GraspXL: Generating Grasping Motions for Diverse Objects at Scale

Hui Zhang, Sammy Christen, Zicong Fan, Otmar Hilliges , Jie Song

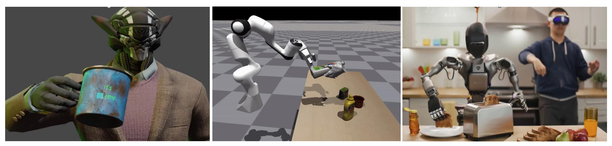
ETH zürich · aitlab · MAX PLANCK INSTITUTE FOR INTELLIGENT SYSTEMS

Overview

1 Motivation

Goal: generate 3D hand-object grasping motions for general objects

Q1: Why do we need grasping motions?



[Blender]

Realism of VR/AR animation

[Christen et al., 2023]

Simulate humans for human-robot interactions

[NVIDIA GROOT]

Expert data for robots

Q2: Why do we generate grasping motions?

Motion capture is

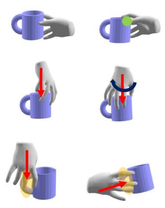
- Expensive
- Not scalable
- Not real-time & adaptive

2 Task: objective-driven grasping motion synthesis at scale

General objects

Objaverse [Deitke et al, 2022]

Multiple objectives



Different hand models

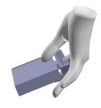
3 Challenges

Challenge 1: Generalization among object shapes, multiple objectives, and diverse hand morphologies

Question: A general framework for objects, motions, and hands?

Challenge 2: Precise objectives & Stable grasping

Flip the object when trying to achieve the objectives



Grasp the object without satisfying the objectives

Question: Stable grasping + adhering to objectives?

Project Page

Personal Website

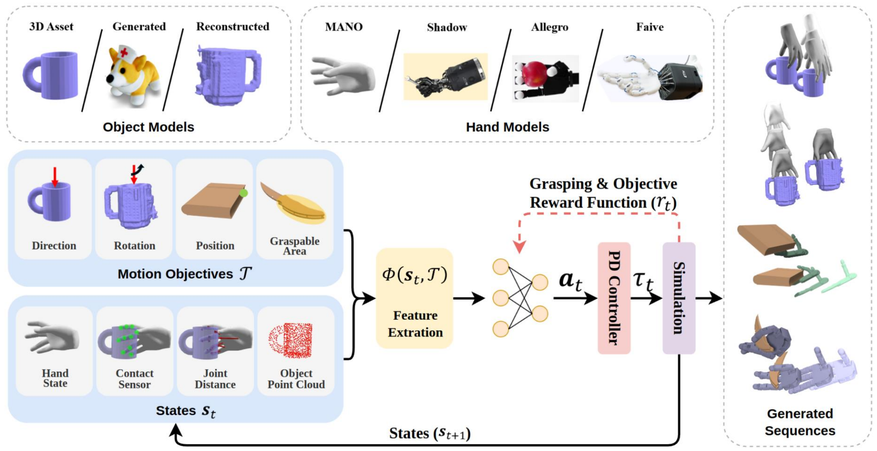
Correspondence

huizhang@ethz.ch

Framework

1 Framework : Reinforcement Learning + Physics Simulation

Trained with 58 objects and randomly sampled objectives



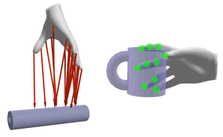
Dataset released!

Diverse grasping **motions** for **500k+ objects** with **different hands**

Method

1 General pipeline (Challenge 1)

General objects



Joint distance + Contact sensors

Multiple objectives

Objective-driven wrist guidance

Different hand models

$$r = r_{\text{grasp}} + r_{\text{objective}}$$

Encourage stable grasping

Encourage precise objectives

General reward function (model-free)

2 Learning curriculum (Challenge 2)

Fixed

$r = r_{\text{grasp}} + r_{\text{objective}}\uparrow$

Setting: Fixed object + Larger objective rewards

Goal: Simple contact & Precise finger motion for objectives

Moving

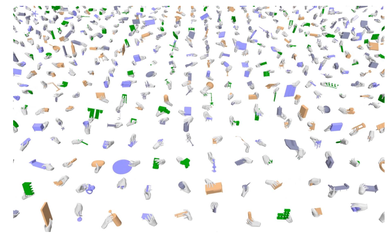
$r = r_{\text{grasp}}\uparrow + r_{\text{objective}}$

Setting: Moving object + Larger grasp rewards

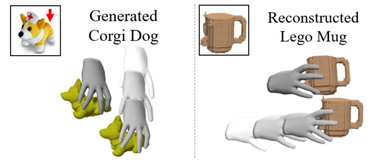
Goal: Precise wrist motion (without hitting/flipping)

Results

1 Generated motions

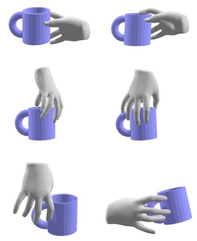


500k+ objects



Generated/reconstructed objects

Different dexterous hands



Different objects

Applications

1 Motion generation

GOAL [Taheri et al., 2022]

[Braun et al., 2024]

As a plug-in for whole-body motion generation methods (Animation)

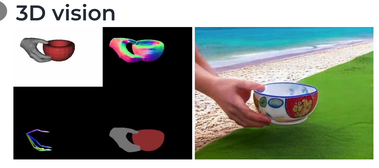
Frog wearing a sweater

Eggshell broken in two with an adorable chick standing next to it

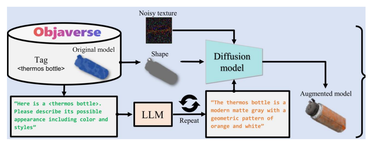
DreamFusion [Poole et al., 2022]

Zero-shot text-to-motion generation (with text-to-mesh methods)

2 3D vision



HOIDiffusion [Zhang et al., 2024]



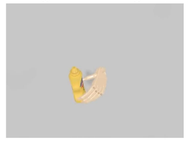
FoundationPose [Wen et al., CVPR 2024]

Generating large scale RGBD grasping motion data with pseudo ground-truth labels (with texture generation methods) and downstream applications like training general pose estimation models

3 Robotics

Human grasping demo

Retargeted robot grasping



DexMV [Qin et al., ECCV 2022]

As expert demonstrations with ground-truth motion labels for robots

SynH2R [Christen et al., ICRA 2024]

Large-scale human motion simulation for human-robot interaction