Decoupling Observation and Motion:

Object-Centric Representations for Enhanced Manipulation

Goal: To manipulate articulated objects with the ability to generalize across camera views, object placements , robot kinematics

Decoupling ***Observation*** and Motion:

Object-Centric Representations for Enhanced Manipulation

Goal: To manipulate articulated objects with the ability to generalize across ***camera views, object placements***, robot kinematics

Decoupling Observation and ***Motion***:

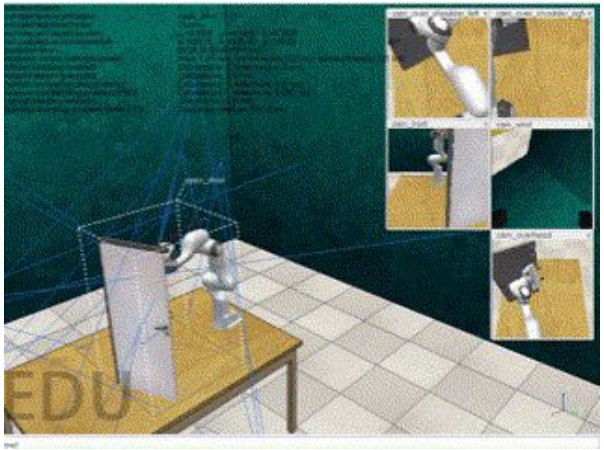
Object-Centric Representations for Enhanced Manipulation

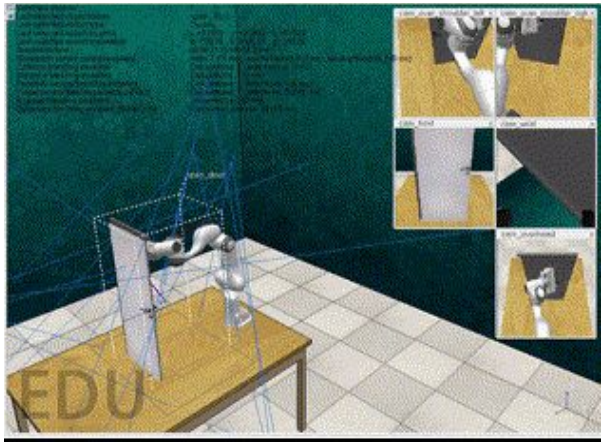
Goal: To manipulate articulated objects with the ability to generalize across camera views, object placements, ***robot kinematics***

1. Test the Baseline

1.1 Data Collection

Observation





Action: end effector positions

1.2 Data Training

Baseline

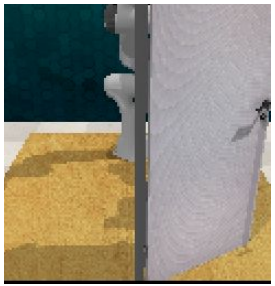


image-based
transformer or diffusion

↓

robot action in base frame

1.2 Data Training

Baseline

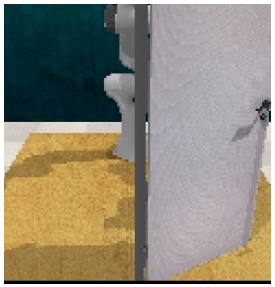
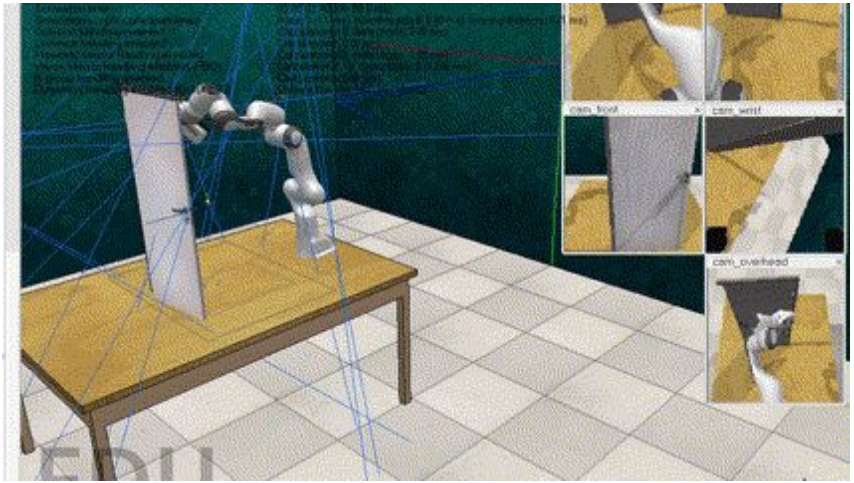


image-based
transformer or diffusion

↓

robot action in base frame

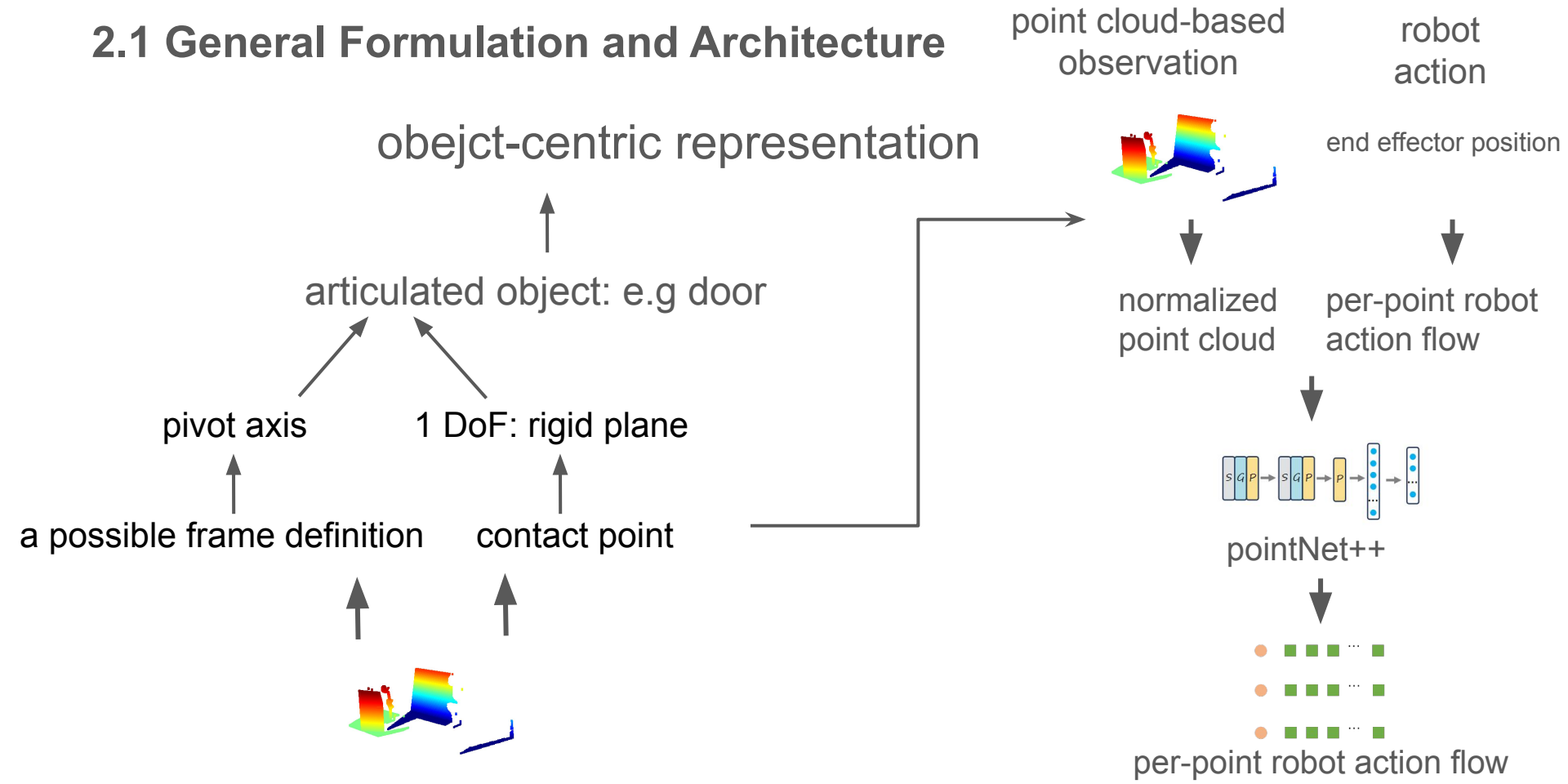
1.3 Inference Result



almost fail due to variations of object placements and camera views

2.Object-centric Pipeline

2.1 General Formulation and Architecture



2.2 Progress on Obejct-centric Representation

obejct-centric representation

articulated object: e.g door

pivot axis

1 DoF: rigid plane

a possible frame definition

contact point

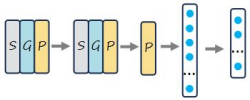
point cloud-based observation

robot action

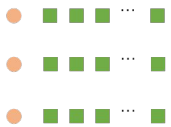
end effector position

normalized point cloud

per-point robot action flow



pointNet++



per-point robot action flow

2.2 Progress on Obejct-centric Representation

obejct-centric representation

articulated object: e.g door

pivot axis

1 DoF: rigid plane

a possible frame definition

contact point

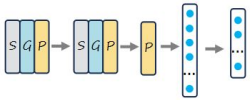
point cloud-based observation

robot action

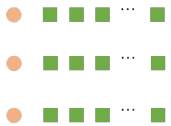
end effector position

normalized point cloud

per-point robot action flow



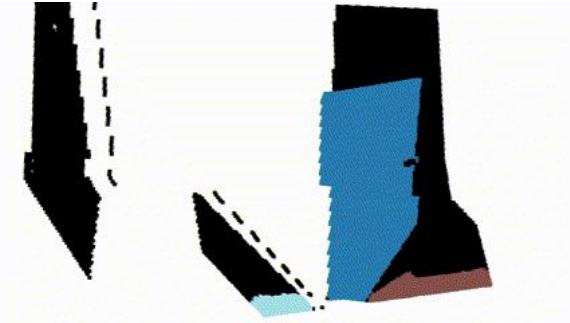
pointNet++



per-point robot action flow

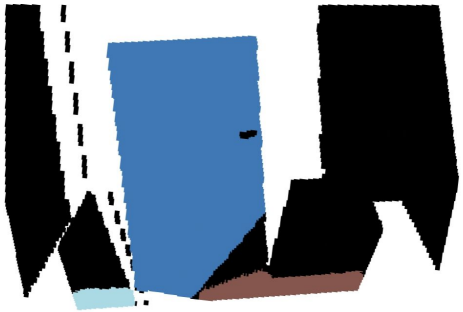
2.2 Progress on Obejct-centric Representation

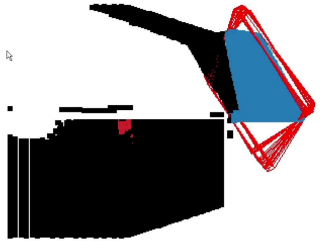
door segmentation



2.2 Progress on Obejct-centric Representation

door segmentation → pose estimation

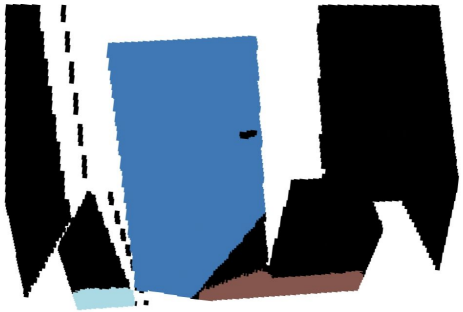


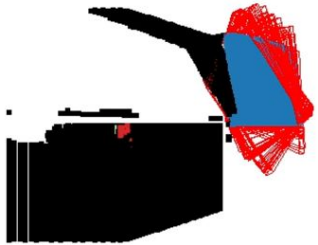


based on oriented bounding box

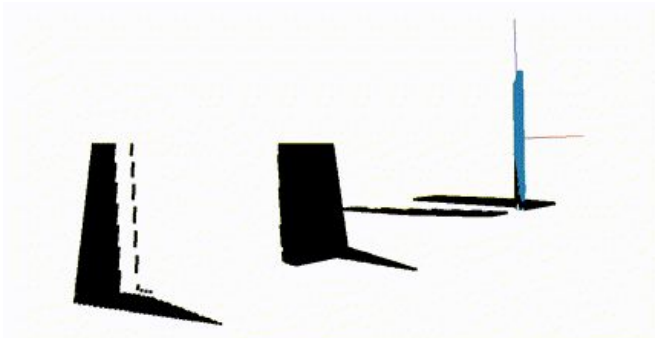
2.2 Progress on Obejct-centric Representation

door segmentation → pose estimation → obejct-centric frame



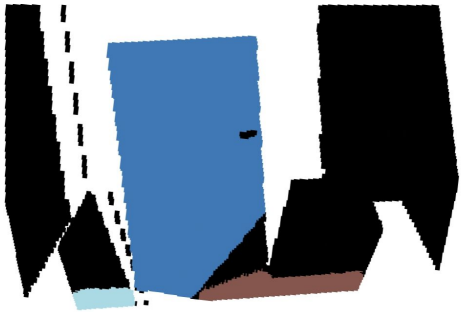


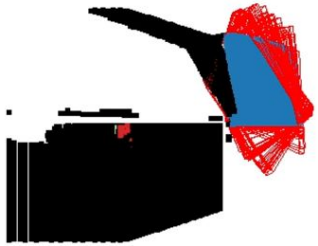
based on oriented bounding box



2.2 Progress on Obejct-centric Representation

door segmentation → pose estimation → obejct-centric frame



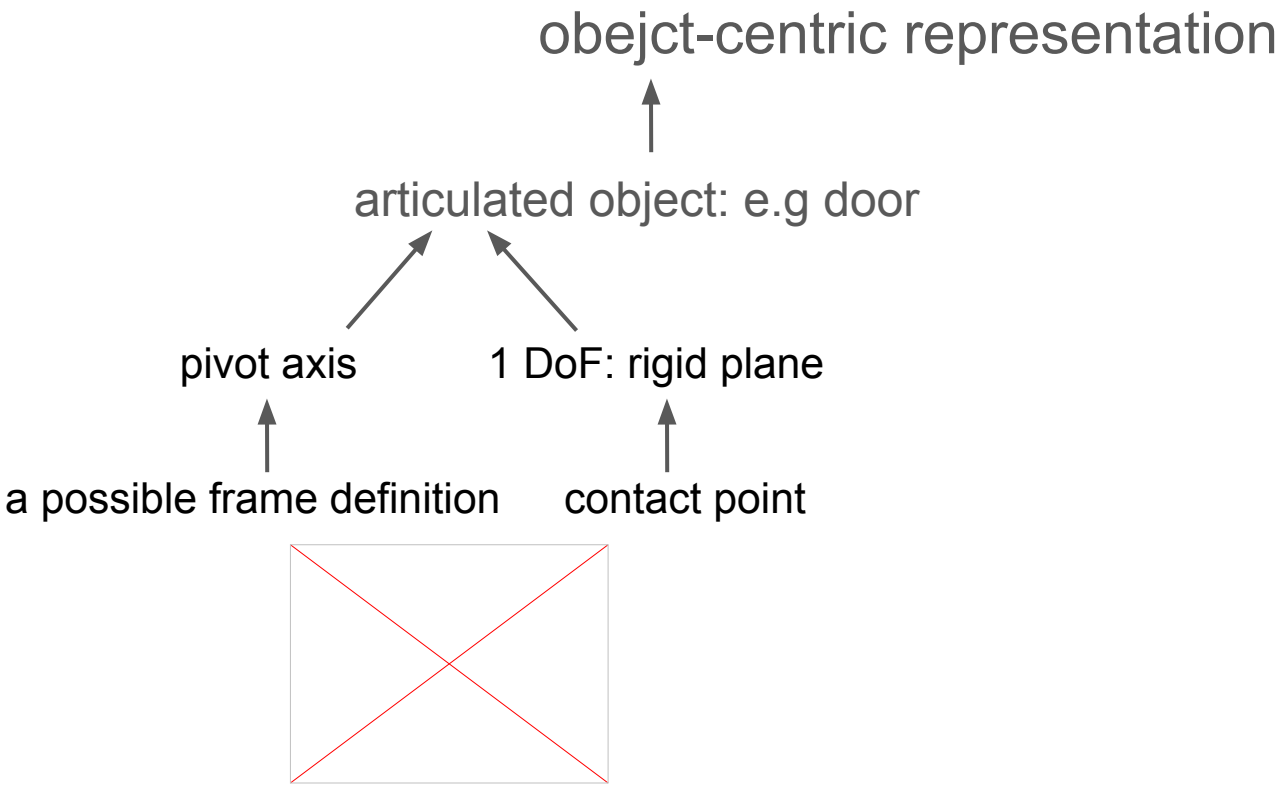


based on oriented bounding box

Next Steps

→ contact detection → per-point action flow representation construction

3. Discussion



3. Discussion

obejct-centric representation

articulated object: e.g door

pivot axis

1 DoF: rigid plane

a possible frame definition

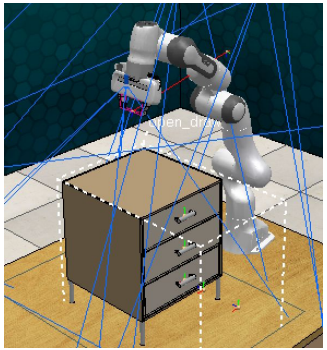
contact point

Other Tasks

Push

Toss

Multi DoF



4.1 Progress on Obejct-centric Representation

obejct-centric representation

articulated object: e.g door

pivot axis

1 DoF: rigid plane

a possible frame definition

contact point

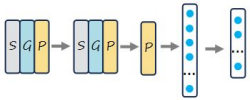
point cloud-based observation

robot action

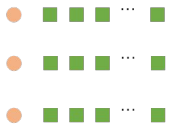
end effector position

normalized point cloud

per-point robot action flow

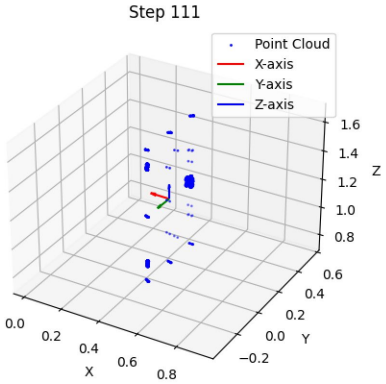


pointNet++



per-point robot action flow

4.1 Progress on Obejct-centric Representation



inconsistent results

point matching

a possible frame definition

contact point

point cloud-based observation

robot action

end effector position

per-point robot action flow

pointNet++

per-point robot action flow