Visual Tactile Grasping

David Watkins-Valls

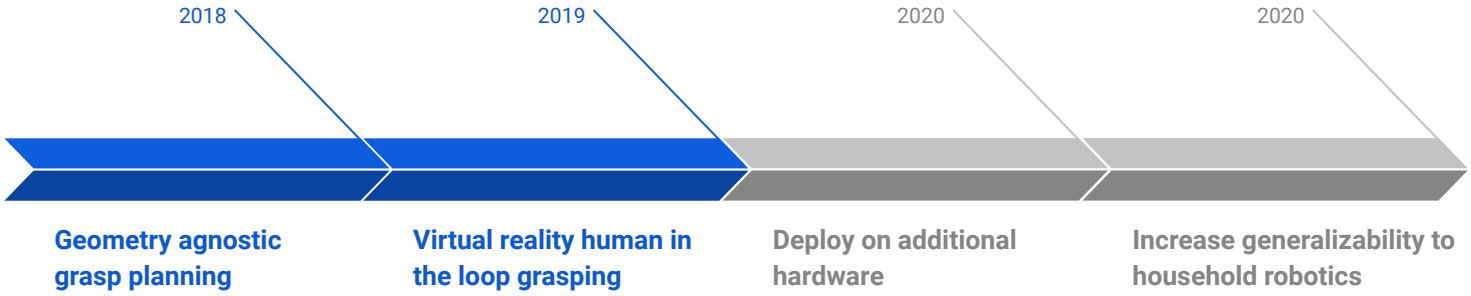
CRL

Source: https://www.youtube.com/watch?v=rNLv7f2wmLQ

Overview

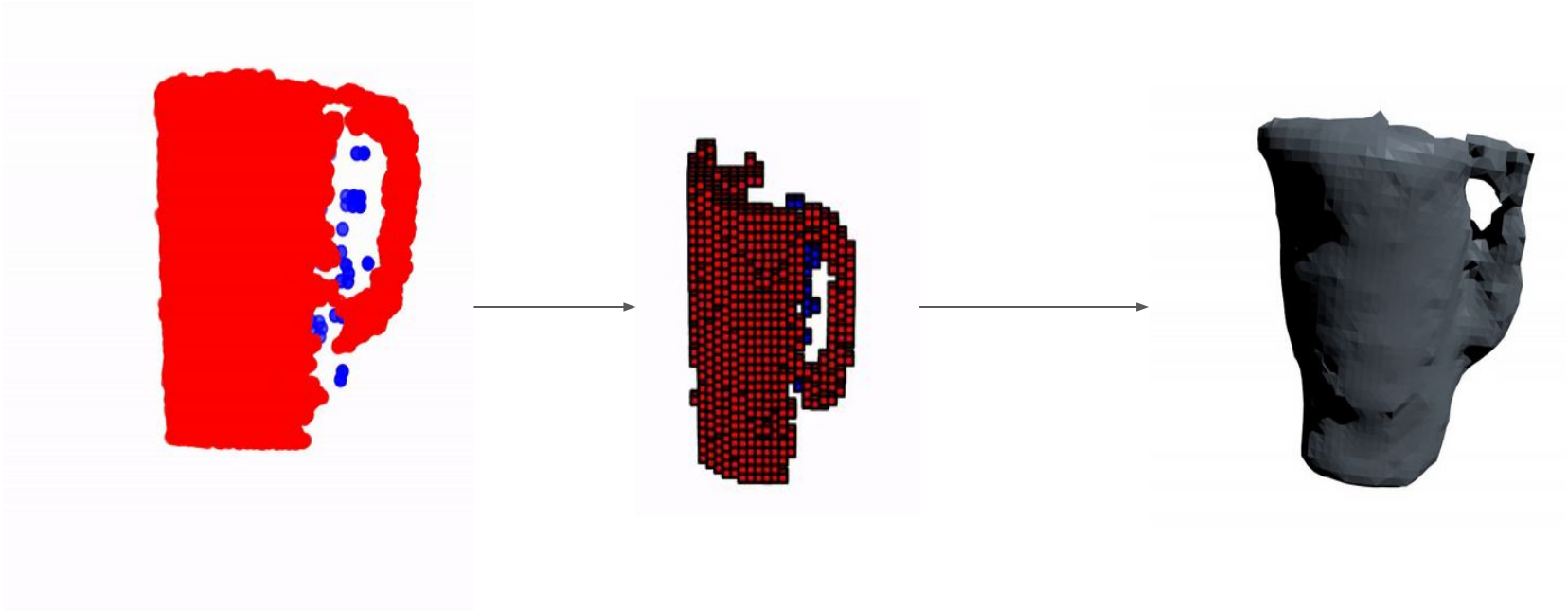
- Dataset generation
- Depth and tactile shape completion
- Grasping results
- VR to enable human in the loop grasping
- Future work

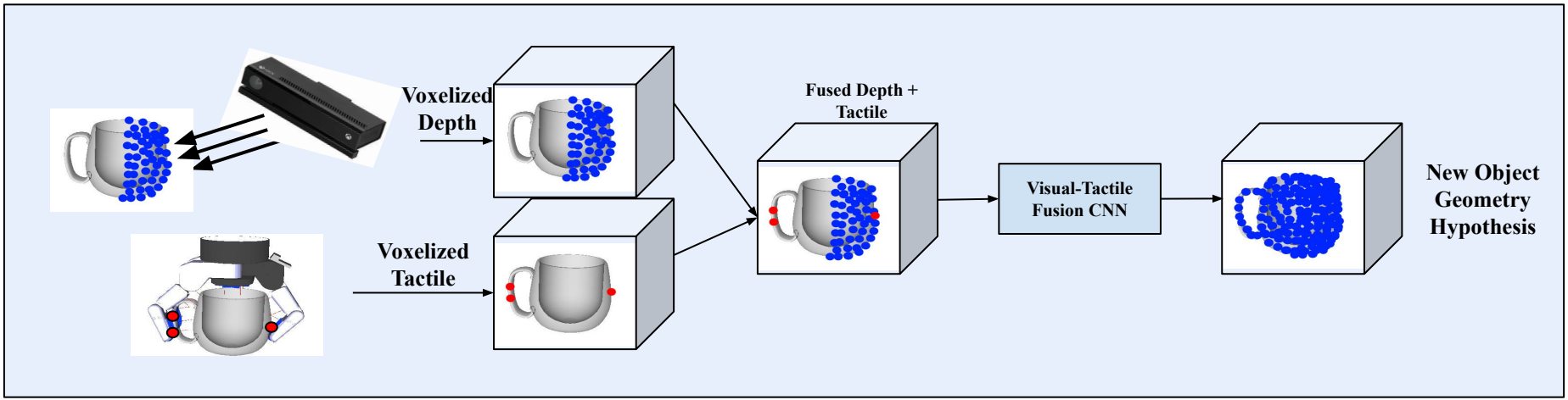
Timeline



Project summary

- Fellowship collaboration with Columbia Robotics lab
- Resources at Samsung
- Publications that are in the works



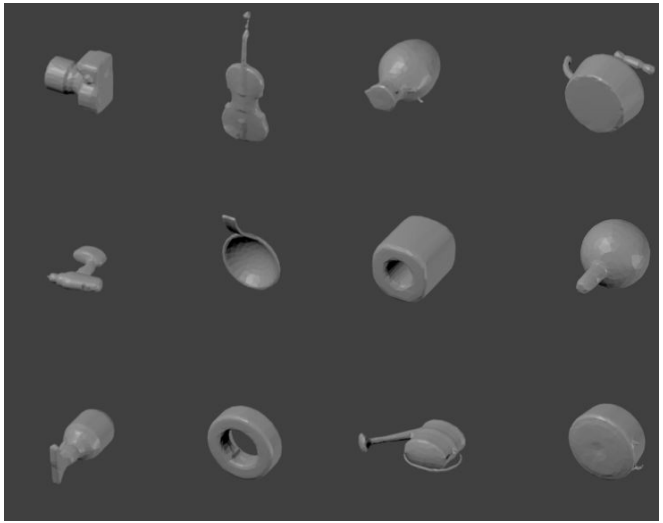


Dataset generation

YCB Object Dataset

28 unique 3D models

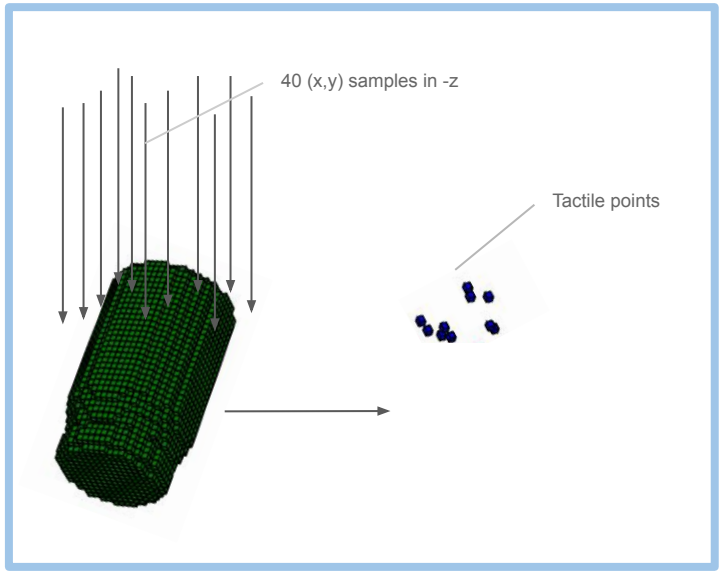
GRASP Dataset



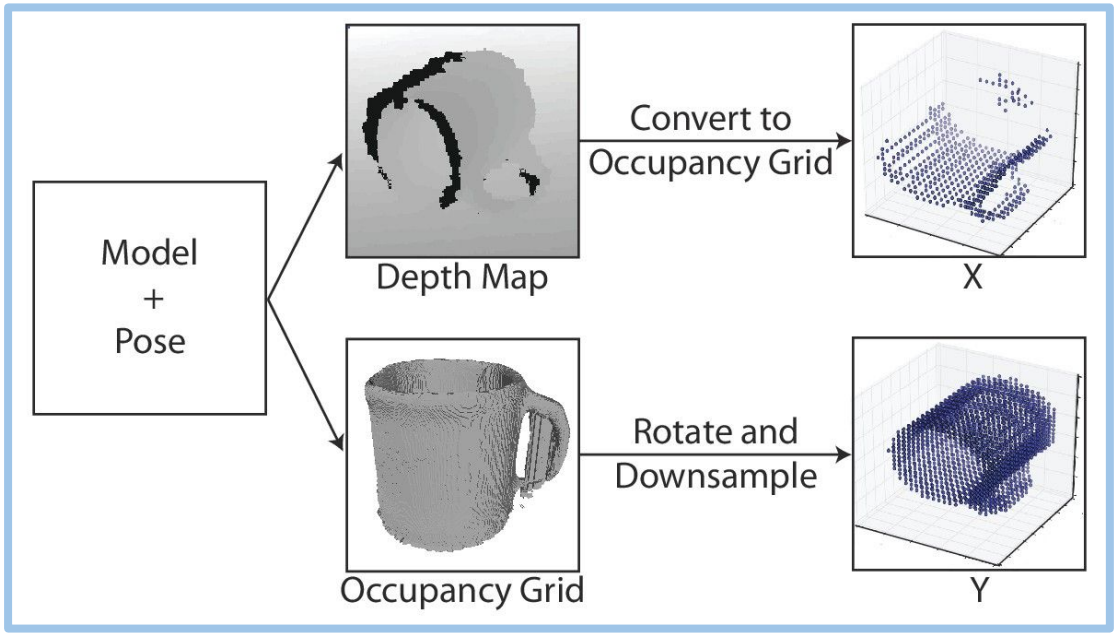
590 unique 3D models

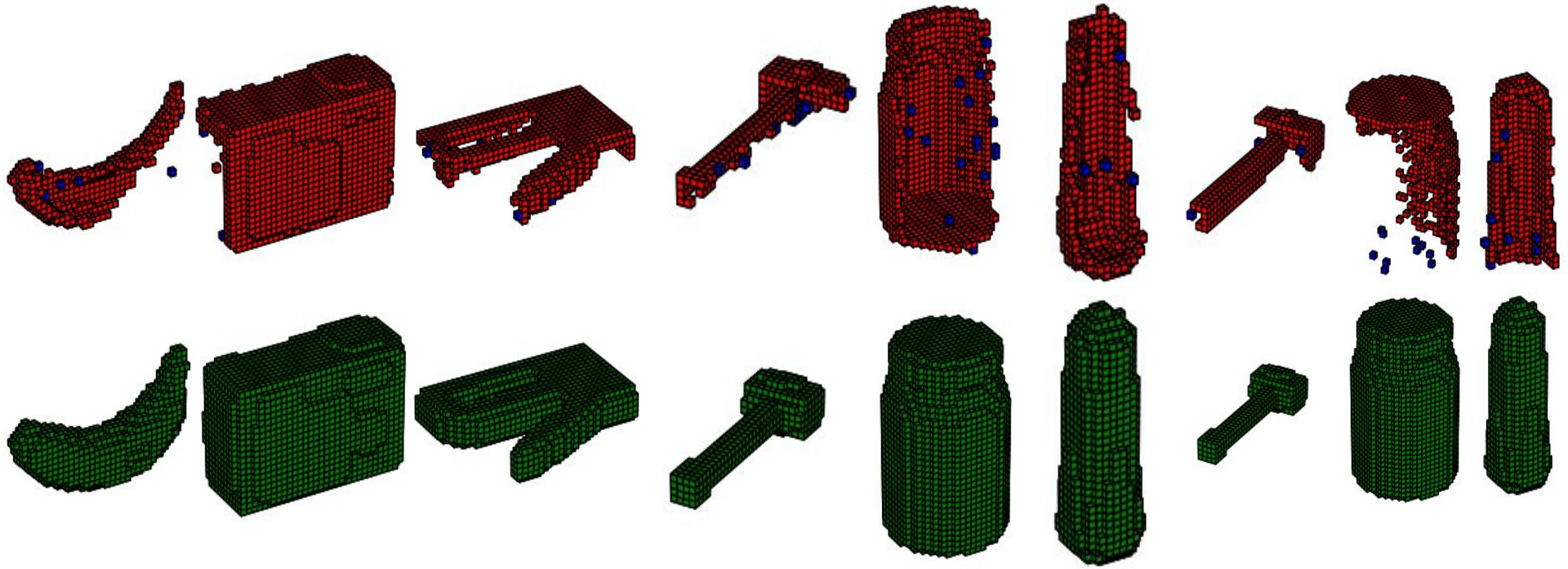
618 objects
726 views per object
448,668 depth images

Capturing tactile samples

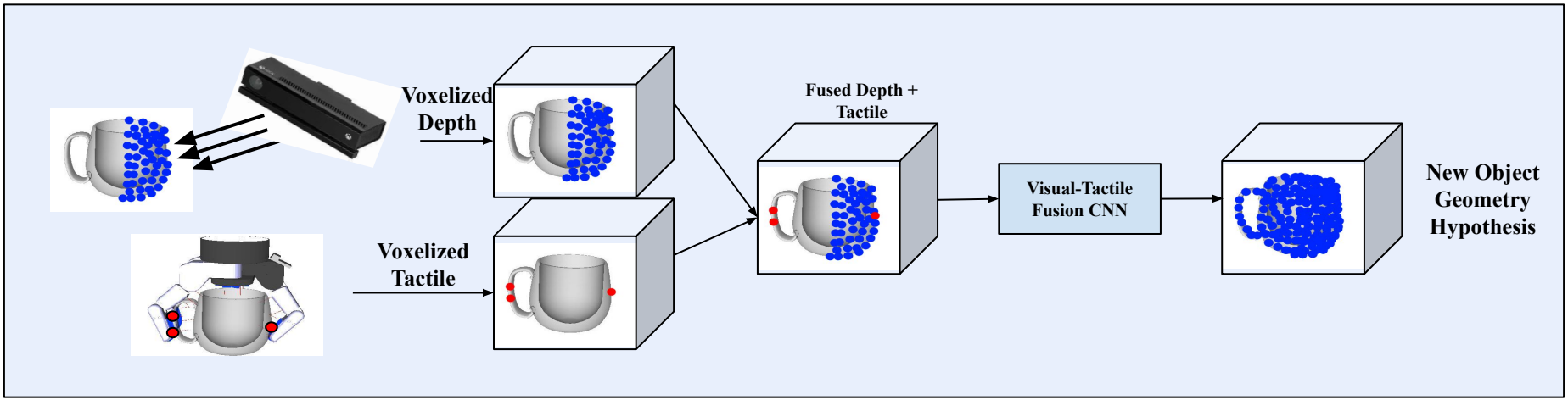


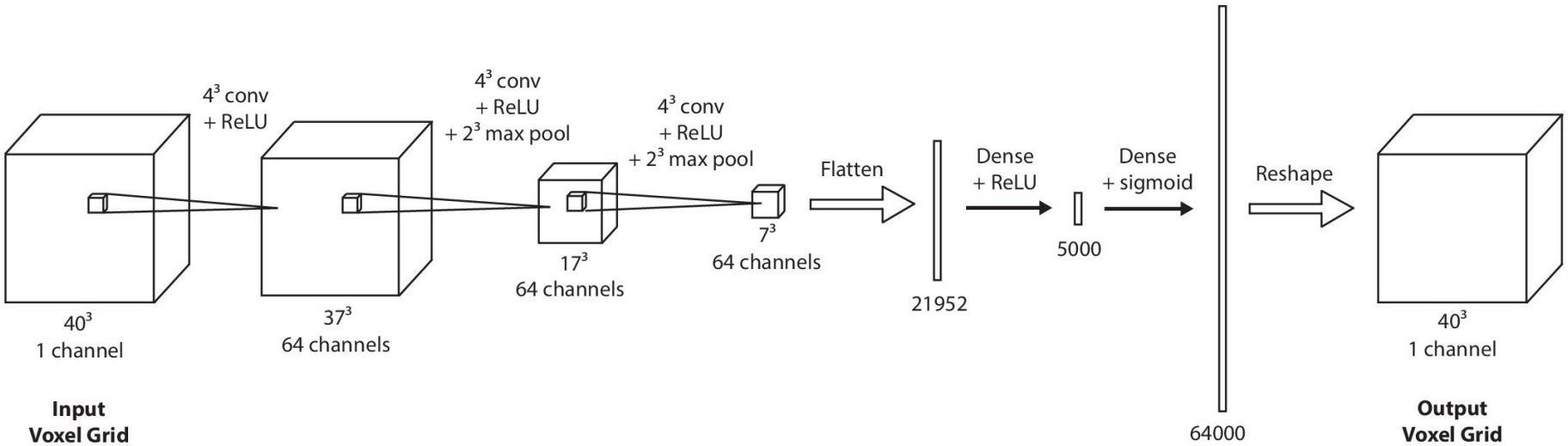
Capturing depth samples



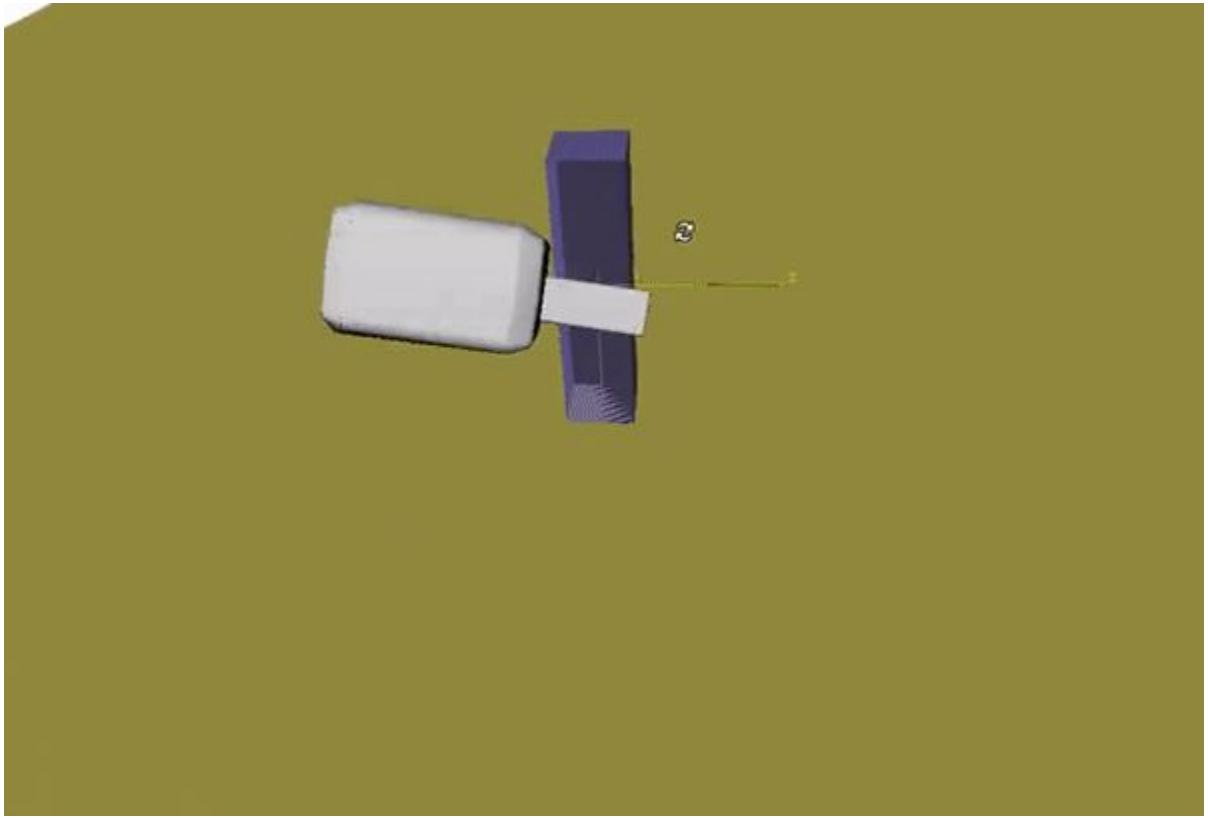


Depth and tactile completion

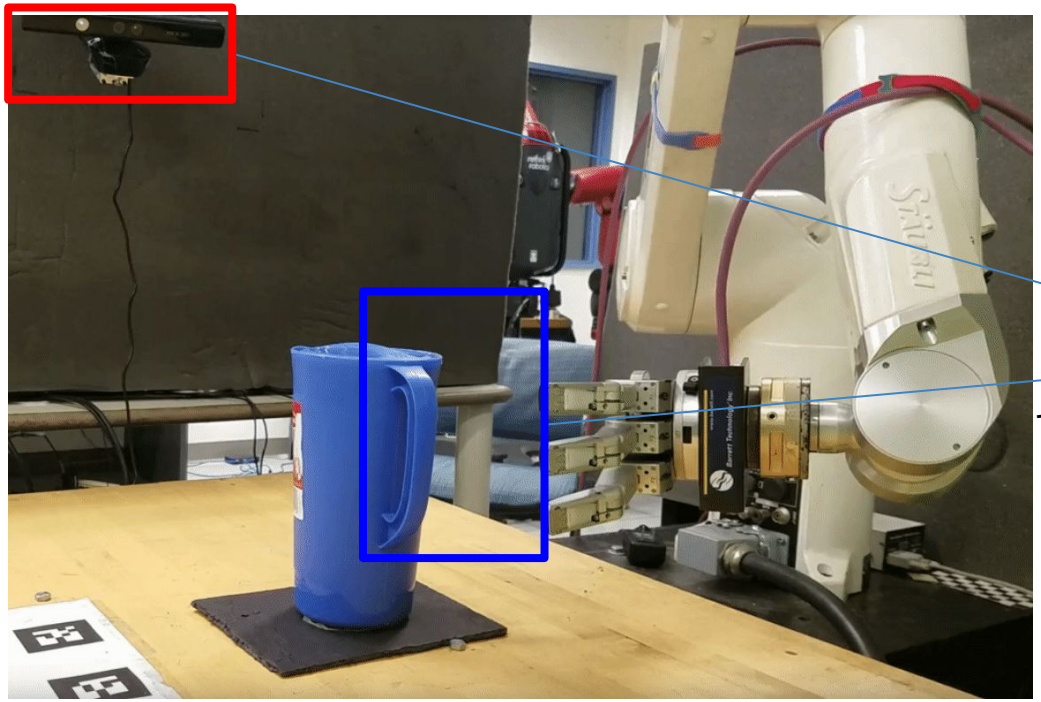




Grasp-It!

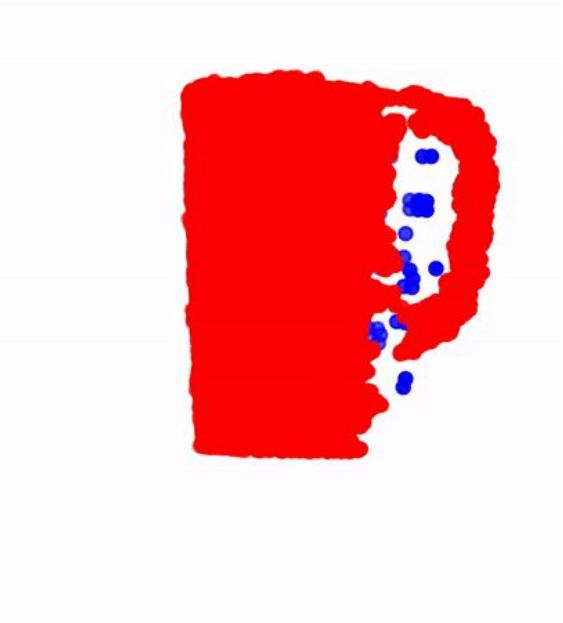


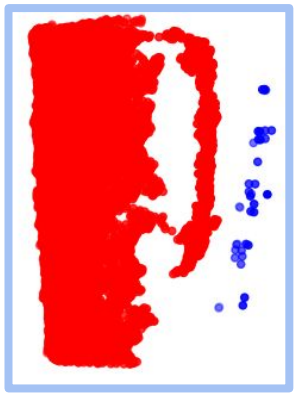
Results

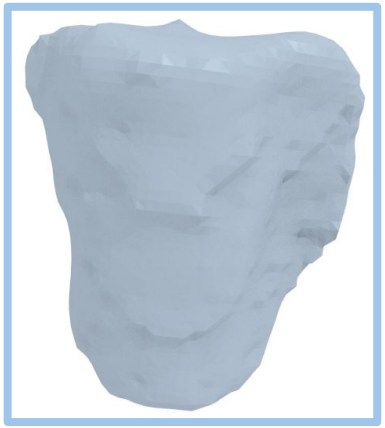


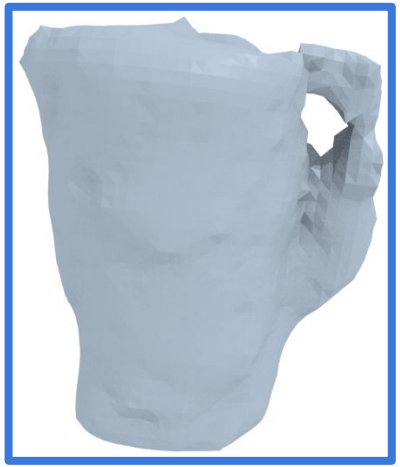
Depth (Red)

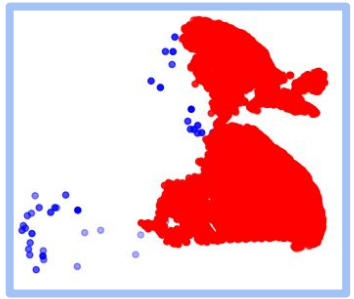
Tactile (Blue)

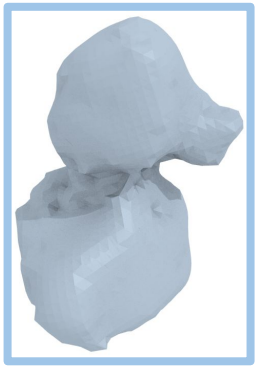


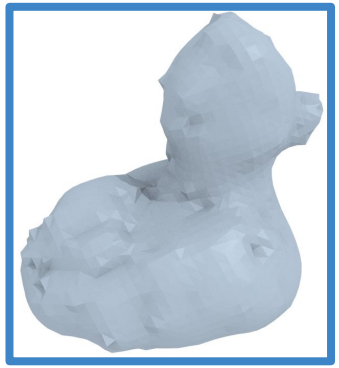












Ground Truth

Depth and Tactile Clouds

Depth Only Completion

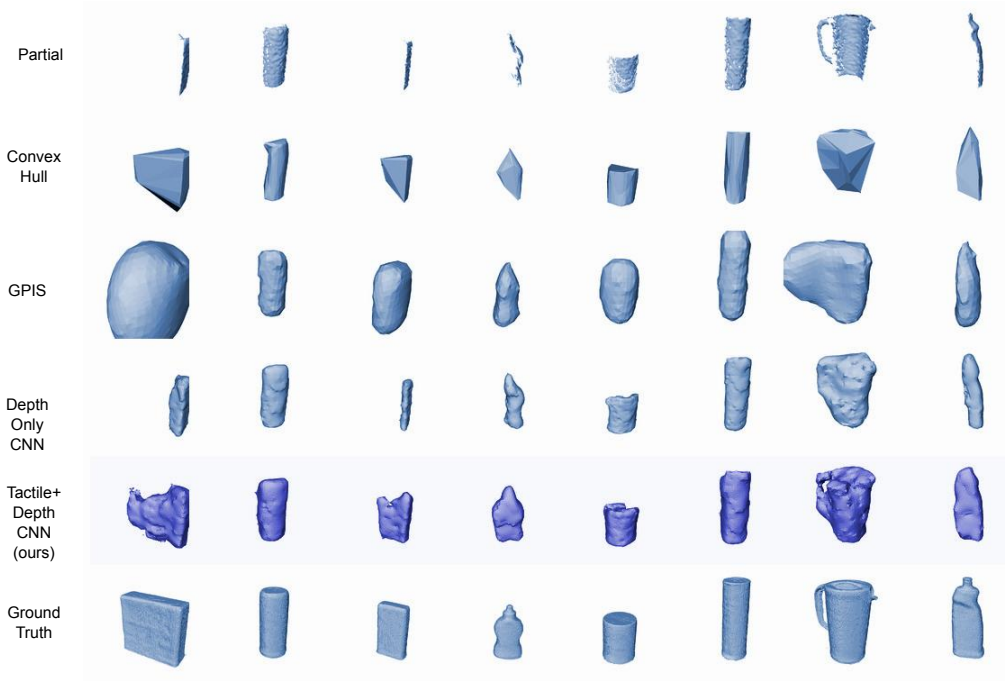
Tactile and Depth Completion (ours)

Live results

RGB

Depth and Tactile Cloud

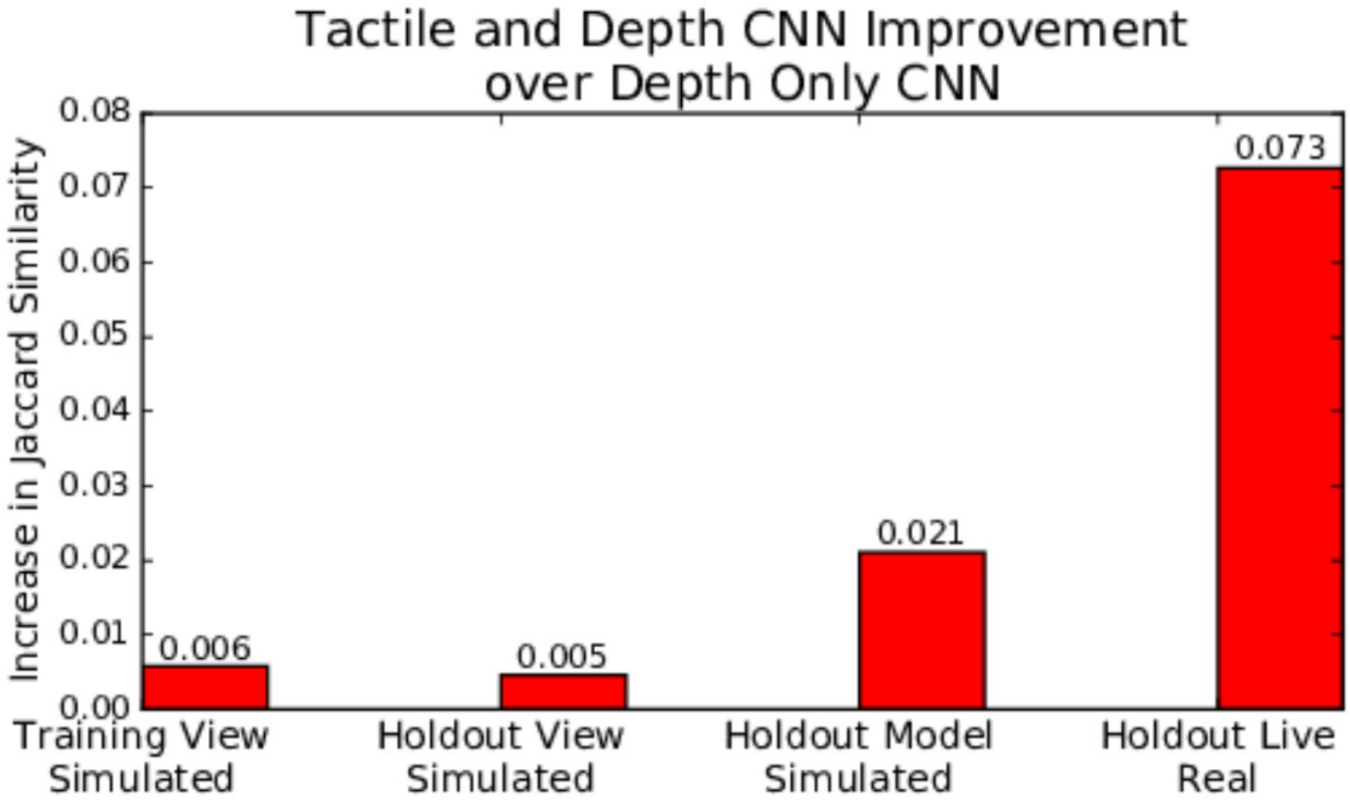
Completion comparisons



Completion results

Completion Method	Train View(Sim)	Holdout View(Sim)	Holdout Model(Sim)	Holdout (Live)
Partial	7.8	7.0	7.6	11.9
Convex Hull	32.7	45.1	49.1	11.6
GPIS	59.9	79.2	118.0	17.9
Depth CNN	6.5	6.9	6.5	16.5
Ours	**5.8**	**5.8**	**6.2**	**7.4**

Hausdorff distance measuring the mean distance in millimeters from points on one mesh to points on another mesh

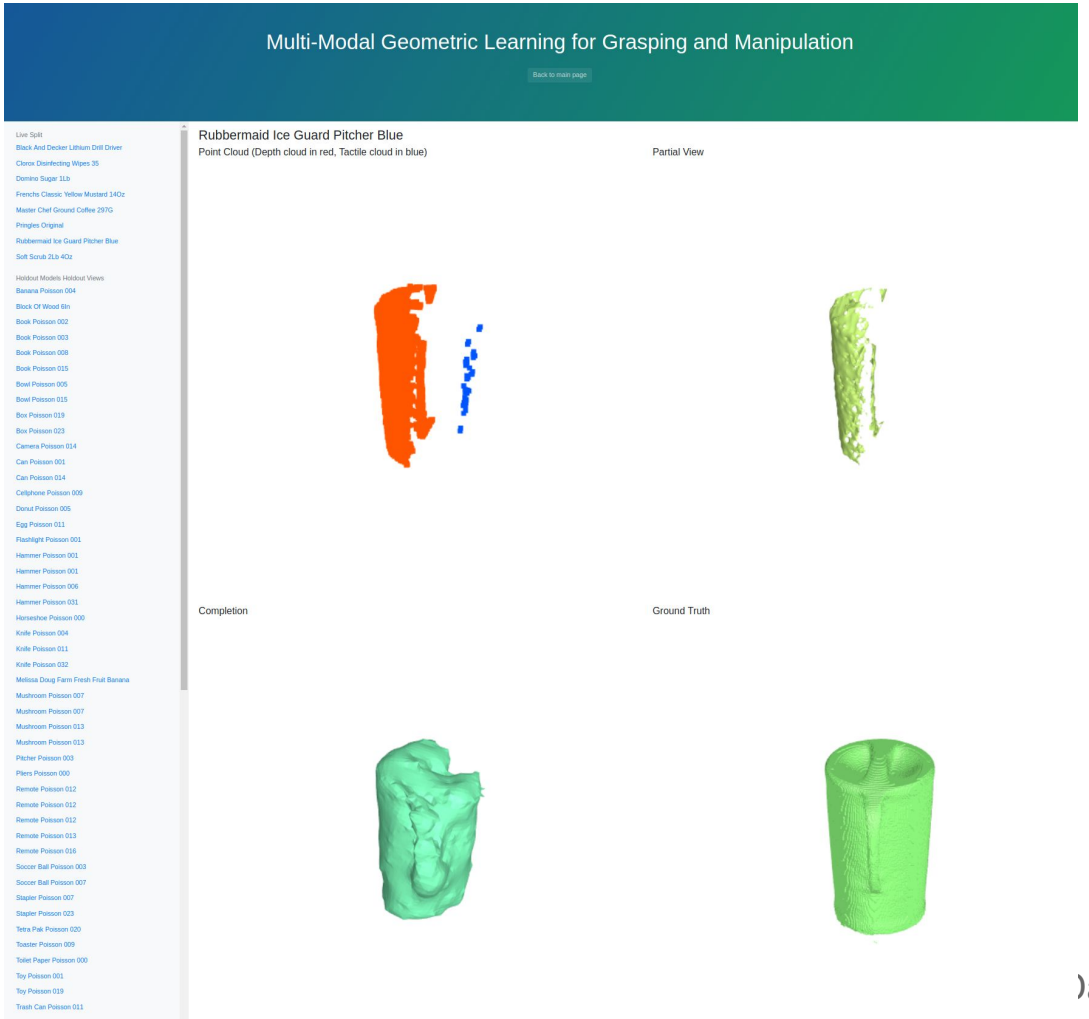


Grasping Results

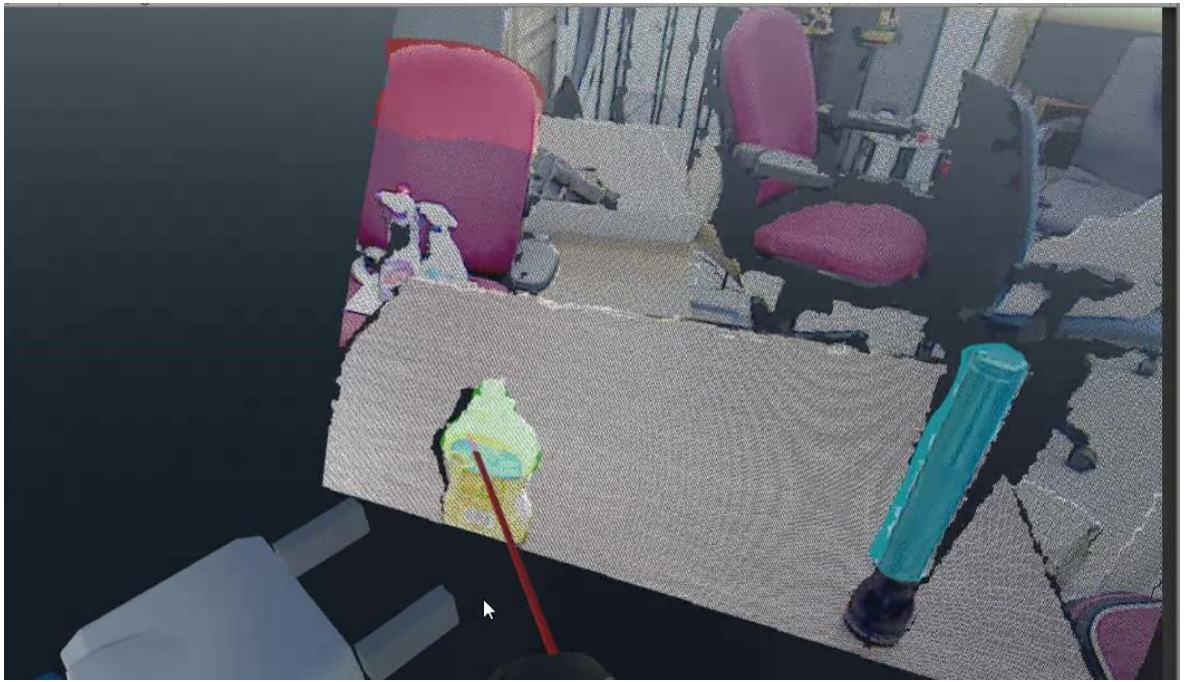
Completion Method	Train View(Sim)	Holdout View(Sim)	Holdout Model(Sim)	Holdout (Live)
Partial	19.9mm	21.1mm	16.6mm	18.6mm
Convex Hull	13.9mm	16.1mm	14.1mm	10.5mm
GPIS	17.1mm	16.0mm	21.3mm	20.8mm
Depth CNN	12.1mm	13.7mm	12.4mm	22.9mm
Ours	**7.7mm**	**13.9mm**	**13.6mm**	**6.2mm**

L2 difference between planned and realized grasp pose
averaged over the 3 finger tips and the palm of the hand

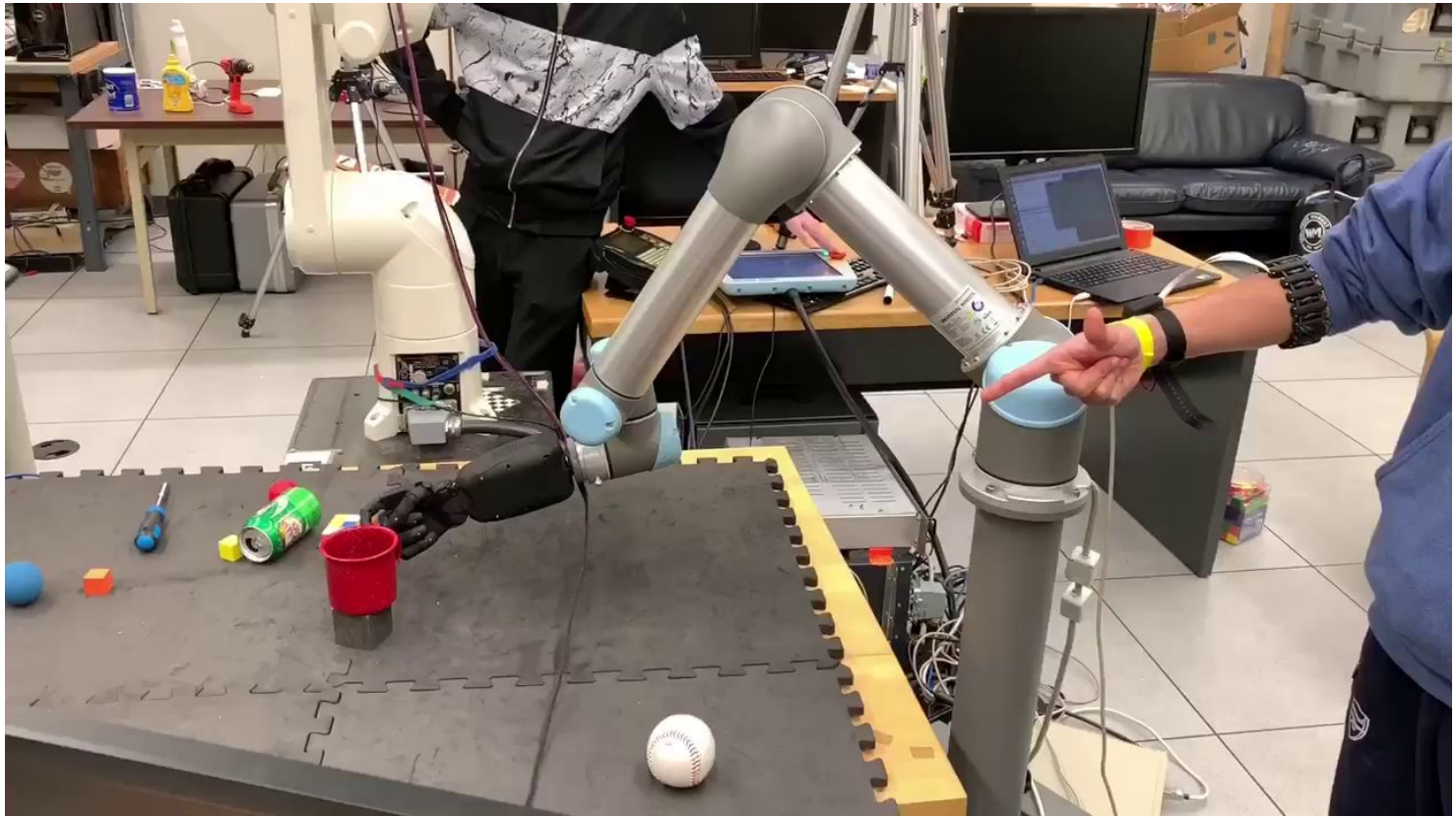
crlab.cs.columbia.edu/visualtactilegrasping



Virtual reality



Remote teleoperation



Next Steps

- RGB voxel grid
- Affordance labeling of output voxels
- Segmentation of resultant voxel grids
- Next best touch

- Higher resolution grasping
- Scene segmentation

Multi-Modal Geometric Learning for Grasping and Manipulation

David Watkins-Valls, Jacob Varley, Peter Allen

We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research
This work is supported by NSF Grant CMMI 1734557