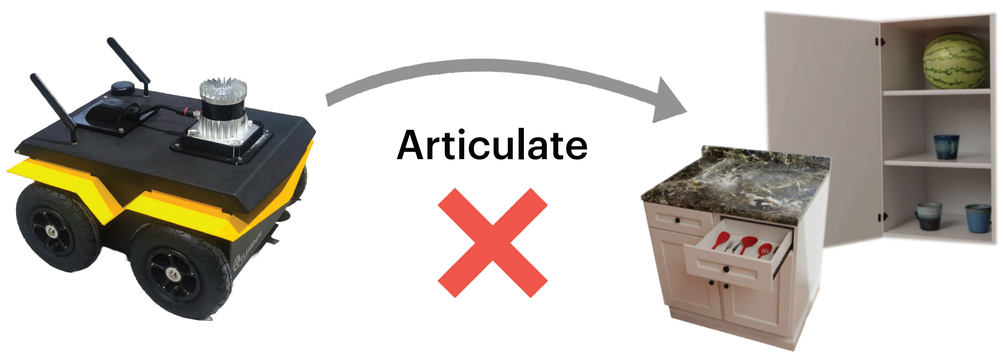
Pandora Articulted 3D Scene Graphs from Egocentric Vision

Alan Yu, Yun Chang, Christopher Xie, Luca Carlone

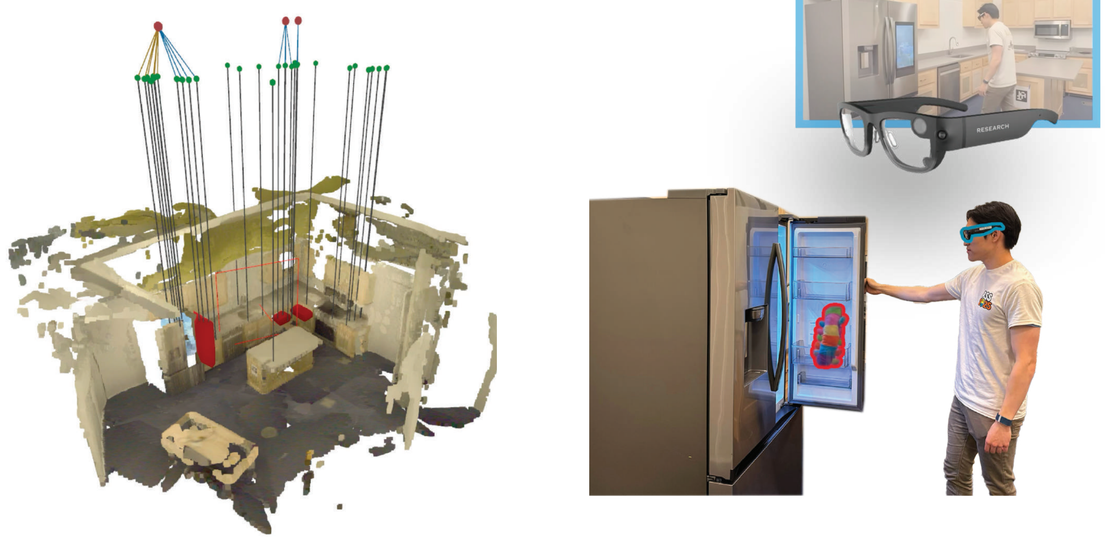
Overview

One common paradigm is to let a robot explore a scene and create a map from its observations as it discovers objects. However, the robot's **embodiment** poses constraints on what can be explored.



Can we delegate the mapping task to the human?

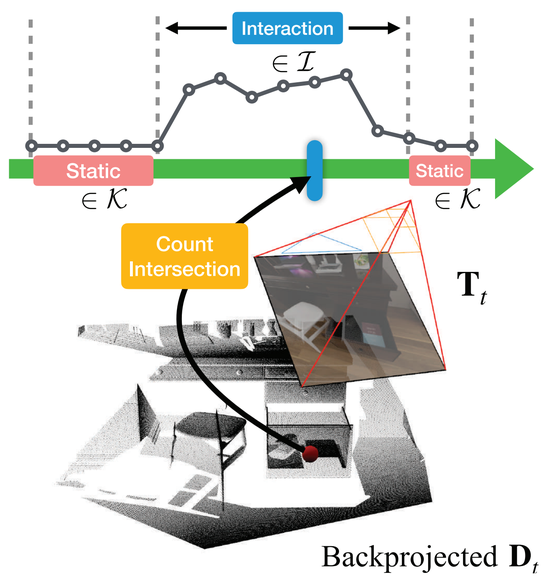
Making Scene Graphs More *Complete*



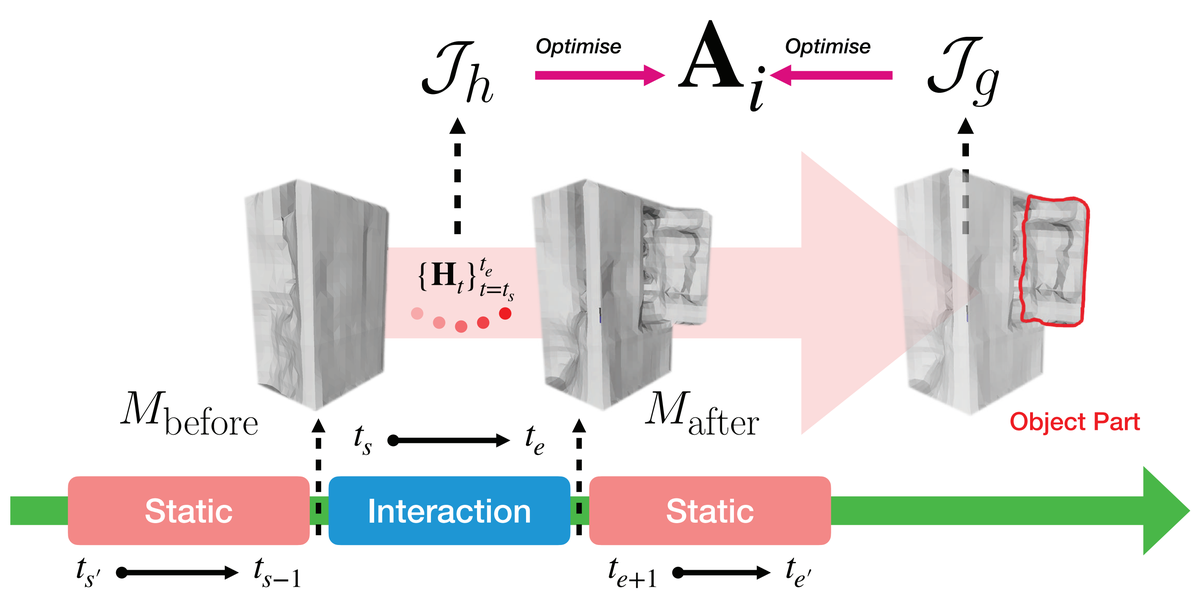
We use Project Aria glasses to record the human's interaction with the scene, and construct an articulated 3D scene graph that can be used for robotic object retrieval.

Inferring Articulation Models from Interaction

(1) Obtaining Interaction Keyframes



(2) Estimating Object Parts



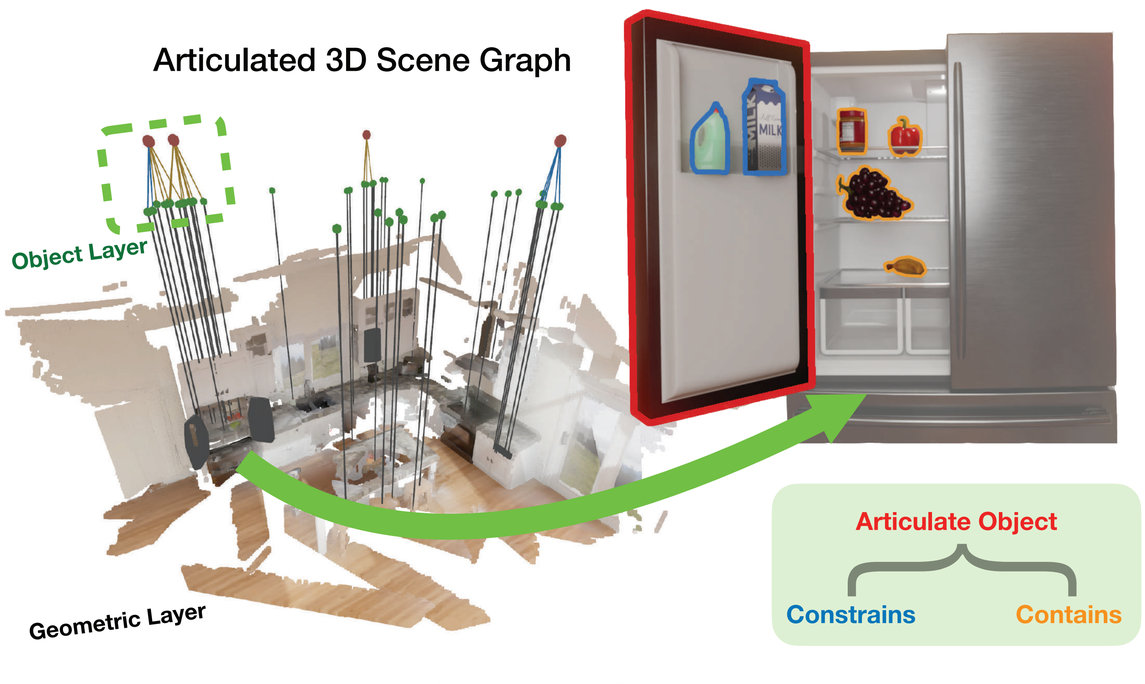
Articulation Model

Type (Revolute, Prismatic)

Pointcloud

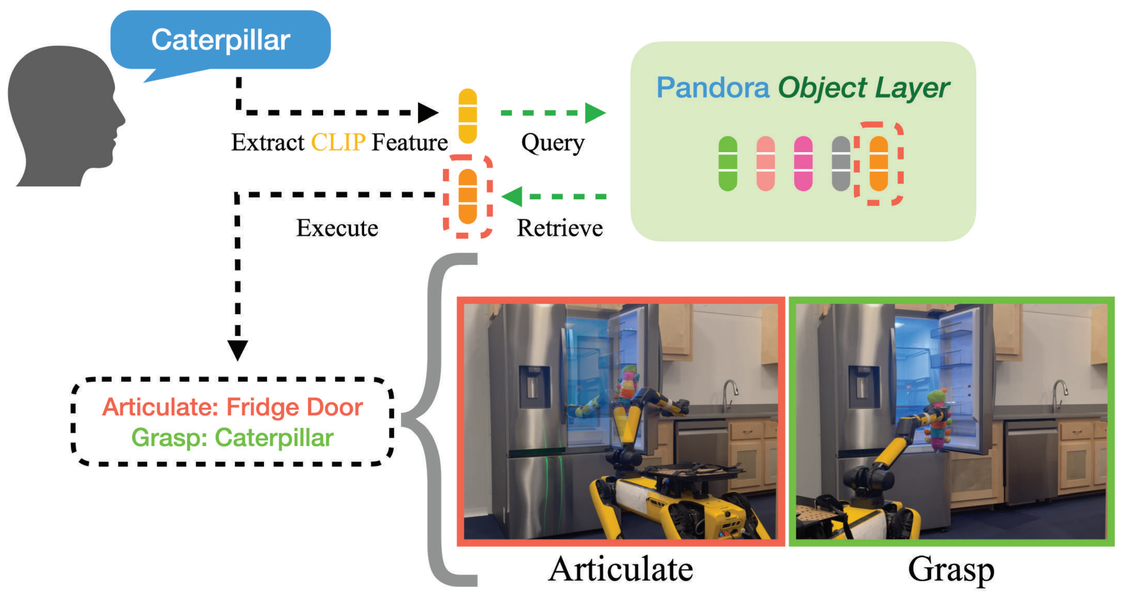
$$(t_{s_1}, t_{e_1}, \Delta\theta_1)$$
$$\vdots$$
$$(t_{s_n}, t_{e_n}, \Delta\theta_n)$$

Scene Graph Construction



The inferred articulation models are used to understand object relationships by distinguishing between contained objects, constrained objects, and static objects.

Downstream Robotic Object Retrieval



We query the scene graph for the object and its container. The robot navigates to, articulates, and grasps the target object with knowledge of the scene graph alone.

Method		Prism.	Revol.	
Name	**Input**	**Axis↓**	**Axis↓**	**Pivot↓**
Art. Anything	RGB Seq.	0.00	0.00	0.22
Ditto	2 PCs	10.82	1.26	0.26
Hands-Only	Hand Seq.	0.02	15.14	0.63
Ours	Ego Seq.	0.19	0.62	0.004

Table 1: Articulated model estimation errors. Hands-only is our method but only using the hand data loss term.

Dataset	Method	Metric (Pre / Rec / F1) ↑			
		Mesh Recovery	Static Objects	Dynamic Objects	Object Relationships
Kitchen	Hydra [12]	88.3 / **50.6** / 64.3	**75.0** / 62.1 / 68.0	–	–
	Khronos [29]	92.7 / **50.6** / **65.4**	66.7 / 46.7 / 54.9	16.8 / 6.9 / 9.8	–
	NoArt	98.5 / 46.0 / 63.0	57.1 / **93.3** / 70.9	–	30.7 / 50.0 / 41.2
	Ours	**98.6** / 46.0 / 62.7	**75.0** / 90.0 / **81.8**	**100.0 / 91.7 / 95.7**	**100.0 / 100.0 / 100.0**
Bedroom	Hydra [12]	90.1 / **68.8** / 78.0	**86.7** / 60.5 / 71.2	–	–
	Khronos [29]	**98.2** / 68.3 / **80.6**	42.6 / 60.5 / 50.0	1.1 / 0.7 / 0.8	–
	NoArt	98.1 / 66.2 / 79.0	61.5 / **93.0** / 74.1	–	41.0 / 60.0 / 48.7
	Ours	98.1 / 66.2 / 79.0	72.7 / **93.0** / **81.6**	**93.5 / 84.1 / 88.5**	**87.5 / 93.3 / 90.3**

Table 2: (Background) mesh recovery and object tracking metrics. Object relationships refer to classification of constrained vs. contained instances.