

1 Dataset Curation

Collection

EG3D Ego4D

EG3D Ego-Exo4D

Apple EgoDex

Ropedia Xperience-10m

EPIC KITCHENS EPIC-KITCHENS

...

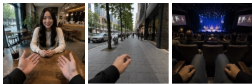
Raw Samples



2 Video selection

Ego-Interaction Filter

Non-interactable Clips



Egocentric Manipulation



Image Captioning-Based Filter

Acquire Action Caption

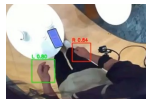
"Open the drawer."
"Grasp the mug and put it on plate."
...

Remove Invalid Caption

(Missing/Incomplete/...)

3 3D Hand Reconstruction

2D Tracking



Left

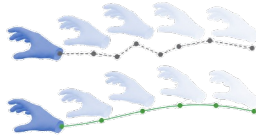
Right

Local Pose Estimation



Estimate Fine-Grained Articulation

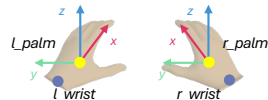
Trajectory Optimization



Temporally consistent global pose

4 Action Parameterization

Robot Action Convention



Human End-effector Equivalents

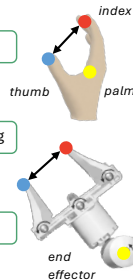
Skeletal Pose



Subject Mapping



Gripper Action



5 Quality Control

Static Filter

Remove low-motion actions

Spike Filter

Remove abnormal spikes



Completeness Filter

Remove invalid rotations and incomplete data

Bimanual Filter

Grasp event ↔ End effector motion coupling

Training Set



- ✓ High-quality
- ✓ Consistent
- ✓ Robot-ready