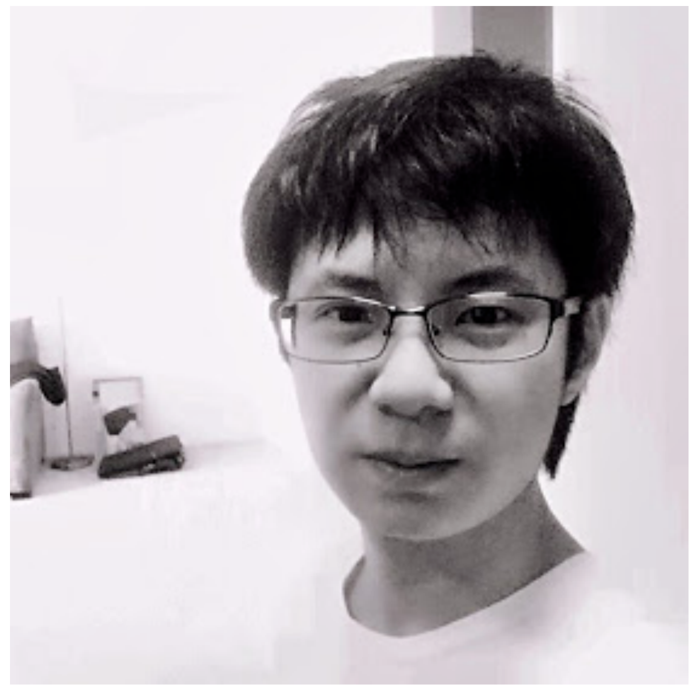
Category-Level Object Pose Estimation

Shuran Song

COLUMBIA | ENGINEERING
The Fu Foundation School of Engineering and Applied Science

Acknowledgment


He Wang


Srinath Sridhar


Jingwei Huang


Julien Valentin

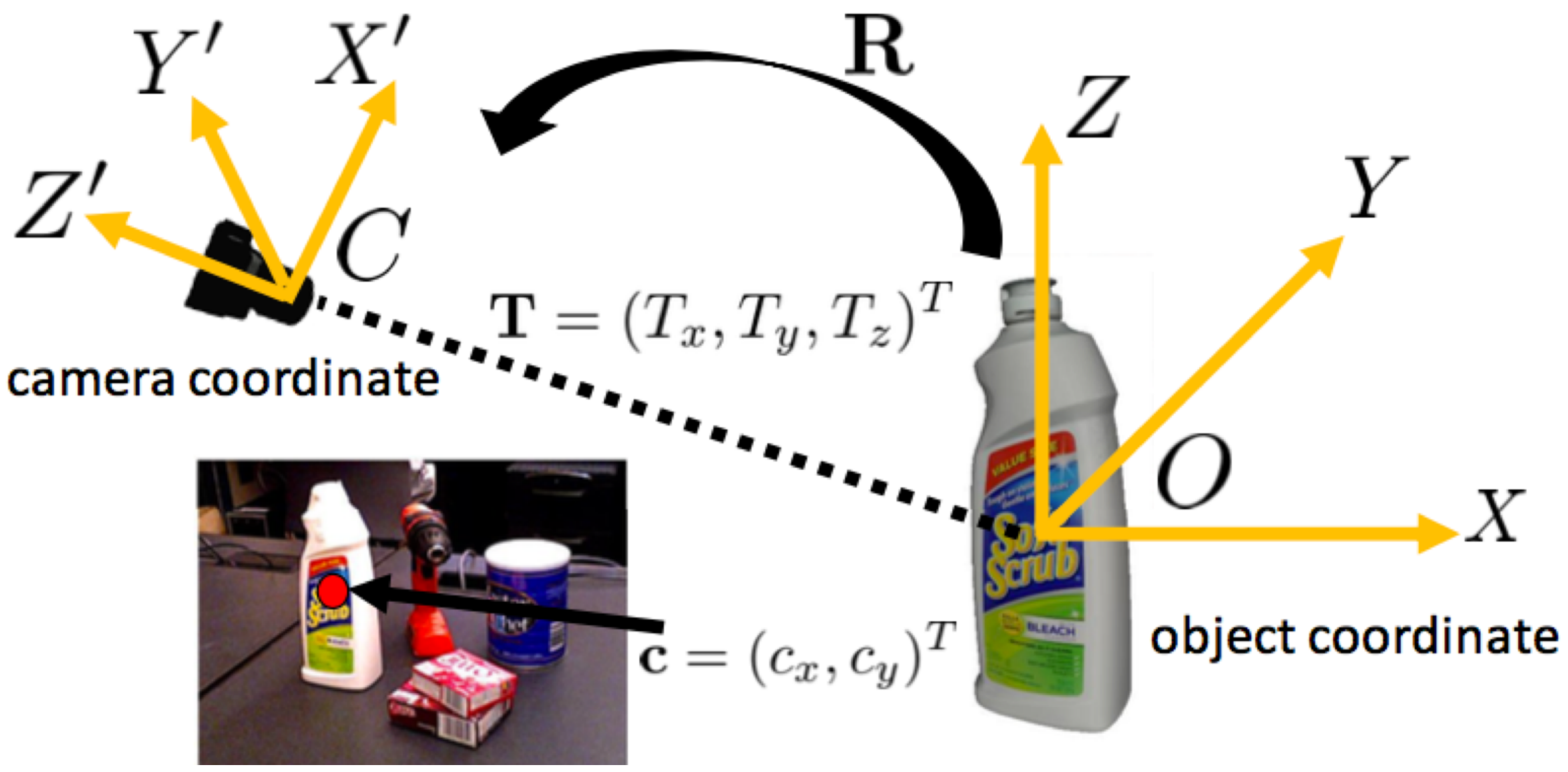

Xiaolong Wang


Yi Li


A. Lynn Abbott


Leonidas Guibas

Instance-level 6DoF Pose Estimation



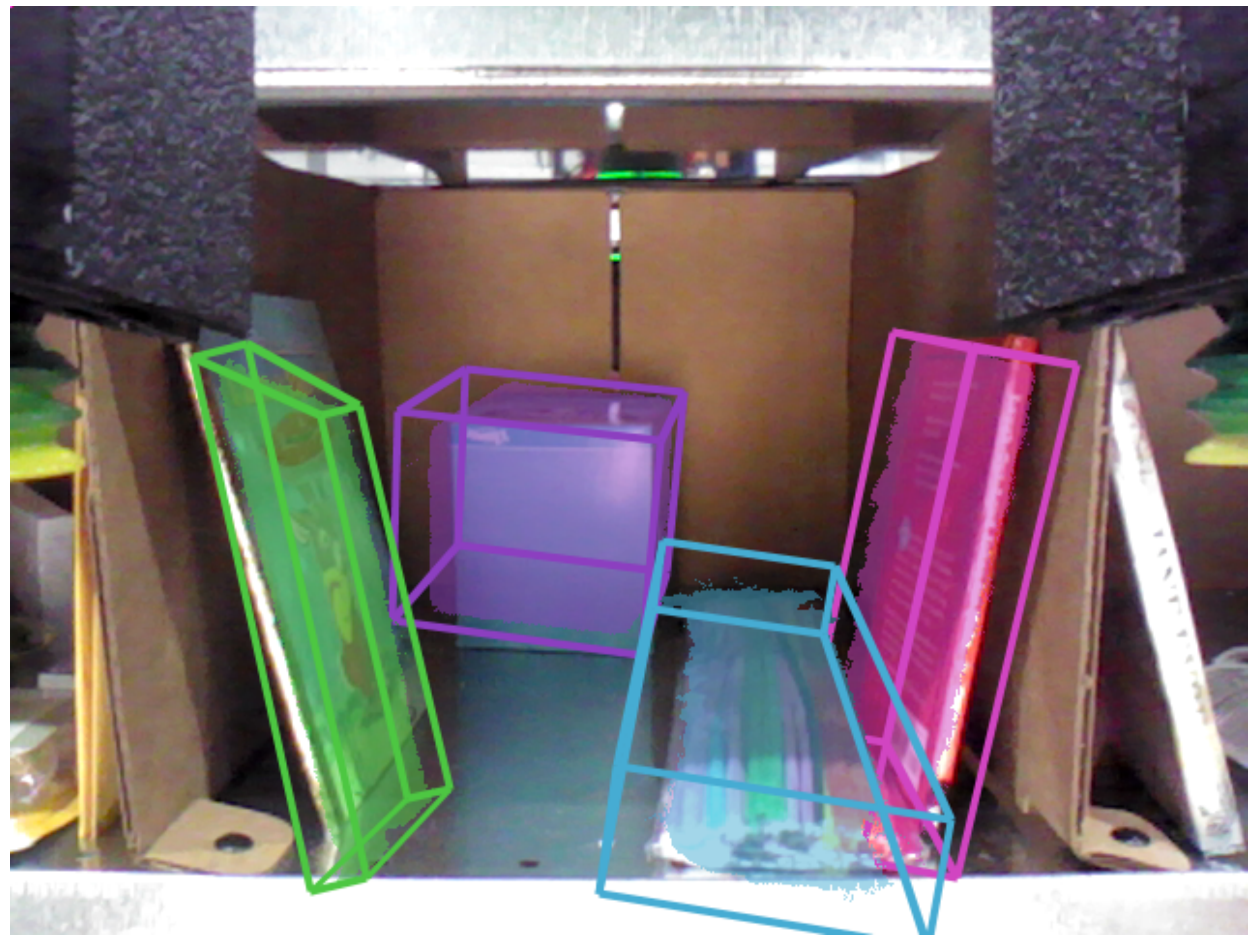
Xiang, et al. 2017

- Instance-level 6DoF pose estimation
  - 3DoF translation
  - 3DoF rotation

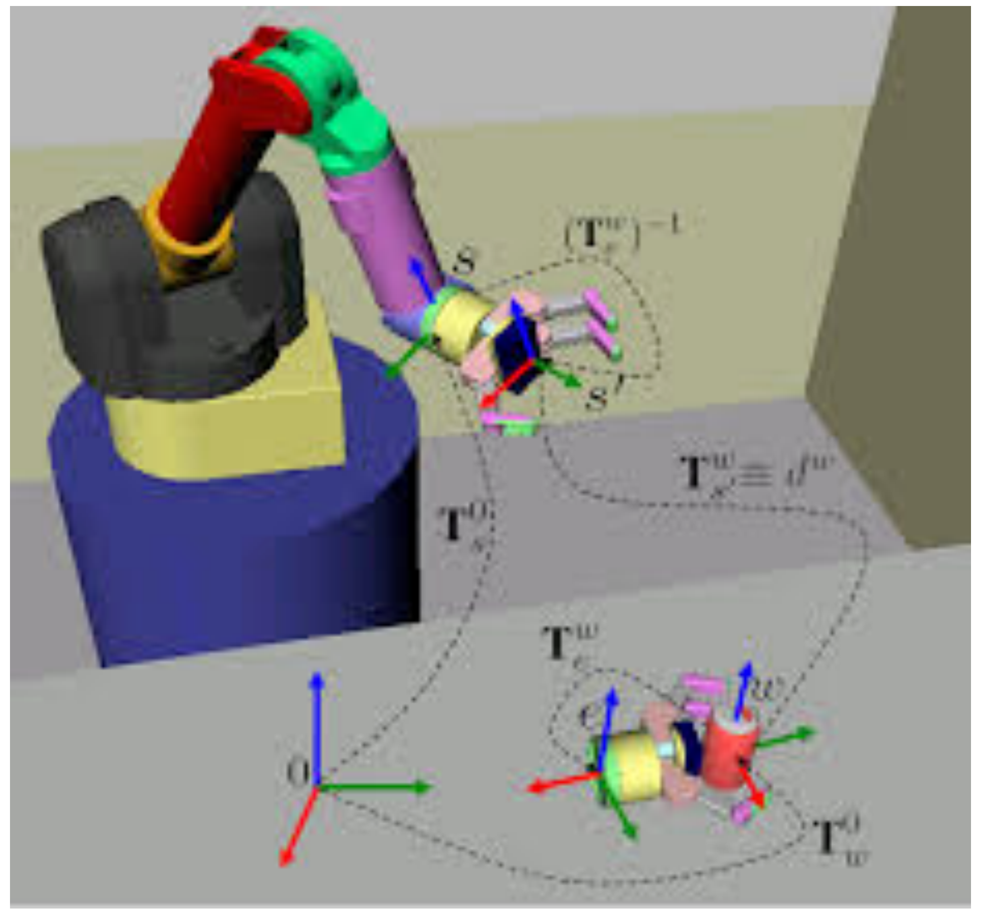
Instance-level 6DoF Pose Estimation

Why it is useful?

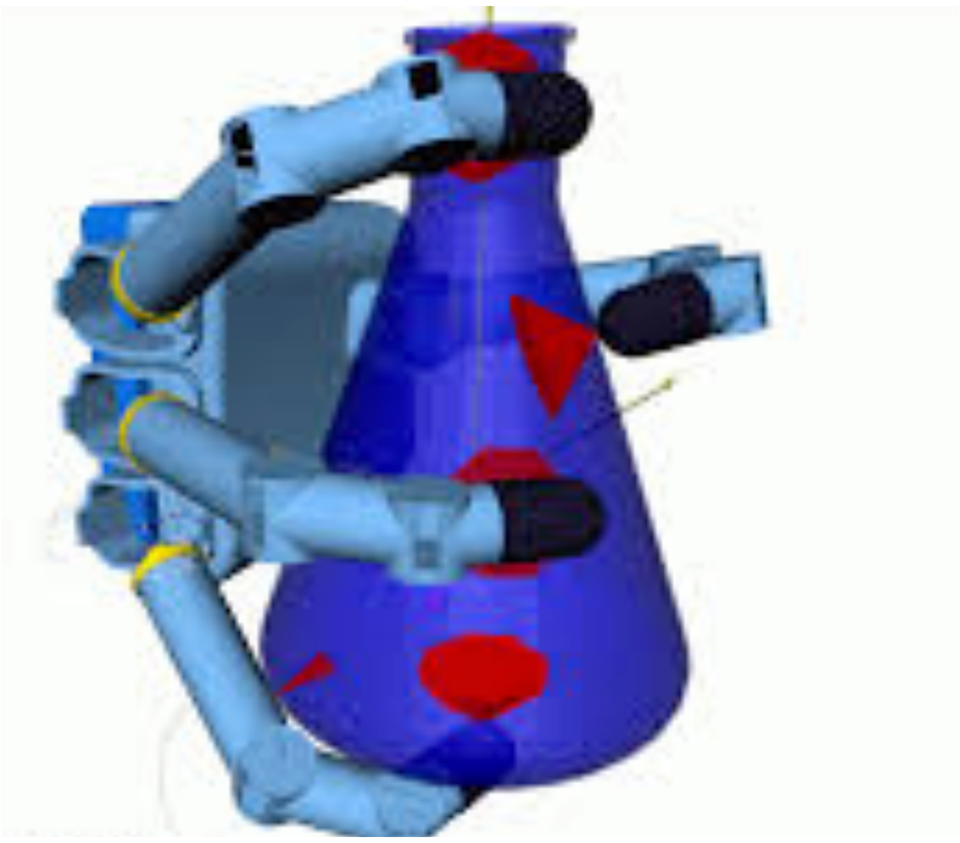
- Concise description of scene and state
- Can be easily used by planning algorithm



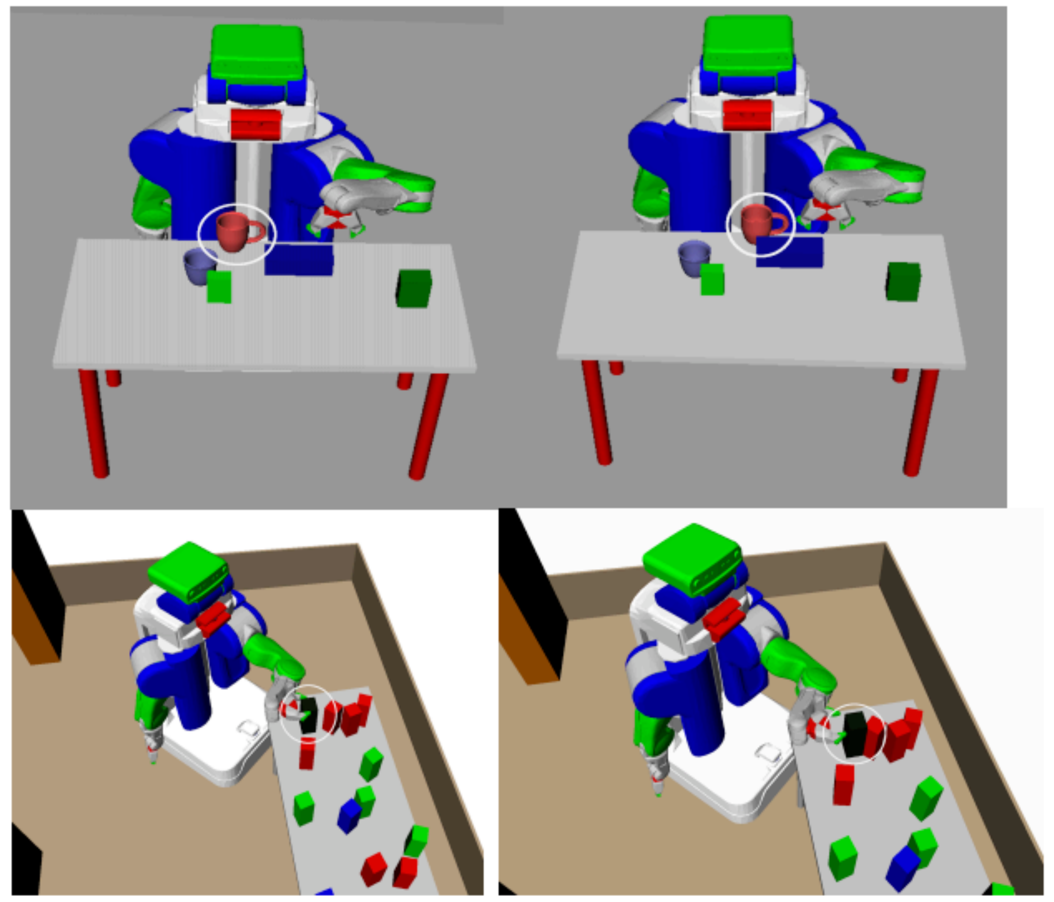
6D Poses Estimation



Berenson et al., 2009



Miller and Allen 2009

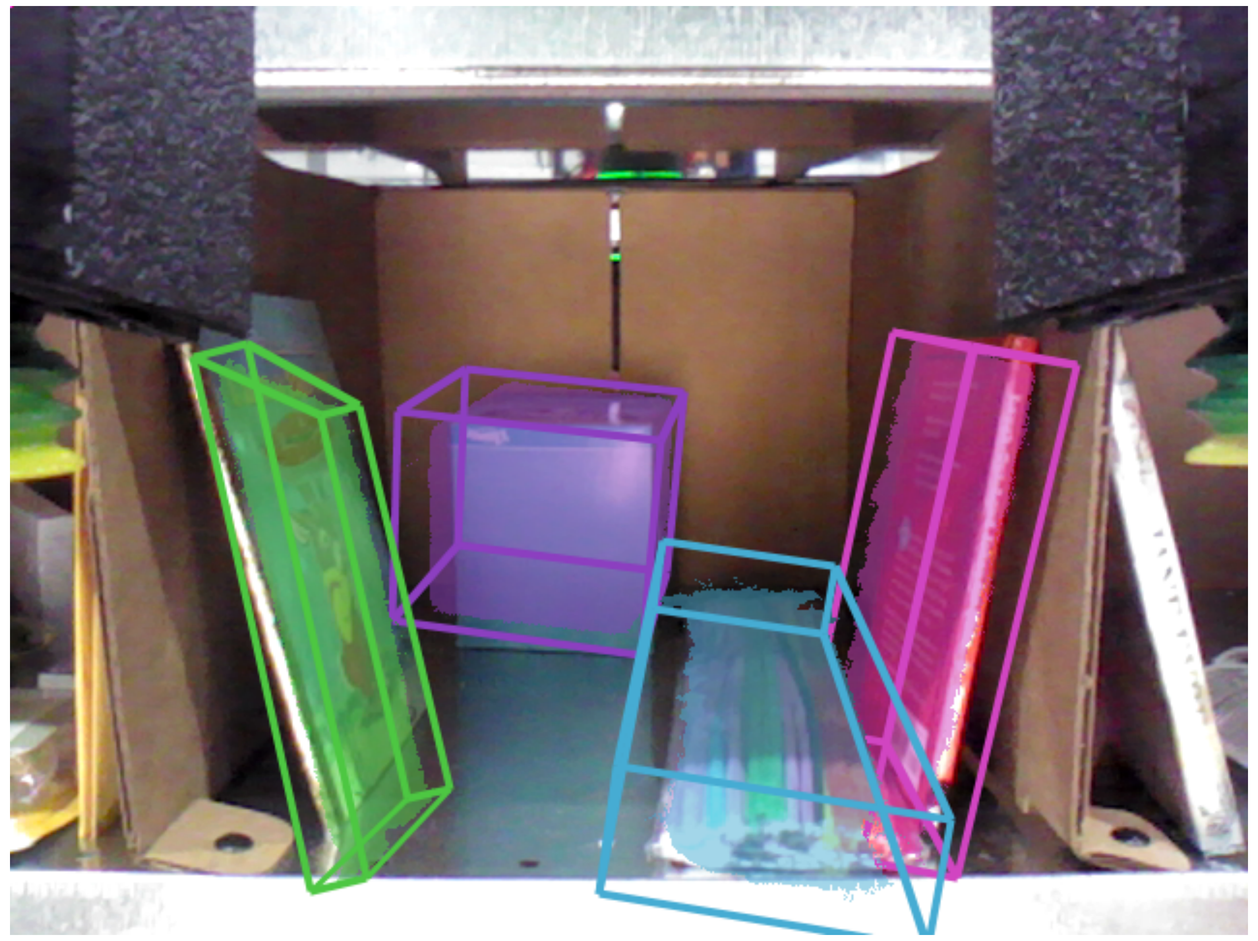


Kimm et al 2019

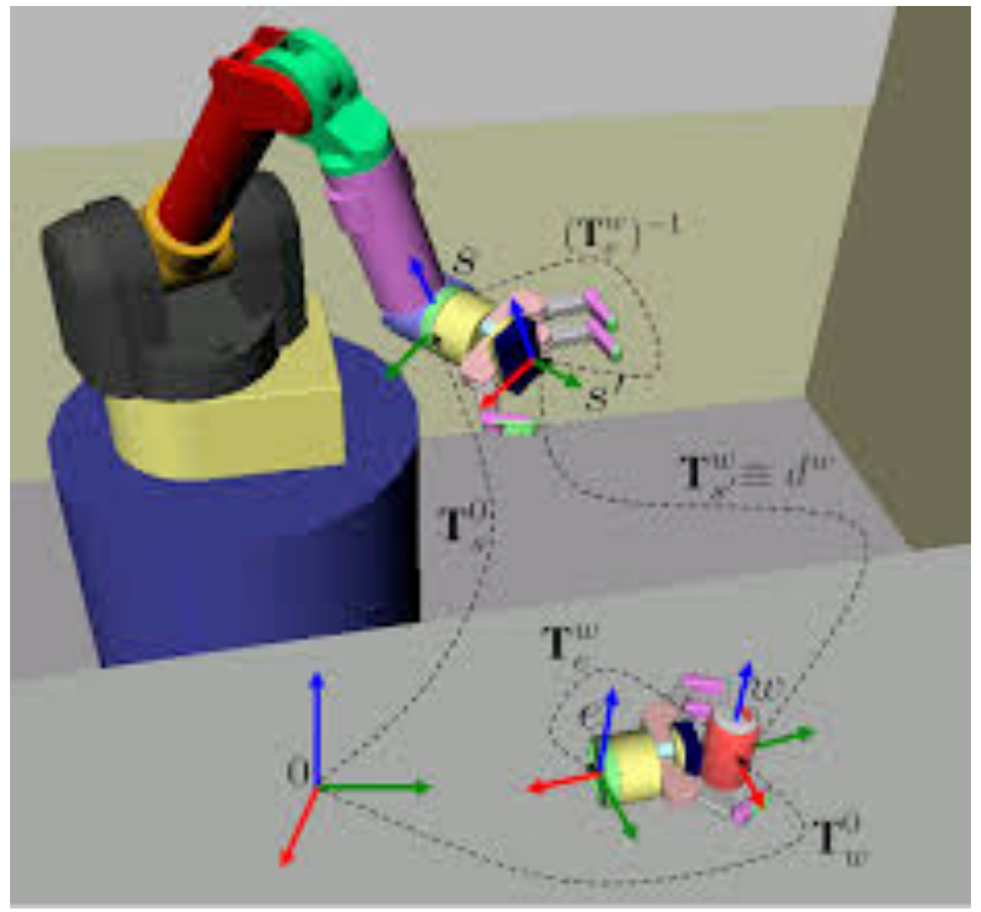
Instance-level 6DoF Pose Estimation

Why it is useful?

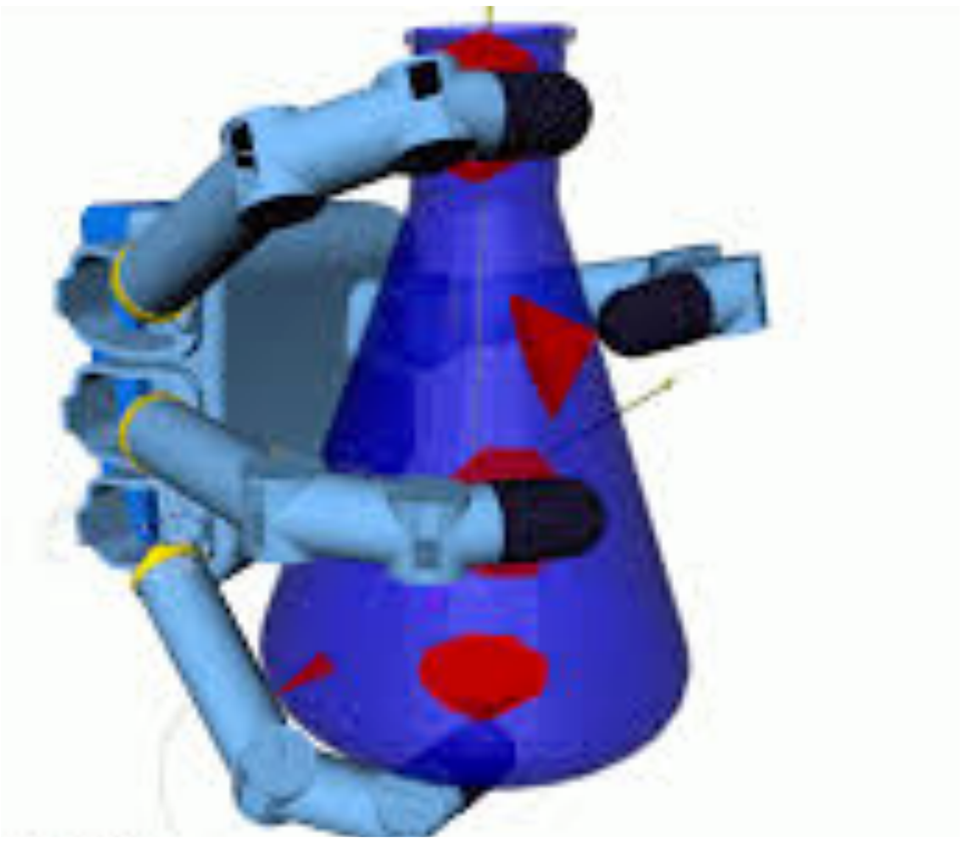
- Concise description of scene and state
- Can be easily used by planning algorithm



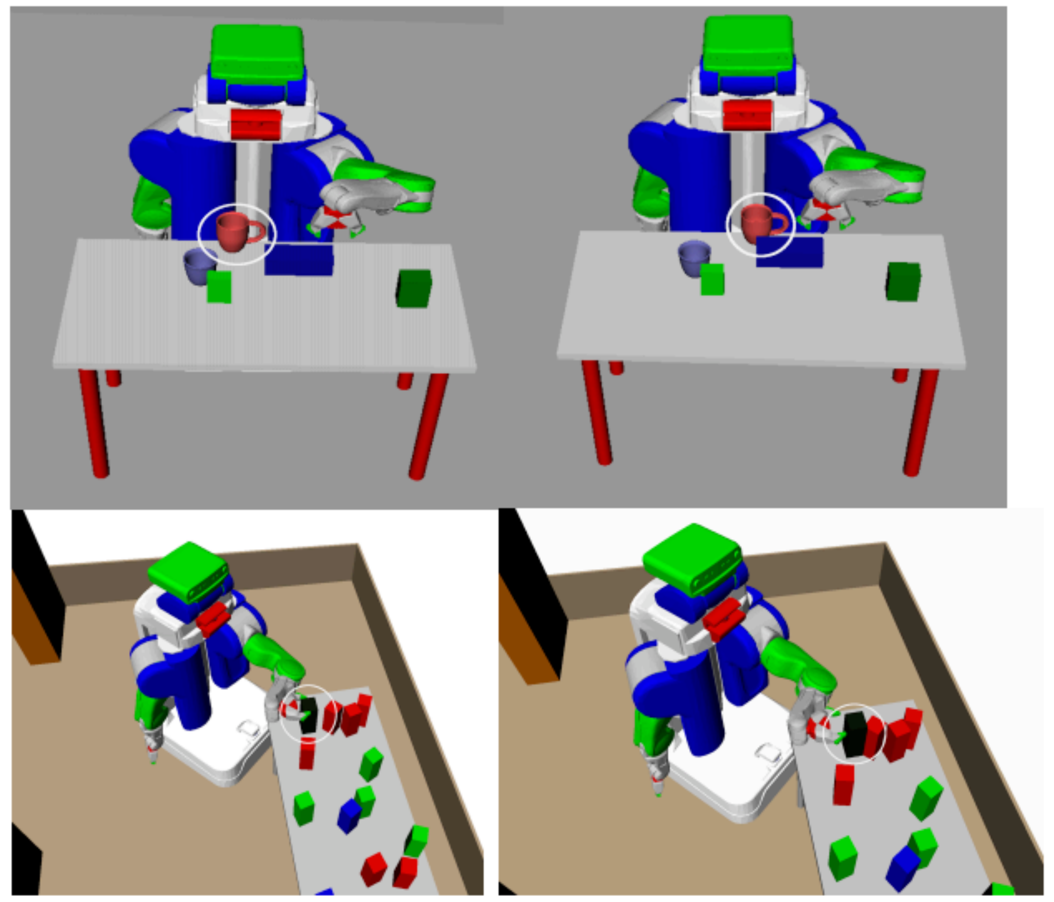
6D Poses Estimation



Berenson et al., 2009



Miller and Allen 2009



Kimm et al 2019

Good for structured known settings with rigid objects

Instance-level 6DoF Pose Estimation

Why it is limiting?

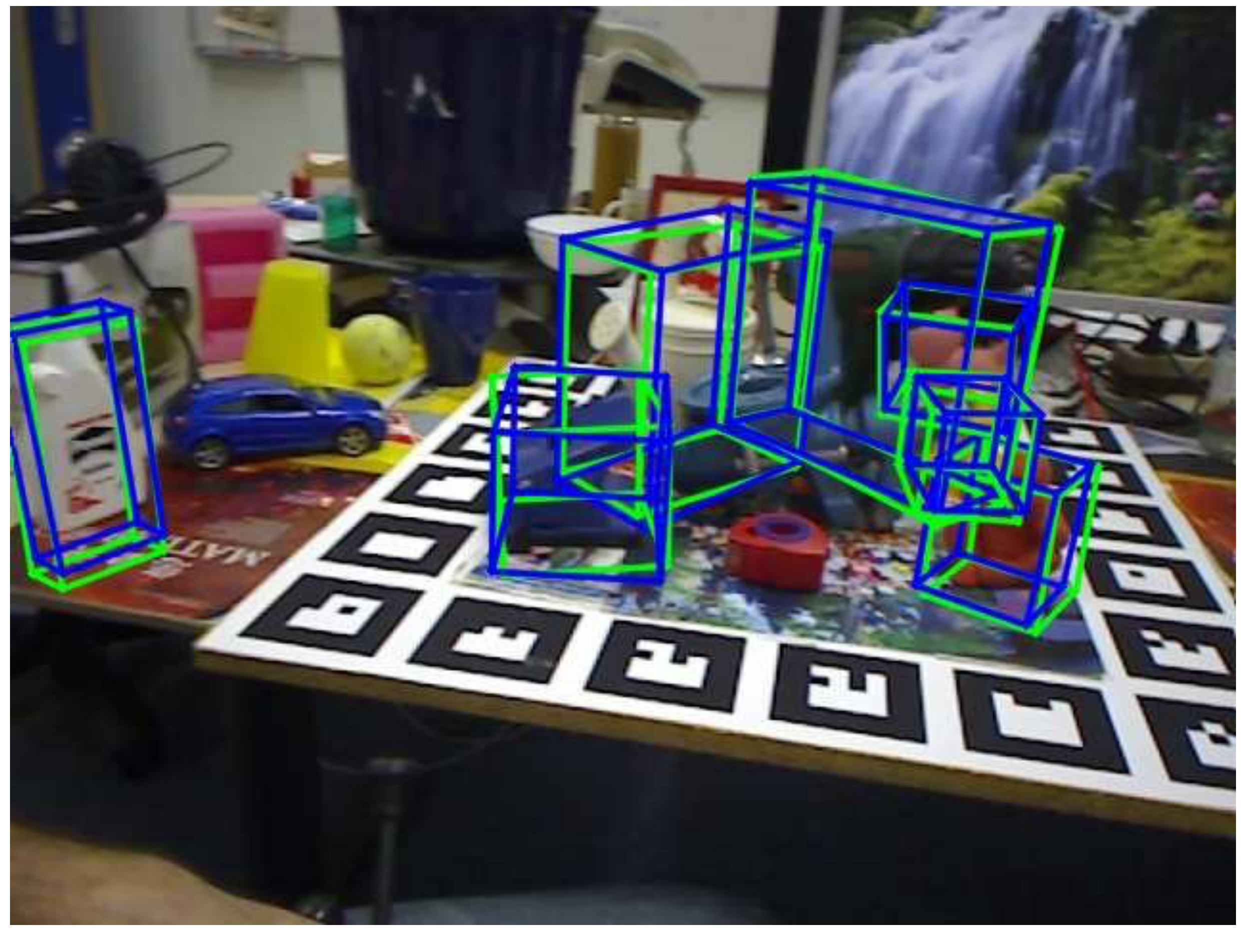
- Need 3D models of the exact objects
- Cannot generalize to <u>unseen</u> objects instances

Instance-level 6DoF Pose Estimation

Why it is limiting?

- Need 3D models of the exact objects
- Cannot generalize to unseen objects instances

- Handles a small collection of known instances.



LINEMOD dataset

YCB dataset

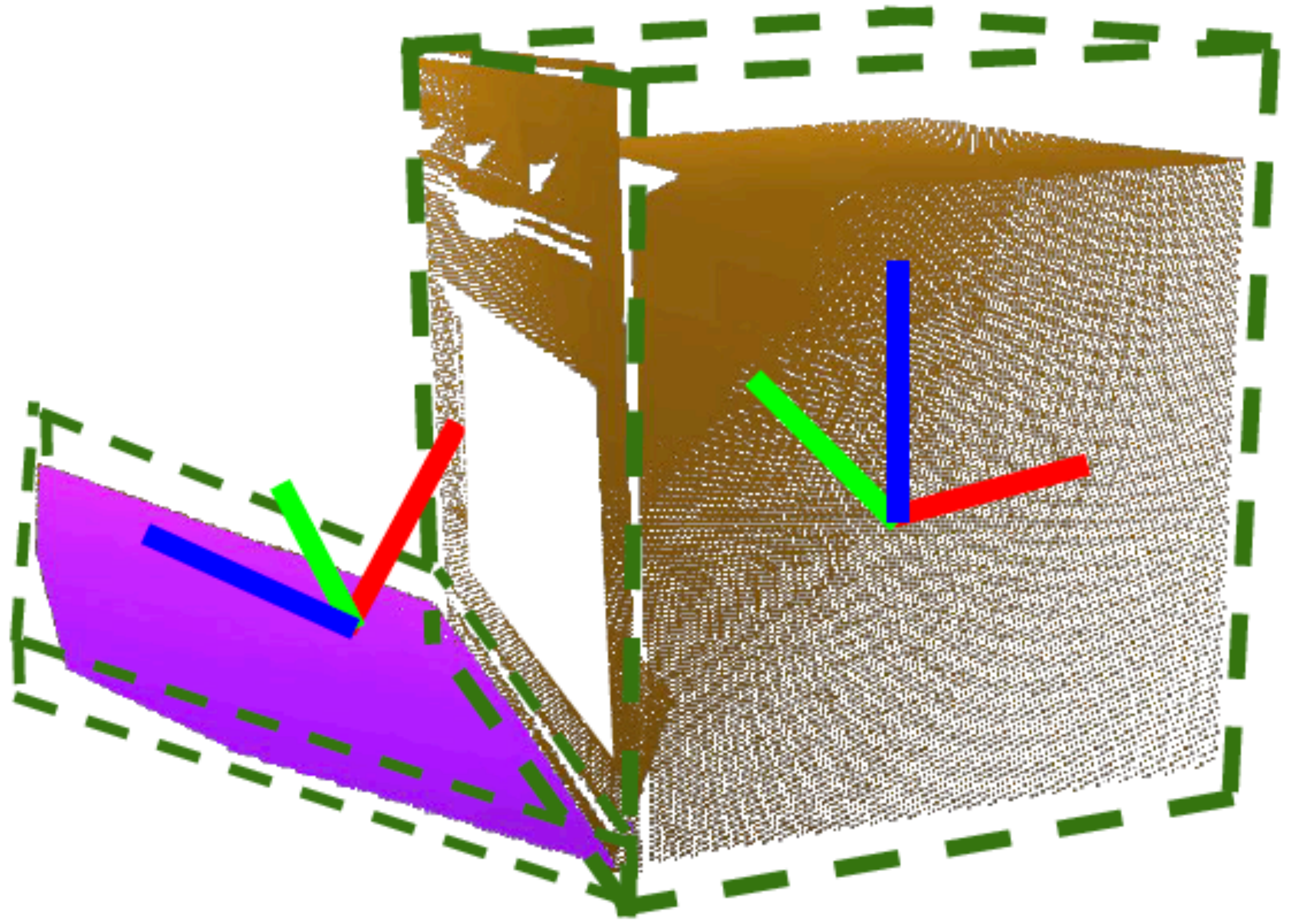
Today's Talk

Category-level pose estimation that enables algorithm to generalize to unseen object instance by leveraging category prior.

Today’s Talk

Category-level pose estimation for

Rigid Object



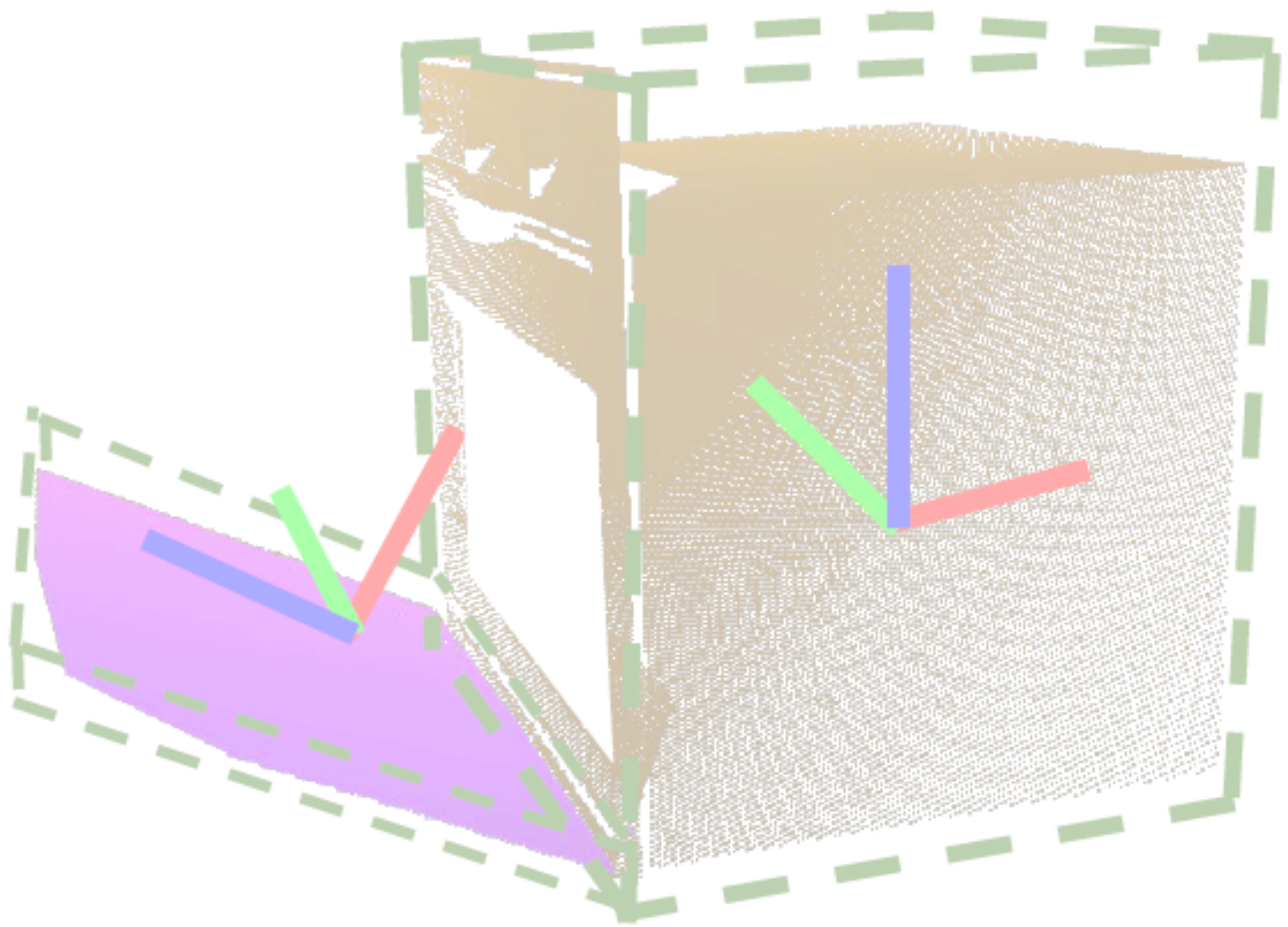
Articulated Object

What’s next?

Today's Talk

Category-level pose estimation for

Rigid Object



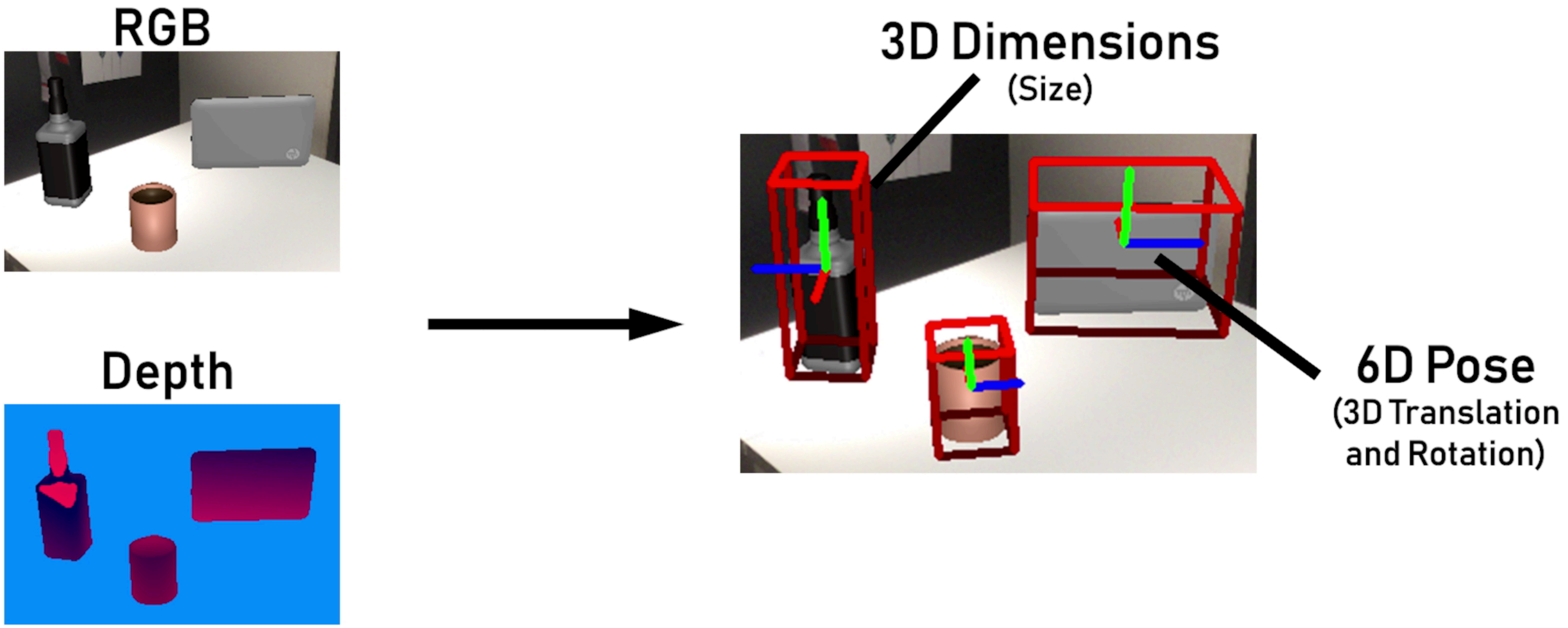
Articulated Object

What's next?

Normalized Object Coordinate Space for Category-Level 6D Object Pose and Size Estimation
He Wang, Srinath Sridhar, Jingwei Huang, Julien Valentin, Shuran Song, Leonidas Guibas CVPR'19

Problem Formulation

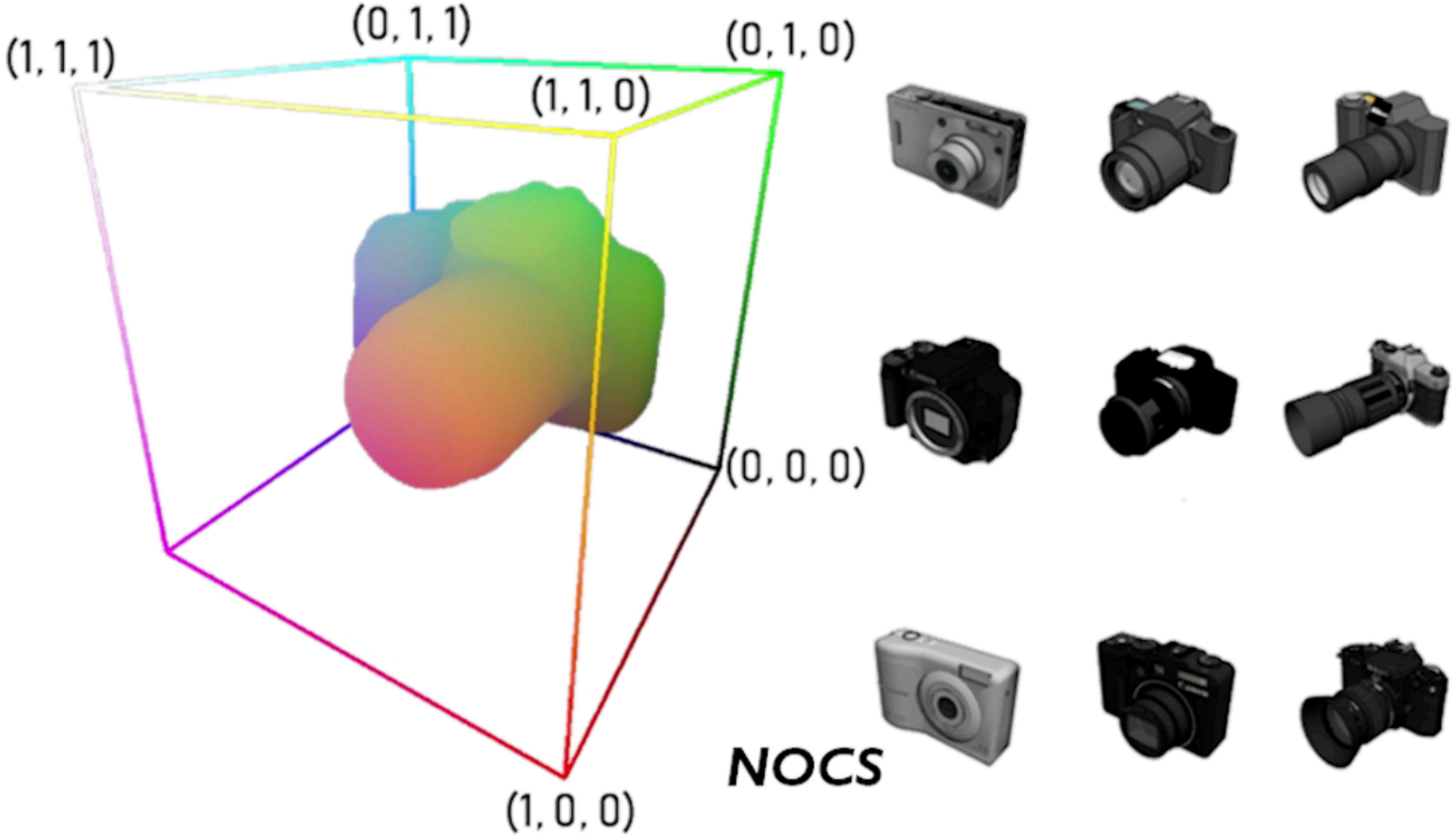
Task: detecting and estimating **6D pose** and **3D size** of **novel** objects in certain categories from RGBD images.



Challenges

- Defining and representing category-level object poses
- Intra-category shape variations
- Training data collection

Normalized Object Coordinate Space (NOCS)



Normalized Object Coordinate Space (NOCS)

Category-level object orientation can be defined for each category up to the limit of global symmetry in the category.

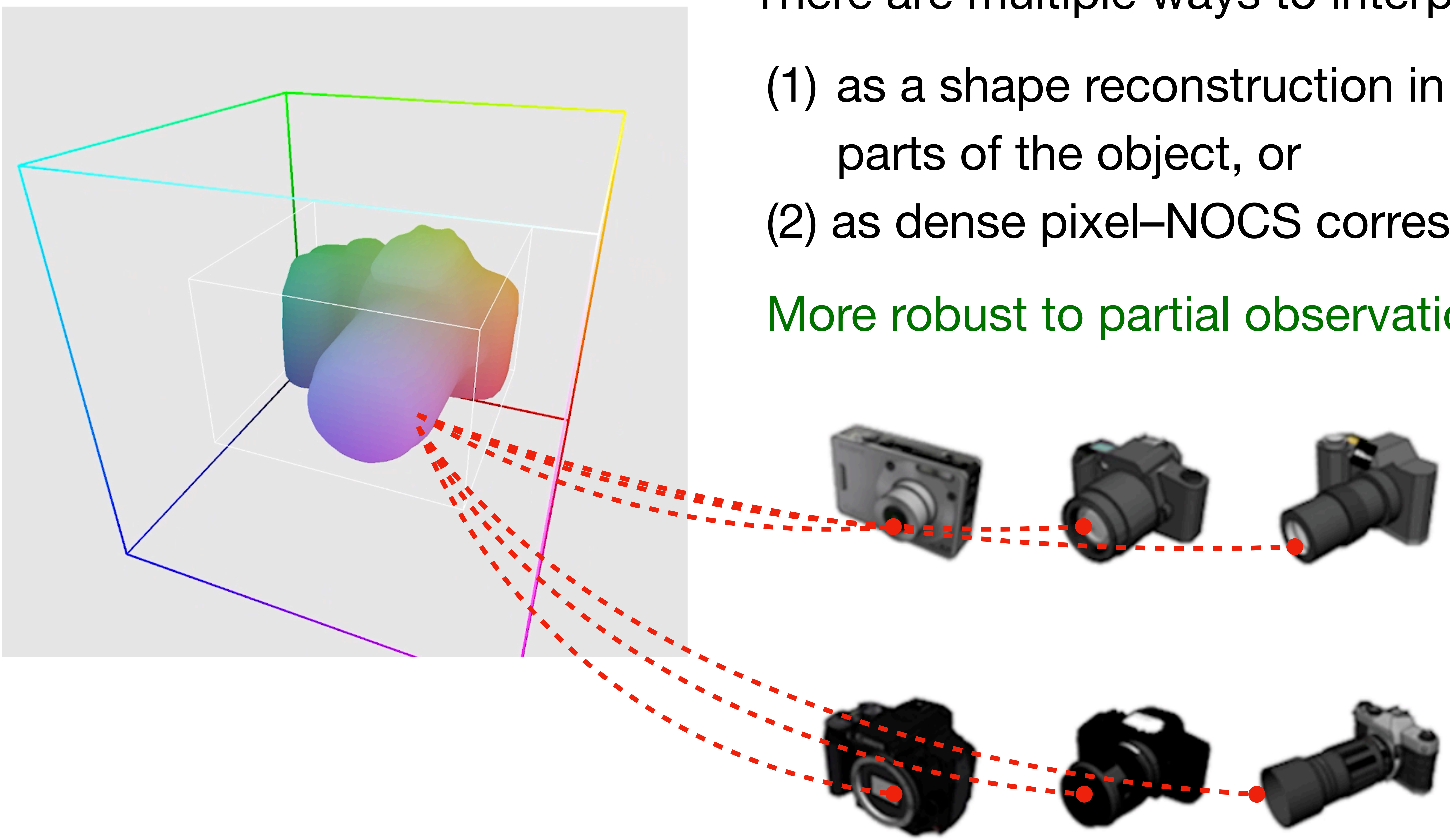
Normalized Object Coordinate Space (NOCS)

There are multiple ways to interpret a NOCS map:

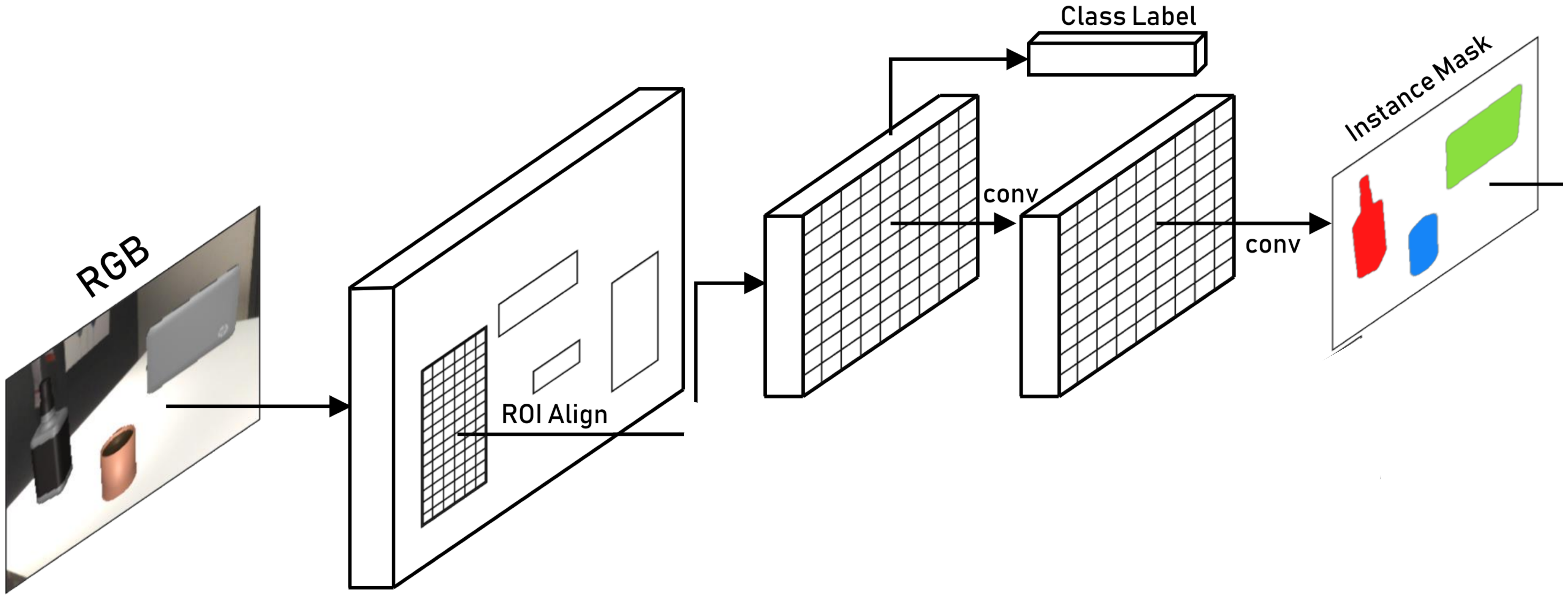
(1) as a shape reconstruction in NOCS of the observed parts of the object, or

(2) as dense pixel–NOCS correspondences.

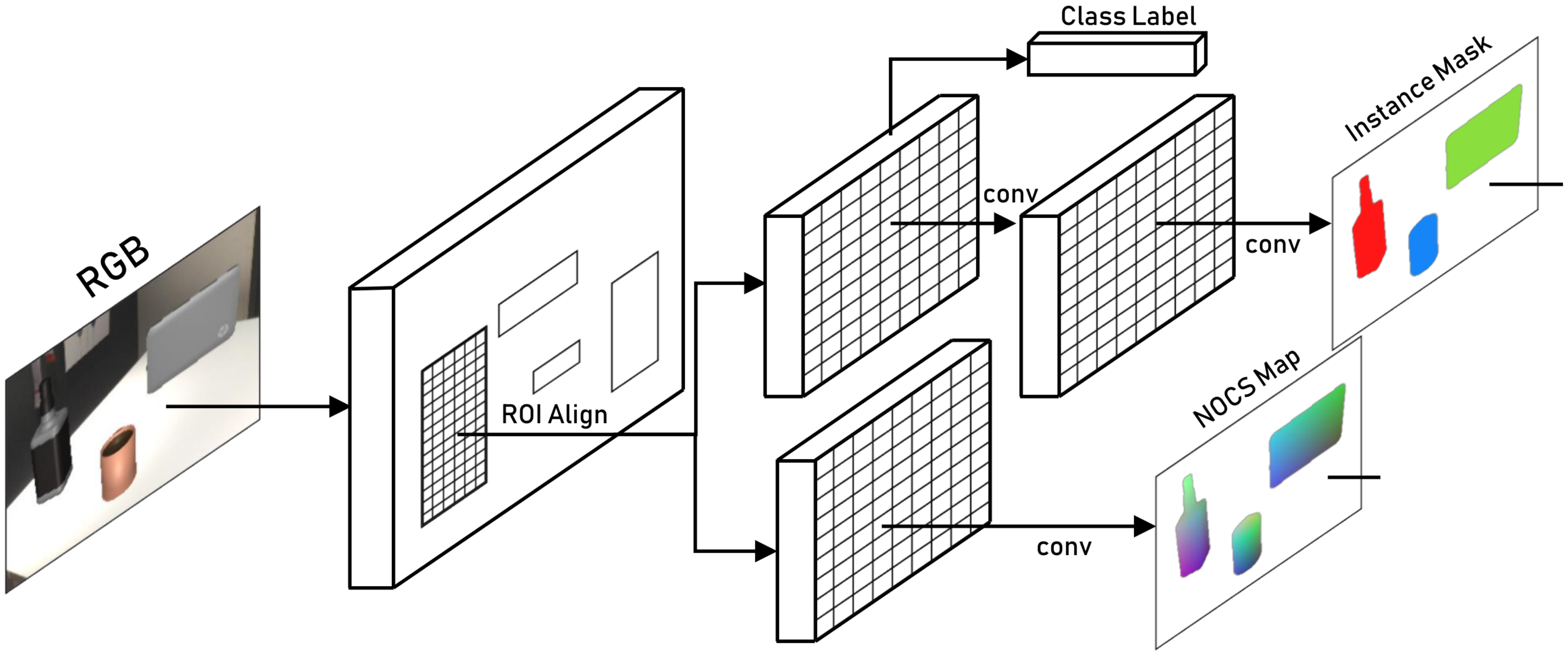
More robust to partial observation



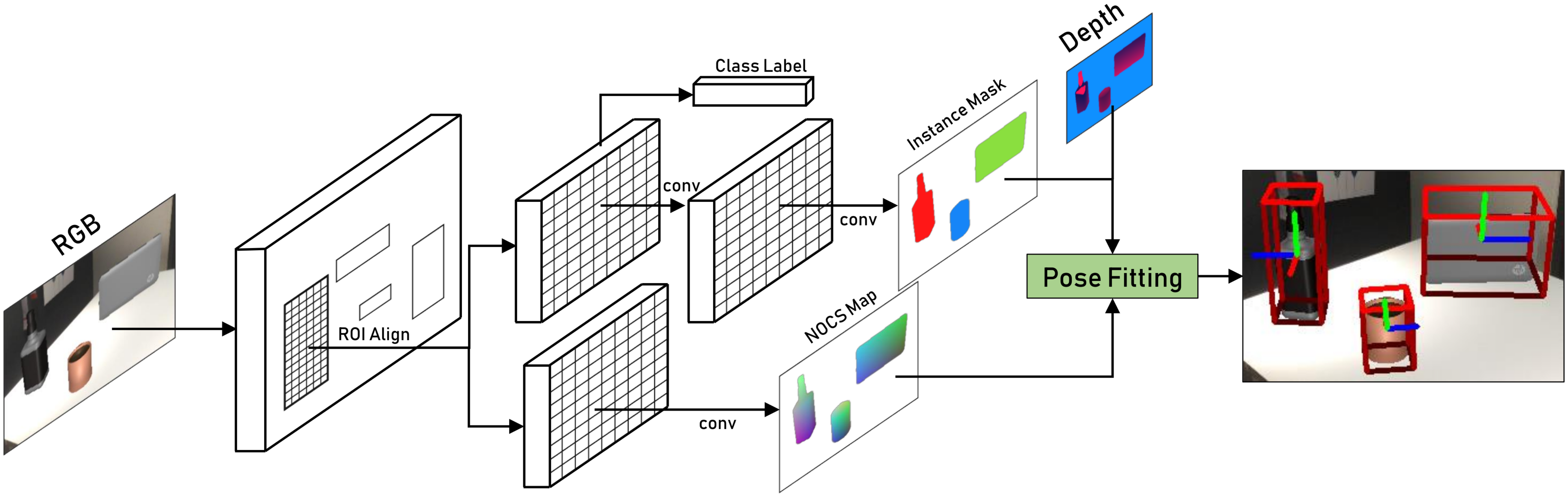
Method Overview



Method Overview



Method Overview

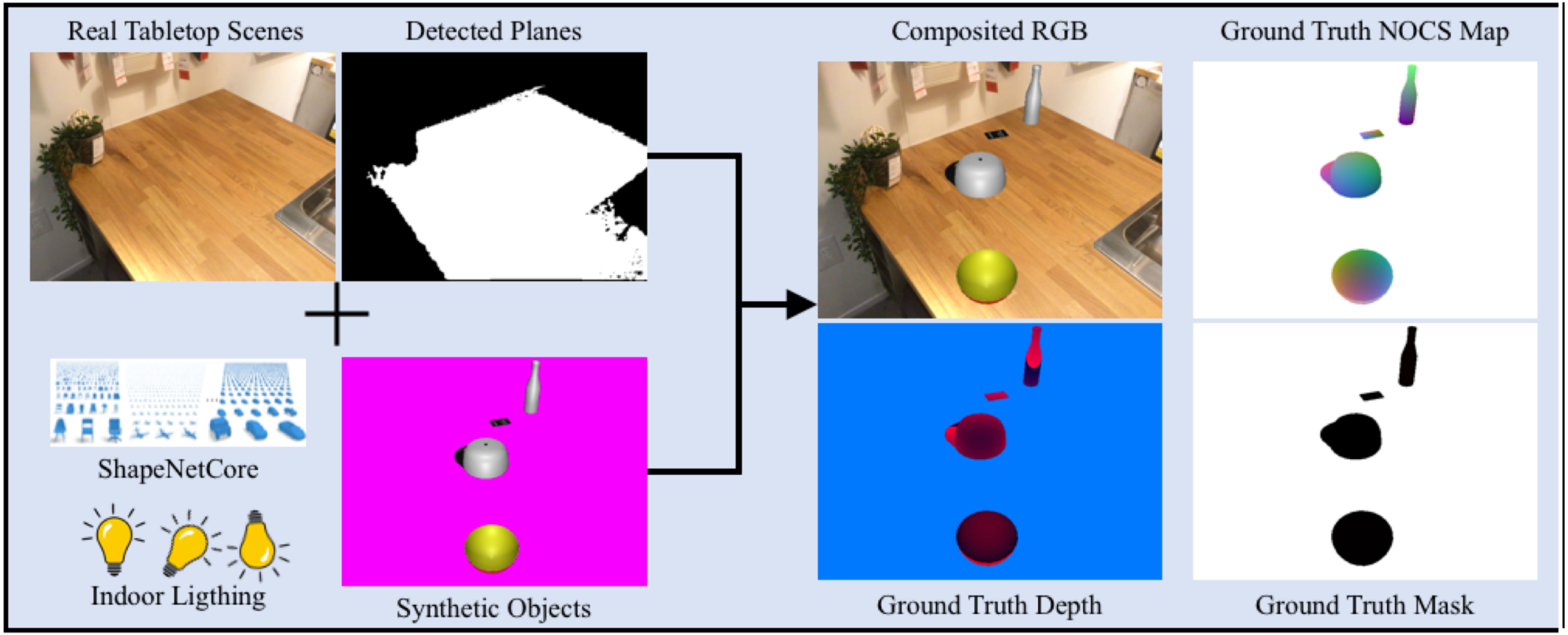


Handling Symmetry

Symmetry-aware loss function: $L_s = \min_{i=1,\ldots,|\theta|} L\left(\tilde{\mathbf{y}}_i, \mathbf{y}^*\right)$

min-of-N loss

Context-Aware MixEd ReAlity (CAMERA)

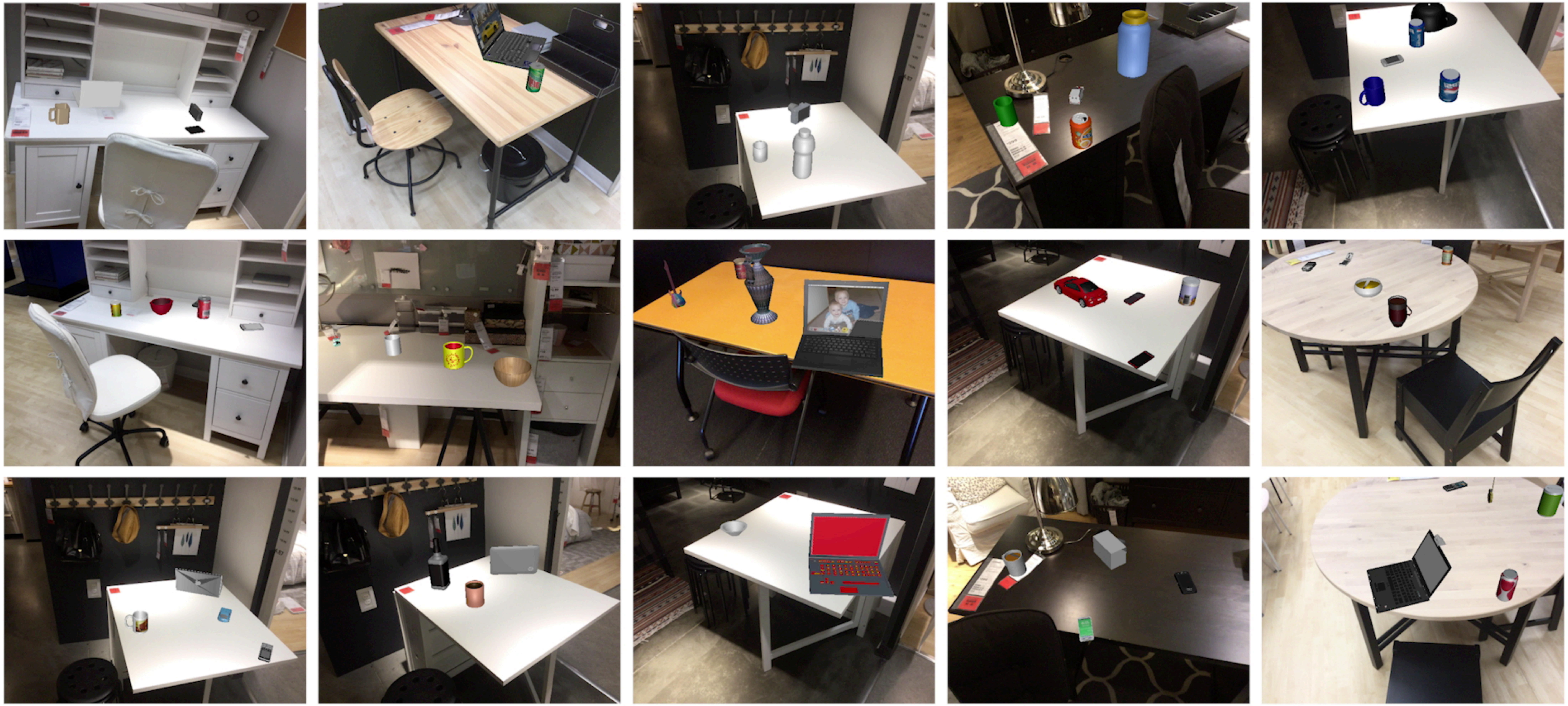


Mixed-reality data generation pipeline

Dataset Statistics

- **Context-Aware MixEd ReAlity**
  - 300K images (275K/25K)
  - 31 real-world tabletop scenes
  - 1085 ShapeNet models

Mixed Reality Data



Real-World Dataset

* markers not used for prediction

Dataset Statistics

- **Real Dataset (Train/Val/Test)**
  - 8K RGB-D (4K/0.95K/3.75K)
  - 18 real scenes (6/6/6)
  - 42 unique instances (3/1/3)
- **6 Object Categories**
  - bottle, bowl, camera, laptop, mug
  - symmetric vs. asymmetric

Results on Real Test Data

Training Data:

- 275K CAMERA
- 20K COCO (without NOCS)
- 4.3K real images

Performance:

3D IoU @ 50%:	76.4%
5°, 5cm:	10.2%
10°, 10cm:	23.1%

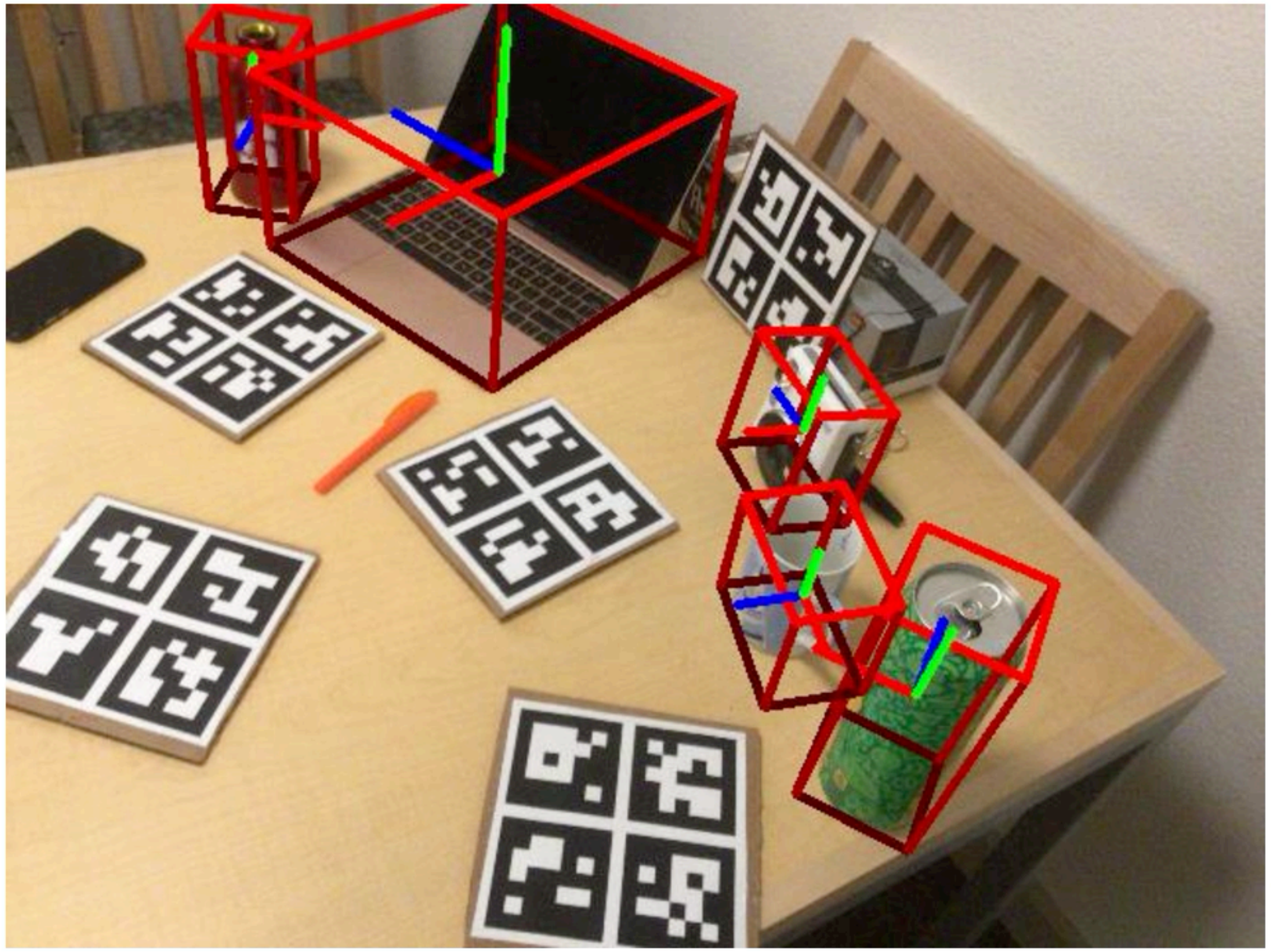
CAMERA data

Real data

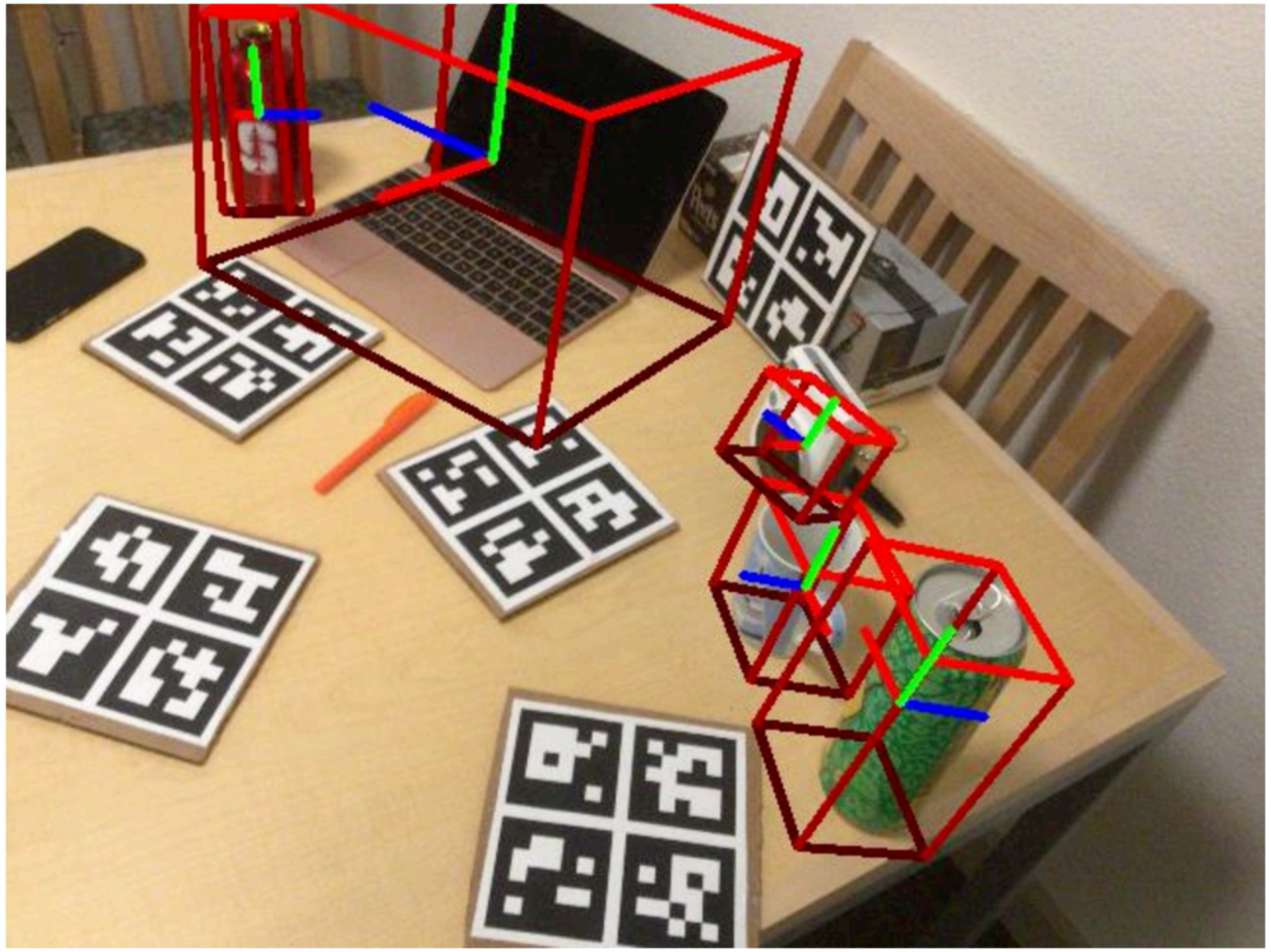
MS COCO data

Result Visualization

Ground Truth



Prediction (Ours)



Recap

Category-level pose estimation for

Rigid Object

Normalized object coordinate space (NOCS)

defines a shared space with consistent object scaling and orientation.

Recap

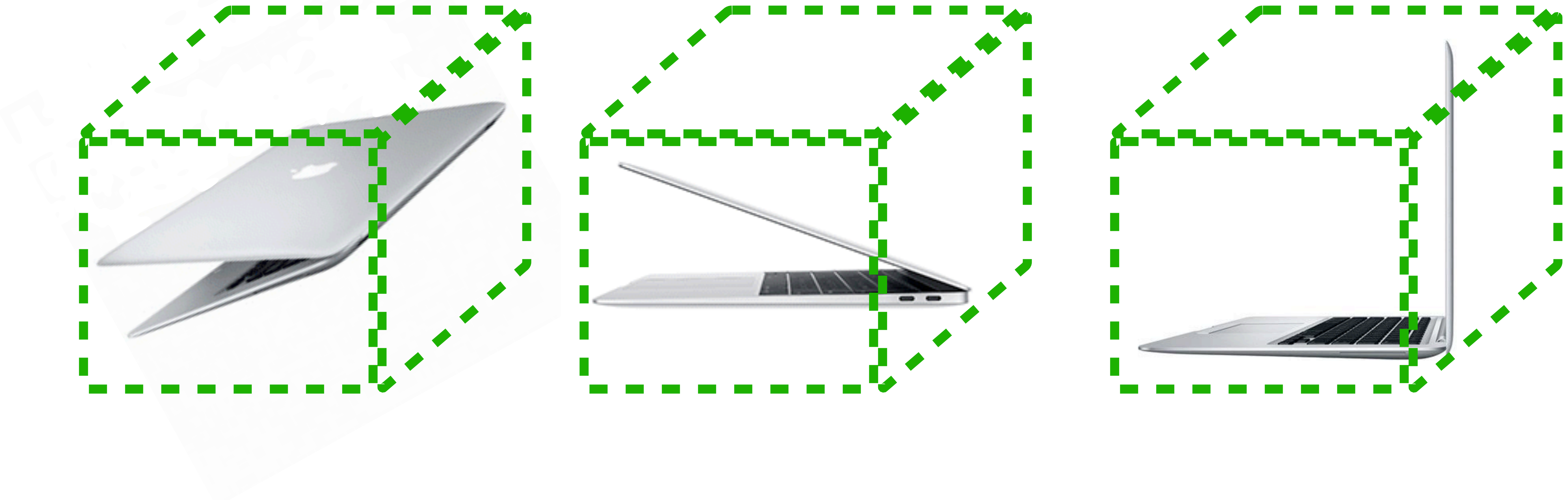
Category-level pose estimation for

Rigid Object

Normalized object coordinate space (NOCS)

defines a shared space with consistent object scaling and orientation.

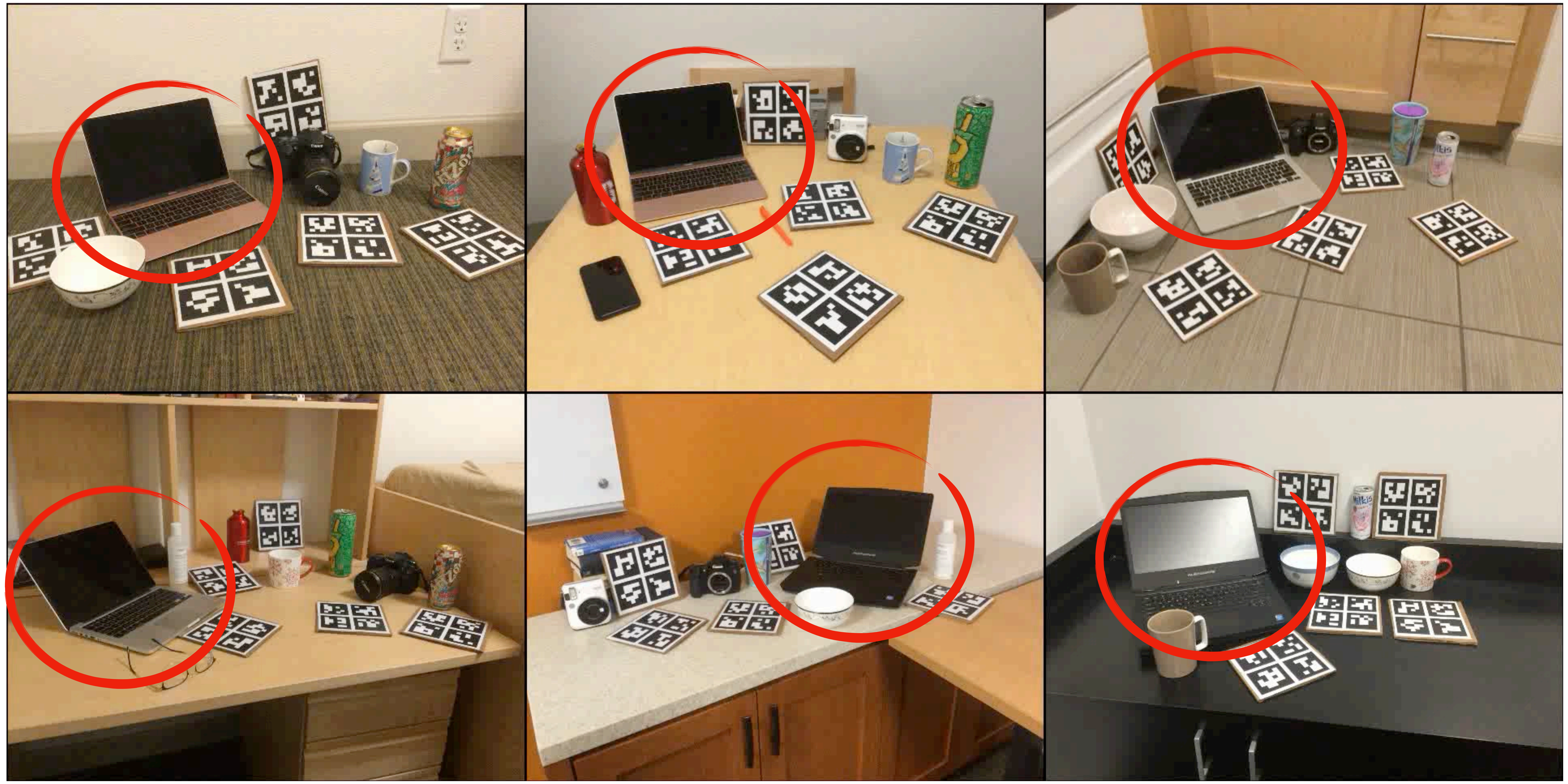
x Rigid object Assumption



Recap

Category-level pose estimation for

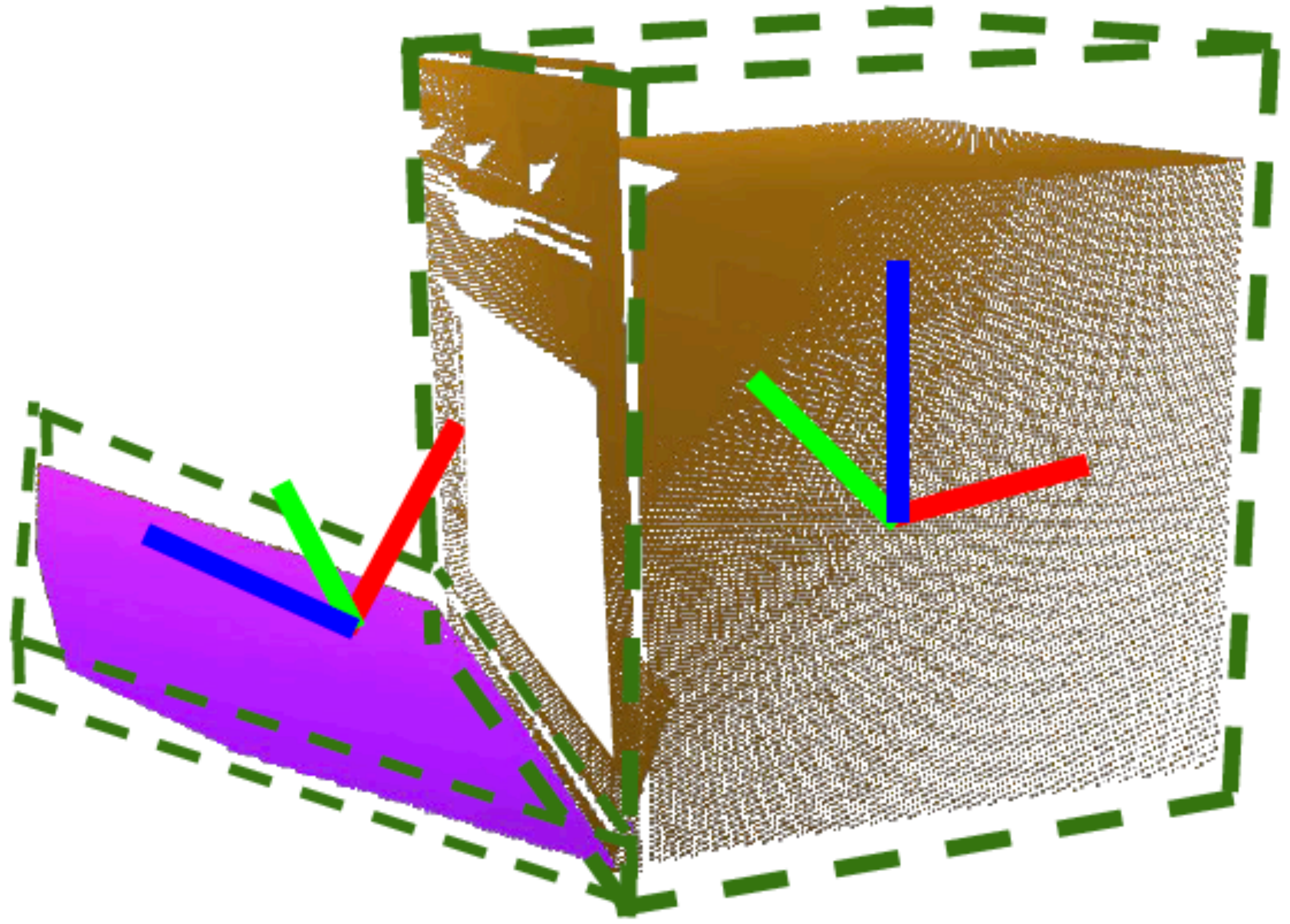
Rigid Object



Today’s Talk

Category-level pose estimation for

Rigid Object



Articulated Object

What’s next?

Category-level Articulated Object Pose Estimation (CVPR’20)

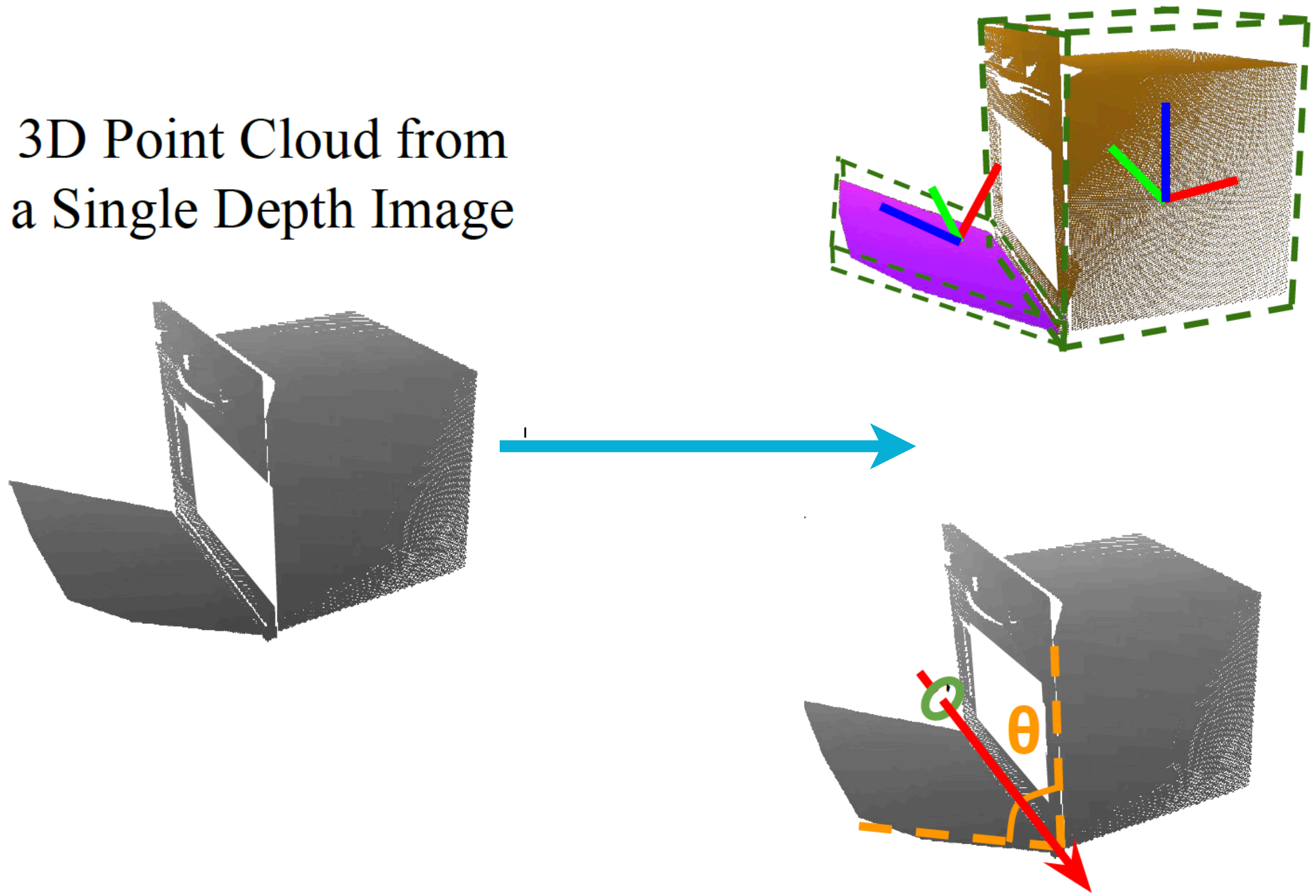
Xiaolong Li*, He Wang*, Li Yi, Leonidas Guibas, A. Lynn Abbott, Shuran Song

Problem definition

Input

3D Point Cloud from
a Single Depth Image

Output



Segmentation

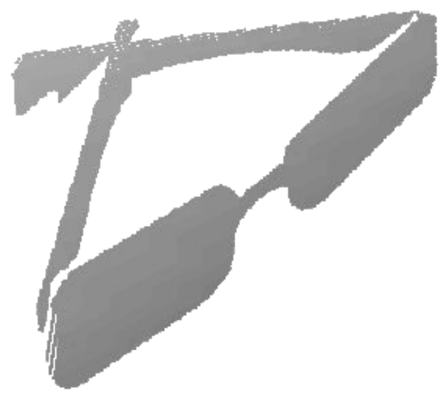
Part Pose + Size

Joint States

Joint Parameters

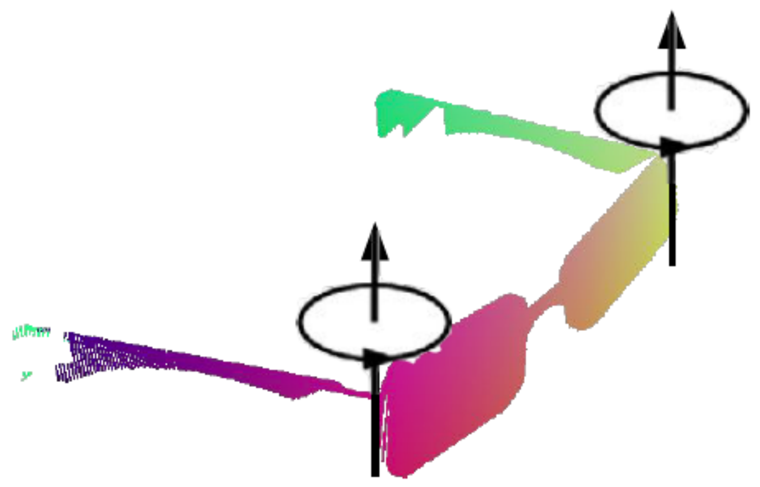
Representation: ANCSH

Articulation-aware Normalized Coordinate Space Hierarchy

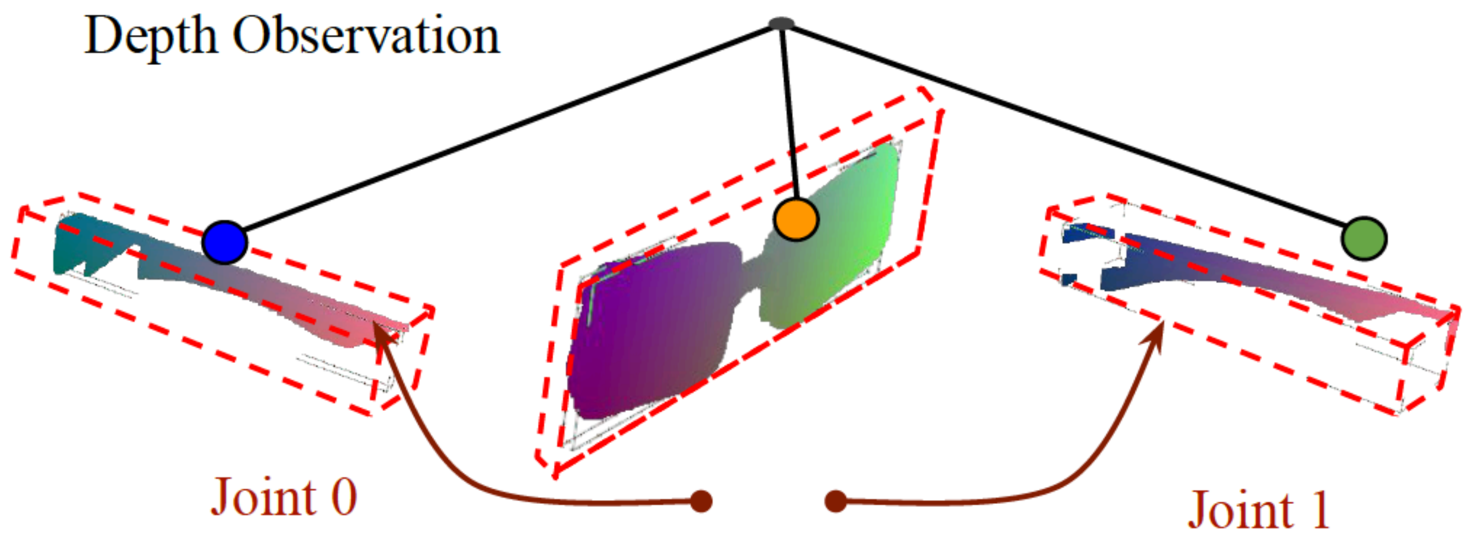


Depth Observation

Normalized Articulated Object Coordinate Space (NAOCS)

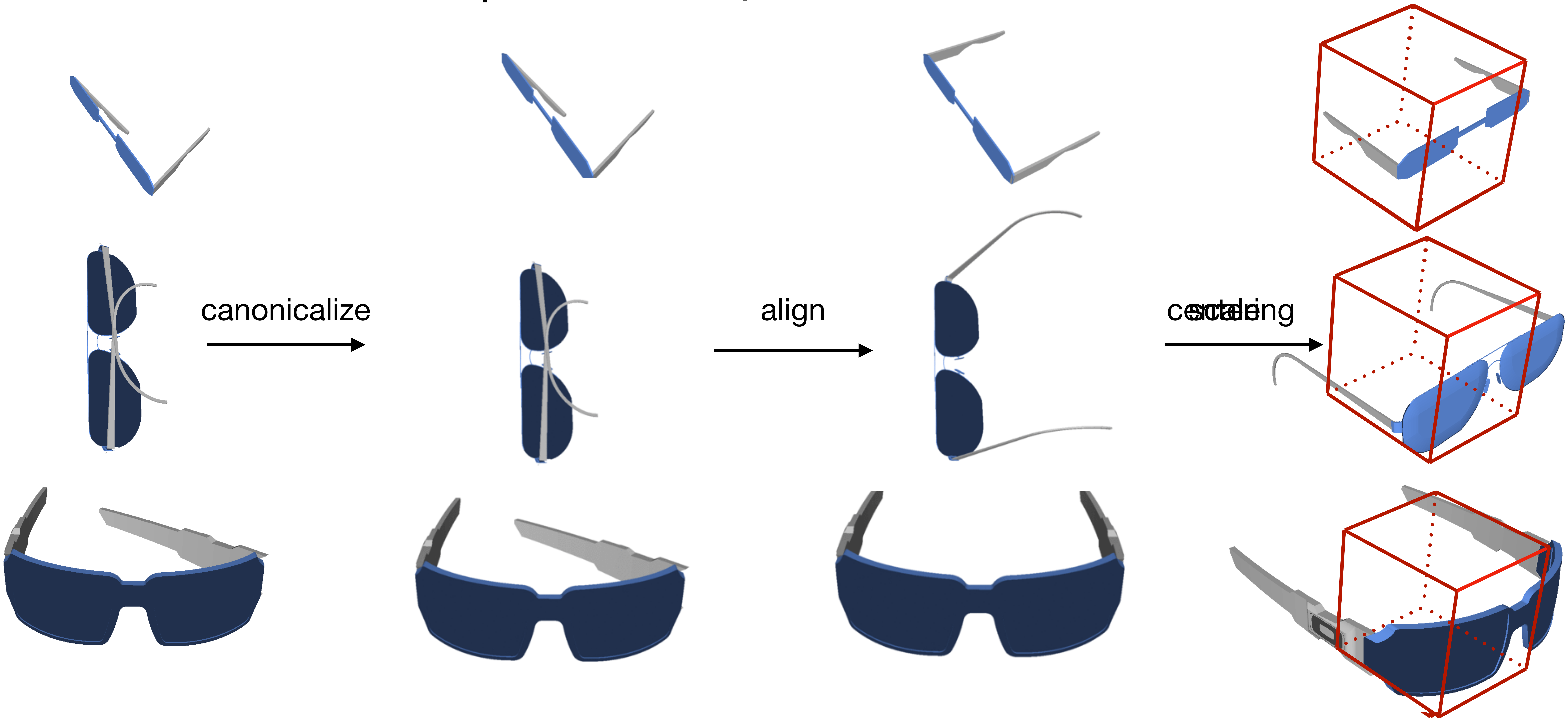


Normalized Part Coordinate Space (NPCS)



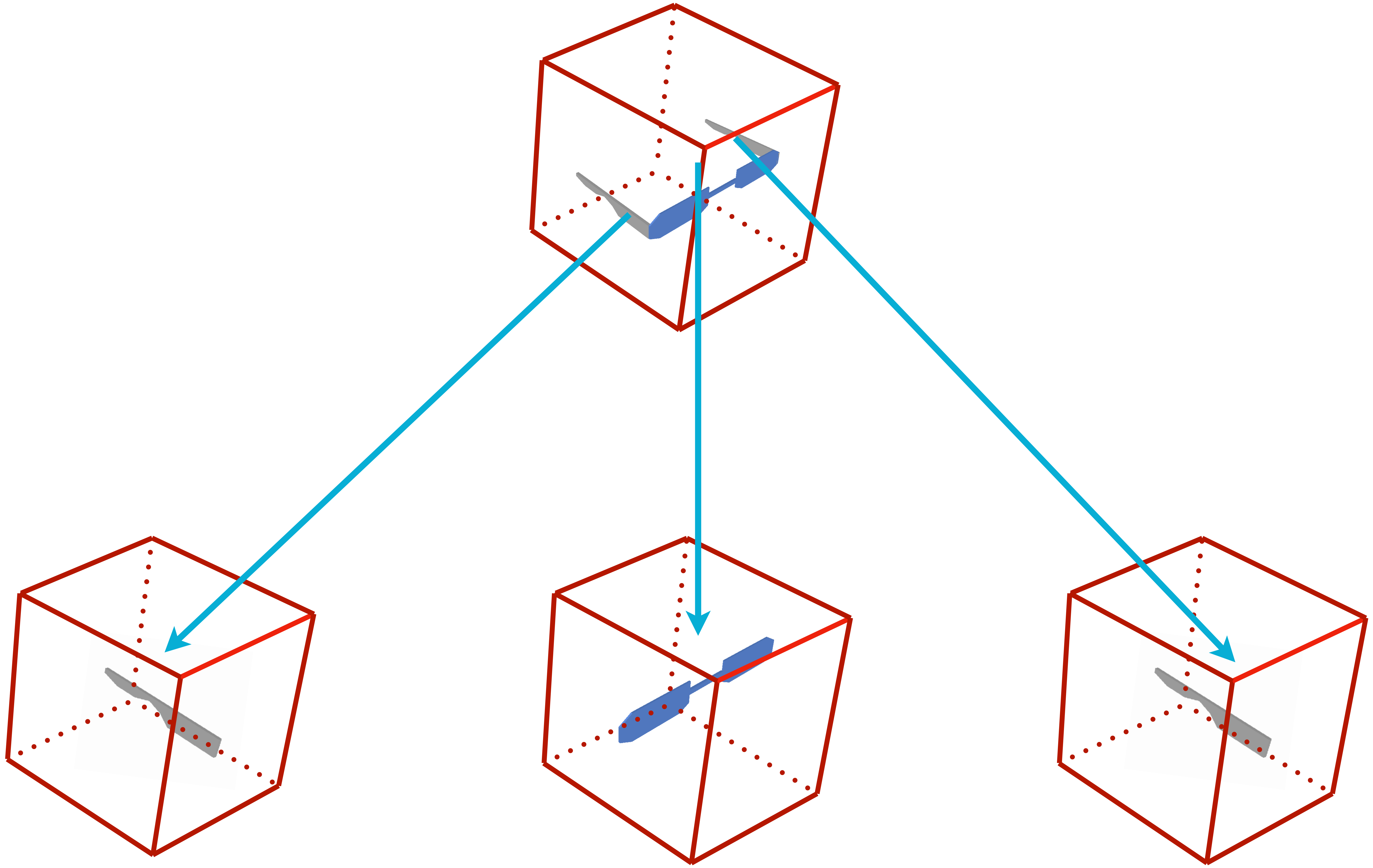
Hierarchical Normalization in ANCSH

Normalized Articulated Object Coordinate Space(NAOCS)

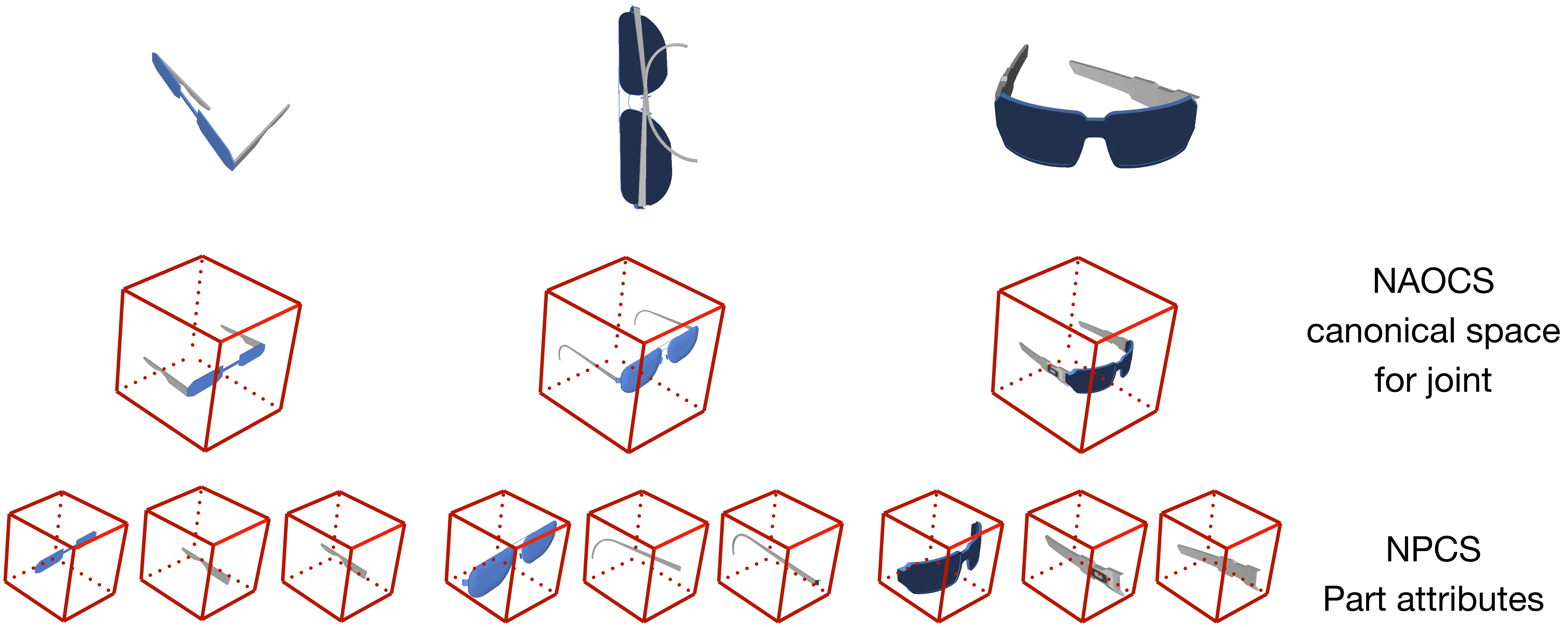


Hierarchical Normalization in ANCSH

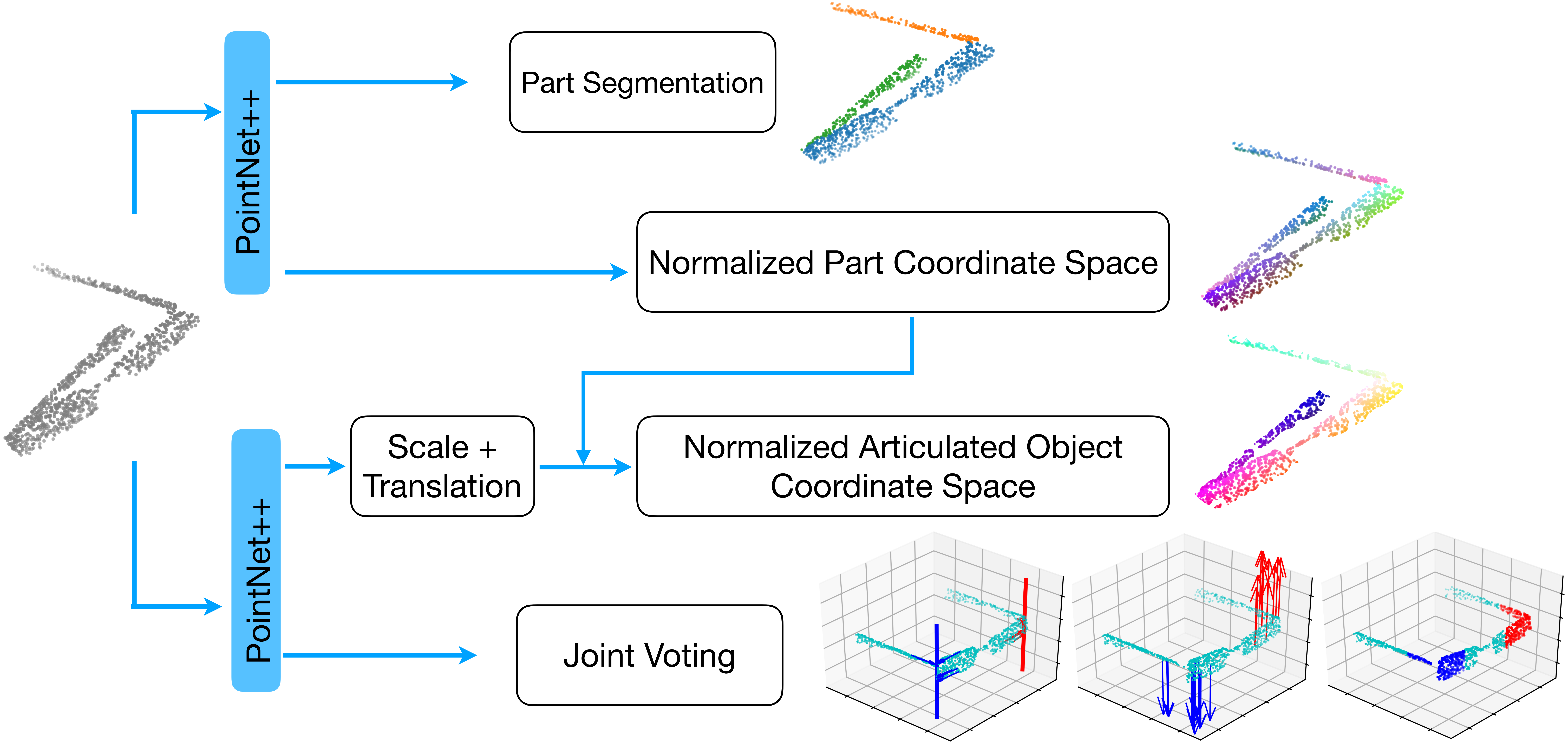
Normalized Part Coordinate Space(NPCS)



Representation: ANCSH

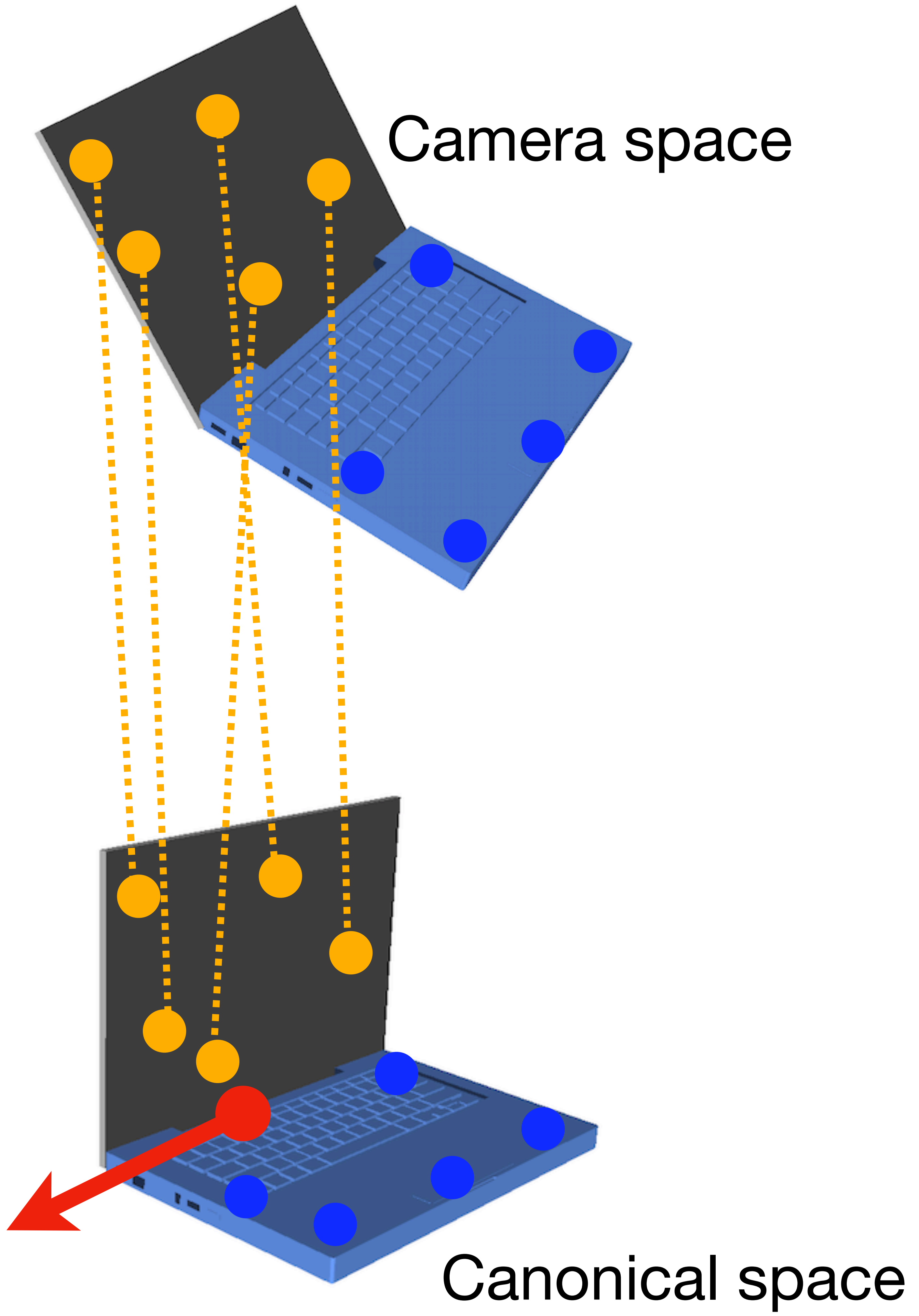


Method: ANCSH Network

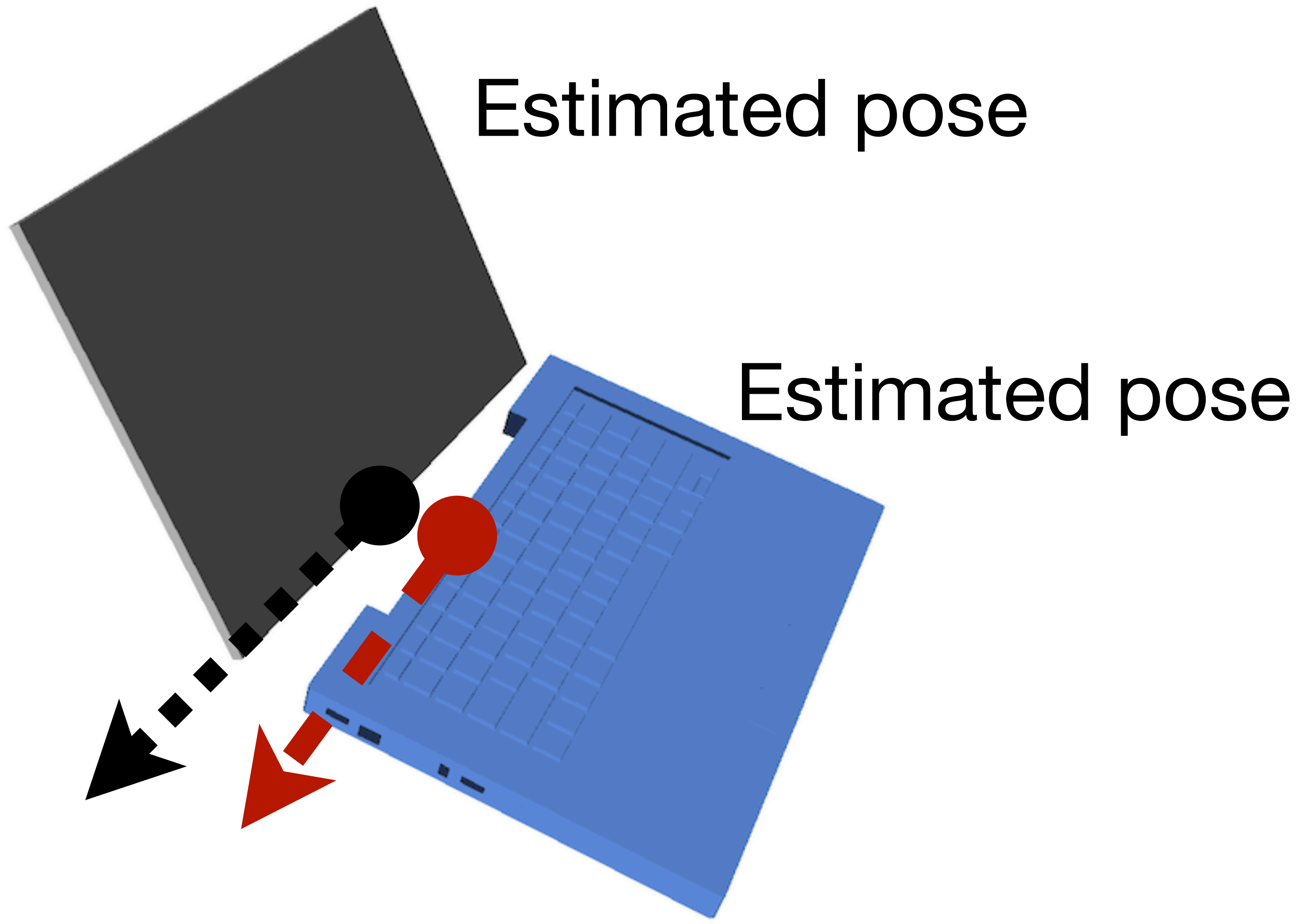


Method: Pose Fitting with Kinematic Chain Constraint

Input

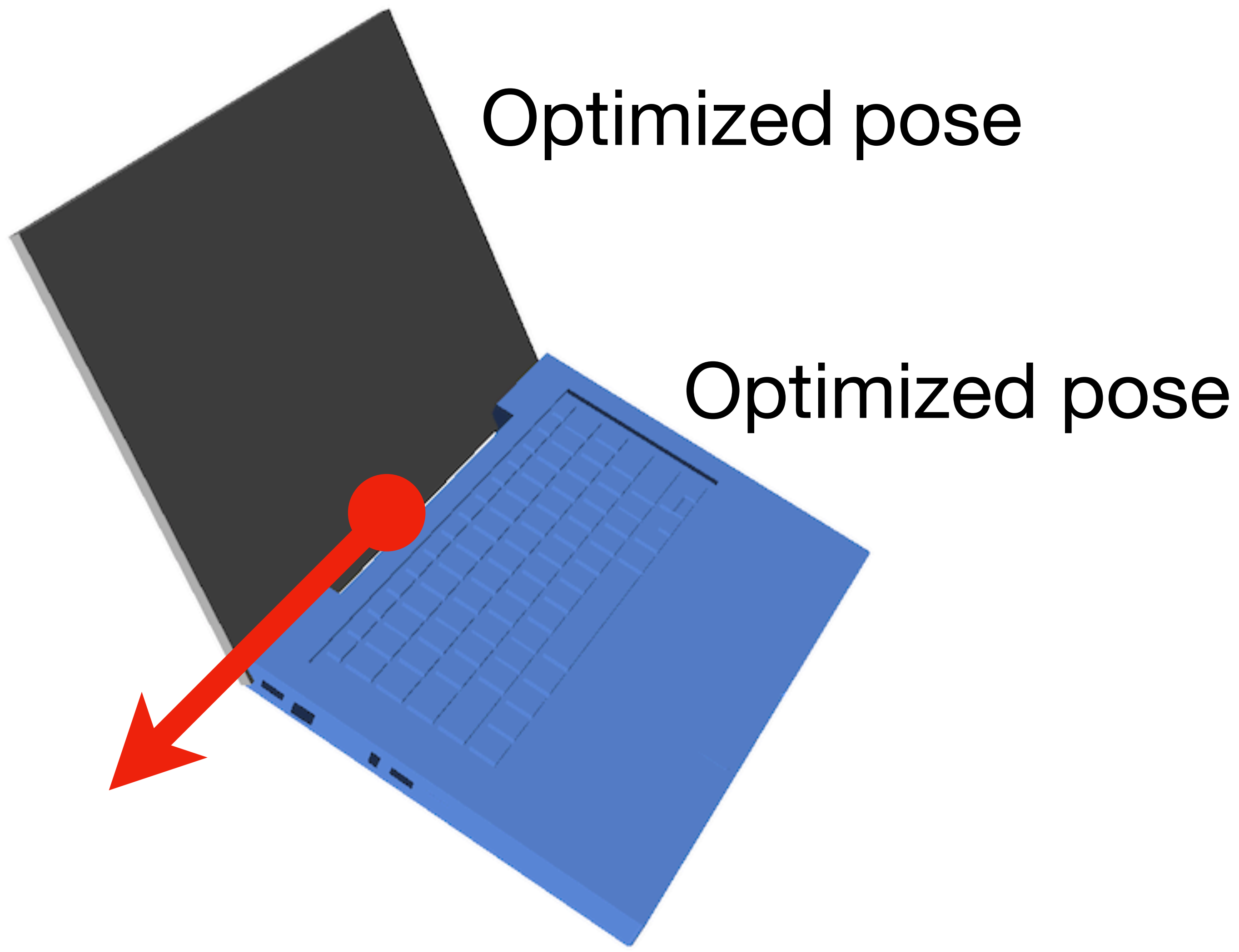


Separate fitting



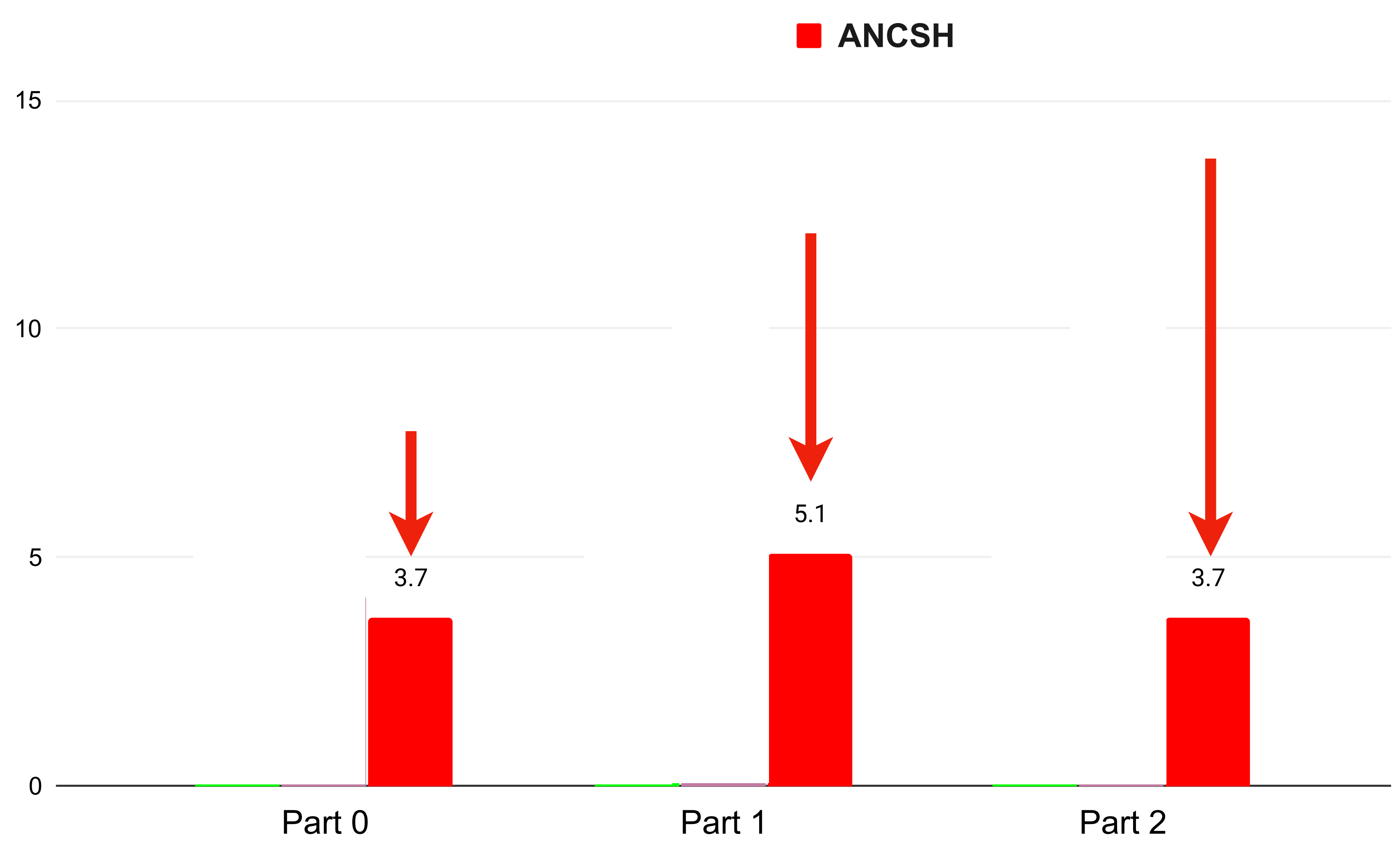
Not consistent!

Joint fitting

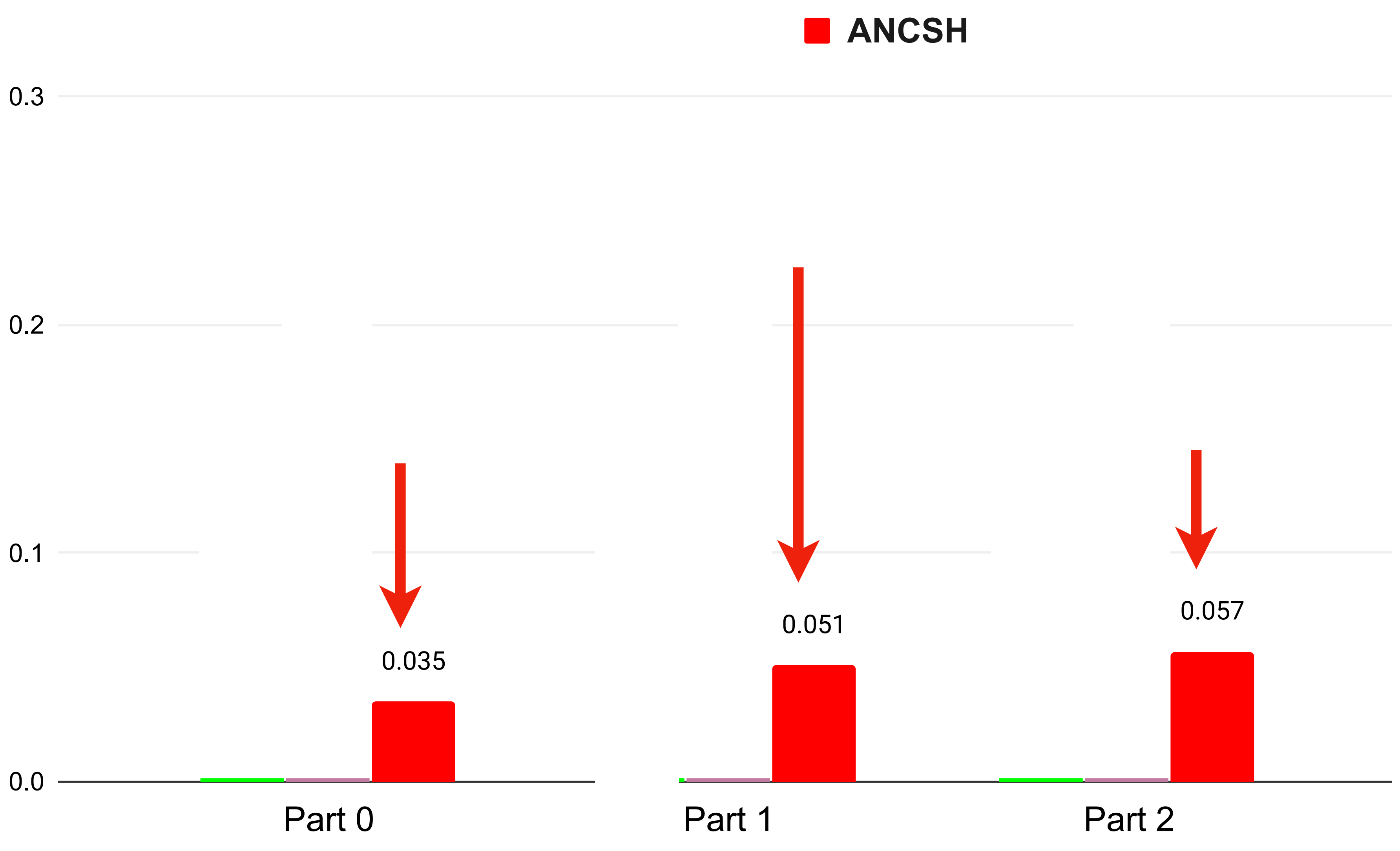


More consistent!

Results: Pose Estimation

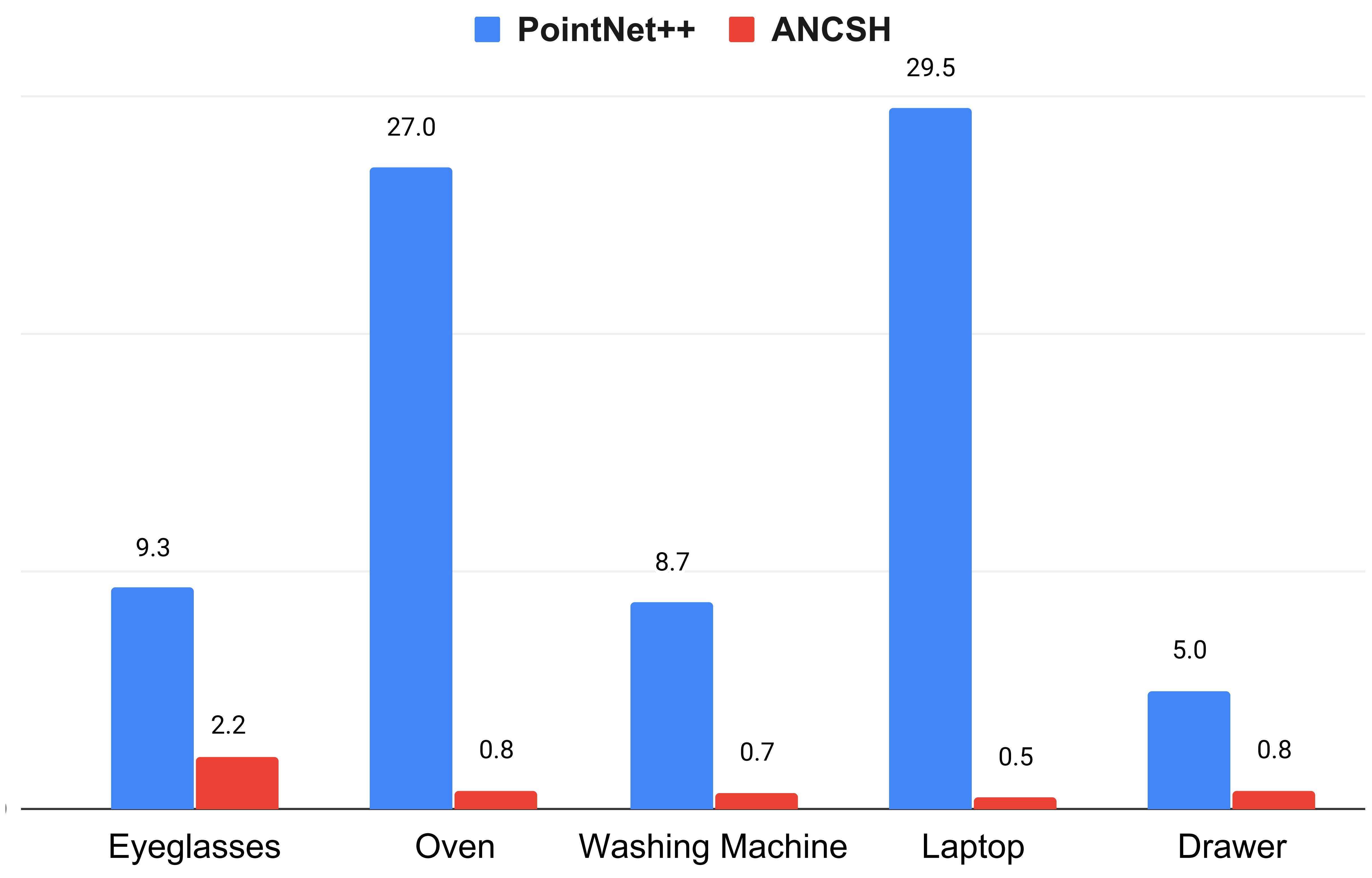


Rotation Error

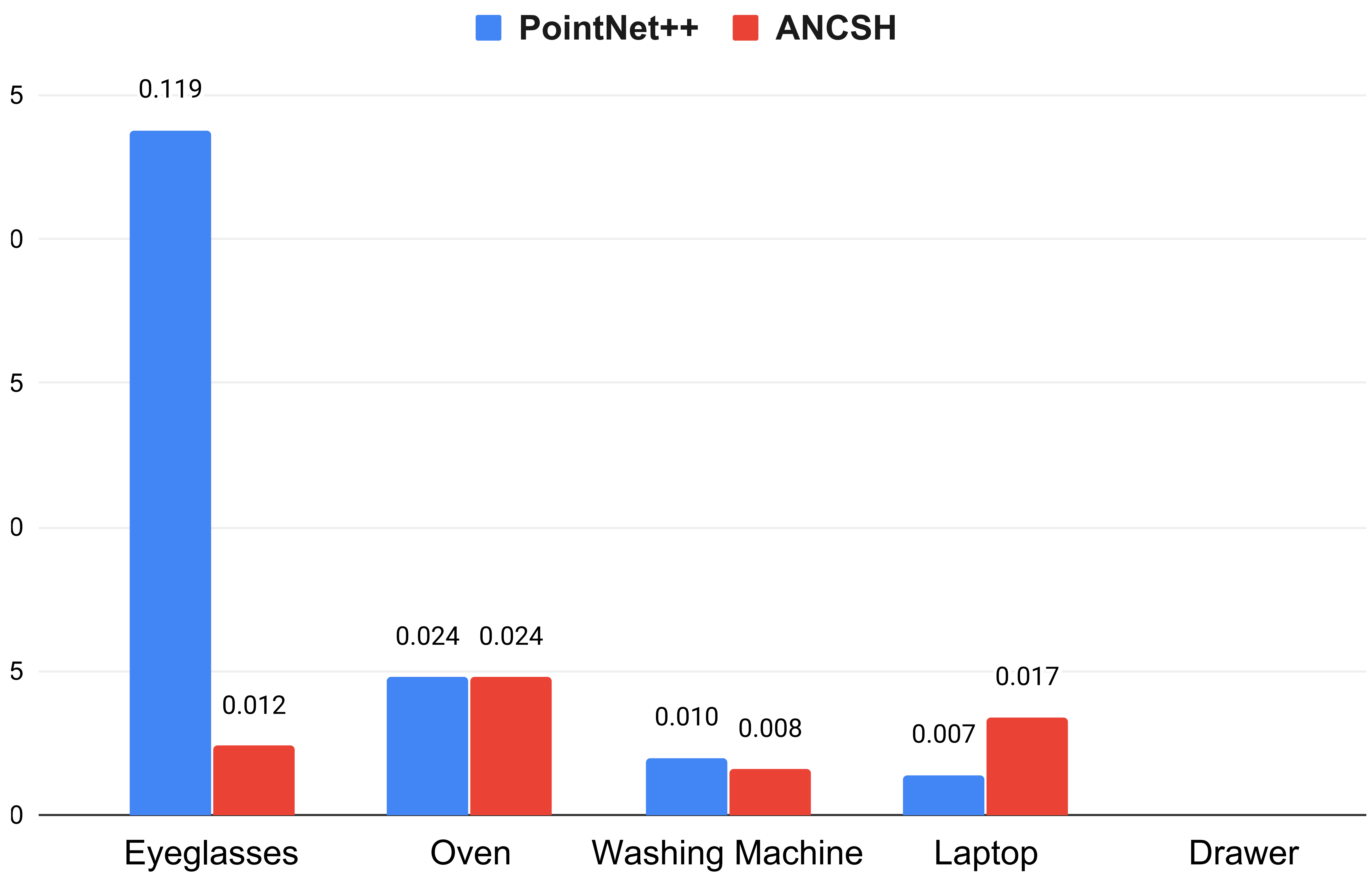


Translation Error

Results: Joint Parameters Estimation



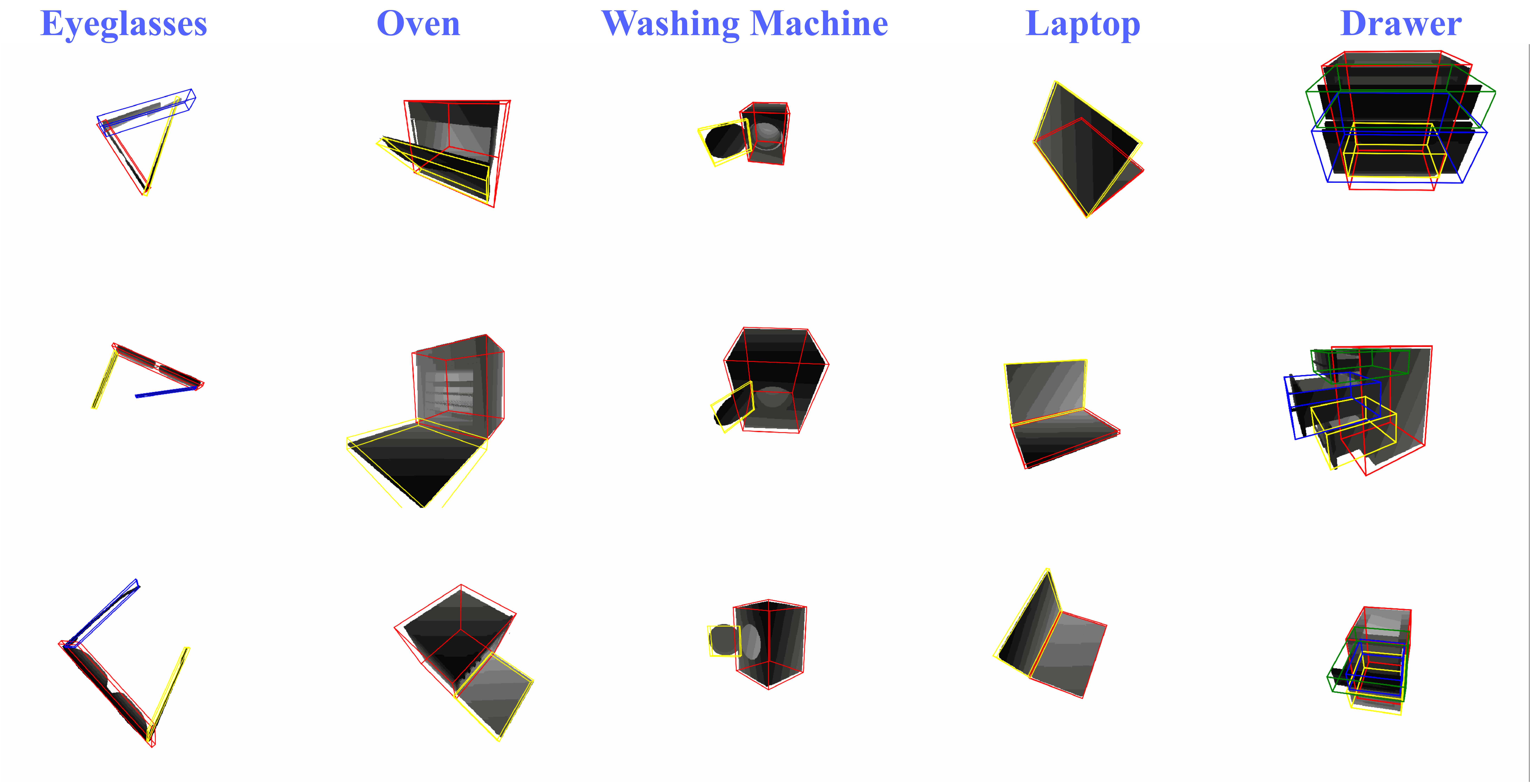
Joint Axis Orientation Error



Axis Location Error

Our method achieves ***high-accuracy*** joint parameters estimation!

Results visualization of Hold-Out Instances



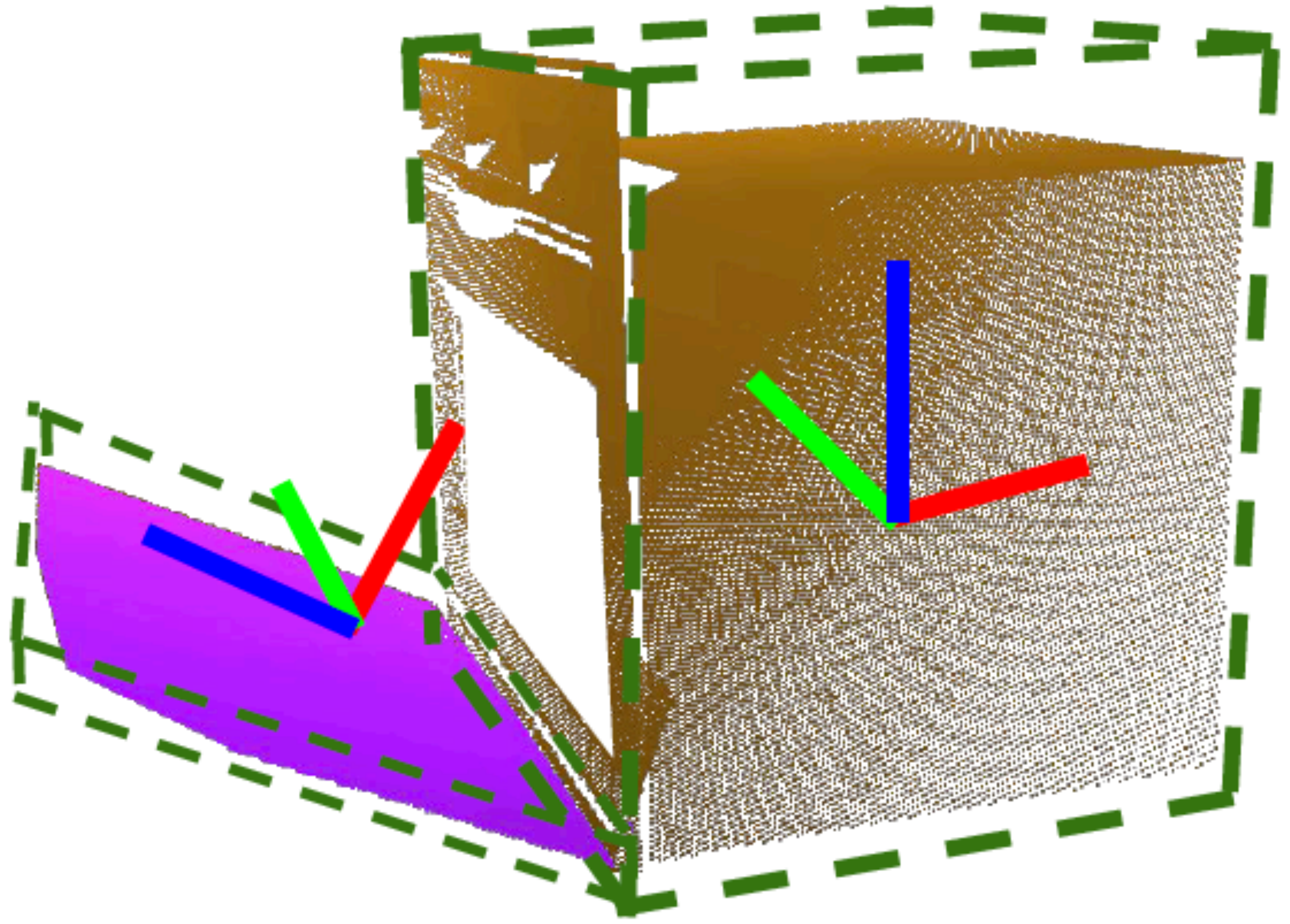
Results: ICCV 2015 Articulated Pose Estimation Dataset

RGB reference | Predicted Pose | Labeled GT Pose

Recap

Category-level pose estimation for

Rigid Object



Articulated Object

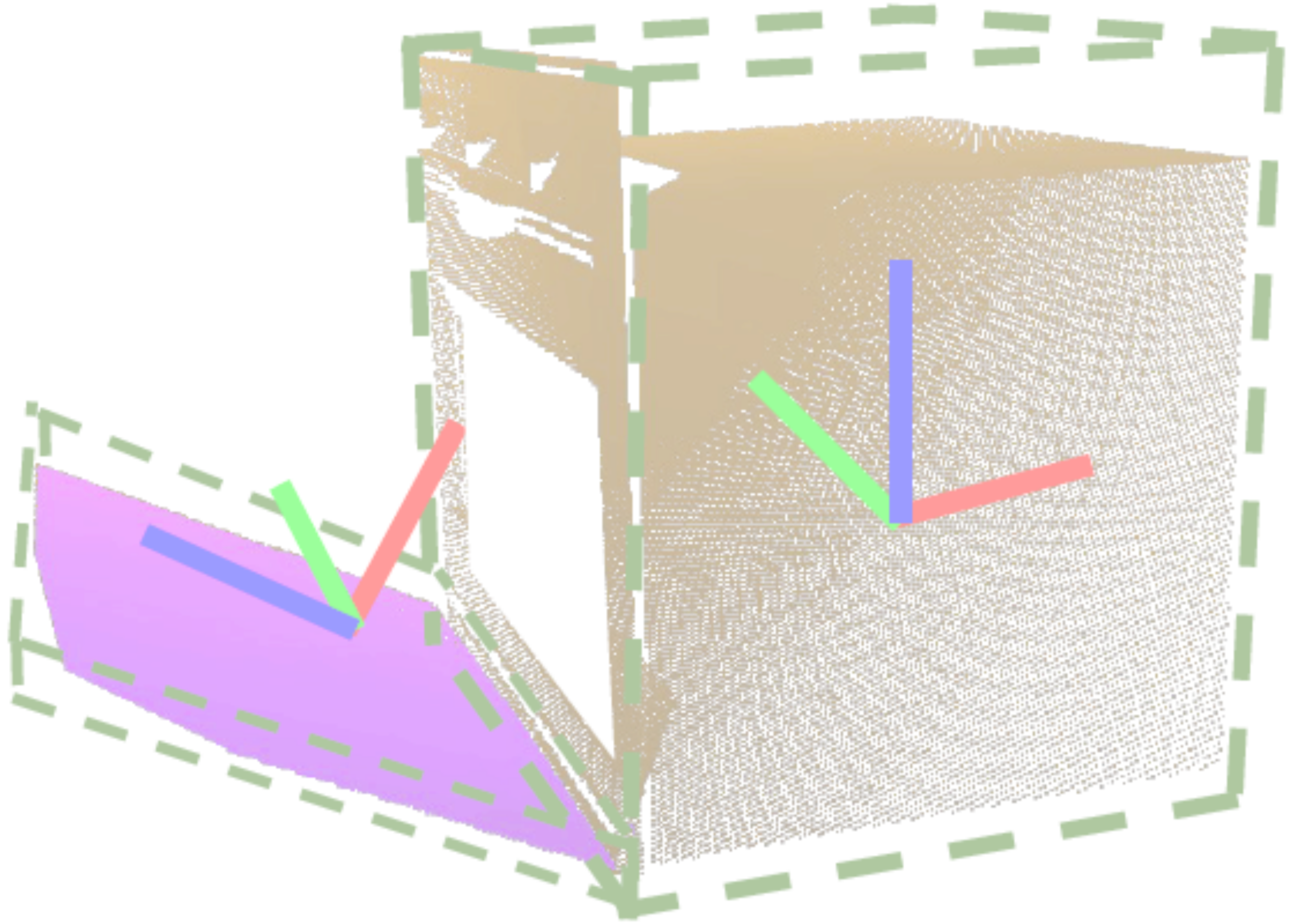
What's next?

- A novel representation for category level articulated objects——**ANCSH**;
- Joint optimization of part poses with kinematic chain constraints improves accuracy

What's next?

Category-level pose estimation for

Rigid Object



Articulated Object

What's next?

Not everything in the world is a rigid body with a canonical pose!

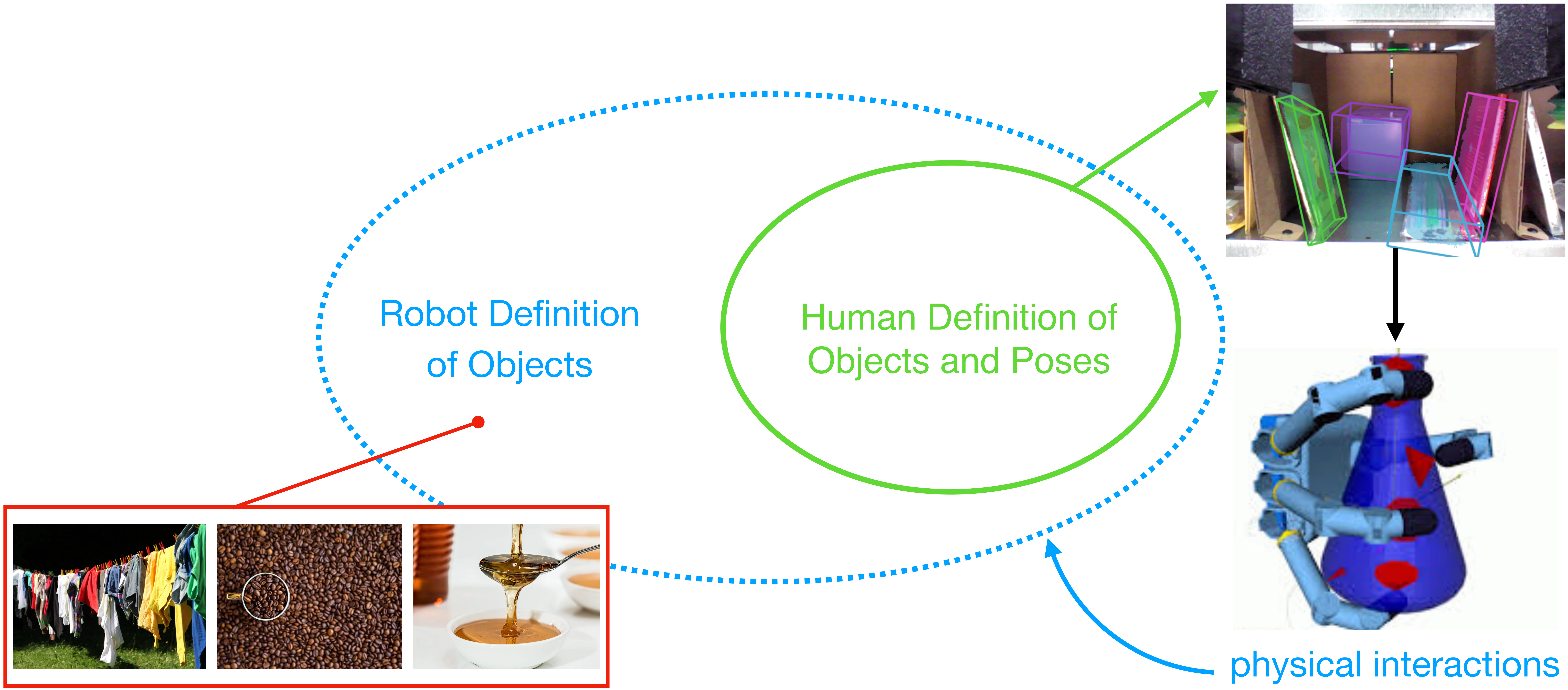
Deformable?


“piles” of stuff?

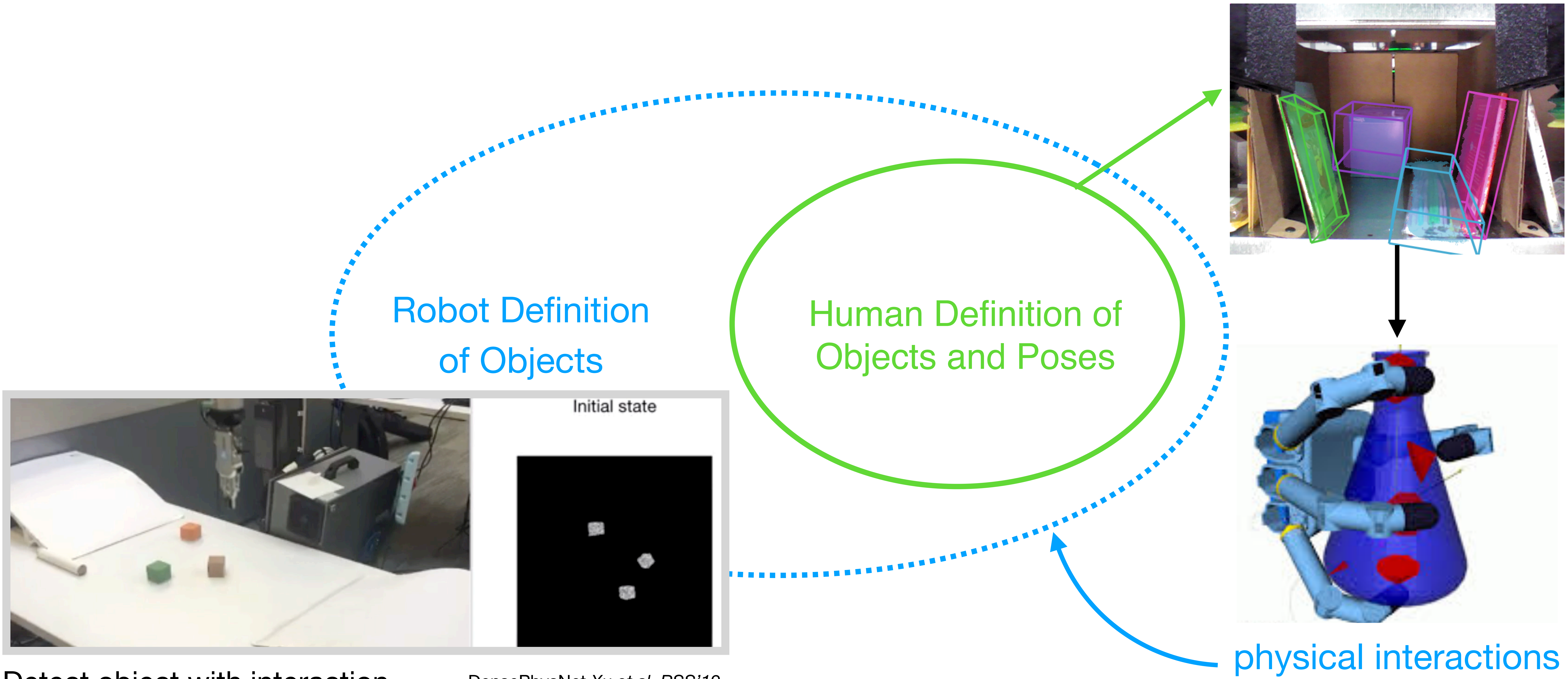
Fluid?

Unknown Category

Beyond object pose estimation



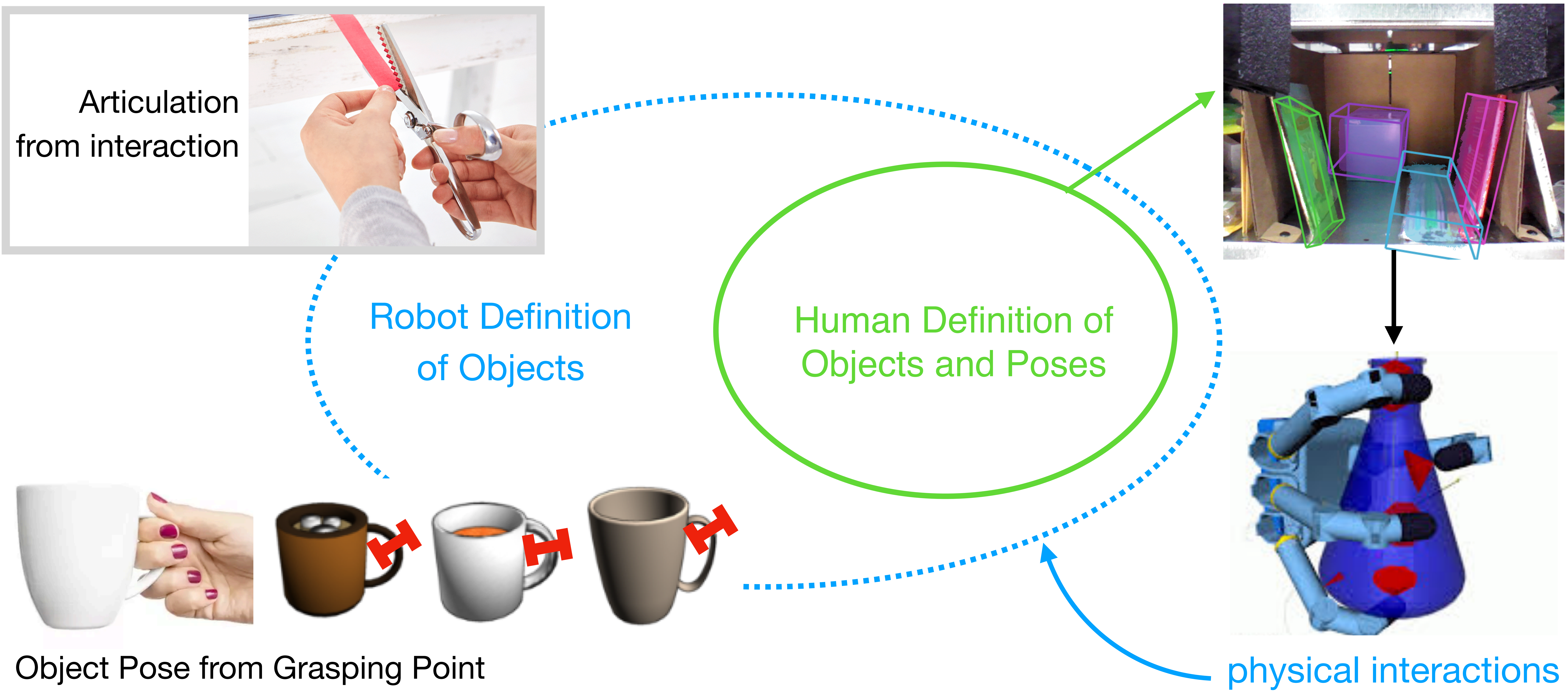
Beyond object pose estimation



Detect object with interaction

DensePhysNet *Xu et al, RSS'19*

Beyond object pose estimation

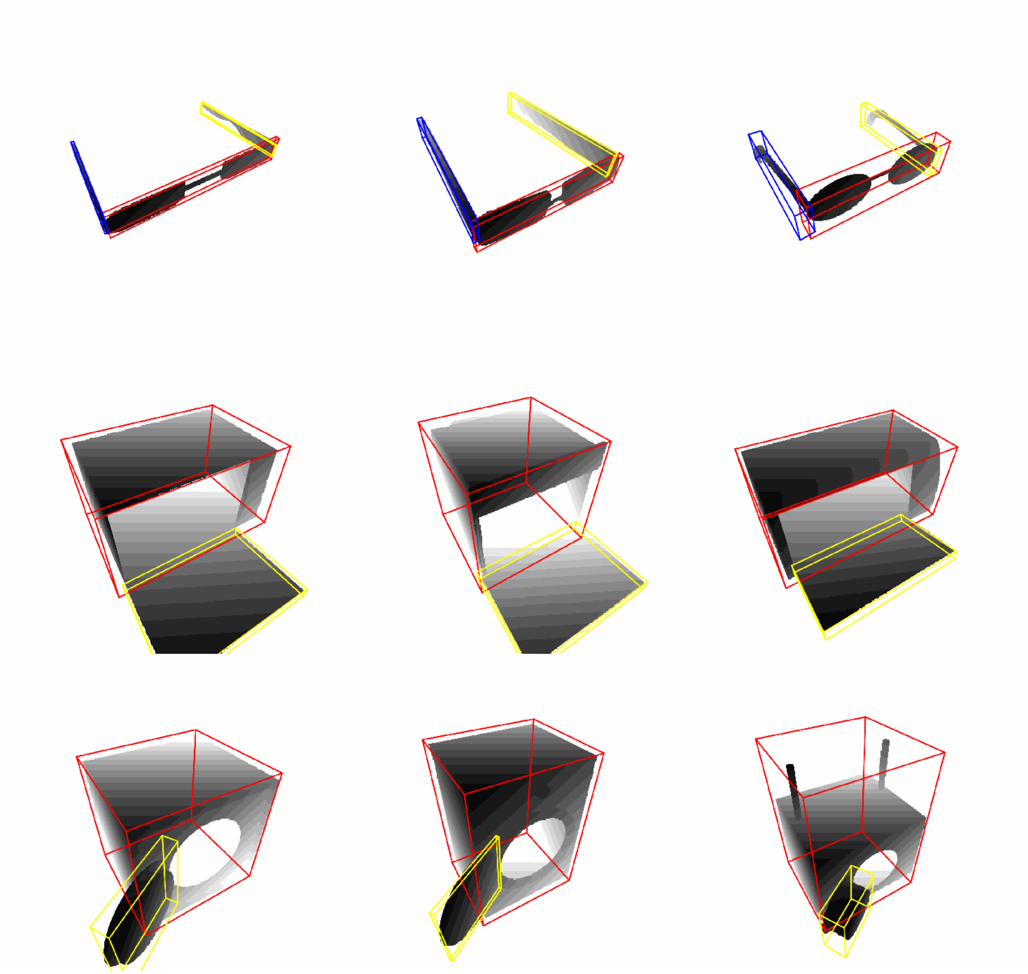


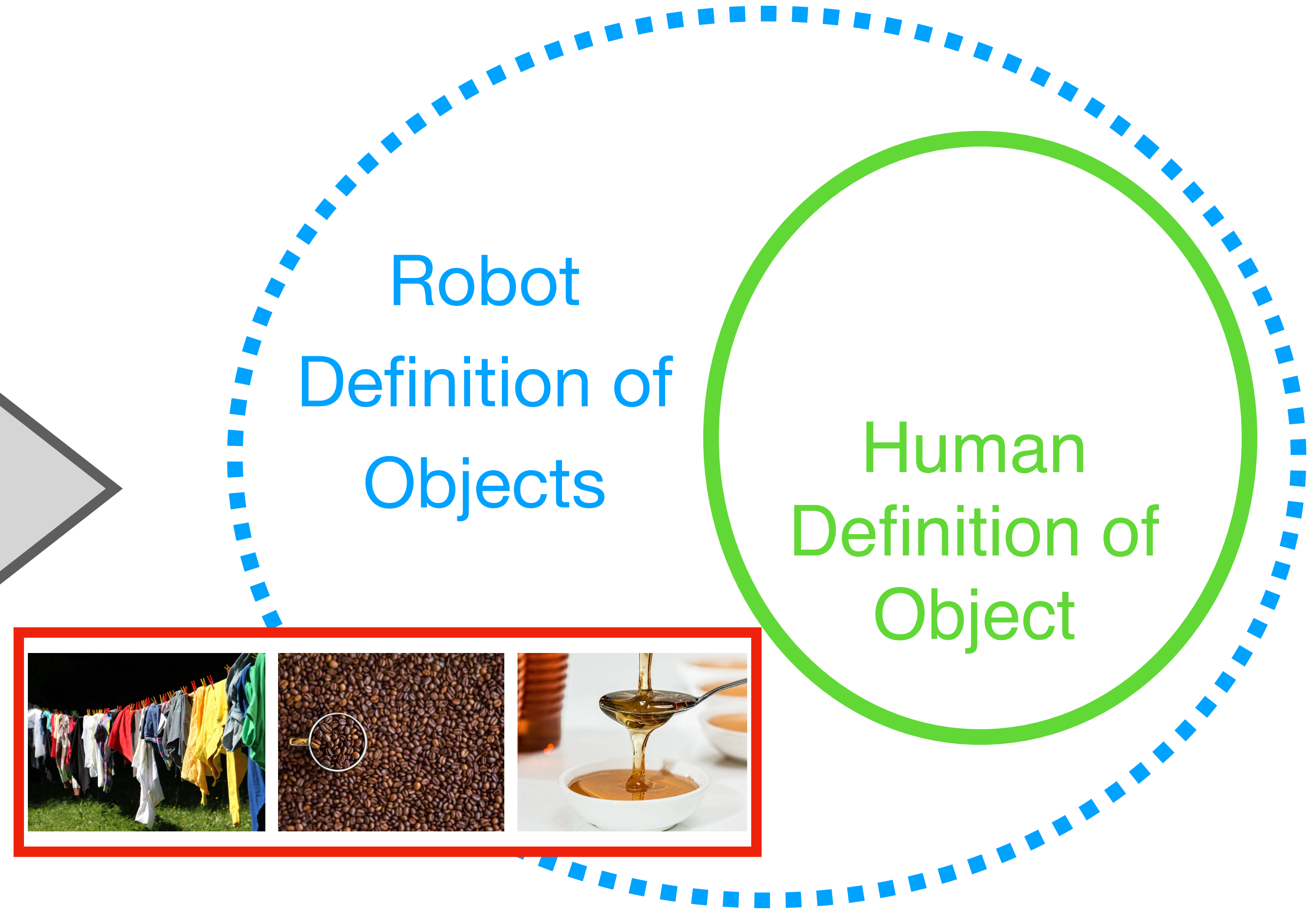
Summary

Instance-level

Category-level Pose Estimation

Beyond Category-level



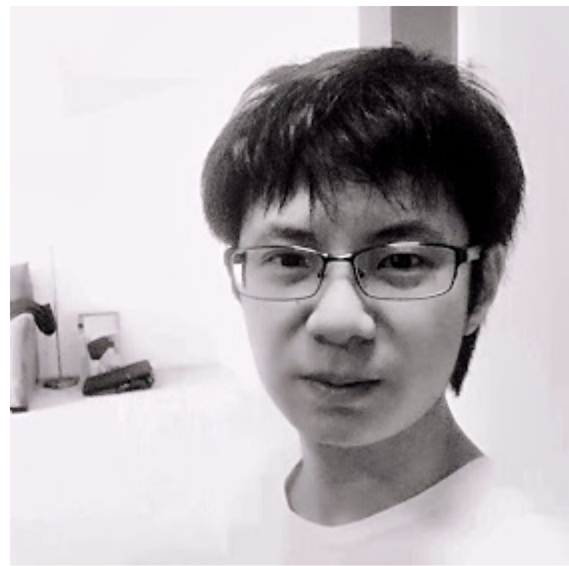


Rigid object

Articulated object

Deformable, Granular, Fluid, unknown category…

Acknowledgment


He Wang


Srinath Sridhar


Jingwei Huang


Julien Valentin


XiaoLong Wang


Yi Li


A. Lynn Abbott


Leonidas Guibas

Normalized Object Coordinate Space for Category-Level 6D Object Pose and Size Estimation
He Wang, Srinath Sridhar, Jingwei Huang, Julien Valentin, Shuran Song, Leonidas J. Guibas
CVPR 2019

Category-level Articulated Object Pose Estimation
He Wang*, Xiaolong Li*, Li Yi, Leonidas Guibas, A. Lynn Abbott, Shuran Song
CVPR 2020

TOYOTA RESEARCH INSTITUTE
amazon robotics
Google AI
NSF