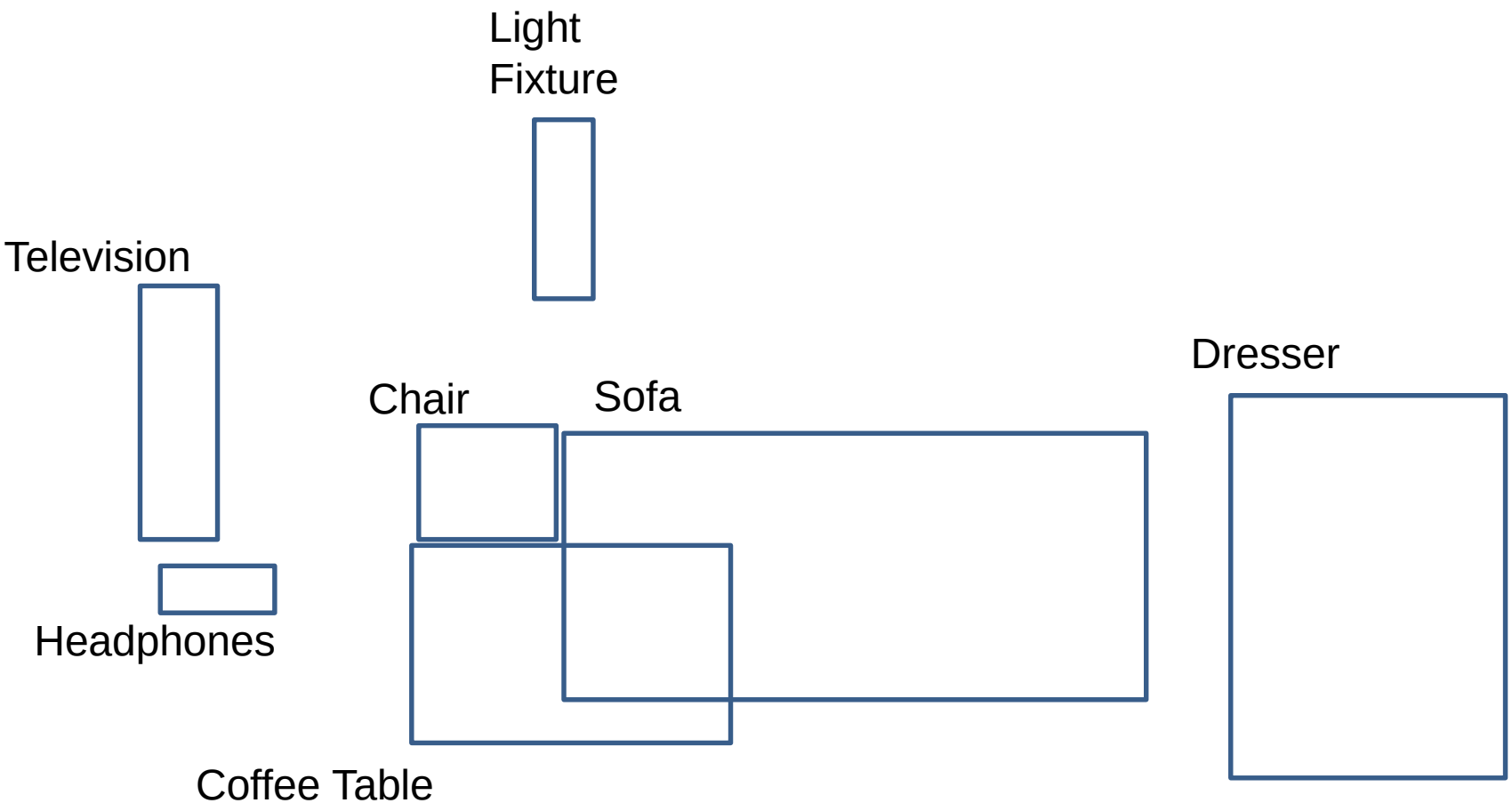
Ren, He, Girshick, Sun, 2015



Object Recognition & Detection

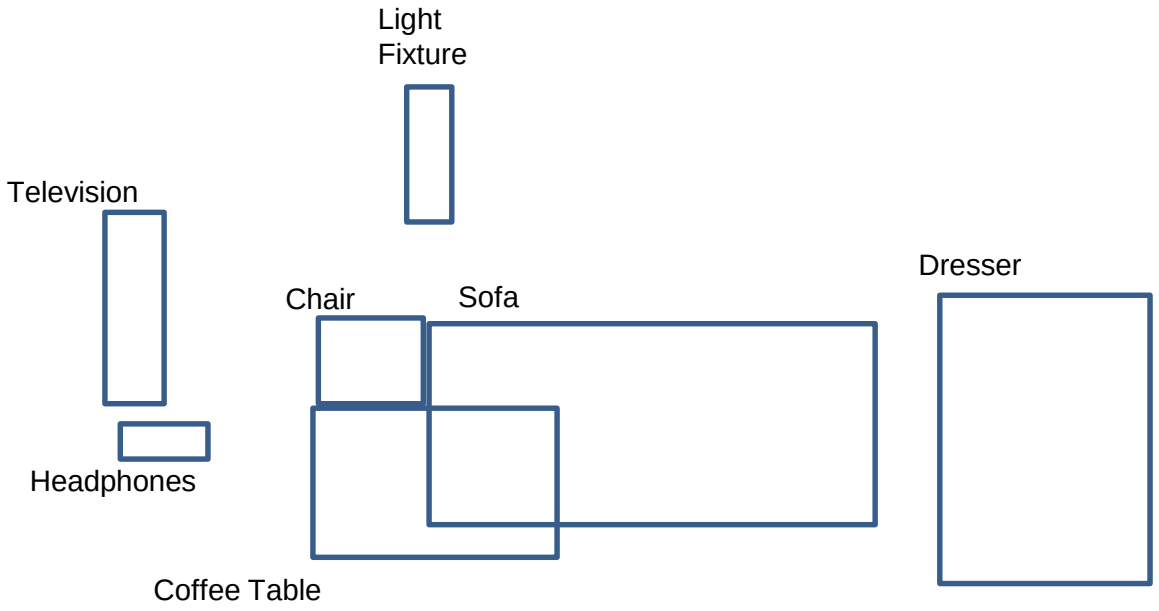


Today’s Capabilities…



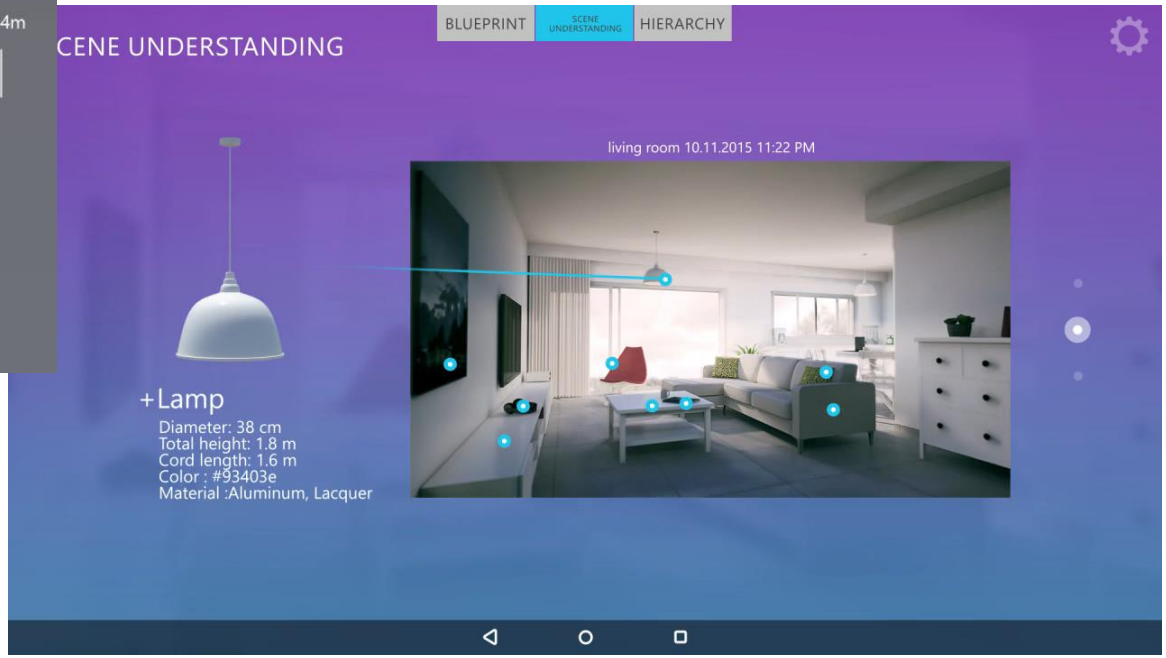
… are not quite sufficient

- “Robot, bring me my headphones.”
- “What would the room look like if I replace the sofa with this loveseat?”
- “Does that chair swivel?”



We want…


6.9m
4.19m
8.14m
3.9m


BLUEPRINT
SCENE UNDERSTANDING
HIERARCHY
CENE UNDERSTANDING
living room 10.11.2015 11:22 PM
+Lamp
Diameter: 38 cm
Total height: 1.8 m
Cord length: 1.6 m
Color : #93403e
Material :Aluminum, Lacquer

Scene Understanding is …

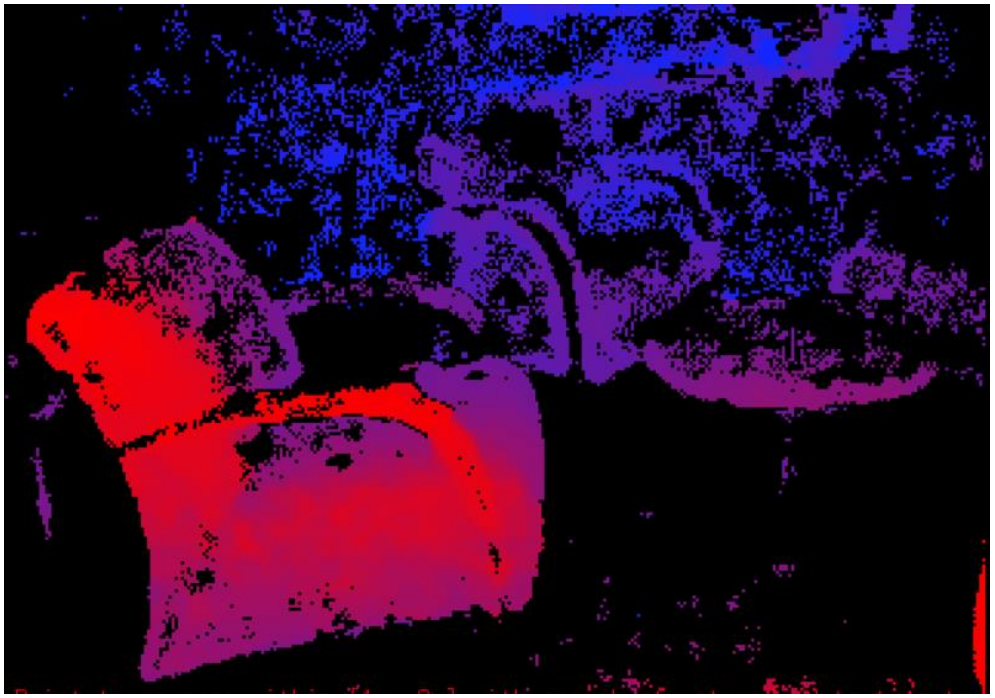
- Identify everything (that I can) in the scene
  - The space it occupies
  - Its boundaries
- Relationships & Interactions
  - Object-object
  - Person-object
- Character & Function
  - Does it move? Is it articulated?
  - What is it composed of?
  - What is it used for?

Step 1: Acquire 3D Geometry

3D Capture

Depth data from a single frame is:

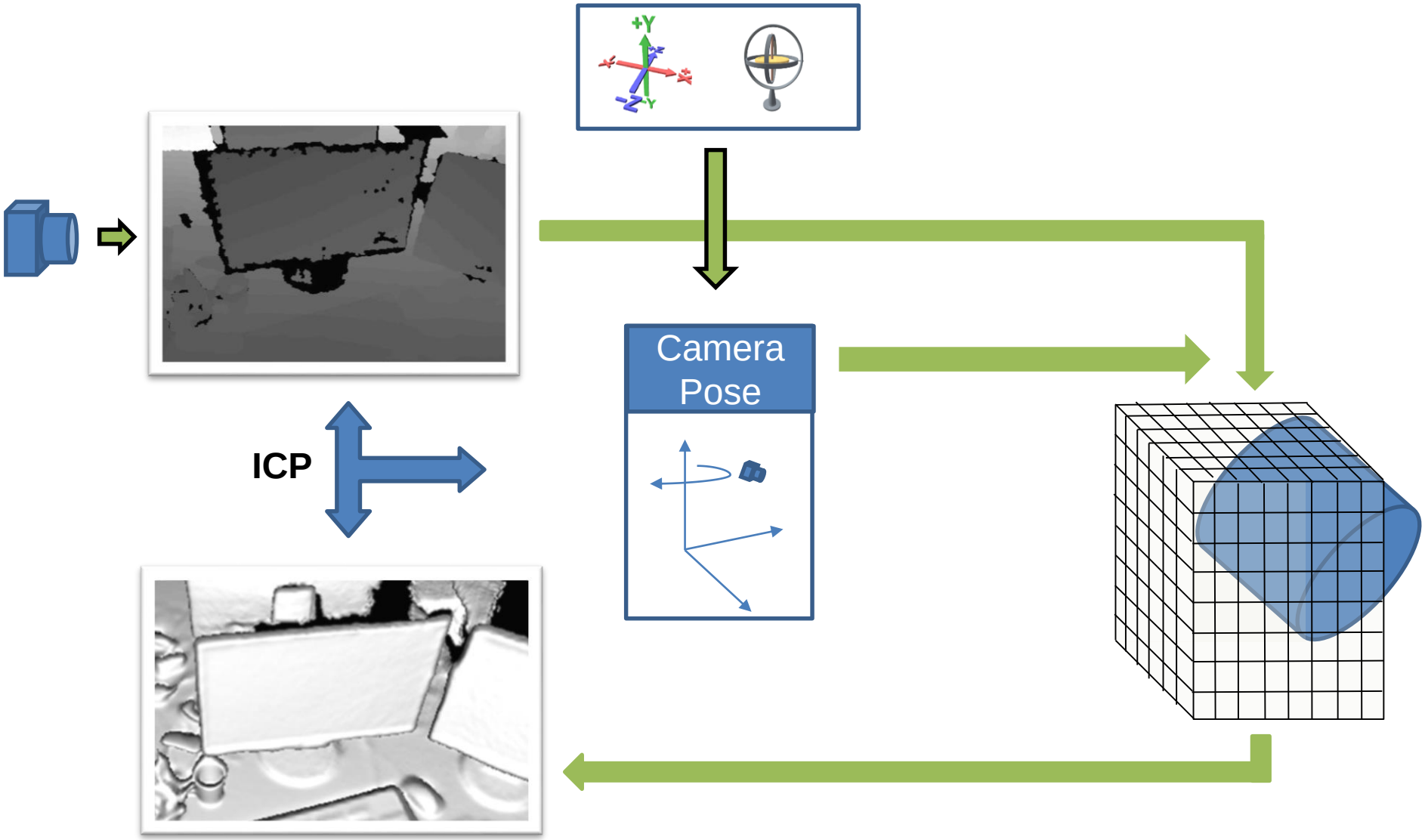
- Noisy
- Incomplete
- Occluded



Solution:

1. Scan the camera and calculate its pose at each frame.
2. Transform each frame and accumulate all the 3D points into a single volume.

Depth Fusion



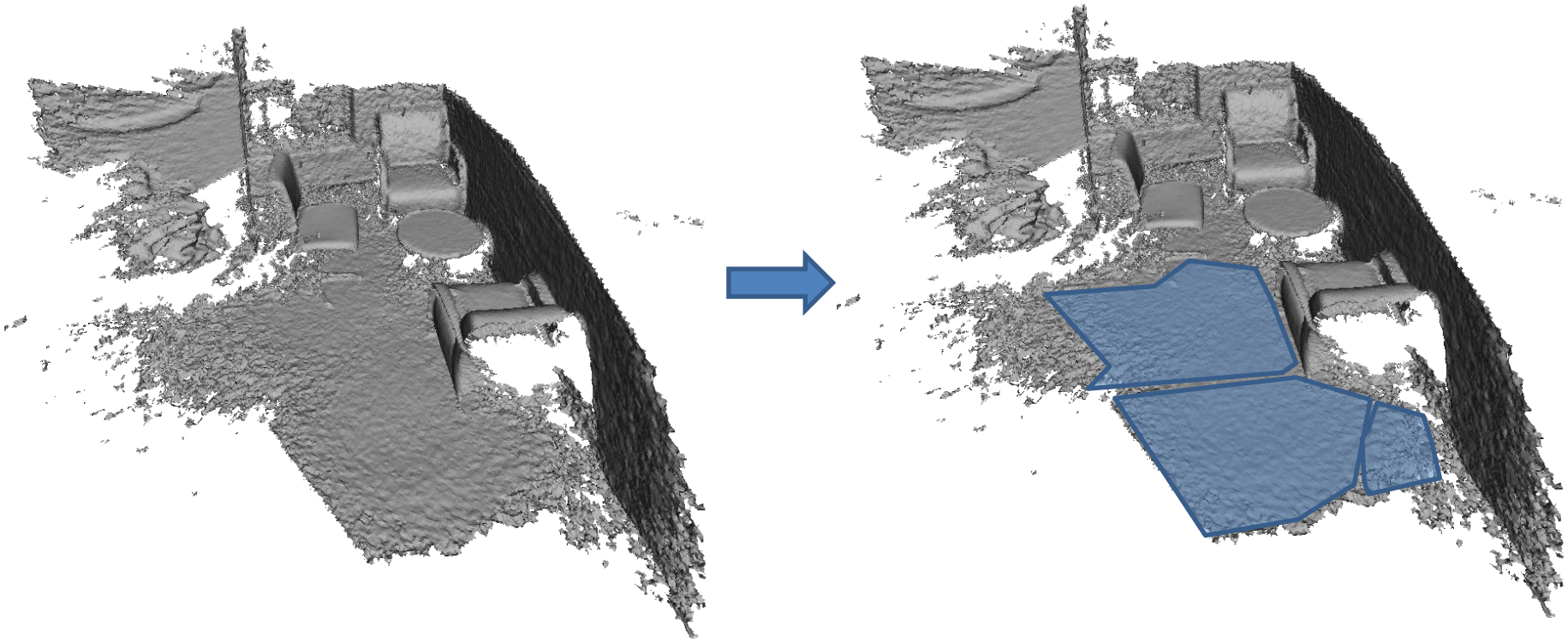
Izadi, et al., “KinectFusion:…”, 2011

1. 3D Capture

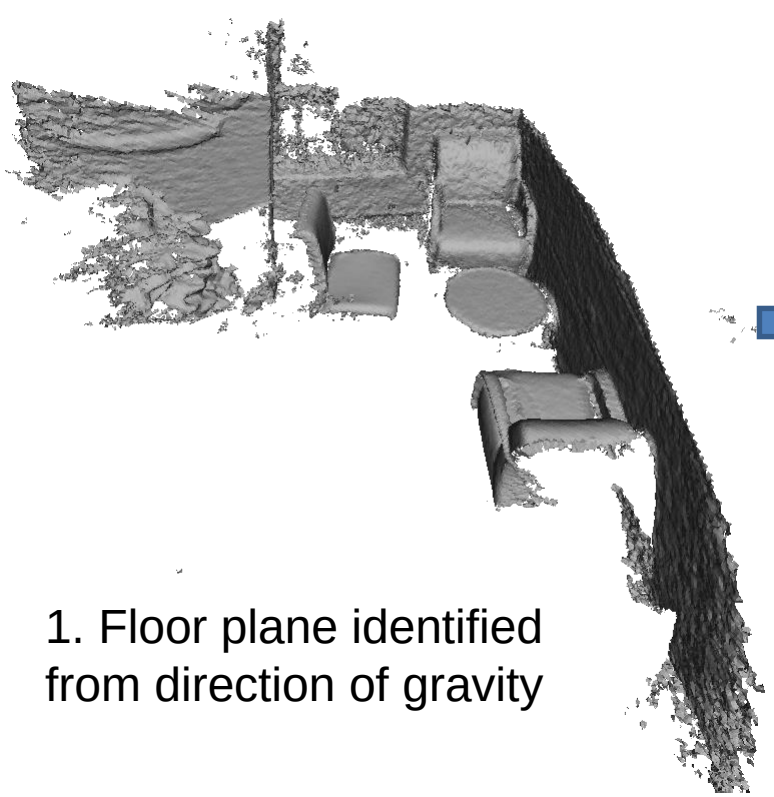
2. Segmentation

- Fit planes to flat surfaces in the scene

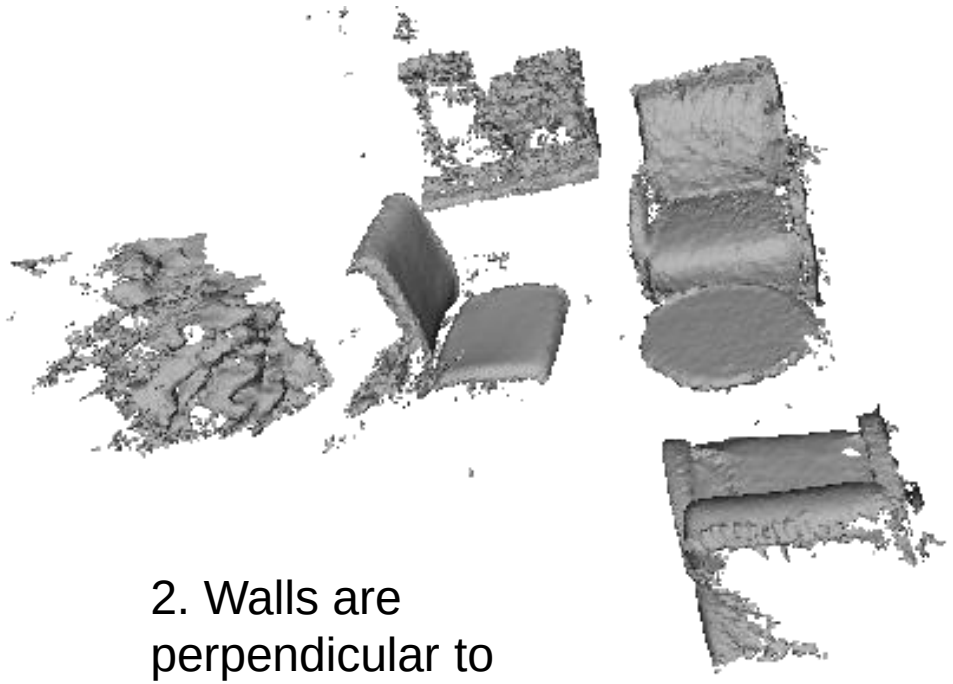
- Merge similar planes:
  - Approximately planar with one another
  - “Close” to one another



Generate Object Candidates

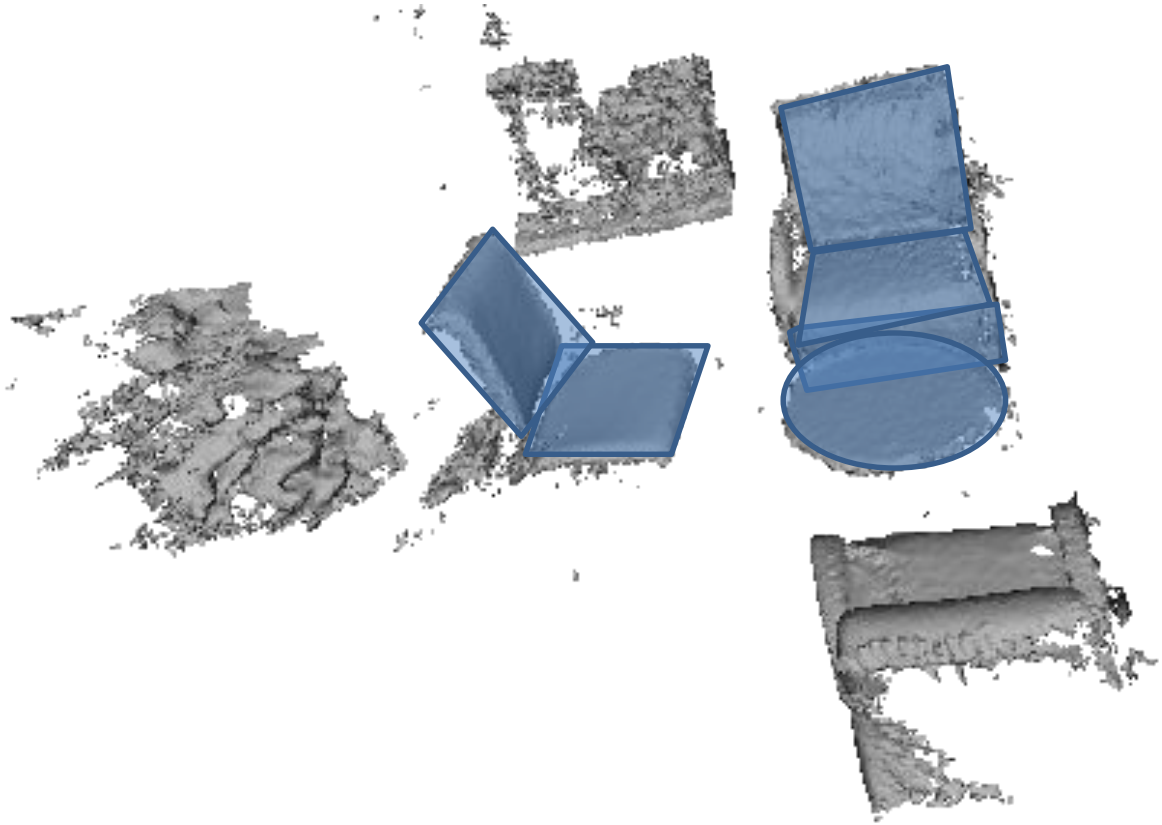


1. Floor plane identified from direction of gravity

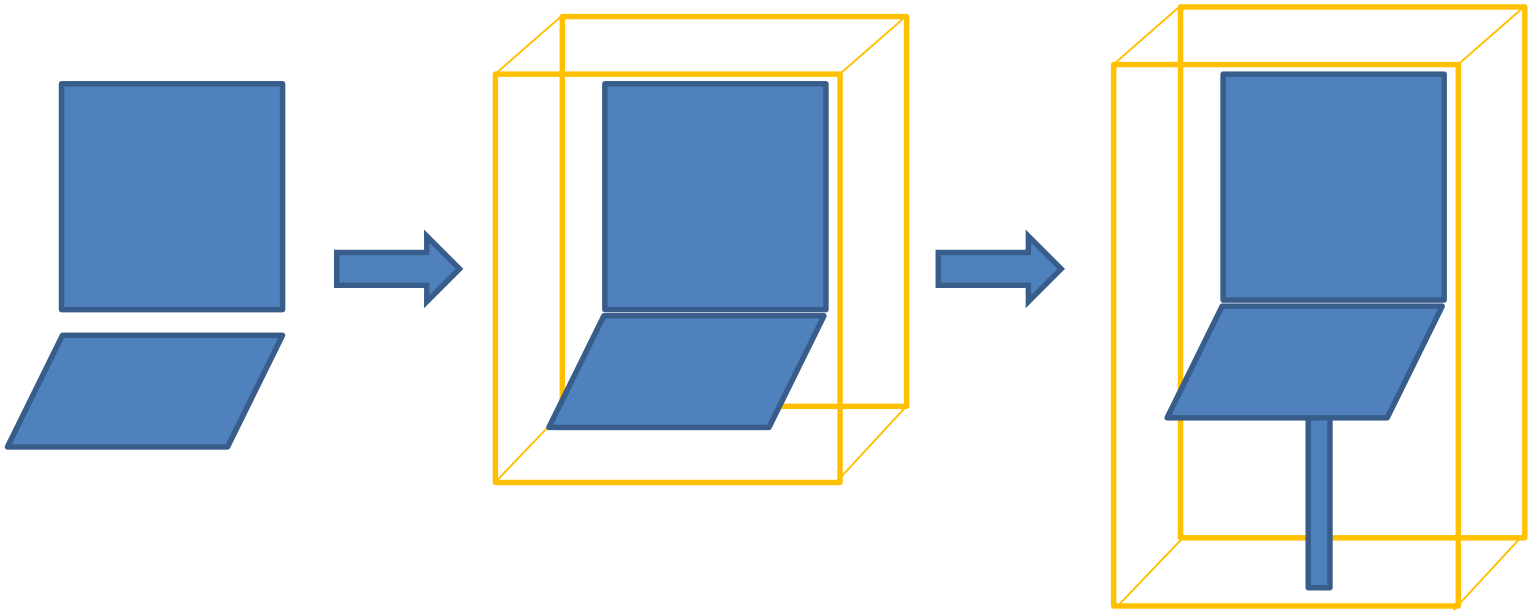


2. Walls are perpendicular to the floor plane

Bound Objects with Plane Primitives

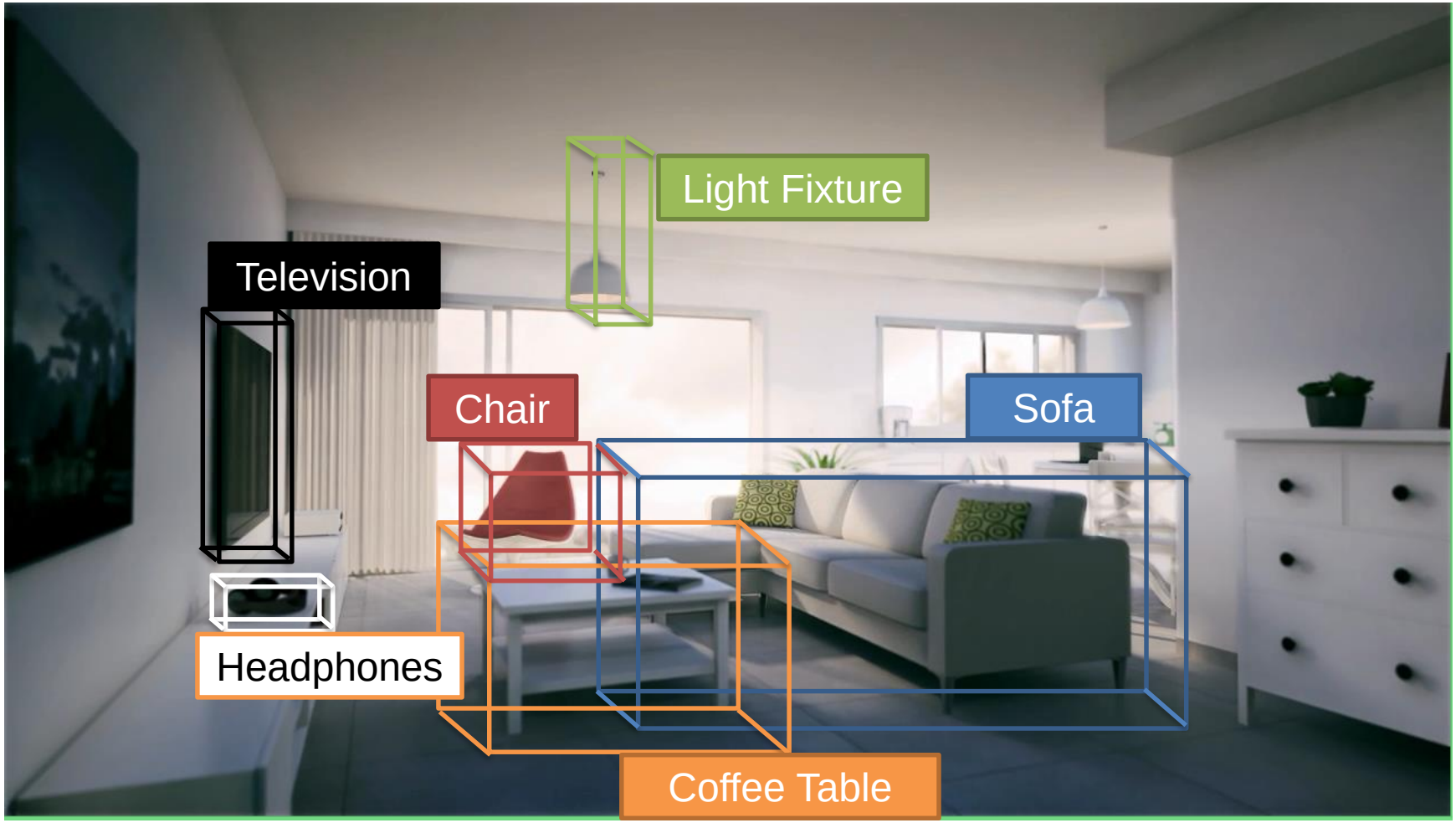


Merge Planes Together to Form Candidates

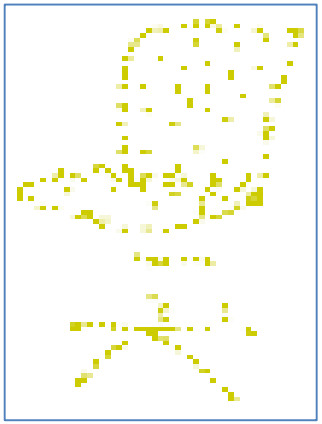


3. 3D Bounding Boxes

4. Object Recognition

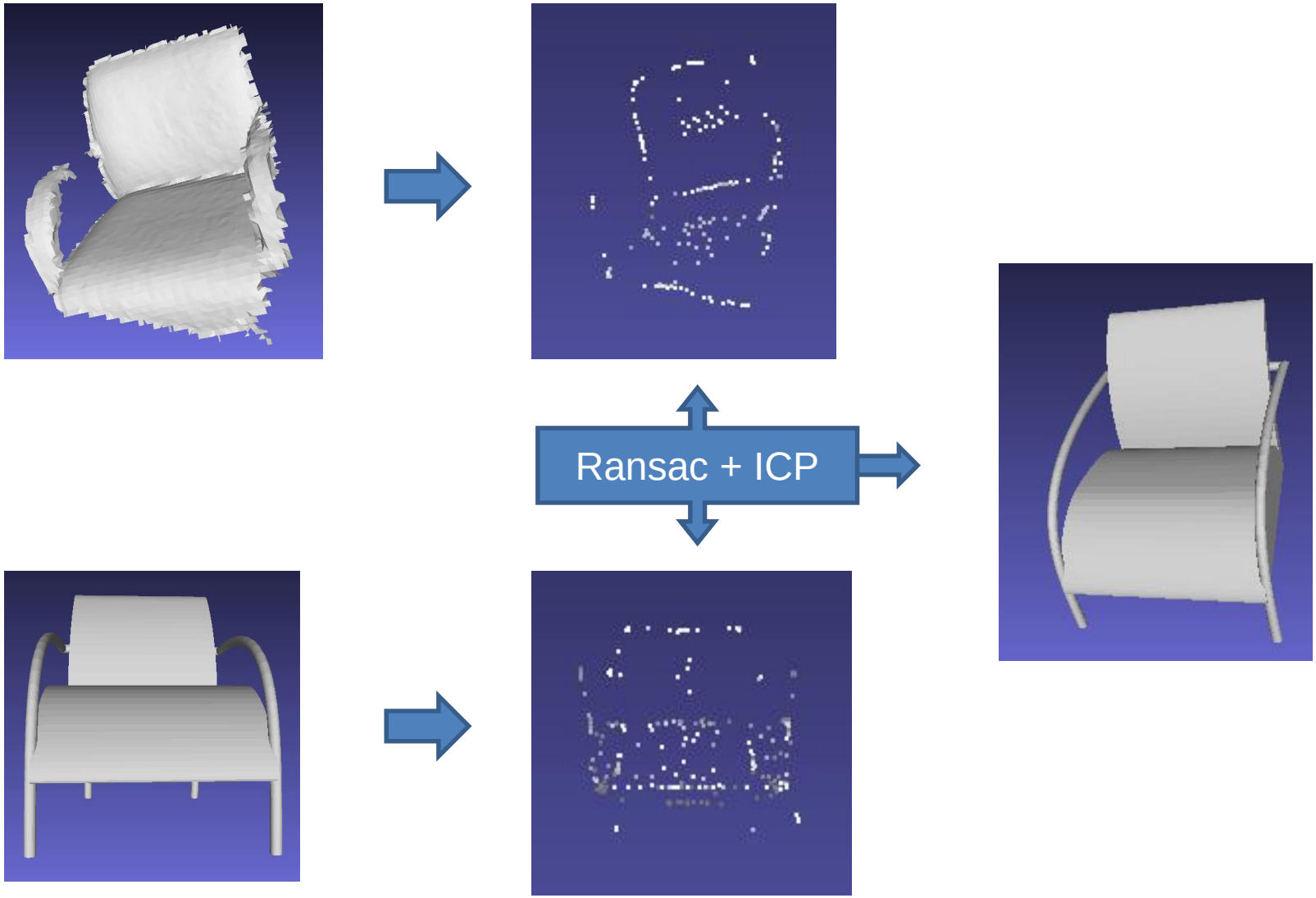


5. Object Retrieval

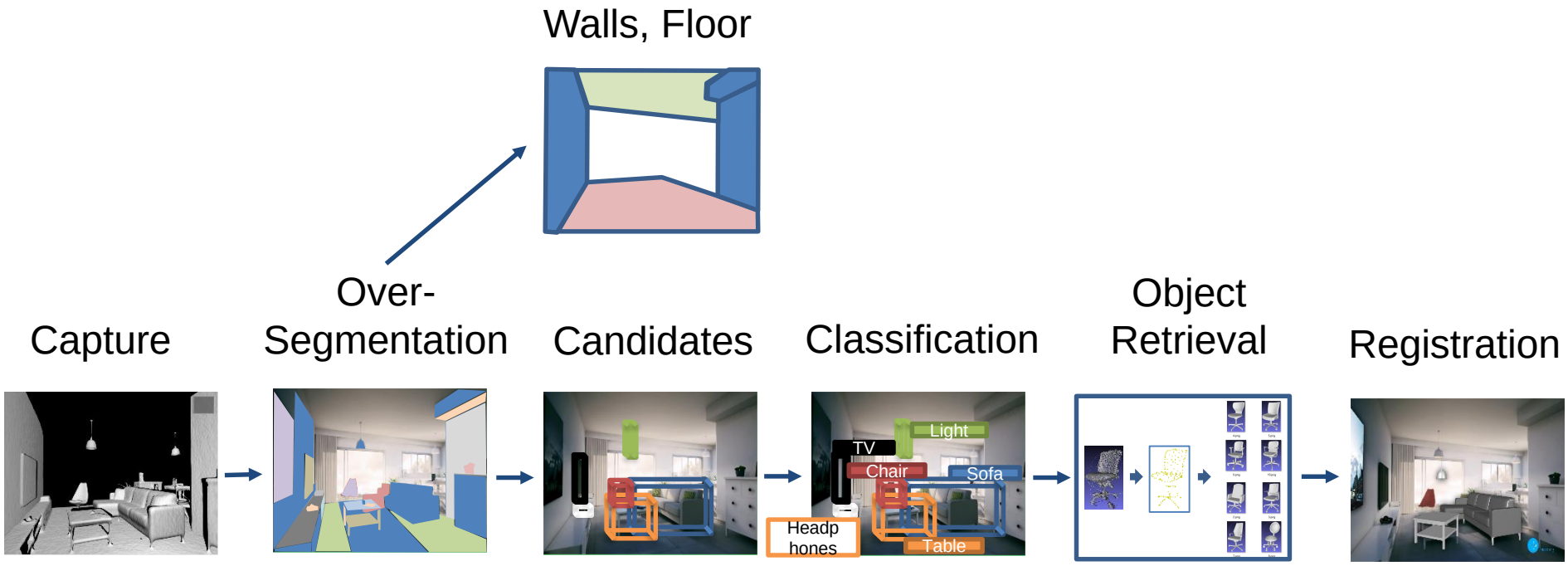


6. 3D Registration

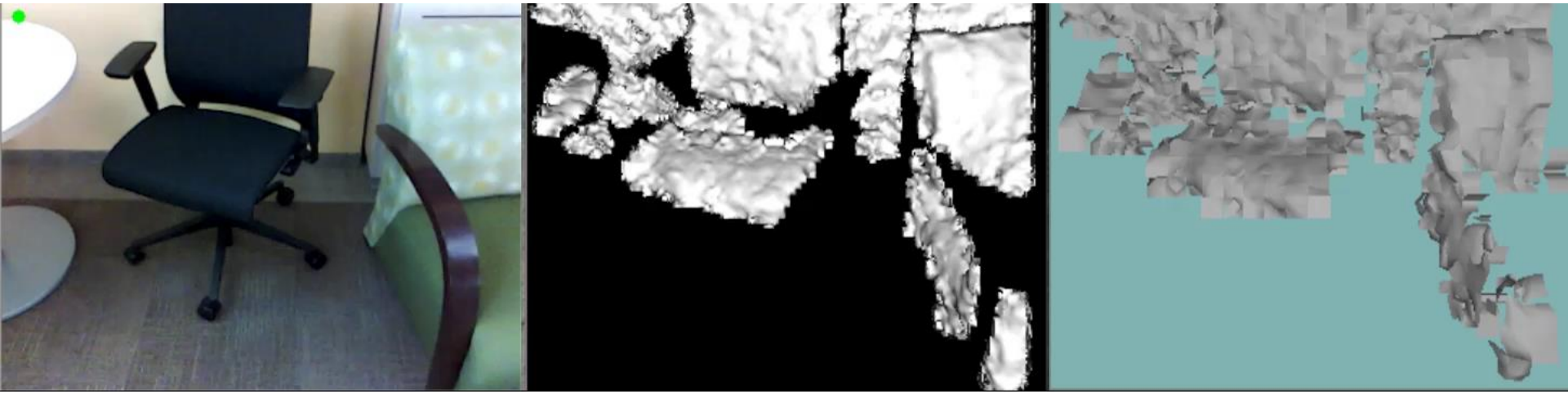
3D Registration



Scene Understanding Pipeline



Move from pixel-level capture to semantic-level



We are hiring …