Cross-Embodiment Robotic Manipulation: Unifying Grippers Across Robots

IRVL
INTELLIGENT ROBOTICS AND VISION LAB
UT DALLAS

Yu Xiang

Assistant Professor

Intelligent Robotics and Vision Lab

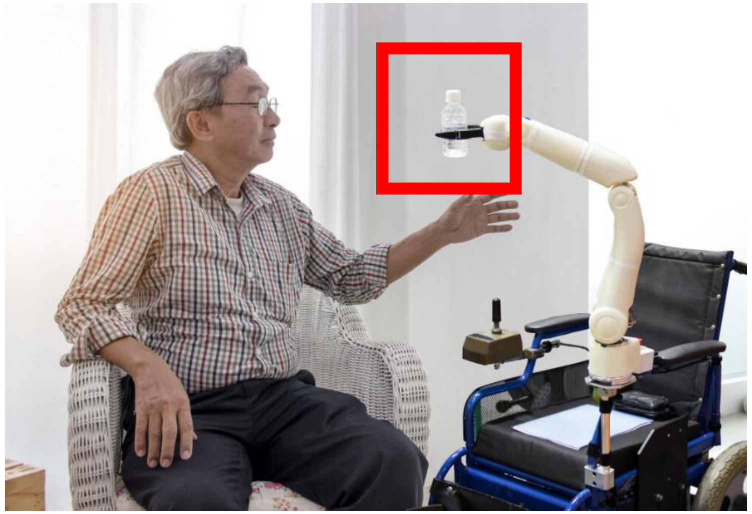
University of Texas at Dallas

4/3/2026

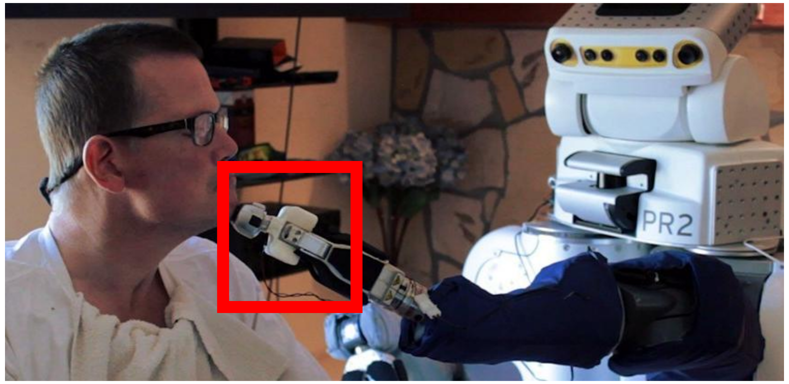
University of Texas at Arlington, TEROS 2026

Future Intelligent Robots in Human Environments

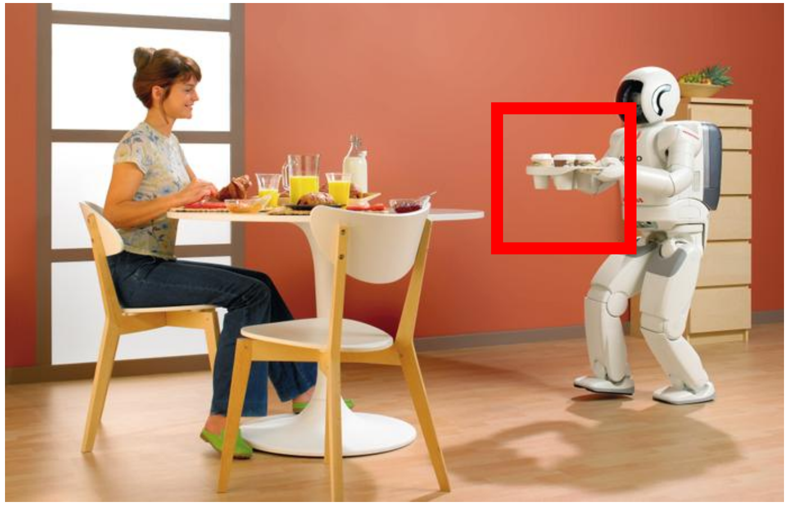
Manipulation



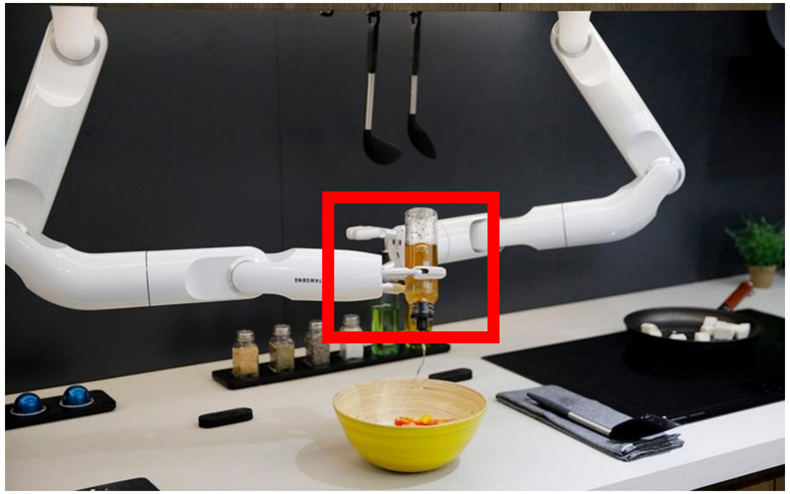
Senior Care



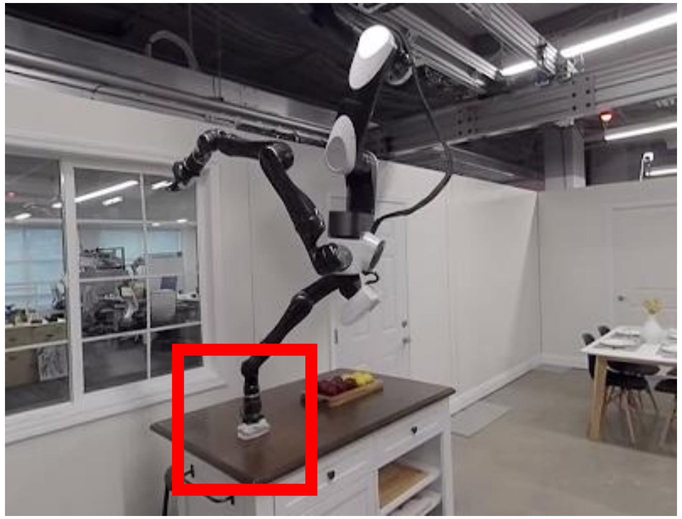
Assisting



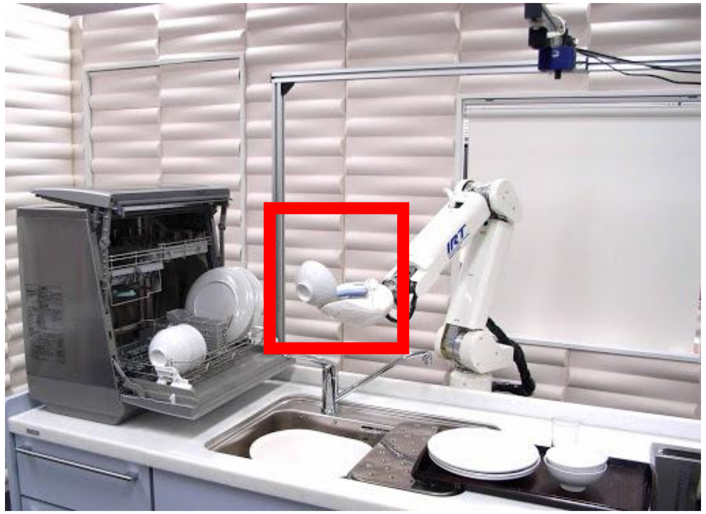
Serving



Cooking



Cleaning



Dish washing

We will have many different robots

Humanoid

Boston Dynamics Atlas

Unitree G1

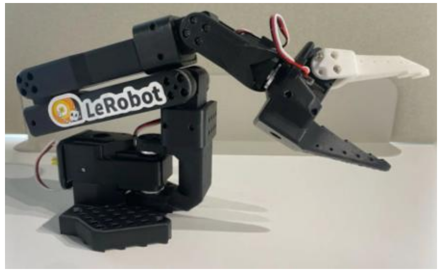
Figure

XPeng Iron

1X

Manipulator

Franka Emika



SO101 arm

Kuka arm

Mobile Manipulator

Fetch

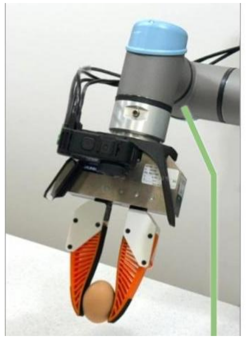
Mobile AI

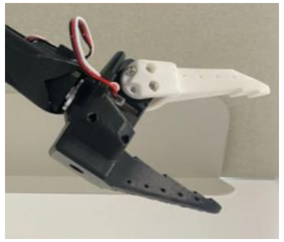
We will have many different grippers/hands


Panda


Fetch


Sawyer


UMI


SO101


Barrett

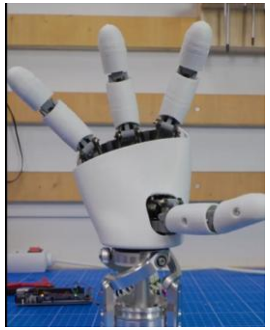
Robotiq


Atlas

Jaco


Allegro


Leap


LeRobot


Shadow


INSIPRE


PAL Hey5 hand

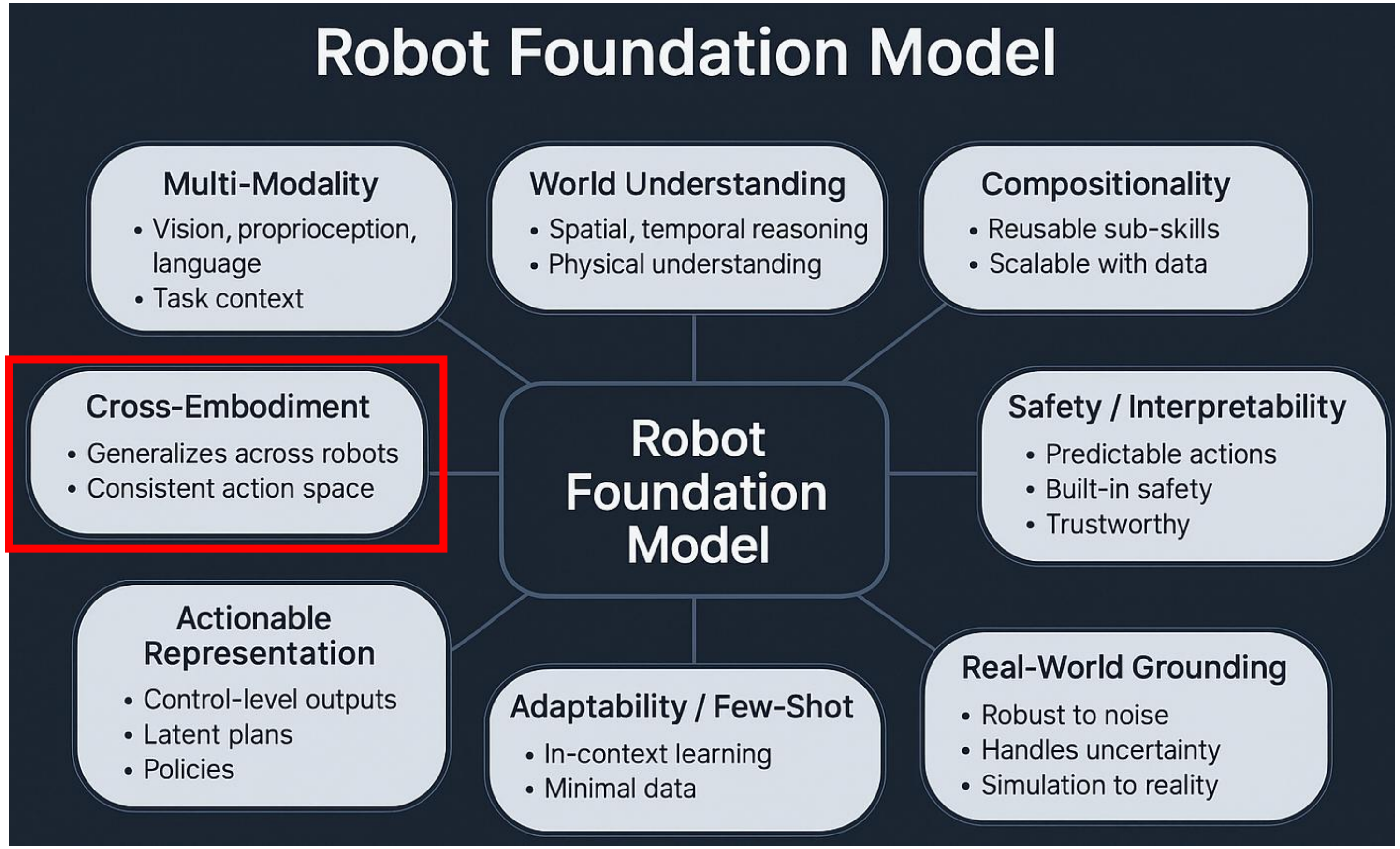

Sharpa


Aero Hand Open

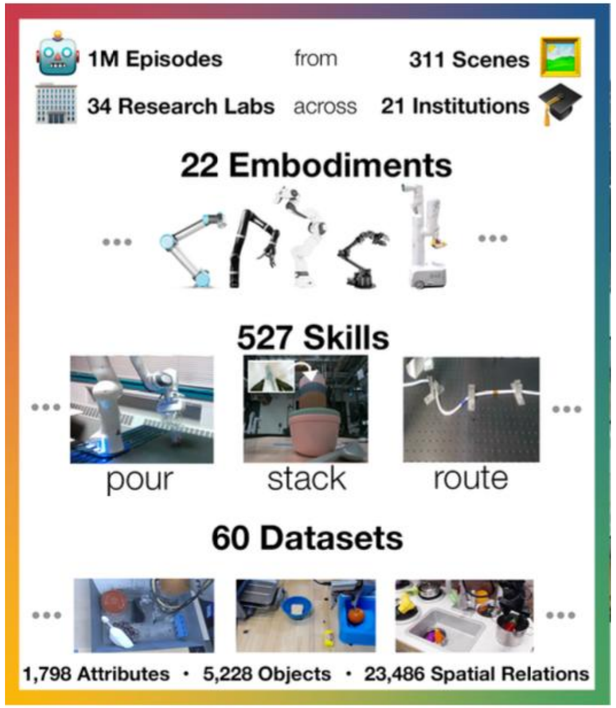
How to make all these robots work?

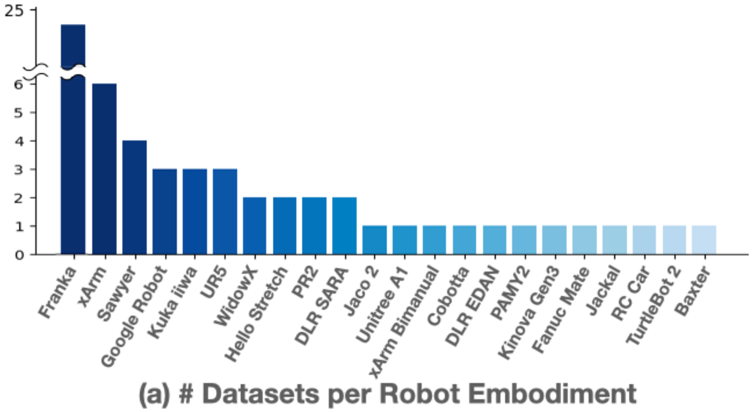
- Robot Foundation Model
  - A **foundation model** in AI refers to a **large, pre-trained model** that serves as a *base* (or "foundation") for building many downstream applications and specialized models.

Our recent focus



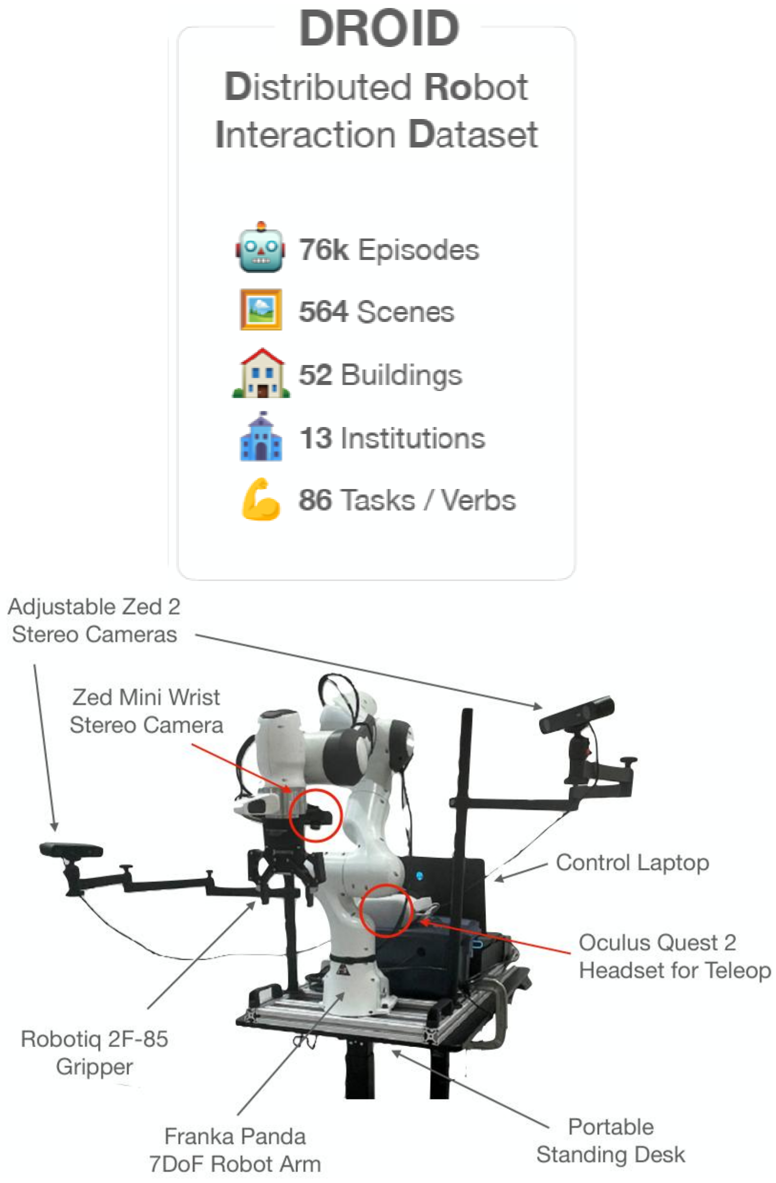
Current Robot Data





Open X-Embodiment

- Franka
- xArm
- Sawyer
- Google Robot
- Kuka iiwa
- UR5
- WidowX
- Hello Stretch
- PR2
- DLR SARA
- Jaco 2
- Unitree A1
- xArm Bimanual
- Cobotta
- DRL EDAN (5-finger)
- PAMY2
- Kinova Gen3
- Fanuc Mate
- Jackal
- RC Car
- TurtleBot 2
- Baxter

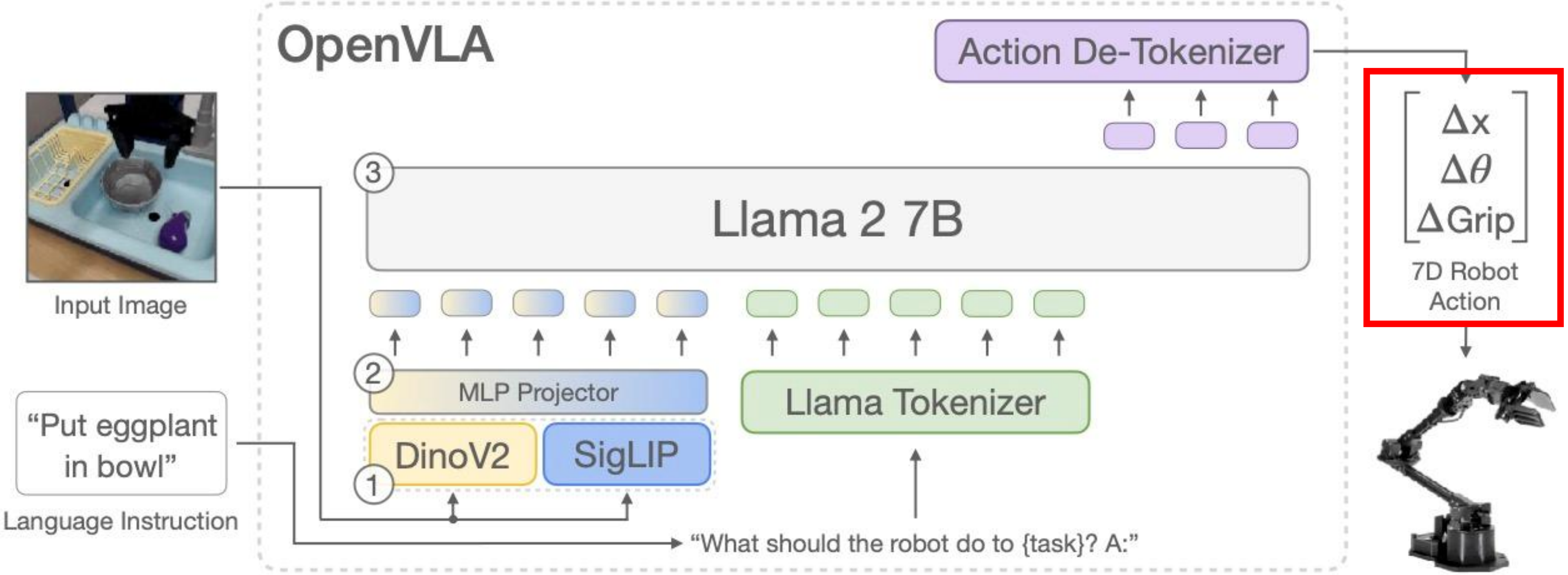


The DROID dataset

Very biased to 2-finger grippers

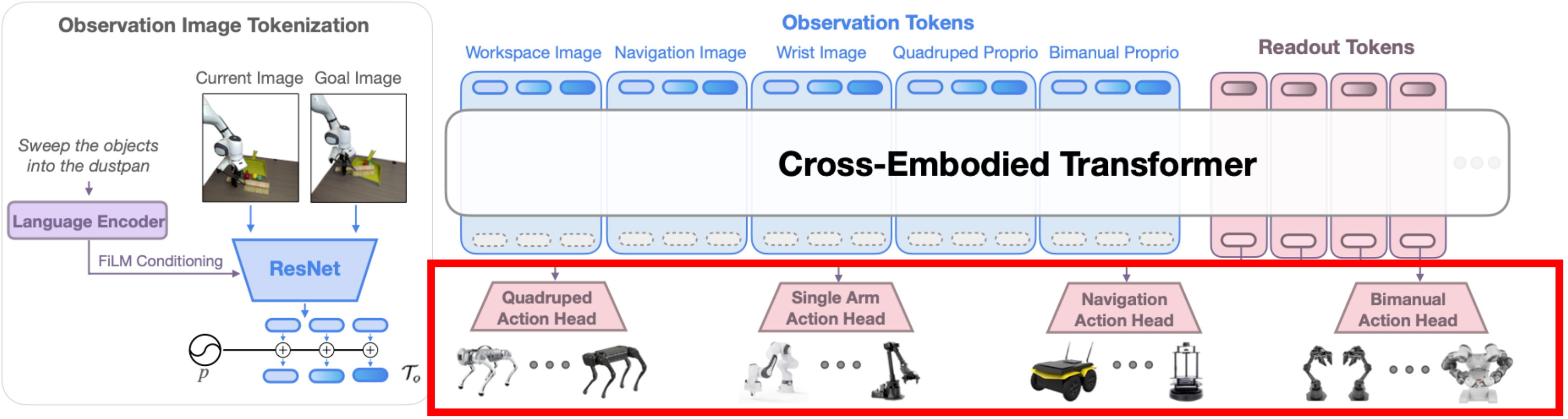
Current Model Architecture

Action
- Gripper pose for two-finger grippers
- Cannot be used for multi-finger hands (no hand joints)

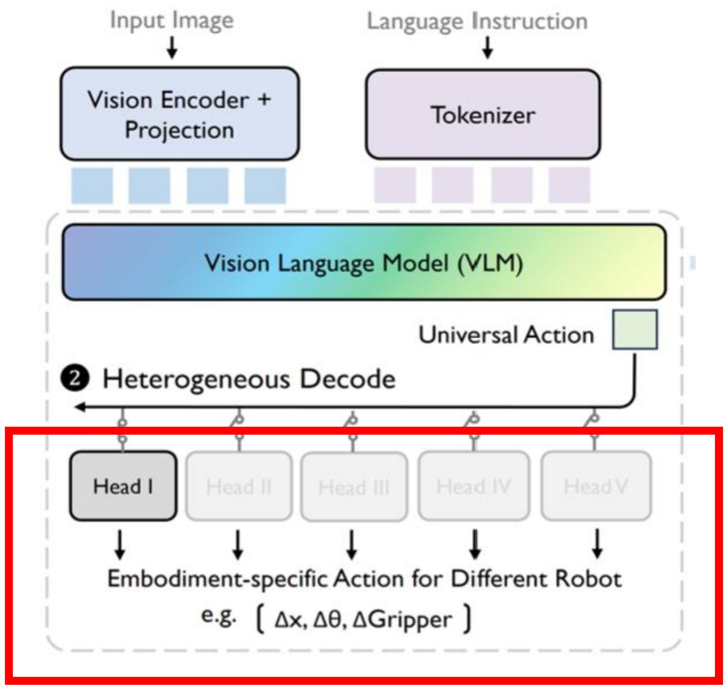


OpenVLA: An Open-Source Vision-Language-Action Model. Kim et al., 2024.

Cross-Embodiment Model Architecture



CrossFormer: Scaling Cross-Embodied Learning for Manipulation, Navigation, Locomotion, and Aviation. Doshi et al., CoRL, 2024.

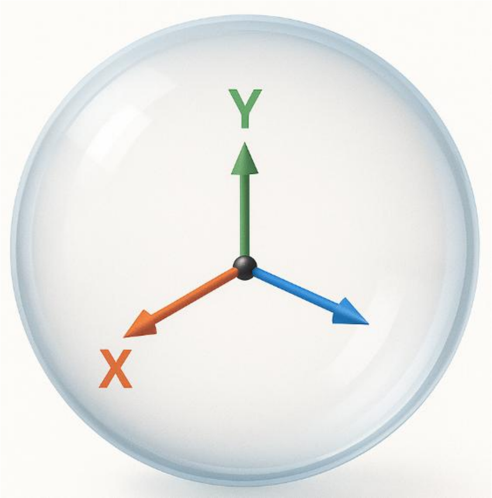


One Action head for each robot type

- Given a new robot, one new head needs to be trained

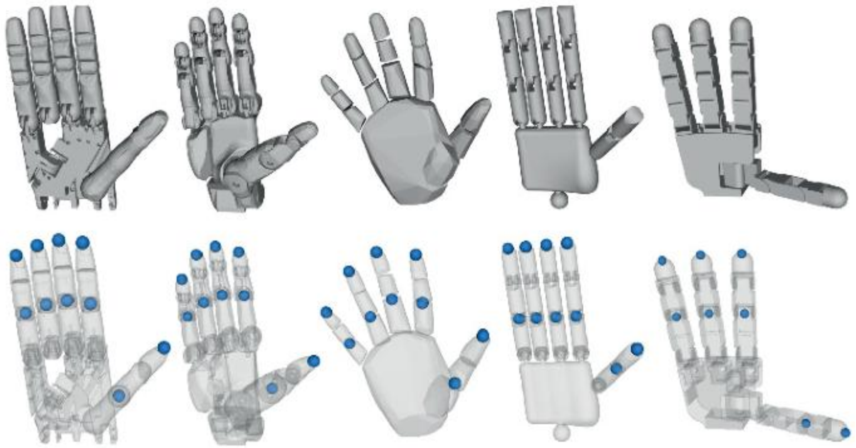
Universal Actions for Enhanced Embodied Foundation Models. Zheng et al., CVPR, 2025.

Can we find a unified action space for different robot grippers?

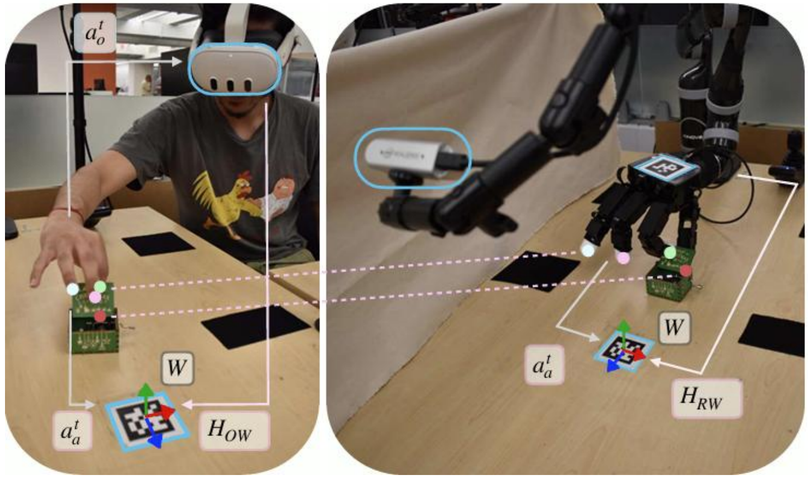


Human Hand
(human data)

Previous work: manual alignment of grippers (retargeting)



Learning Cross-hand Policies for High-DOF Reaching and Grasping (She et al., 2024)



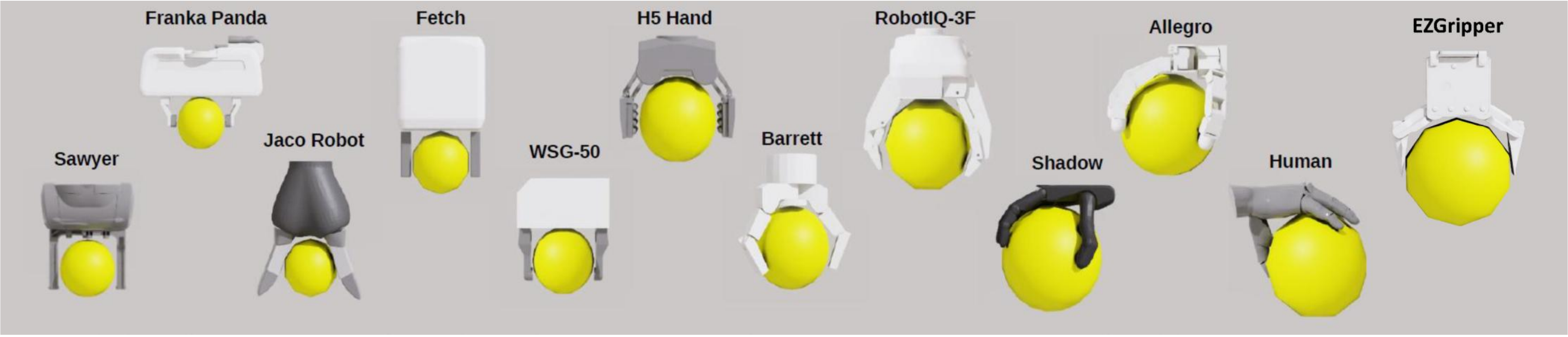
HuDOR, Guzey et al. NYU 2025

https://object-rewards.github.io/

- Manual mapping or hand-designed correspondence
- Hard to deal with different number of fingers
- Cannot handle unseen grippers

Our idea: Let's use a sphere to align grippers

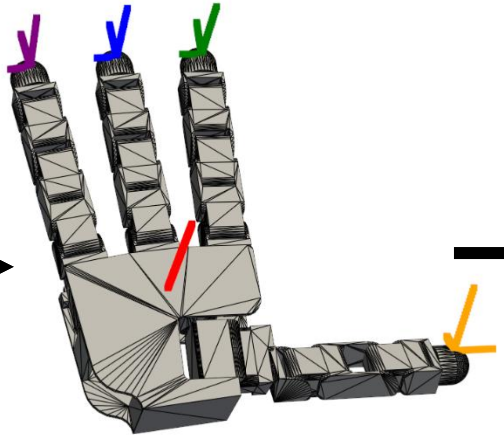
- Because any hand can grasp a sphere! (otherwise, it might not be that useful for manipulation)
- Spheres have some good properties for control (you will see)



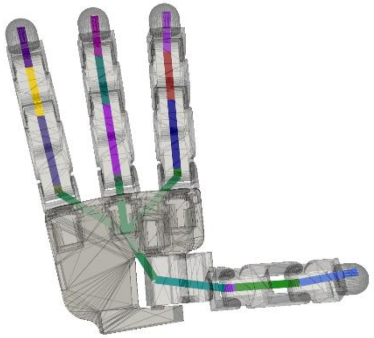
A unified action space for different robot grippers

- Sphere creation

Hand URDF

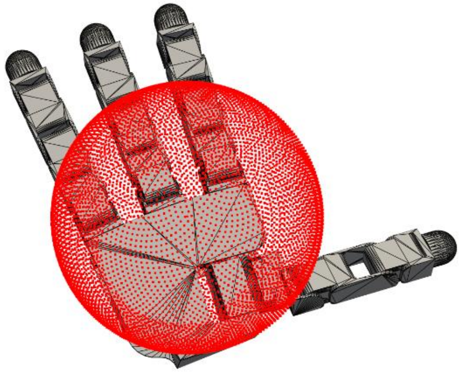


Hand URDF
- Frames for palm center and fingertips

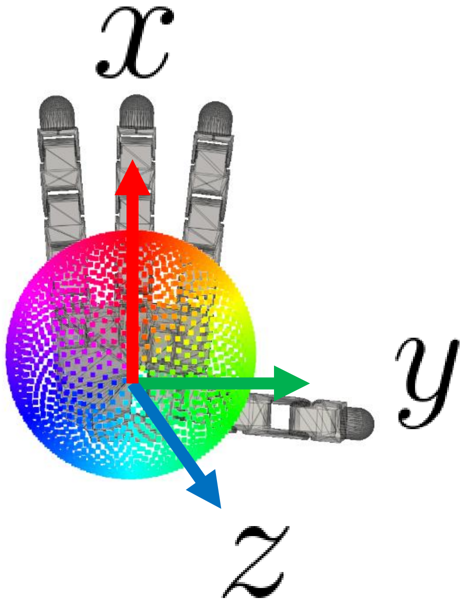


$$l = \frac{l_1 + l_2 + l_3 + l_4}{4}$$

Radius $r = l\frac{2}{\pi}$



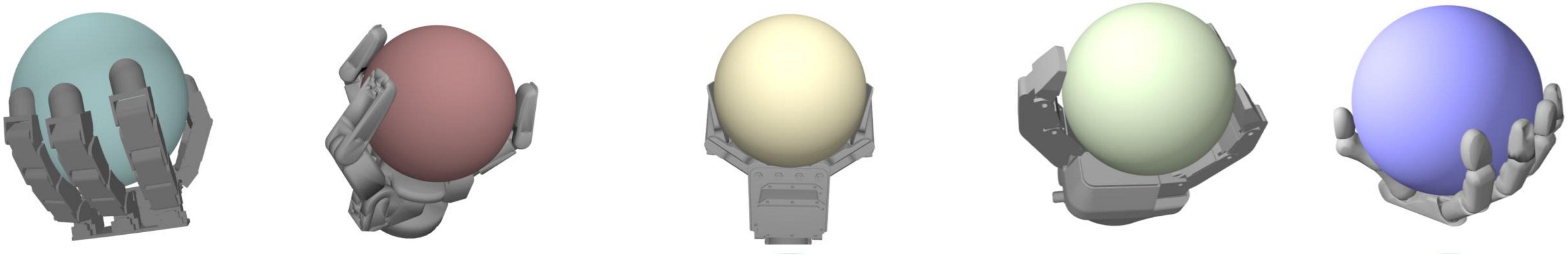
Sphere center above the palm center by r



x-axis towards the middle finger

A unified action space for different robot grippers

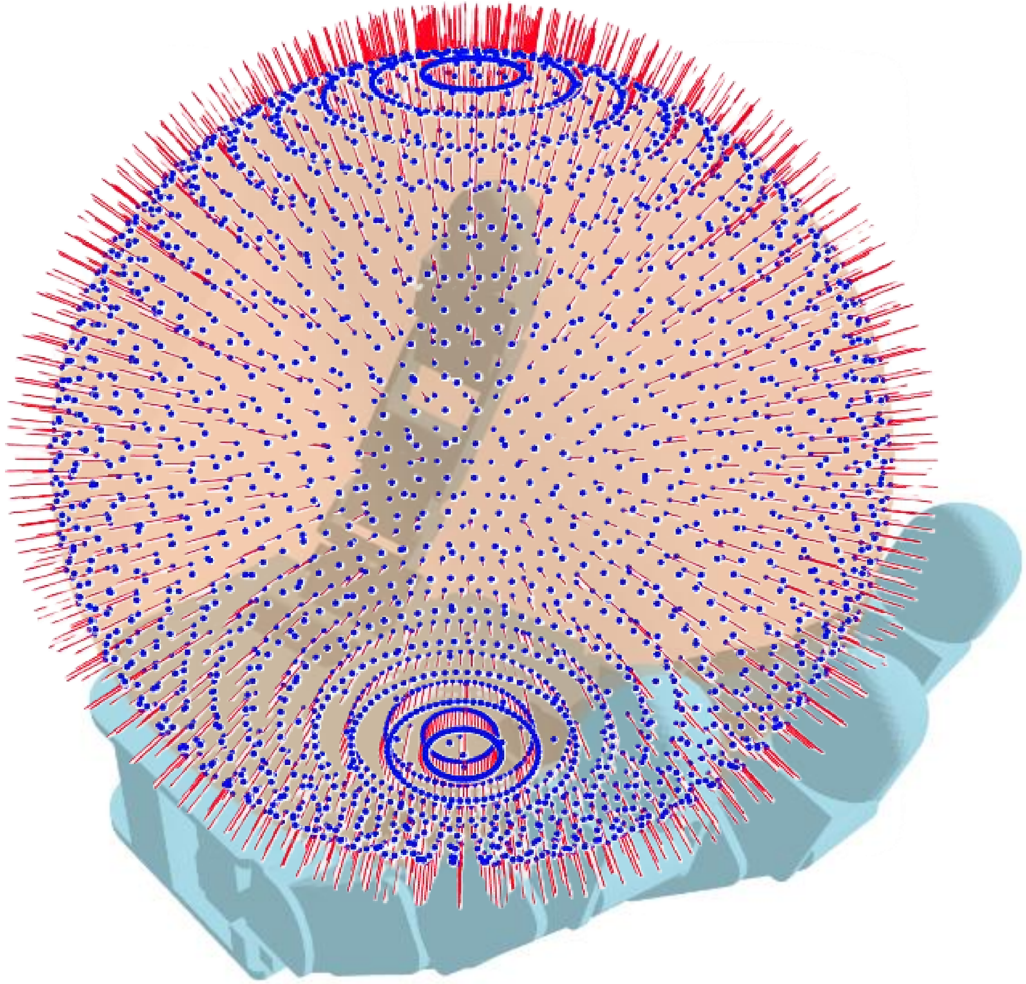
- Sphere creation applies to different grippers
- Close the fingers to grasp the sphere

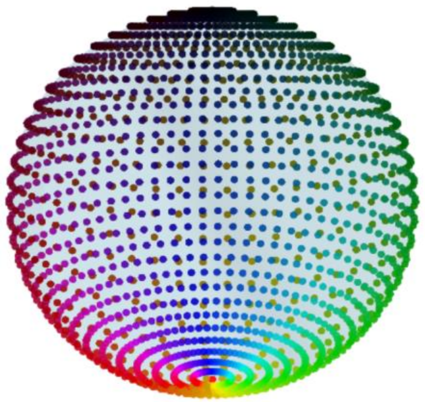


Human Hand

A unified action space for different robot grippers

- Map spherical coordinates to the gripper

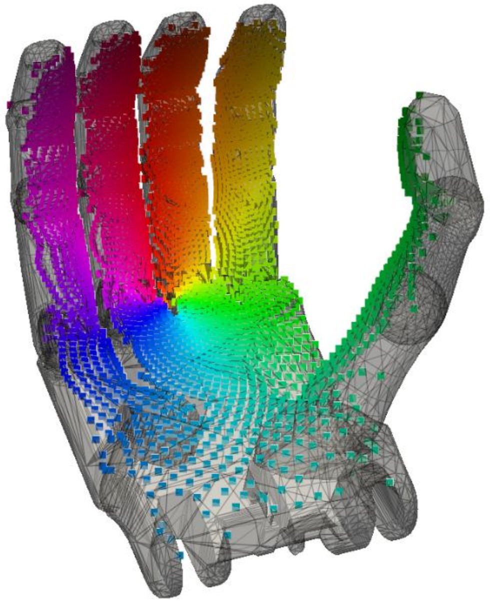


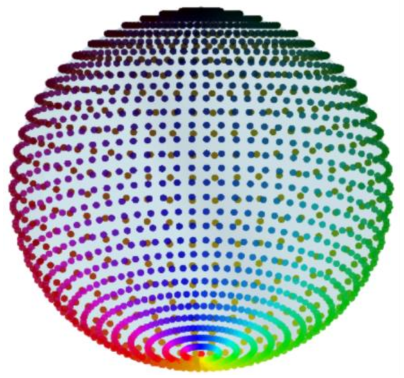


(λ, ϕ)

A unified action space for different robot grippers

- Map spherical coordinates to the gripper (a representation of the gripper)





(λ, ϕ)

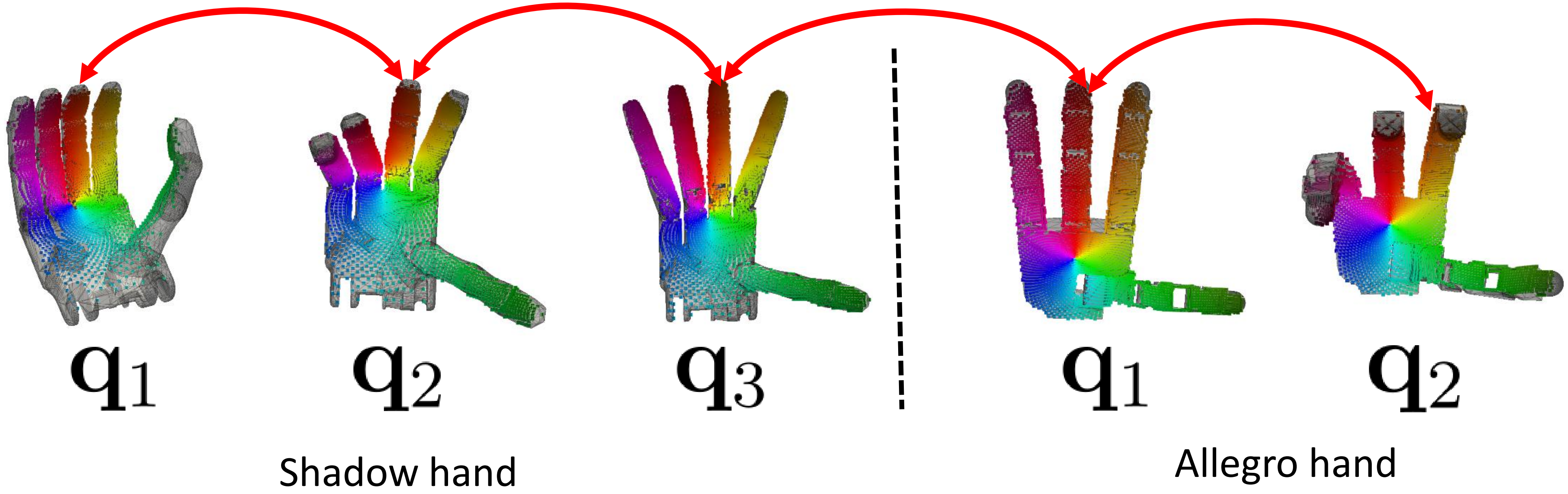
A gripper G is represented by a set of interior points $P_G = \{\mathbf{v}_g \mid \mathbf{v}_g \in \mathbb{R}^3\}$

Each point $\mathbf{v}_g$ is associated with a spherical coordinate (λ, ϕ)

Spherical coordinates $\Phi_G = \{(\lambda_{\mathbf{v}_g}, \phi_{\mathbf{v}_g}) | \mathbf{v}_g \in P_G; \lambda_{\mathbf{v}_g}, \phi_{\mathbf{v}_g} \in [0, 1]\}$

Unified Gripper Coordinate Space (UGCS)

- Property 1: the locations of the gripper points change according to grasp configuration $P_G = \{\mathbf{v}_g \mid \mathbf{v}_g \in \mathbb{R}^3\}$ $P_G(\mathbf{q})$

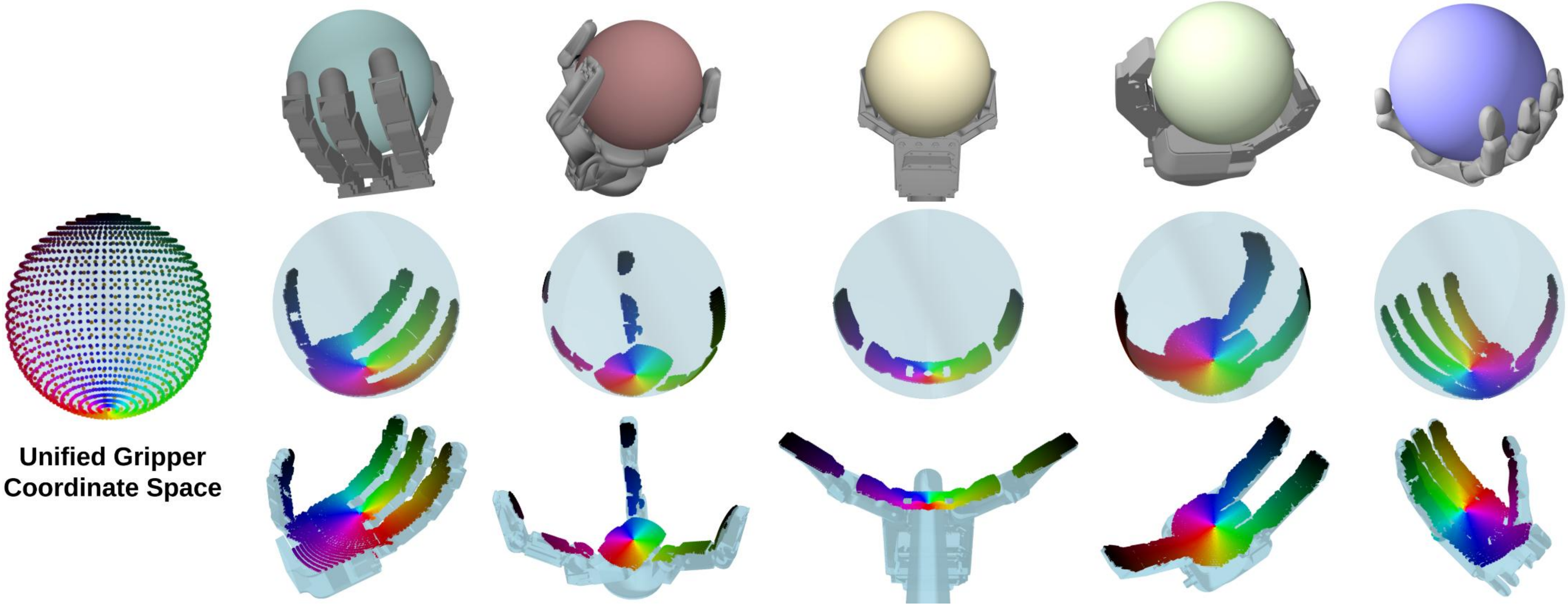


- Property 2: the spherical coordinate for each point **remains the same across configurations and hands** (correspondences!!)

Unified Gripper Coordinate Space

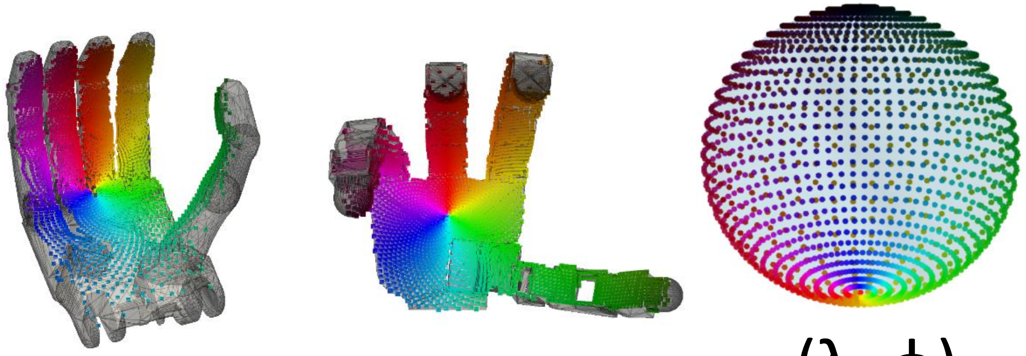
Ninad Khargonkar

Luis Felipe Casas



RobotFingerPrint: Unified Gripper Coordinate Space for Multi-Gripper Grasp Synthesis and Transfer.
Ninad Khargonkar, Luis Felipe Casas, Balakrishnan Prabhakaran, Yu Xiang. In IROS, 2025.

How can we use the UGCS representation for robot manipulation?



(λ, ϕ)

- Two applications in this talk

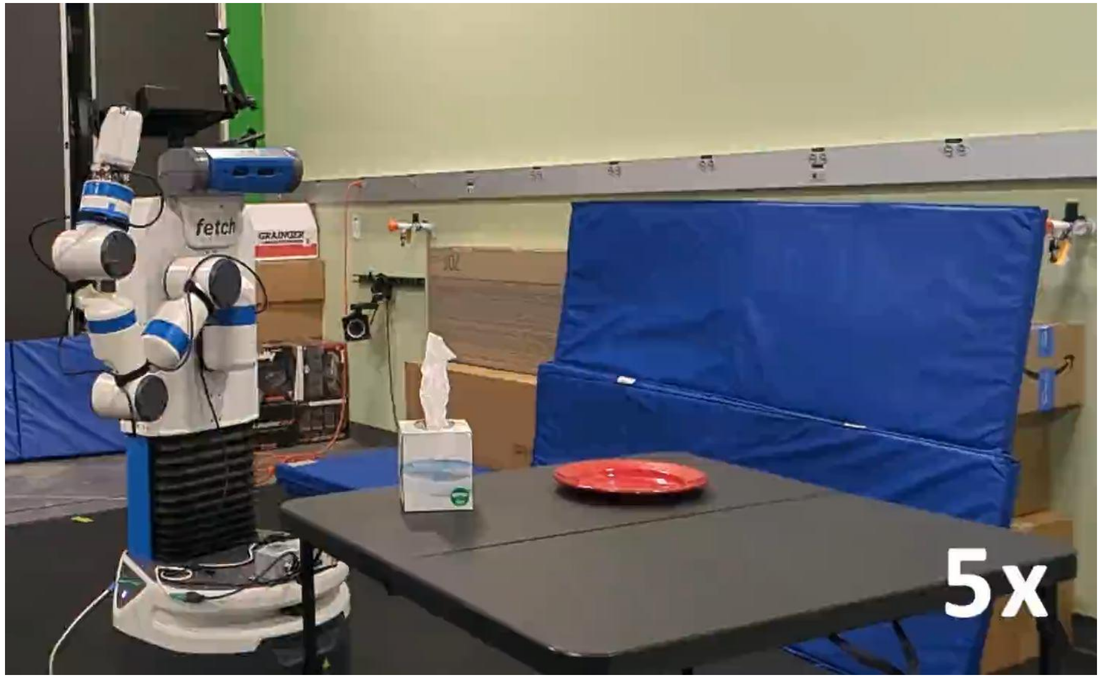
- One-shot human-to-robot trajectory transfer

- Cross-embodiment in-hand manipulation

One-Shot Human-to-Robot Trajectory Transfer

One-shot human demonstration

Robot execution in different environment



Sai Haneesh Allu

Jishnu Jaykumar P

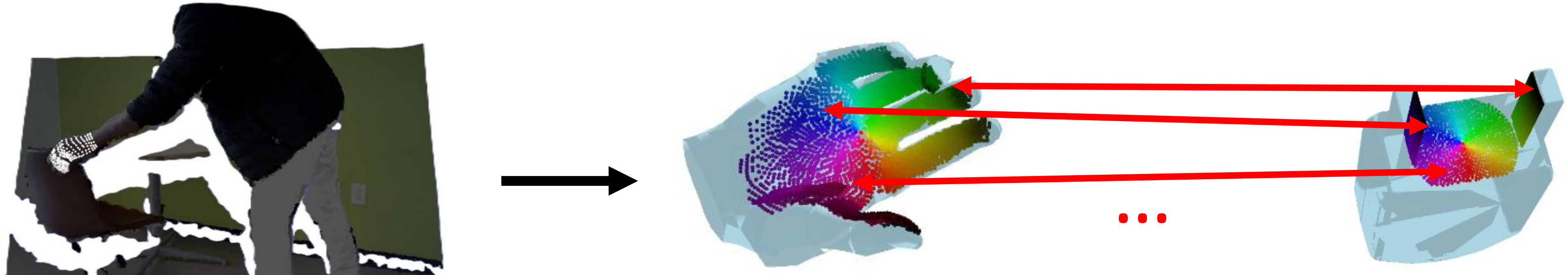
HRT1: Mobile Manipulation via One-Shot Human-to-Robot Trajectory Transfer. https://irvlutd.github.io/HRT1/
Sai Haneesh Allu*, Jishnu Jaykumar P*, Ninad Khargonkar, Tyler Summers, Jian Yao, Yu Xiang. In arXiv, 2025.

Understanding of the Human Demonstration

Hand Pose Estimation (HaMeR)

Optimization using Depth

Grasp Transfer with UGCS



A Human Grasp $\mathbf{q}_H$ A Fetch gripper $\mathbf{q}_F$

Correspondences from UGCS $P_H^c \subset P_H, P_F^c \subset P_F, |P_H^c| = |P_F^c|$

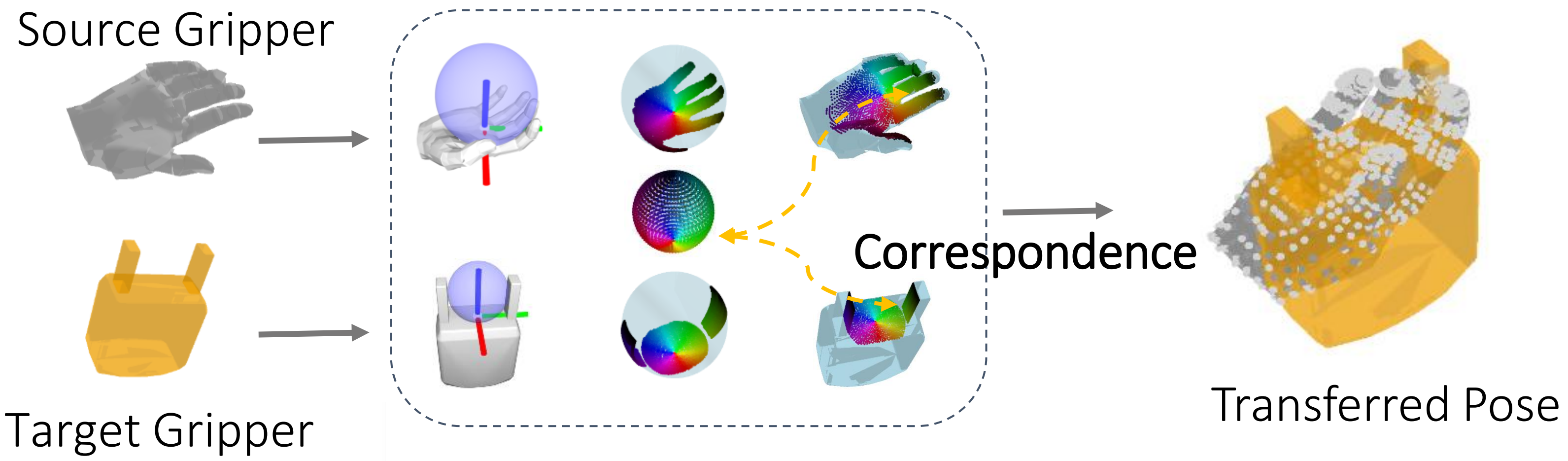
Optimize the target grasp using the

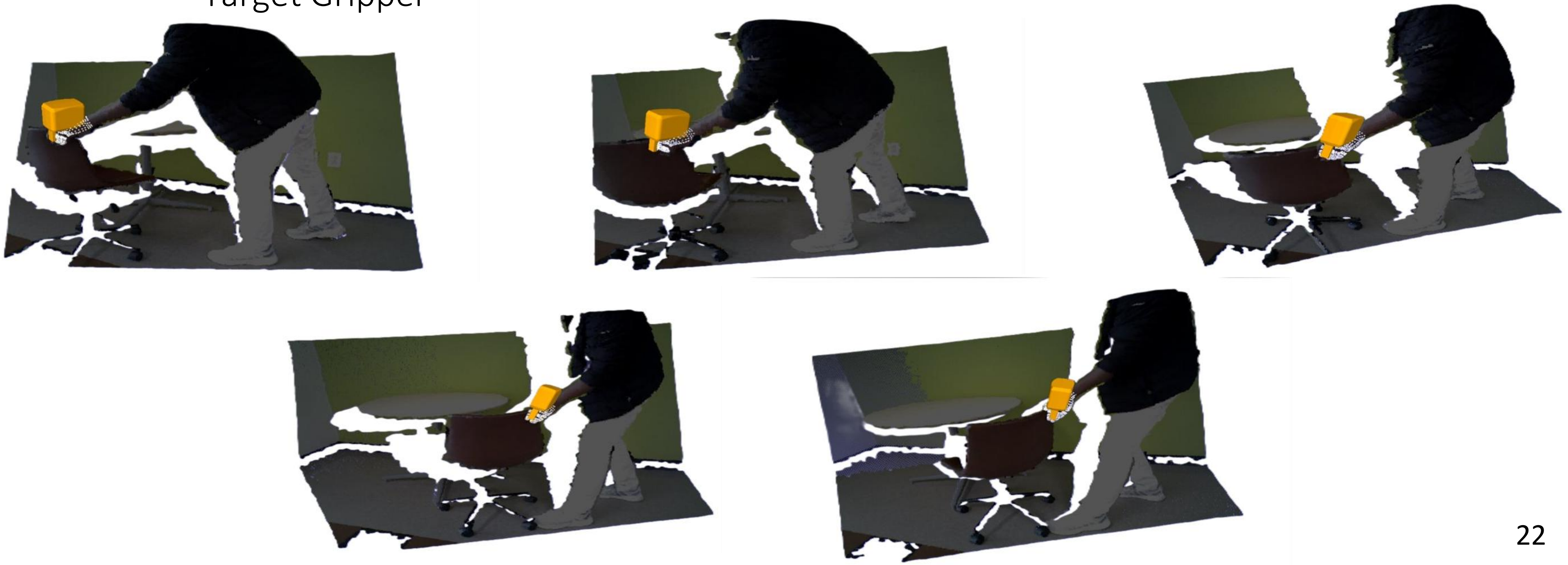
$$\mathbf{q}_F^* = \arg\min_{\mathbf{q}_F} \; E_{\mathrm{dist}}(P_H^c(\mathbf{q}_H), P_F^c(\mathbf{q}_F)) + E_{\mathrm{n}}(\mathbf{q}_F)$$

Reference grasp

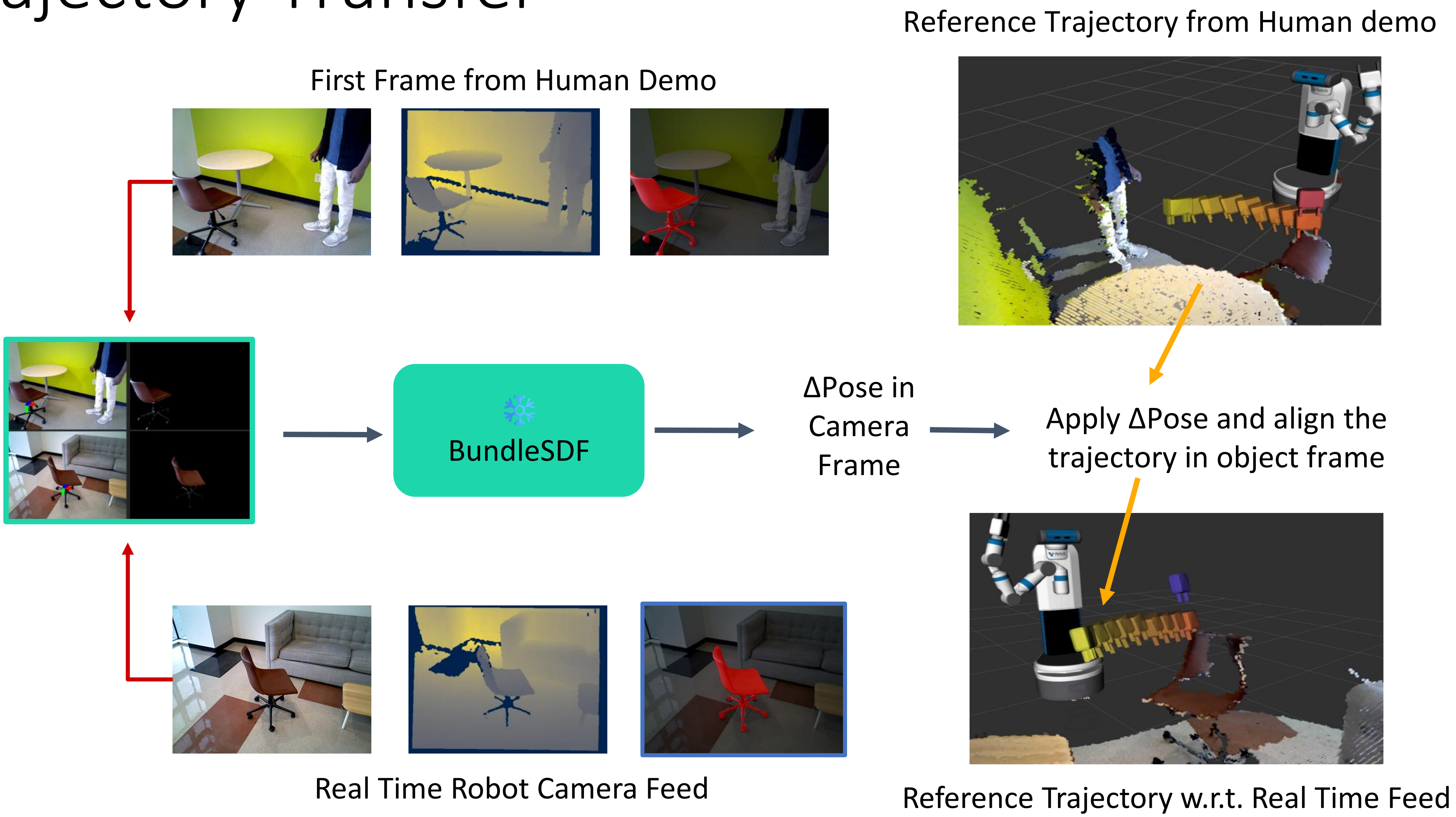
Joint limits

Understanding of the Human Demonstrations





Trajectory Transfer

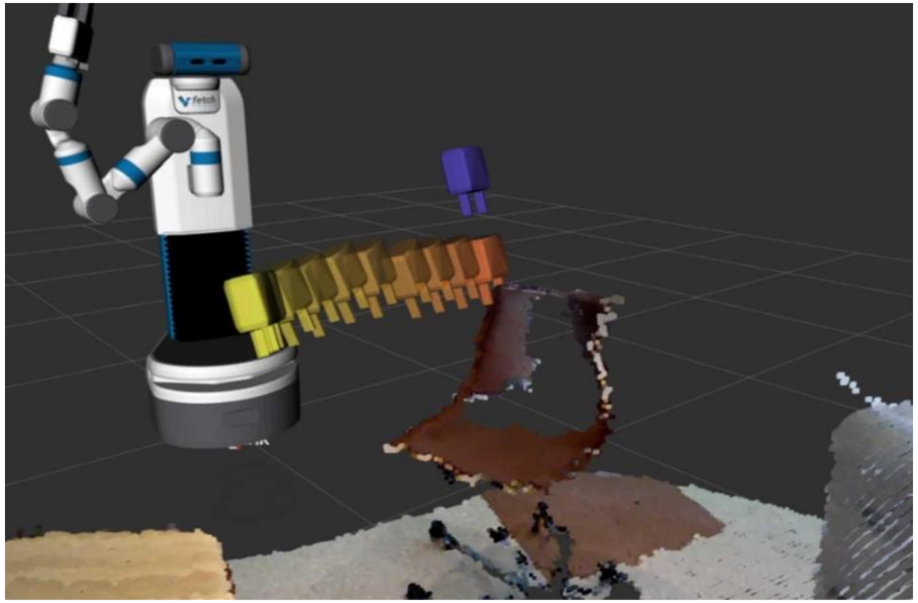


Trajectory Transfer

- How to follow the transferred gripper trajectory?

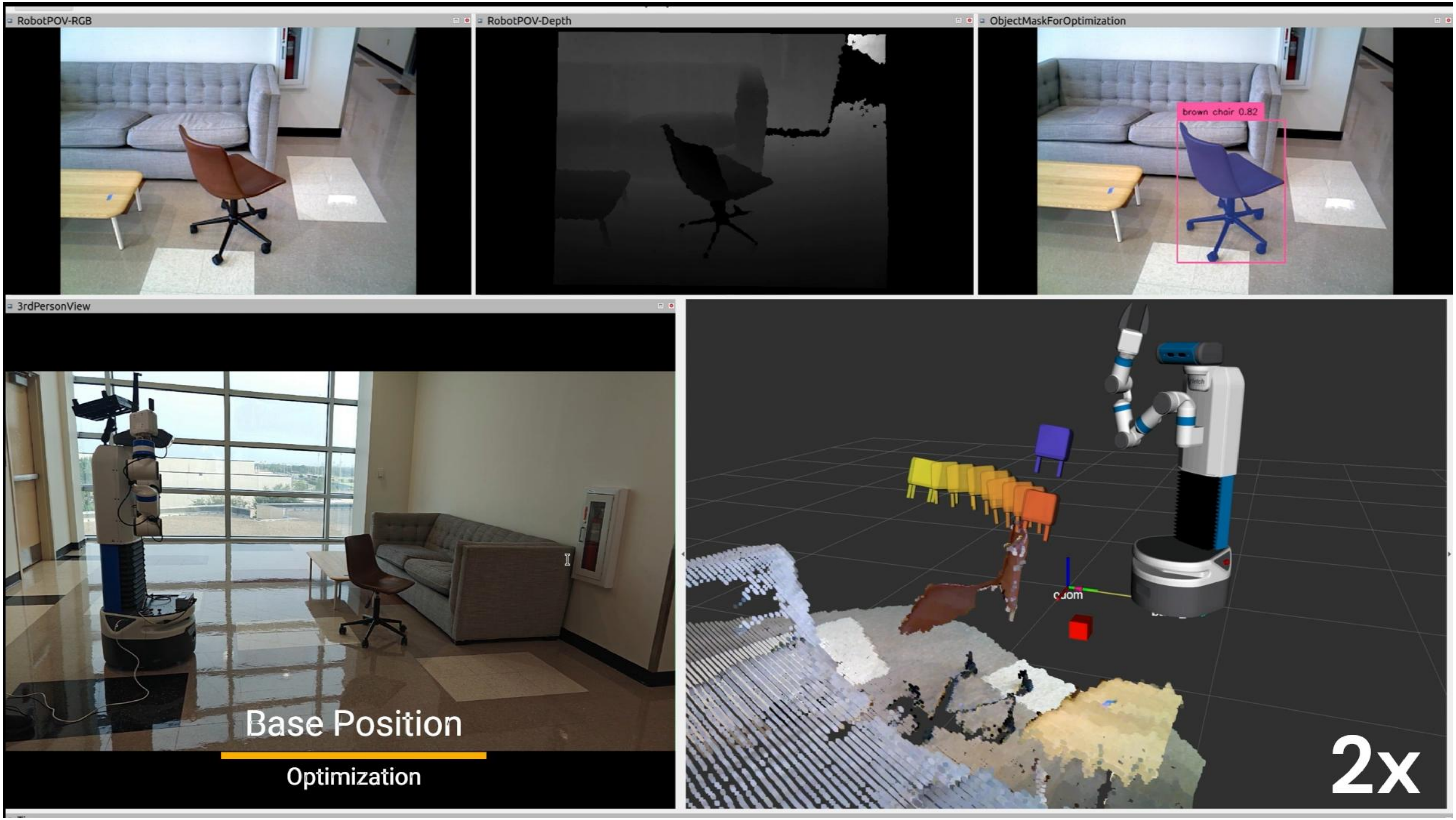
Task Space

Robot View

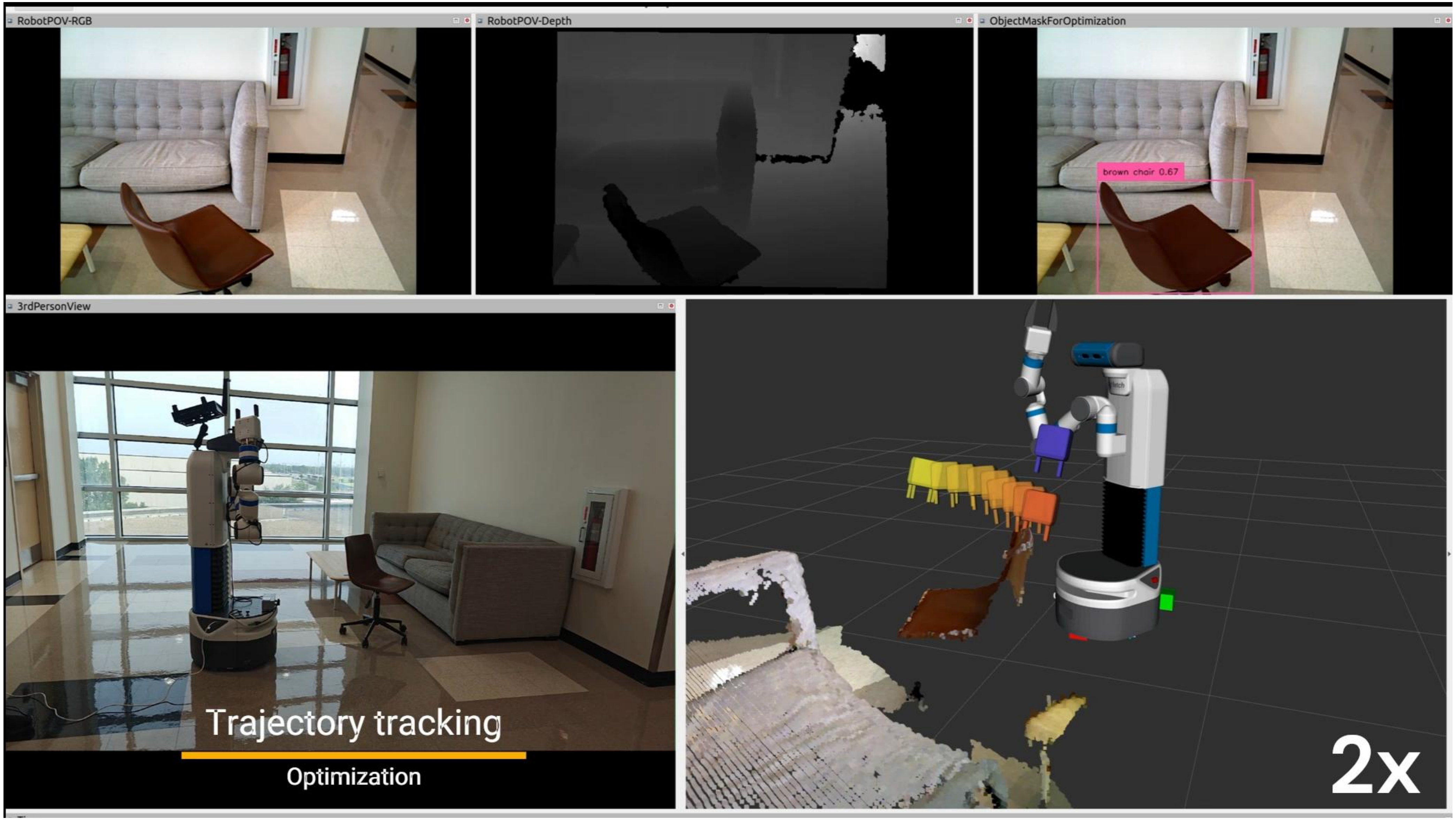


Reference Trajectory w.r.t. Real Time Feed

Optimizing the Robot Base Location

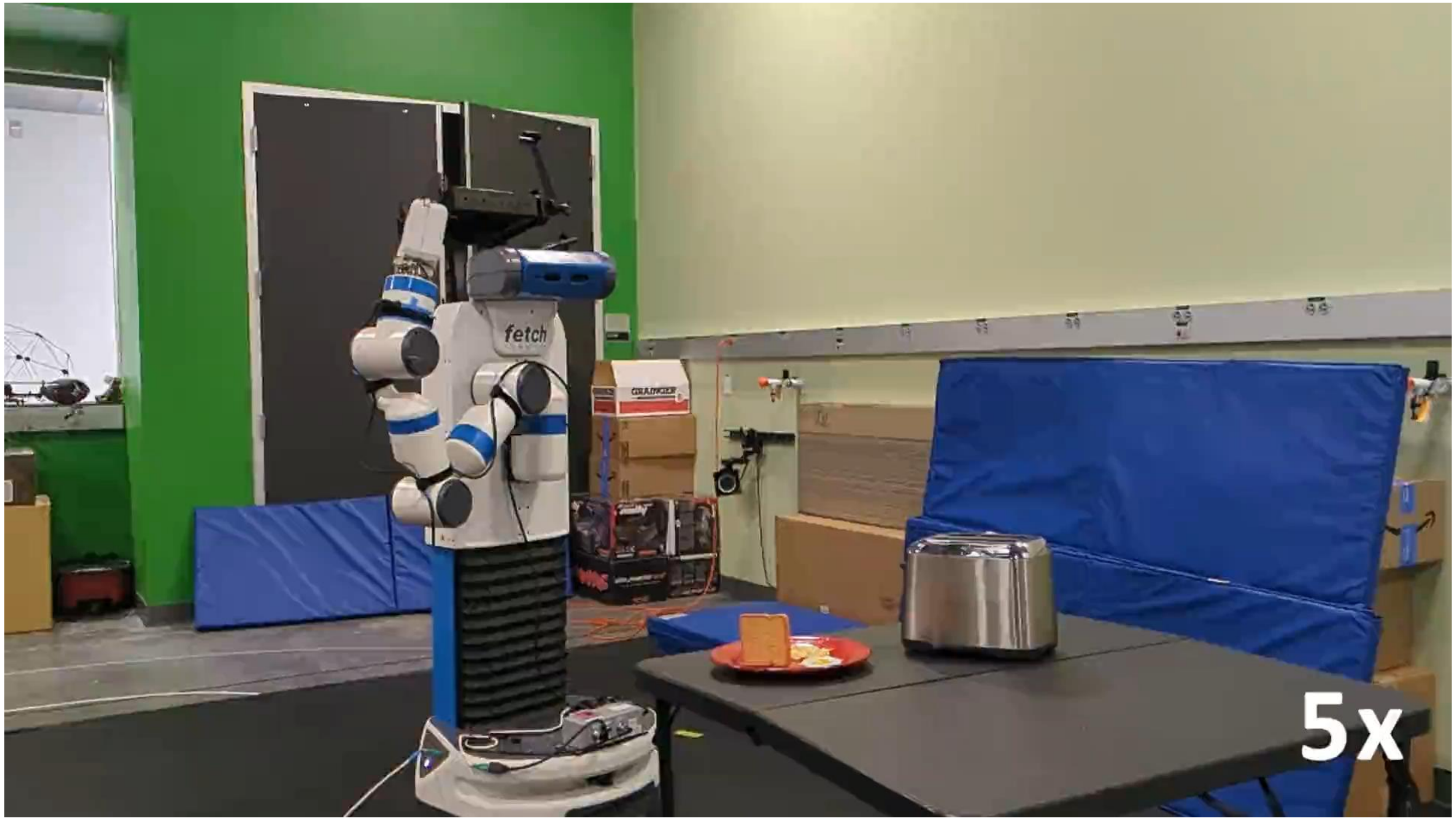


Optimizing the Robot Trajectory



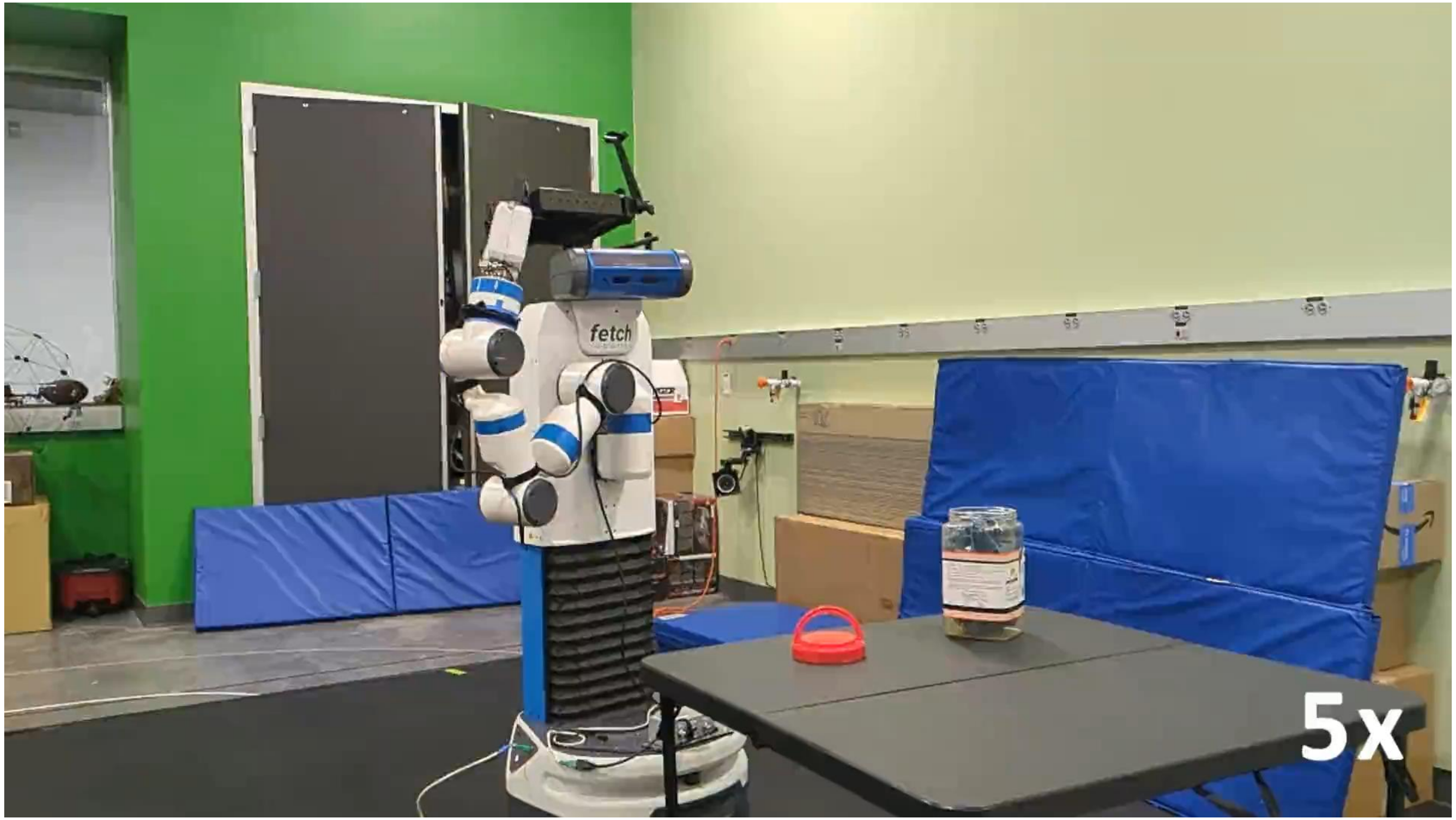
One-Shot Human-to-Robot Trajectory Transfer

Put bread in the toaster



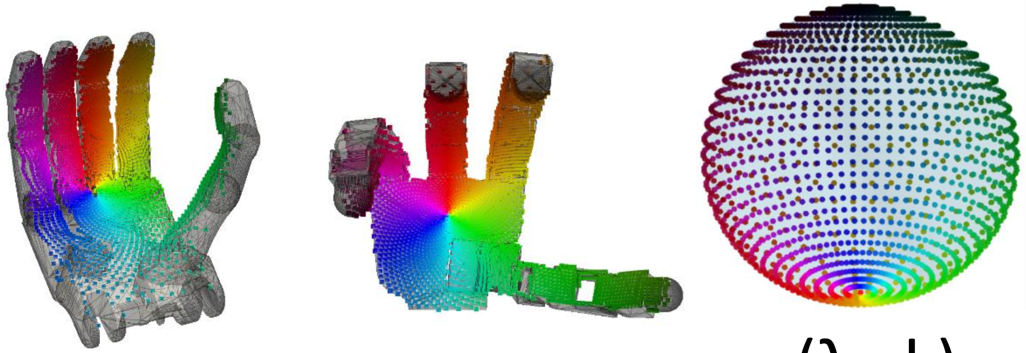
Failure Example

Close jar with a lid



The system heavily depends on object pose estimation accuracy

How can we use the UGCS representation for robot manipulation?



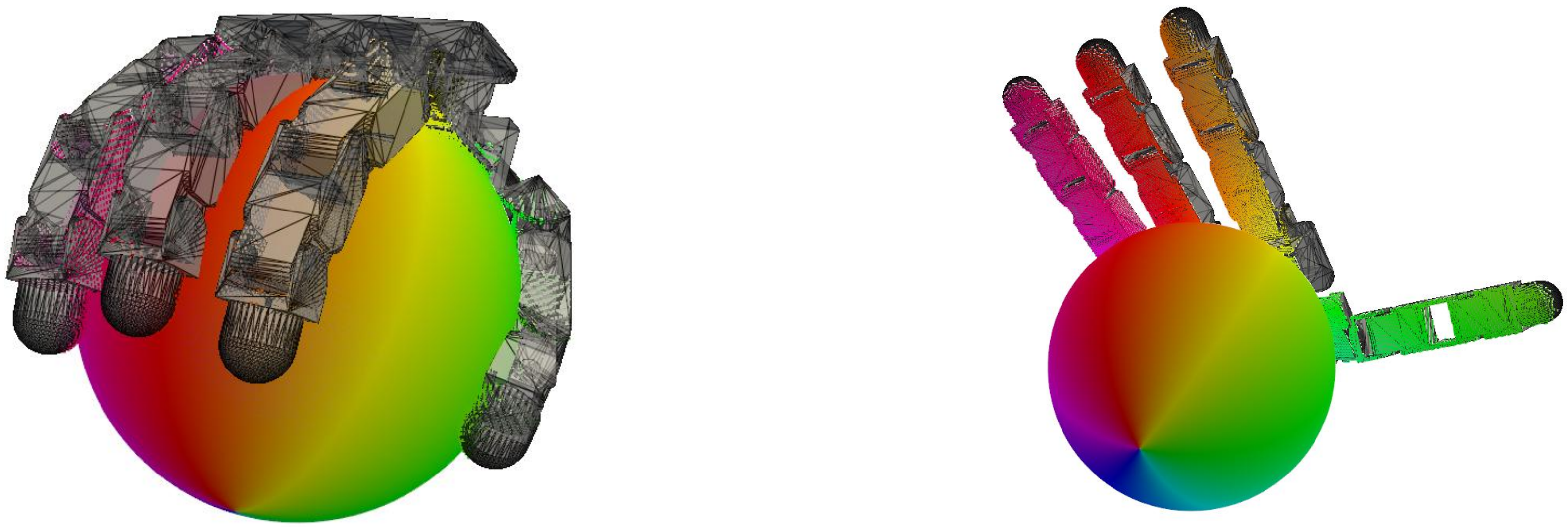
(λ, ϕ)

- Two applications in this talk

- One-shot human-to-robot trajectory transfer

- Cross-embodiment in-hand manipulation (on-going work)

Unified Gripper Action Space (UGAS)



Unified Gripper Coordinate Space

Can we use this sphere to control any robotic gripper/hand?

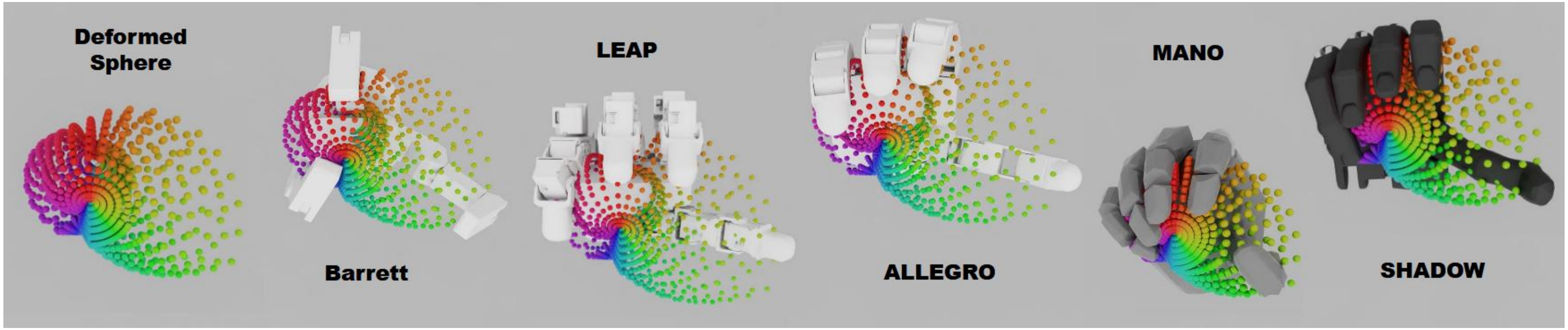
Luis Felipe Casas

Robert Teal

Keval Shah

Unified Gripper Action Space (UGAS)

- **Our Idea:** the deformation of the sphere will drive the movement of the hand (the hand should touch the deformed sphere correctly)
- Action space: deformation of the sphere (shared by any hand!)

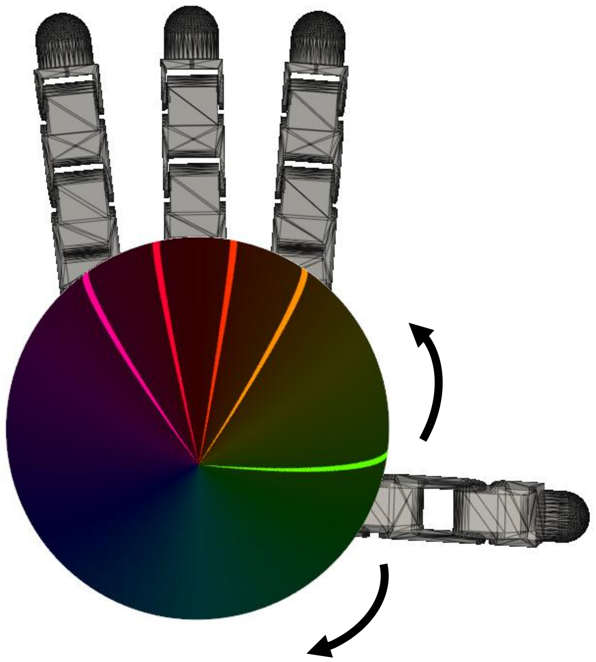


Different Grippers with the same deformed sphere

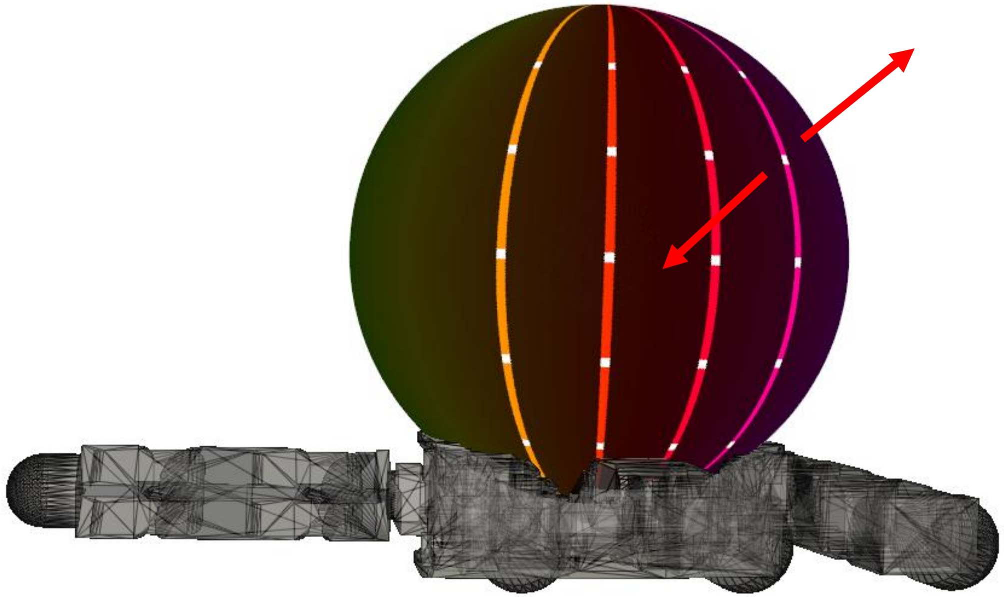
UGAS: deforming the sphere

- Deforming every point on the sphere is too expensive for control

Define several "driving planes"

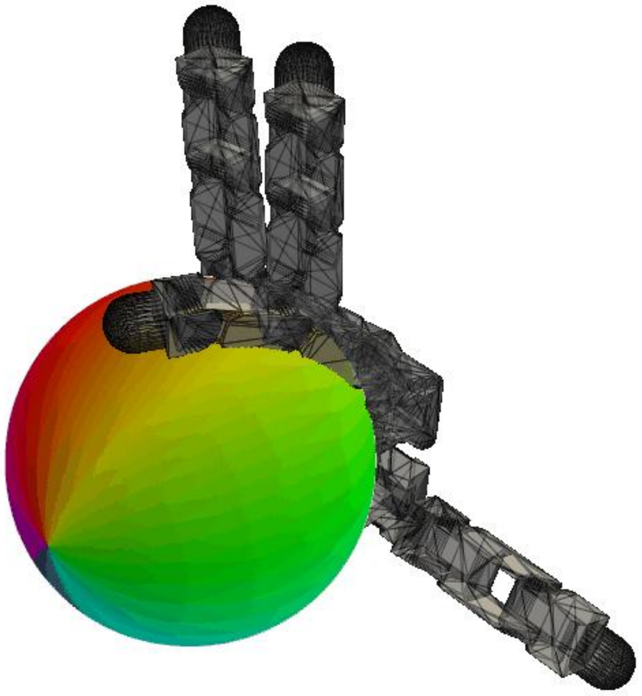


Define several "driving vectors" on each driving plane

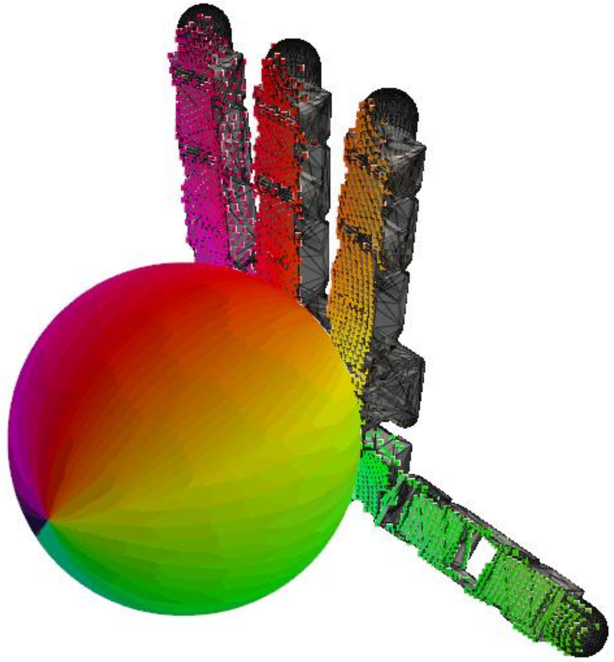


UGAS: Cascaded Inverse Kinematics (CIK)

- Given a deformed sphere, we solve IK to obtain the hand configuration

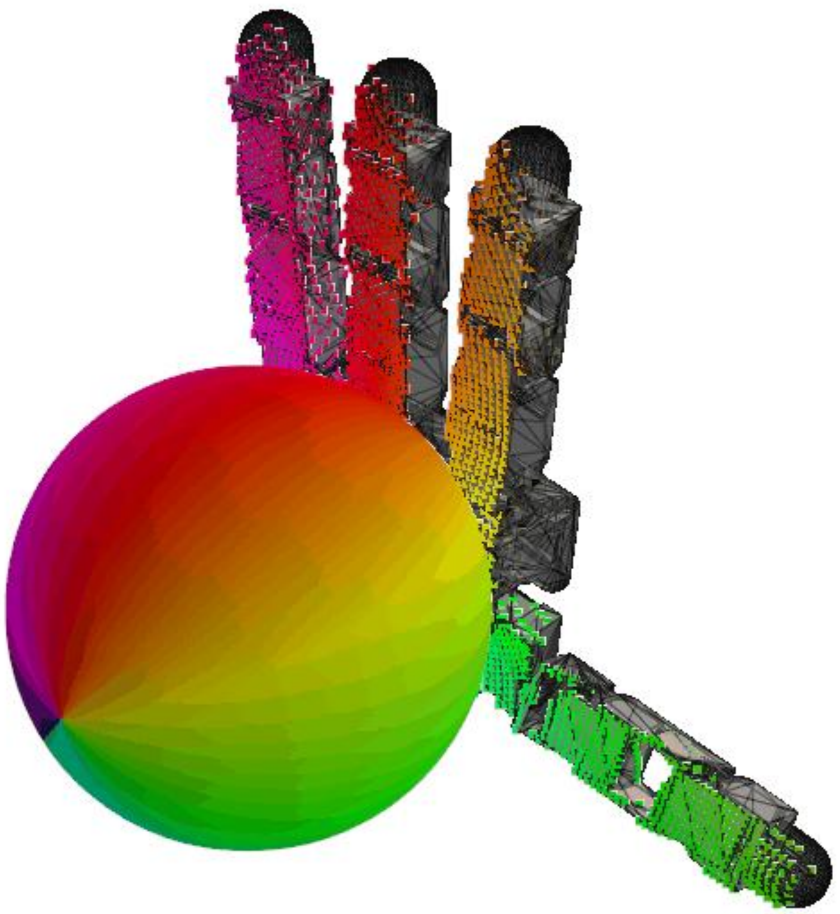


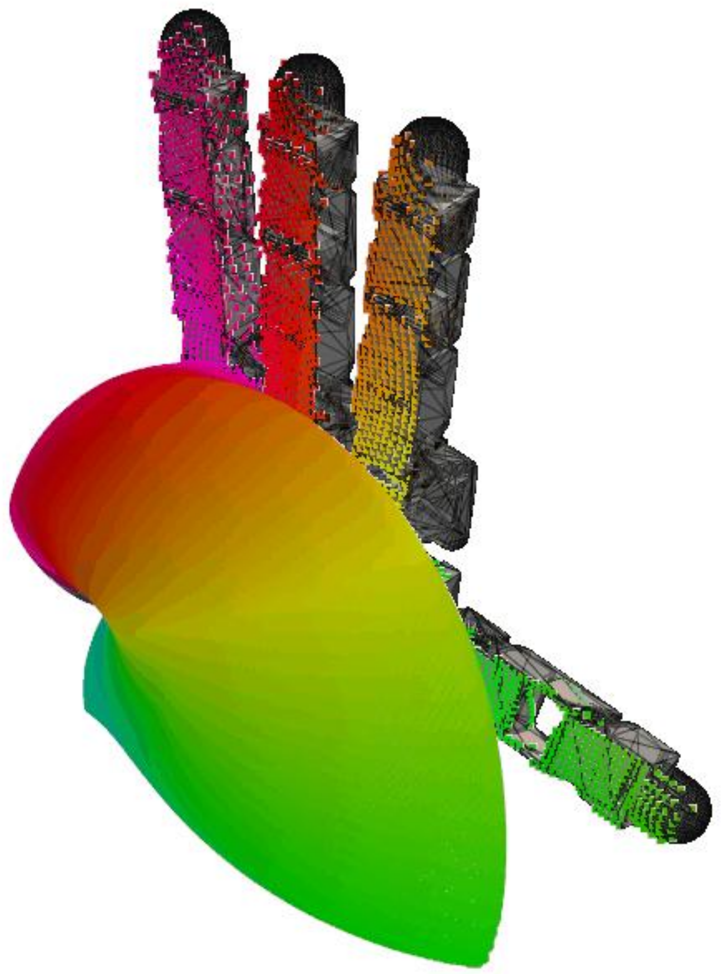
1. Solve lateral joints



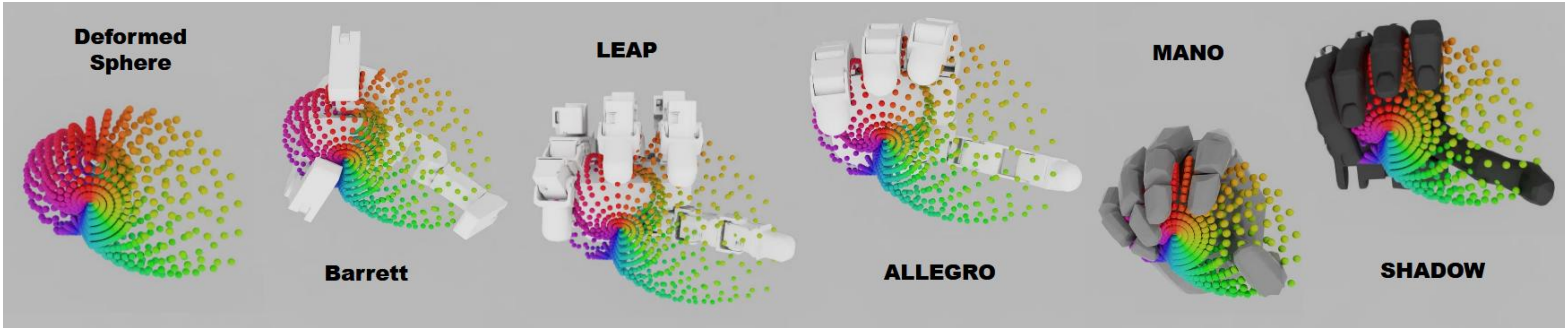
2. Solve encompassing joints
We solve for each joint one at a time, in the order of the kinematic tree.

UGAS: Cascaded Inverse Kinematics (CIK)





Unified Gripper Action Space (UGAS)



Control actions

Driving planes $\Delta\theta$

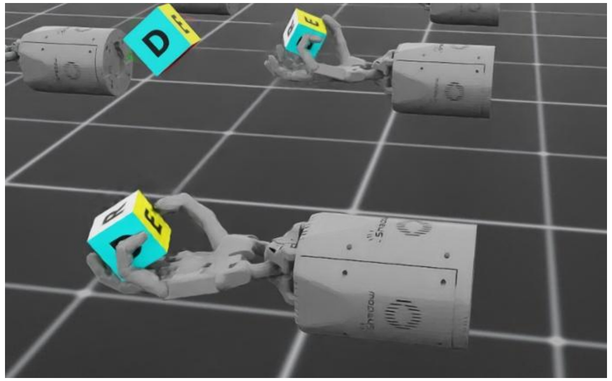
Driving vectors Δr

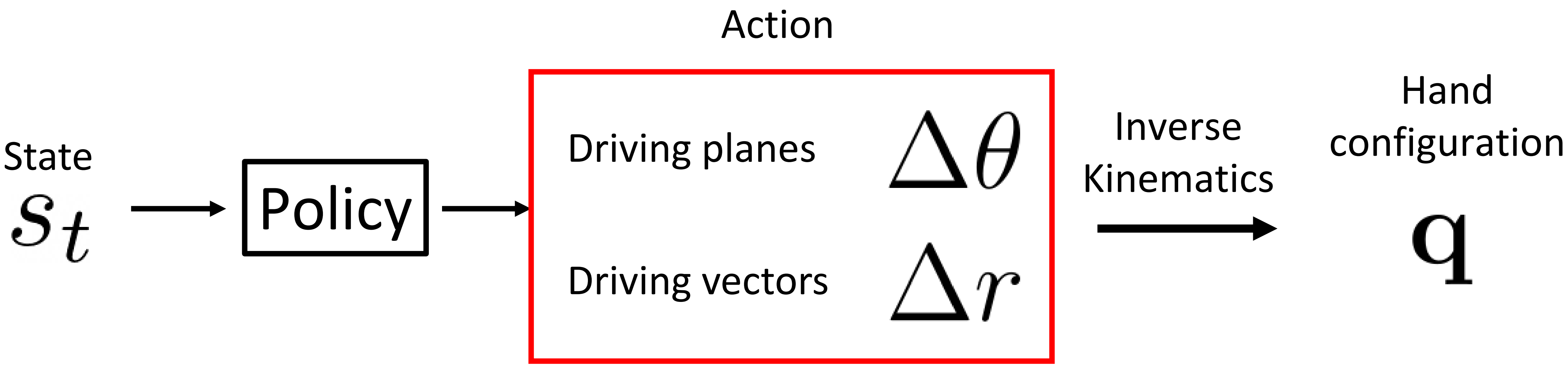
Inverse Kinematics →

Hand configuration $\mathbf{q}$

UGAS for In-hand Manipulation

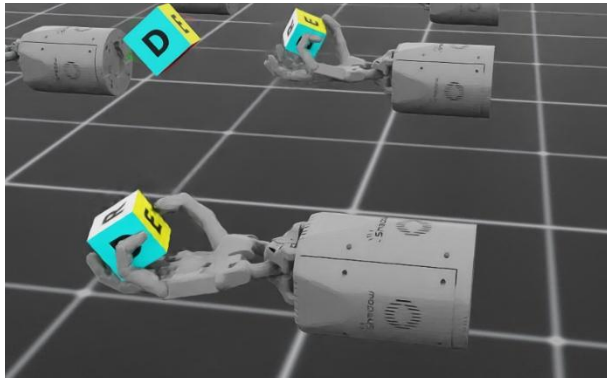
- Task: repose a cube to a target orientation
  - 10 consecutive reposing within 30 seconds
  - RL training in Isaac Lab with PPO and our sphere controller

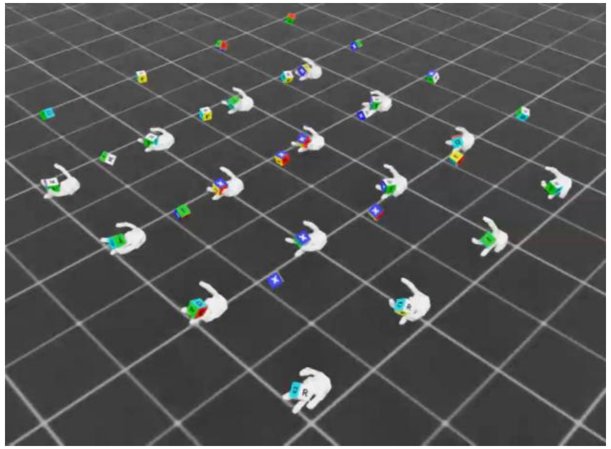




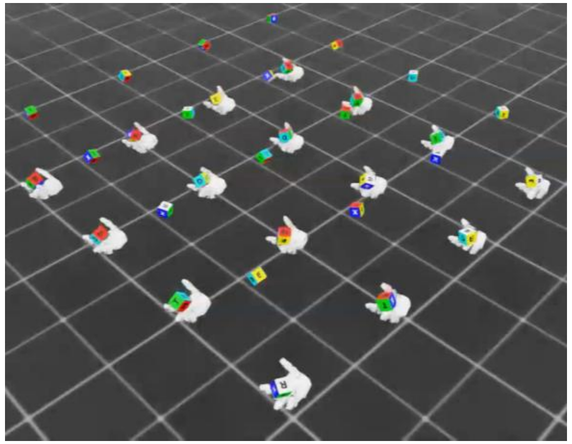
UGAS for In-hand Manipulation

- Task: repose a cube to a target orientation
  - 10 consecutive reposing within 30 seconds
  - RL training in Isaac Lab with PPO and our sphere controller

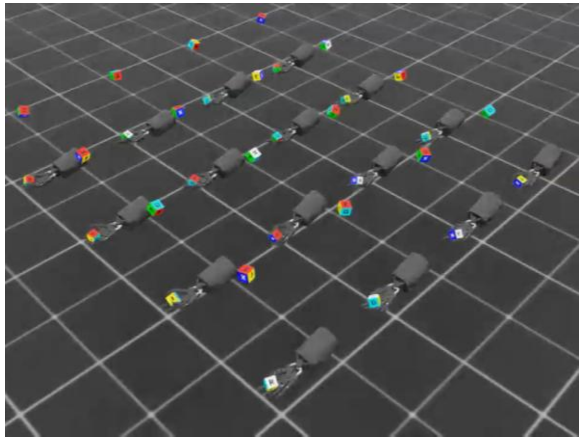




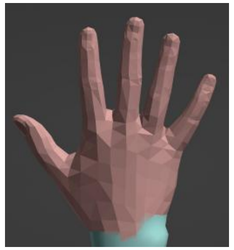
Allegro (4 fingers)
9.6

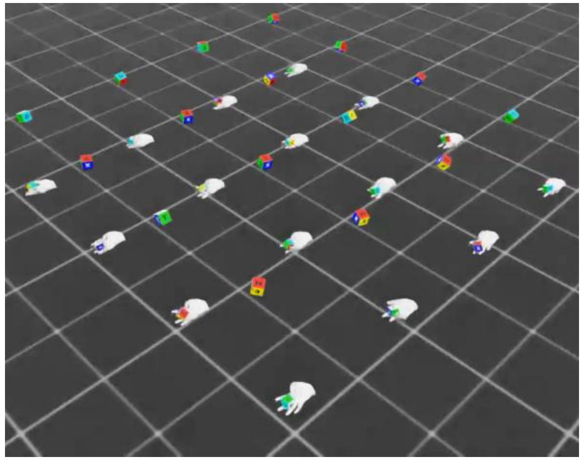


Leap (4 fingers)
9.80



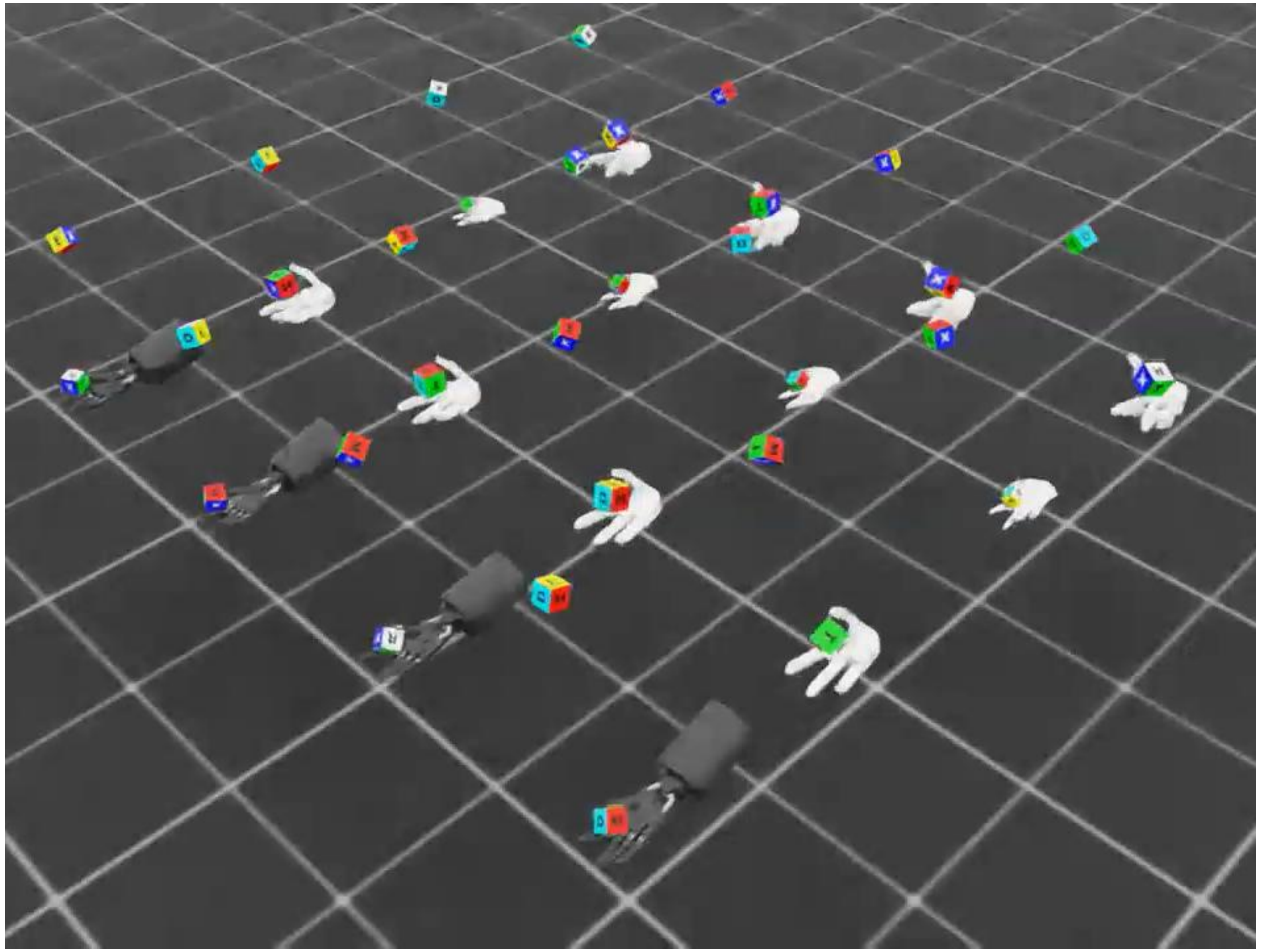
Shadow Hand (5 fingers)
9.75

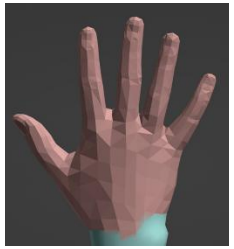




MANO human (5 fingers)
9.87

UGAS for In-hand Manipulation





Training and testing with all hands
9.92

Zero-Shot Policy Transfer (3 In Domain -1 Out of Domain)

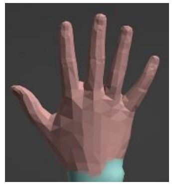
Training

Testing

Allegro (4 fingers)
8.7

Leap (4 fingers)
8.1

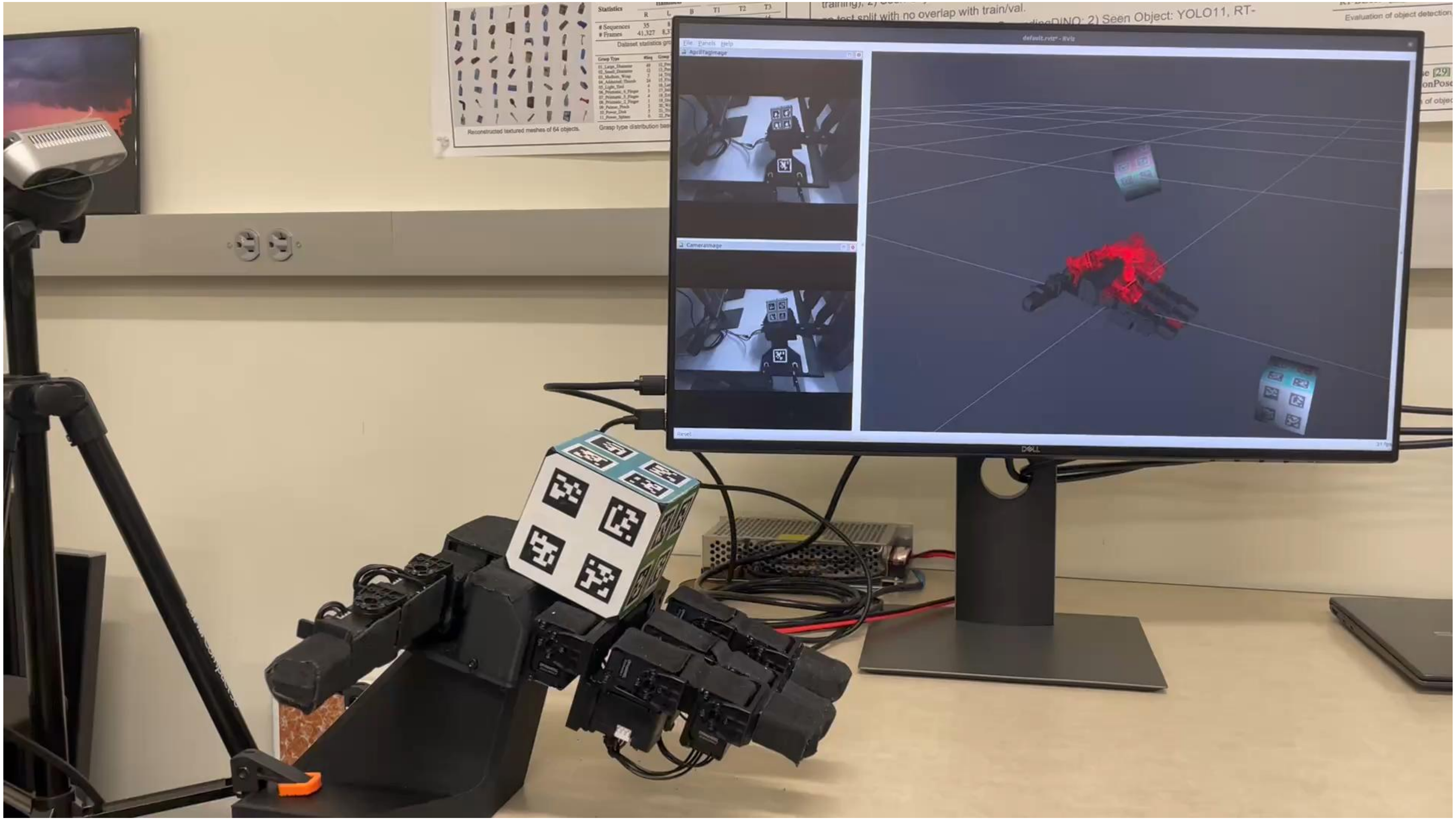
Shadow Hand (5 fingers)
3.2



MANO (5 fingers)
9.0

Real-world in-hand manipulation

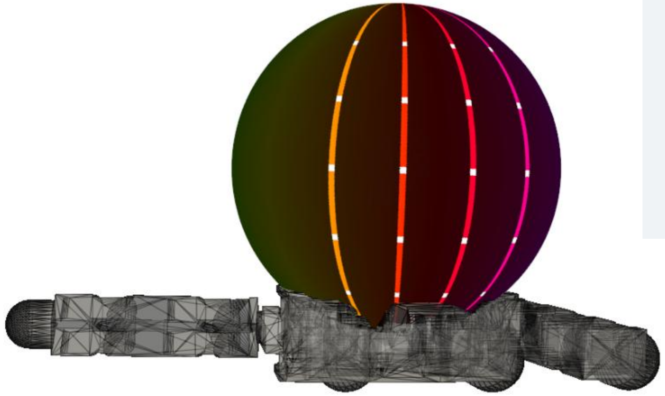
On-going effort

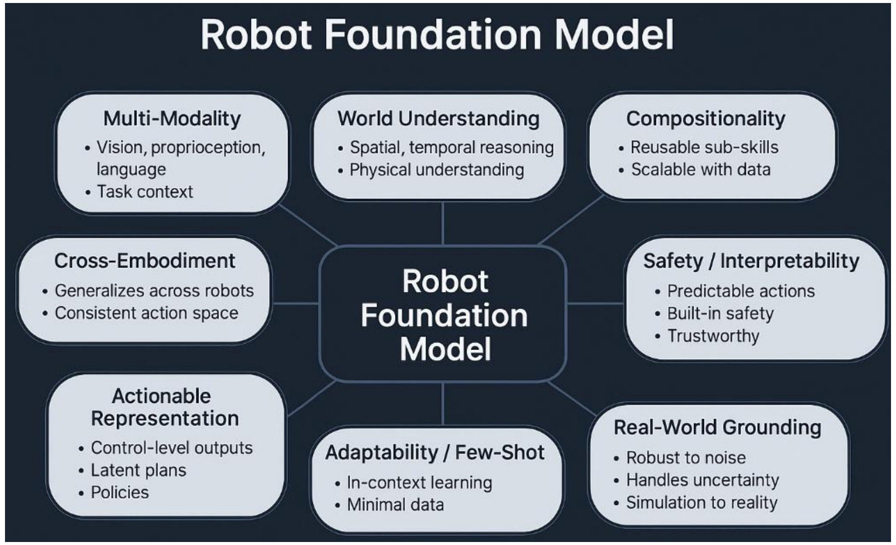


Summary

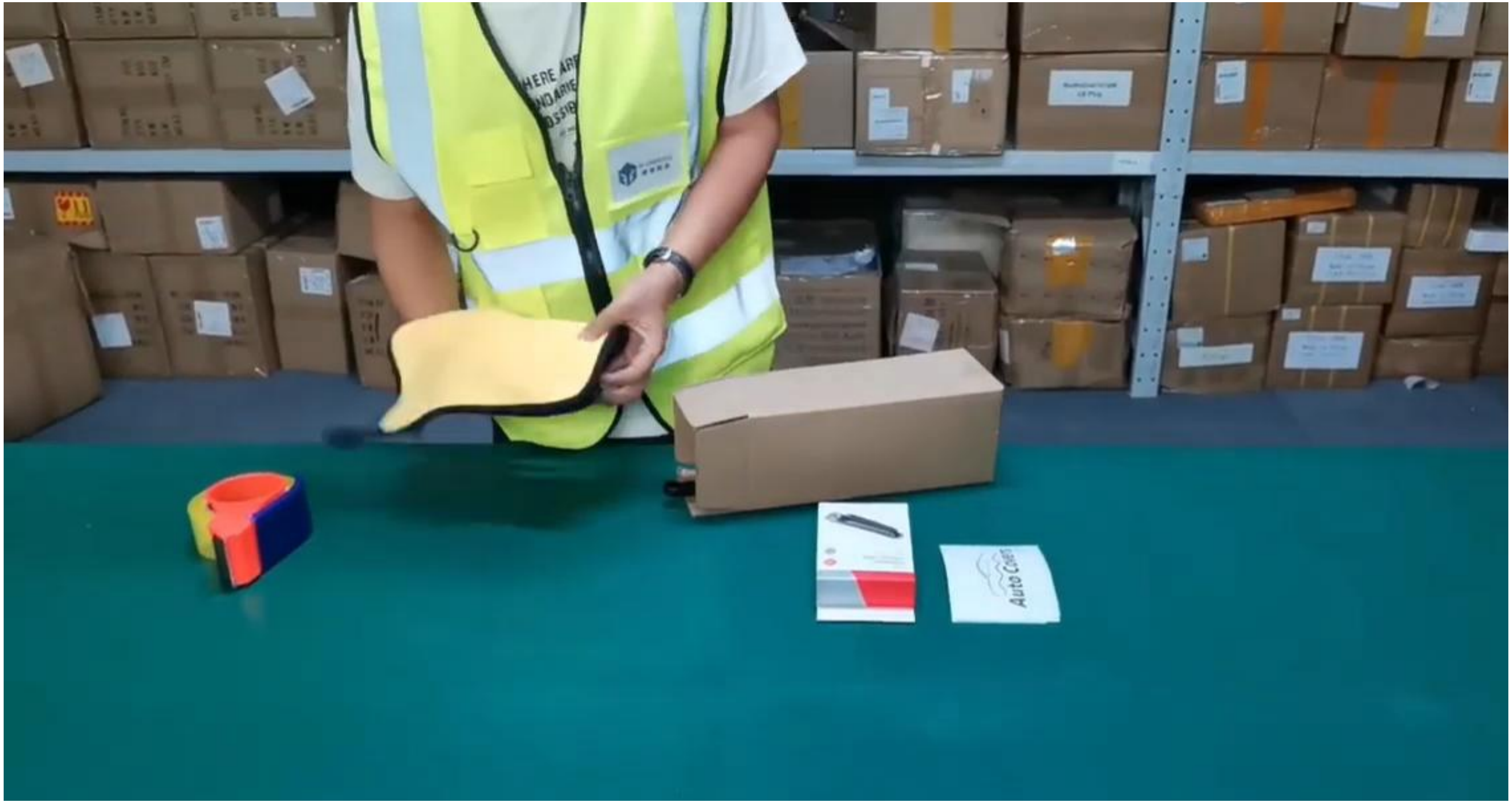
Use data from all robots and human for learning

Unified Gripper Action Space

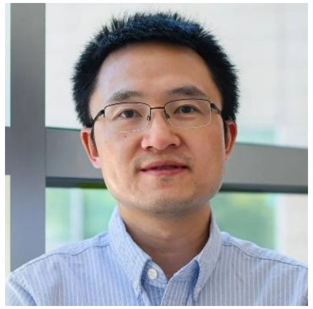


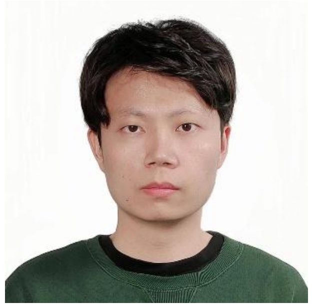


Robot Manipulation is still an Open Challenge

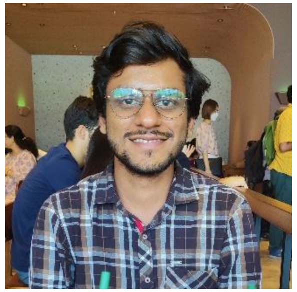


Intelligent Robotics and Vision Lab (IRVL)











https://labs.utdallas.edu/irvl/

Assisted by
Ms. Rhonda Walls

Thank you!