# A Telemanipulation Suite for Deformable Object Manipulation

Bonggyeong Park[1], Chanyoung Ahn[1] and Deahyung Park[1]

[1] KAIST, School of computing, Daejeon, Korea. {iampbk, coldyoung, daehyung}@ kaist.ac.kr

## Abstract

Our aim is to create a human-expert manipulation dataset of deformable objects toward learning deformable manipulation. However, the data collection is challenging due to the limited sensing capability against to the high number of degree-of-freedom and complicate deformation/contact of soft bodies. Further, the self occlusion restricts the observation in grasping. In this work, we introduce a haptic-telemanipulation suite by adopting a haptic glove with a state-of-the-art physics simulator, IsaacGym from NVIDIA. The suite enables users to obtain realistic visual-and-tactile feedback as well as collect any part of object states while teleoperating a robotic gripper. We evaluate the suite by building one hundred demonstrations of dataset given five shape-and-property deformable objects.

***Keywords***— *Deformable Object, Manipulation, Data Collection, Teleoperation*

## I. Introduction

We present a problem of creating deformable object manipulation (DOM) dataset toward modeling of dexterous manipulation skills as human experts. However, the dataset requires observing the state of deformable object during the human demonstrations, though the limited sensing capability restricts recording detailed states such as the high degree-of-freedom (DOF), the complicate deformation, and the precise contact of soft bodies [1]. Further, the sensor occlusion is often another restriction of recording states in that either the target objects or fingers hide each other from vision sensors. To resolve the incomplete sensing in demonstrations, we need a new data collection method that is a task-and-object agnostic framework and also does not intervene the interaction between a human expert and a target object for the integrity of the dataset.

We propose a virtual reality-based telemanipulation suite that allows an expert hand to manipulate simulated deformable objects while recording complete observations

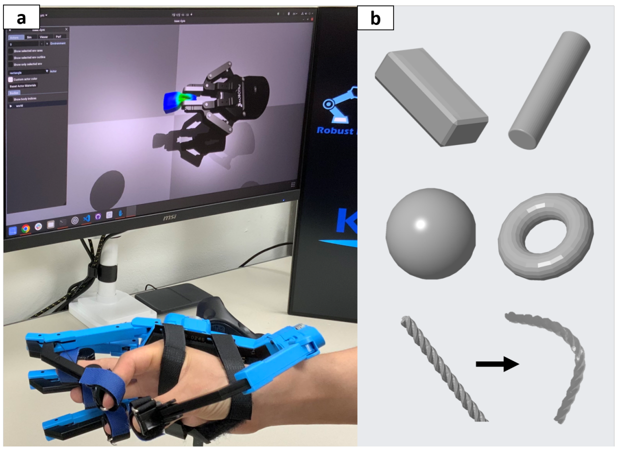


Fig. 1. A capture of telemanipulation task and five objects used in data collection, where a human operator wears the haptic glove (blue) and picks up a simulated deformable object (i.e., a cuboid) with haptic and force feedback.

via a state-of-the-art deformable physics simulator, IssacGym from NVIDIA [2]. The suite consists of telemanipulation and data-collection components. The former aims to transfer the real hand movement, observed via an haptic glove (i.e., SenseGlove DK1), to a simulated gripper (i.e., Robotiq 2F-85) by retargeting the configurations. The latter records the gripper, the object, and their interaction states by precisely simulating on the FLEX physics engine. We performed 100 times of pick-up experiments with 5 types of deformable object. The high-success rate of experiment result shows the suite is suitable for collecting human demonstrations toward learning for DOM.

## II. Method

The proposed telemanipulation suite allows a human operator to grasp a simulated deformable object in IssacGym environment via a master-slave robotic grasp system. The master device is a wearable exoskeleton glove that can sense the 4-DOF movement of the human fingers and return force/haptic feedback to the fingers (see Fig. 2). The slave device is a simulated 2-DOF parallel-jaw gripper that can sense the finger-tip contact and grasp deformable objects with physics engine, FLEX. To transfer the human hand motion to the parallel-jaw gripper, we retarget not

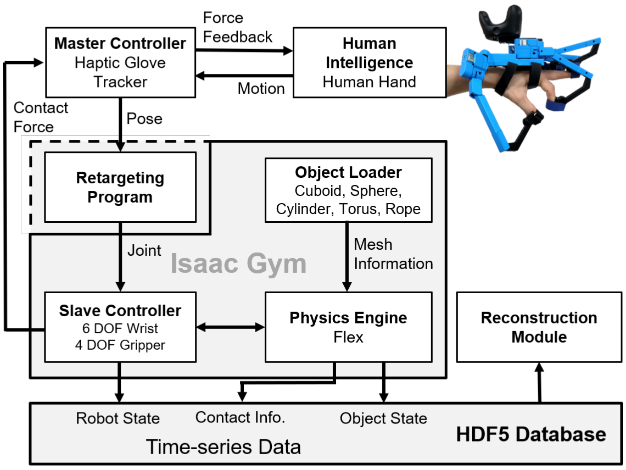


Fig. 2. Overall architecture of the proposed telemanipulation suite. The telemanipulation part exchanges the human behavior observed on the master robot glove to the slave robot gripper by retargeting the mechanical configuration. The data collection part records a set of time-series observations including the slave gripper, the object, their interaction states.

only the master's wrist pose but also its pinch-grasp motion (i.e., distance $d$ between the thumb and the index fingers) to the slave. To do that, we compute the desired transformation of the slave frame $T^{O'}_{slave,t+1}$ at time step $t+1$ with respect to the simulation origin $O'$ after measuring the master frame $T^{O}_{master,t}$ at time step $t$ with respect to the real-world origin $O$. The *master* and *slave* frames are the 6-DOF HTC VIVE tracker frame attached on the wrist and the *wrist-base* frame of the simulated gripper, respectively (see Fig. 3). Given human motion $T^{master}_{slave,t}$ at time step $t$, the updated master-tracker pose $T^{O}_{master,t+1}$ returns the desired slave pose in the simulation given the offset frame $T^{O'}_{O}$ between the real and the simulated origins:

$$\overbrace{T^{O'}_{slave,t+1}}^{\text{Slave motion}} = \overbrace{T^{O'}_{O}}^{\text{Offset}} \overbrace{T^{O}_{master,t}}^{\text{Tracker motion}} \overbrace{T^{master}_{slave,t}}^{\text{Human motion}} \quad (1)$$

$$= T^{O'}_{O} \; T^{O}_{master,t+1}, \quad (2)$$

where $T^{O'}_{O}$ is a pre-defined offset in this work.

After retargeting the master motion to the slave in the Cartesian space, we compute the desired joint configuration of the gripper

Wrist frame to the simulator global frame $T^{wrist}_{sim}$, the inverse matrix of $T^{sim}_{wrist}$, is a input of the inverse kinematics solver for the wrist robot. By setting a $T^{tracker}_{wrist}$ as a identical matrix, we calculated wrist joint angles $\overrightarrow{\theta}_{wristrobot}$.

$$\overrightarrow{\theta}_{wrist\ robot} = \text{InverseKinematics}((T^{wrist}_{sim})^{-1}) \quad (3)$$

Second, the teleoperation framework retargets the slave gripper from the haptic glove. The forward kinematics solver calculate fingertip position $P_{thumb}$ and $P_{index}$ from

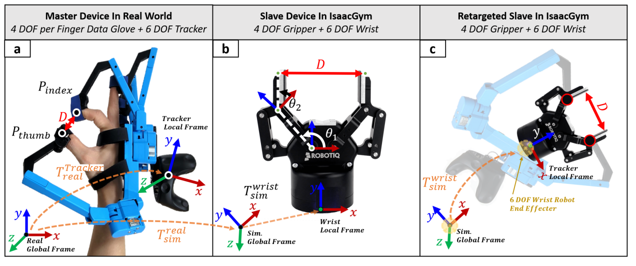


Fig. 3. An illustration of the retargeting process. We map the wrist pose and fingertip distance of the master glove to those of the slave gripper.

the haptic glove. Using the Euclidean distance $D$ of the two point vectors, the inverse kinematics solver compute gripper joint angle $\overrightarrow{\theta}_{gripper}$ by geometrical method.

$$\overrightarrow{\theta}_{gripper} = \text{InverseKinematics}(P_{thumb}, P_{index}) \quad (4)$$

Fig. 3 explains the kinematics and mapping between master devices and slave robots.

## A. *Data Collection*

Data collection framework aims to store information in simulator-deformable object state, slave device state, and contact information. The database format is HDF5, detailed in Table 1. Initial pose means a pre-defined local frame of object with respect to the global frame in the simulator. Elastic modulus and Poisson ratio determine object stiffness and deformability, respectively. Based on these properties, Flex calculates the object state with the telemenipulated gripper. Data collection framework stores the state information every time step on simulation. Notably, the frequency of the data collection is equal to the simulation time step-about 10-20 frames per second. This synchronization provides advantages to reconstructing the state information perfectly, potentially offering a baseline for off-policy learning. We demonstrate the data collection experiment in the Section 3.

| Category | Exploration | Type | Unit |
|---|---|---|---|
| **General Description** | | | |
| **Initial Pose (Pre-defined local frame)** | Object position | FloatArray[3] | m |
| | Object Orientation (Quat.) | FloatArray[4] | - |
| **Object Property** | Density | Float | $kg/m^3$ |
| | Elastic Modulus | Float | Pa |
| | Poisson Ratio | Float | - |
| **Object Mesh Information** | Triangle Mesh Edge Connection | IntArray[$N_1$, 3] | - |
| | Tetrahedron Mesh Edge Connection | IntArray[$N_2$, 4] | - |
| **Per Time Step** | | | |
| **Simulation** | Time | Float | sec |
| **Object State** | Node Position | FloatArray[$N_3$, 3] | m |
| | Force on Node | FloatArray[$N_3$, 3] | N |
| | Von-mises Stress on Triangle element | FloatArray[l] | Pa |
| **Kinematics** | Wrist Position | FloatArray[3] | m |
| | Wrist Orientation (Quat.) | FloatArray[4] | - |
| | Gripper Jaw Length (D) | Float | m |
| **Contact** | Normal Force on the Left Gripper Tip | Float | N |
| | Normal Force on the Right Gripper Tip | Float | N |

Table 1. Simulation setup for the environments. $N_1$, $N_2$, and $N_3$ is the number of connection edge, tetrahedron elements, and nodes, respectively.

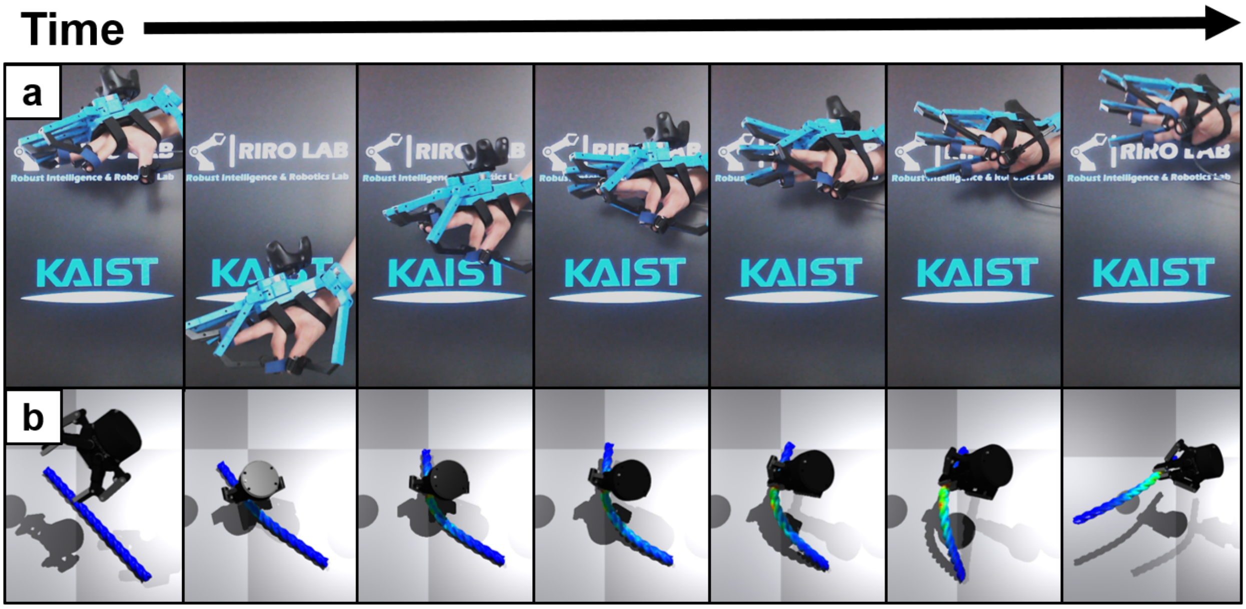


Fig. 4. Demonstration on the robot pick-up scenario, where a user virtually grasps the blue-color of simulated rope in IssacGym by observing its state via a 2D screen. In the simulation, the gripper is mounted on the invisible 6 DOF wrist robot.

## III. RESULT

To evaluate the feasibility of the telemanipulation framework, we collected one hundred data with five objects $\in$ {Cuboid, Sphere, Torus, Cylinder, Rope} and four elastic modulus $E \in \{10^4, 10^5, 10^6, 10^7\}$. The purpose of the dataset is to acquire complete state information containing human intelligence how to manipulate a deformable object. We set experiment for the pickup task, pick and move the object from $[0,0,0]m$ to $[0,0.6,0]m$. The demonstrations are in Fig. 4. Fig. 4a and 4b indicate the master device and slave robot, respectively. Our teleoperation framework retargets the slave robot from the master device pose and kinematics, manipulating rope.

Fig. 5 shows the stored data as a HDF5 format. Each node composes object position, so we defined object feature pose $\overrightarrow{P}_{feature}$ represents the mean of the $N_3 \times 3$ object node position vector $\overrightarrow{P}_{(N_3,3)}$.

$$\overrightarrow{P}_{feature} = \frac{1}{N_3}\sum_{i=1}^{N_3}\overrightarrow{P}_{(i,3)} \quad (5)$$

Importantly, the wrist pose in the *Posing* region changed carefully, meaning that the expert implemented pick action from his optimal policy. It means that our dataset internally contains human intelligence for the manipulation and could be appropriate for the off-policy learning dataset. After the *Posing* region, the expert closes the gripper and performs the pickup task. Gripper length affects the contact force and object stress (Fig. 5bcd), changing the magnitude of a force feedback to the expert. If the expert perceives the force feedback, The grasping state can define stable and moves deformable objects to the goal height $0.6m$ .

We evaluated the success rate in Table 2. The average of success rate is 93%. Success rate shows a dependency with a elastic modulus. As the elastic modulus decreases, the gripper deforms the object with less force, decreasing the contact force. In this case, the gripper cannot apply a enough friction force corresponding to gravity. At the same time, objects swing unpredictably, creating disturbances and increasing instability. This is why deformable object manipulation and data collection is more challenging than

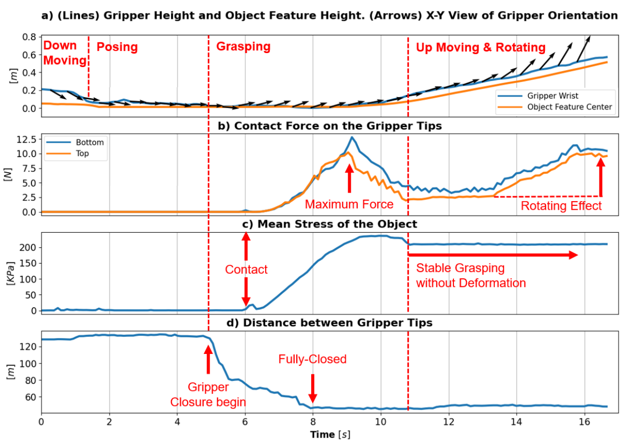


Fig. 5. Stored data on the HDF5 database. a) Gripper orientation (arrows) and the height of the $\overrightarrow{P}_{feature}$ and gripper (lines) $[m]$. b) Contact forces $[N]$. The force on the bottom tip is silightly higher, due to the object weight. c) mean Von-mises stress $[KPa]$, and d) the distance between gripper tip $[mm]$.

a rigid body. However, in our framework, the human expert completes the pickup task in $E = 10^4$ (similar to tofu), with an average success rate 80% within 15 seconds.

| **Young's Modulus** | **Cuboid** | **Sphere** | **Torus** | **Cylinder** | **Rope** | **Net** |
|---|---|---|---|---|---|---|
| $10^4$ | 0.83 | 1 | 0.63 | 0.83 | - | 0.8 |
| $10^5$ | 0.83 | 1 | 0.67 | 1 | 1 | 0.87 |
| $10^6$ | 1 | 1 | 1 | 1 | 1 | 1 |
| $10^7$ | 1 | 1 | 1 | 1 | 1 | 1 |
| **Net** | 0.91 | 1 | 0.78 | 0.95 | 1 | 0.93 |

Table 2. The success rate of the experience. We succeed experiment 100 times with the 108 trials. Higher elastic modulus $E$ means less elasticity

## IV. Conclusion

REWRITE: We aim to create a human-expert manipulation dataset of deformable objects for learning deformable manipulation. In this work, we introduced a haptic-telemanipulation suite by adopting a haptic glove with a physics simulator. We evaluated the suite of the dataset from a human expert by grasping five shape-and-property deformable objects. In the future, we will use the dataset for learning DOM.

## Acknowledgement

This work was partly supported by Institute of Information communications Technology Planning Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2022-0-00311, Development of Goal-Oriented Reinforcement Learning Techniques for Contact-Rich Robotic Manipulation of Everyday Objects) and National Research Foundation of Korea (NRF) grants funded by the Korea government (MSIT) (No.2021R1A4A3032834, Cooperative Intelligence Laboratory for Heterogeneous Robots).

## References

[1] Veronica E Arriola-Rios, Puren Guler, Fanny Ficuciello, Danica Kragic, Bruno Siciliano, and Jeremy L Wyatt. Modeling of deformable objects for robotic manipulation: A tutorial and review. *Frontiers in Robotics and AI*, page 82, 2020.

[2] Viktor Makoviychuk, Lukasz Wawrzyniak, Yunrong Guo, Michelle Lu, Kier Storey, Miles Macklin, David Hoeller, Nikita Rudin, Arthur Allshire, Ankur Handa, et al. Isaac gym: High performance gpu-based physics simulation for robot learning. *arXiv preprint arXiv:2108.10470*, 2021.