DOGlove: Dexterous Manipulation with a Low-Cost Open-Source Haptic Force Feedback Glove

Han Zhang[1,2], Songbo Hu[1], Zhecheng Yuan[1,2,3], Huazhe Xu[1,2,3]
[1] Tsinghua University, [2] Shanghai Qi Zhi Institute, [3] Shanghai AI Lab

https://do-glove.github.io/

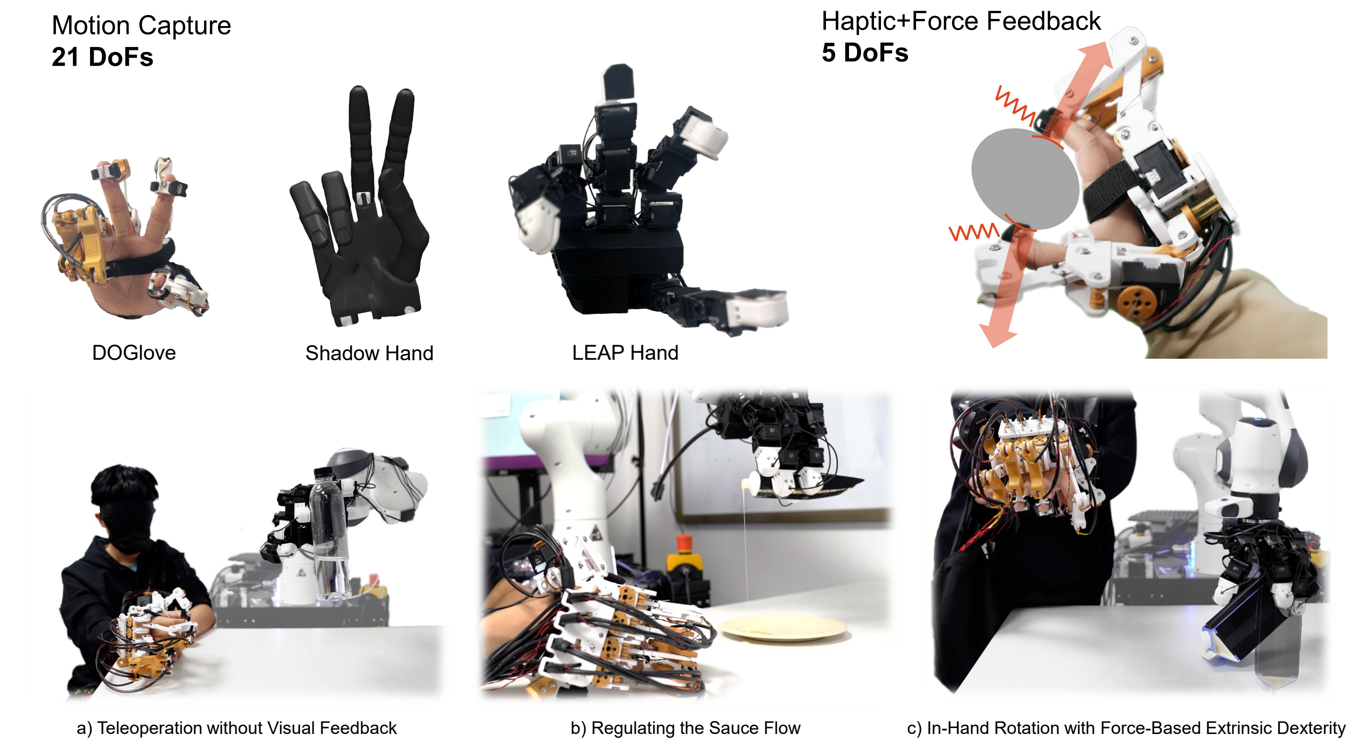


Fig. 1: **DOGlove**, a haptic+force feedback glove designed for precise teleoperation and dexterous manipulation. It features 21-DoF motion capture and 5-DoF haptic+force feedback. By leveraging action and force retargeting, it enables the teleoperation of dexterous hands for complex, contact-rich tasks, including: a) without visual feedback, adjusting contact force with a bottle during teleoperation, b) regulating the flow of condensed milk, and c) performing in-hand rotation by using haptic+force feedback to adjust friction.

***Abstract*—Dexterous hand teleoperation plays a pivotal role in enabling robots to achieve human-level manipulation dexterity. However, current teleoperation systems often rely on expensive equipment and lack multi-modal sensory feedback, restricting human operators' ability to perceive object properties and perform complex manipulation tasks. To address these limitations, we present DOGlove, a low-cost, precise, and haptic force feedback glove system for teleoperation and manipulation. DOGlove can be assembled in hours at a cost under 600 USD. It features a customized joint structure for 21-DoF motion capture, a cable-driven torque transmission mechanism for 5-DoF force feedback, and a linear resonate actuator for 5-DoF fingertip haptic feedback. Leveraging action and haptic+force retargeting, DOGlove enables precise and immersive teleoperation of dexterous robotic hands, achieving high success rates in complex, contact-rich tasks. DOGlove's hardware and software are fully open-sourced at https://do-glove.github.io/.**

I. Introduction

Imitation learning (IL) has shown significant promise in addressing complex manipulation tasks [6, 7, 41, 40]. However, it often necessitates a substantial amount of task-specific data to train a generalizable learning policy. Efficiently collecting and ensuring the high quality of such demonstrations remains a persistent and challenging problem for the robotic community.

Teleoperation is among the most commonly used methods for collecting demonstrations, often involving the development of a wide range of devices tailored to meet diverse data acquisition requirements. These devices enable human manipulation behaviors' transfer to various robotic platforms [12, 11, 5, 9, 14]. However, when it comes to dexterous hands, their high degrees of freedom (DoFs) and inherent complexity impose even stricter demands on operational precision and the accuracy of human motion capture. Hence, it is crucial to design an intuitive, responsive, and highly precise device specifically suited for dexterous hand teleoperation applications.

Vision-based methods are primarily used for tracking the human hand in dexterous teleoperation. A simple approach utilizes RGB cameras [21, 31], but the accuracy of hand gesture capture is questioned and further limited by visual obstacles during hand-object interactions. Motion capture (MoCap)

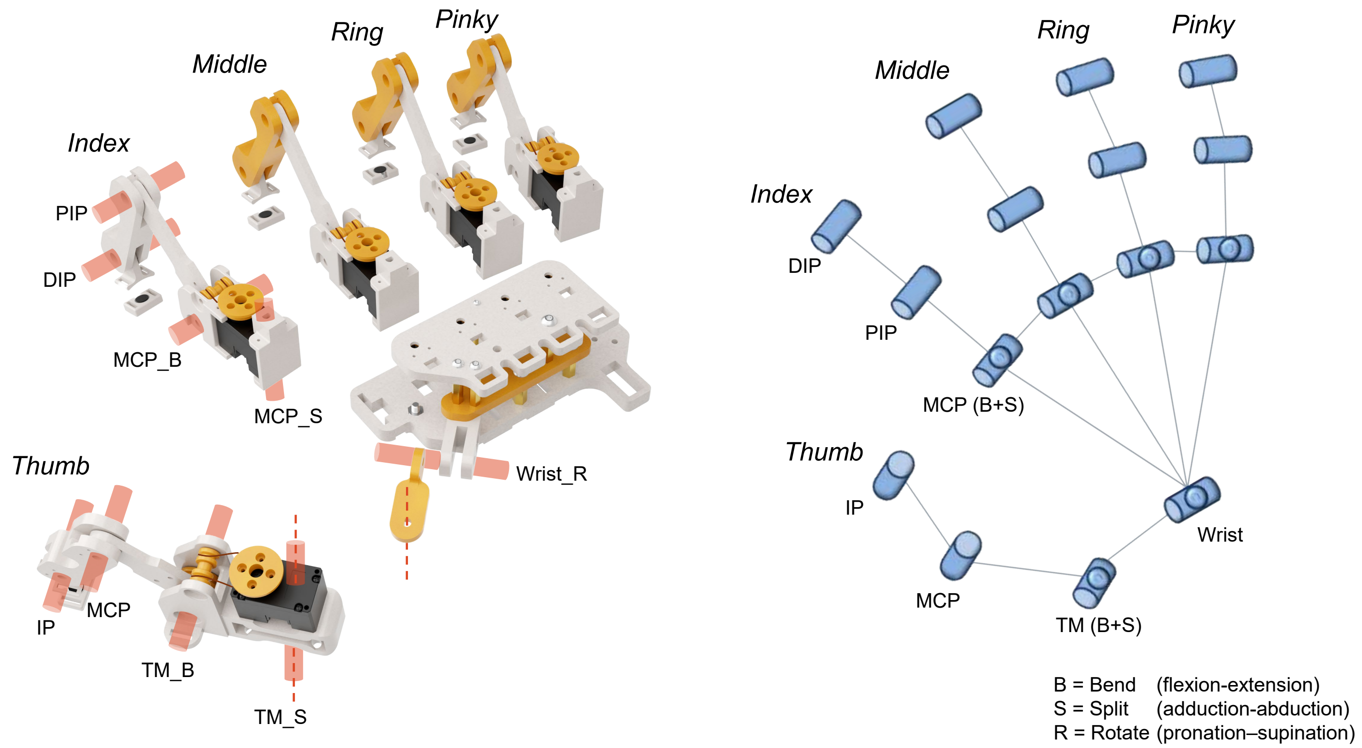


Fig. 2: **The kinematic structure of DOGlove**, designed to replicate the kinematics of a human hand. The MCP (B+S) and TM (B+S) joints are modeled as ball joints using a combination of two rotary joints. The right figure from [3] illustrates the simplified human hand kinematics.

systems [19, 32, 23, 34] provide stable hand tracking, while relying solely on visual feedback for teleoperation makes intuitive control challenging for the human operator.

Haptic+force feedback offers additional perceptual characteristics beyond those provided by vision, such as the ability to sense an object's weight, friction, and softness. Integrating haptic feedback into teleoperation can enrich the feedback available during interaction and enable the completion of more challenging tasks. Recently, some commercialized force feedback gloves [8, 25, 17, 13] have shown promise for enabling intuitive teleoperation. However, these solutions are often prohibitively expensive and require significant integration efforts to work with existing robot learning frameworks.

In this paper, we introduce **DOGlove**, a low-cost, open-sourced, and easy-to-manufacture haptic+force feedback glove for dexterous manipulation. Key features of DOGlove include 21-DoF motion capture, 5-DoF haptic+force feedback, action retargeting pipeline, and haptic+force retargeting strategy.

II. Hardware System

A. Kinematic Design

The kinematic design of DOGlove refers to the use of constraints to achieve desired movements, emulating the natural motion of a human hand. To ensure precise MoCap capabilities and a comfortable wearing experience, DOGlove are designed to closely resemble the anthropomorphic structure of the human hand. Several studies [3, 4] model the hand skeleton as a kinematic chain, represented by a hierarchical structure of rigidly connected joints. As shown in Figure 2, the kinematic structure of the human hand primarily consists of two types of joints: hinge joints and ball joints.

To implement these joints in DOGlove, hinge joints are modeled as two linkages connected by a rotary joint, with a joint encoder installed on the rotary axis to capture motion. Ball joints are designed as a combination of two orthogonal rotary joints, each equipped with a joint encoder on its respective rotary axis.

B. Finger Design

As shown in the exploded-view in Figure 2, DOGlove is composed of the thumb, index, middle, ring, and pinky finger assemblies, along with the palm base structure. The design of each finger assembly follows a modular approach, ensuring consistent structural elements across all fingers.

Given the limited space on the back of the human hand, the `Dynamixel XC/XL330` servo motors were selected as the actuators for force feedback.

1) *Joint Encoders:*

To integrate joint encoders into the finger linkages, the encoders must be compact while maintaining high precision. Additionally, as 16 encoders are required in combination with 5 servo motors to achieve 21-DoF MoCap capabilities, the cost of each encoder needs to be affordable. Considering these constraints, we selected the `Alps RDC506018A` rotary sensor as the joint encoder. This compact encoder (W11 mm $\times$ L14.9 mm $\times$ H2.2 mm) is easily integrated into the 3D-printed joint structures. The encoder operates as a variable resistor, changing its resistance as the shaft rotates.

*2) **Cable-Driven Force Feedback Structure**:*

To provide force feedback on the human fingers, the output torque of the Dynamixel servo must be transmitted to the glove's finger linkage system. As illustrated in Figure 2, the rotary axis of the servo and the rotary axis of the MCP_B joint are misaligned. Consequently, a transmission mechanism is required to transfer the torque effectively.

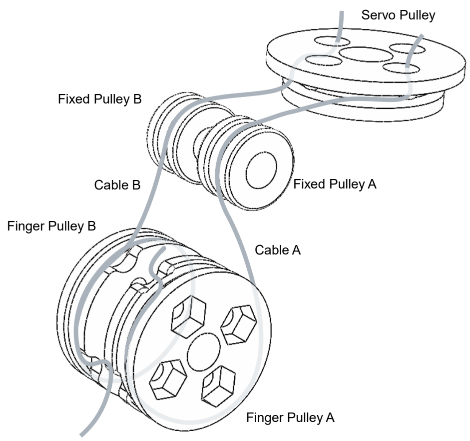


Fig. 3: **Pulley system** of the cable-driven mechanism.

DOGlove utilizes a pulley system to provide the bi-directional force feedback, as shown in Figure 3. DOGlove uses a 0.6 mm stainless steel braided wire as the cable, chosen for its strength and durability. The *Servo Pulley* connects the servo to the finger linkage via the *Finger Pulley*, maintaining a 1:1 transmission ratio. To minimize friction during transmission, the *Fixed Pulley* is used to redirect the cable's path. When the *Servo Pulley* rotates clockwise, the tension on *Cable B* increases, causing *Finger Pulley* to rotate clockwise. The extra slack on the *Cable A* side is taken up by the *Servo Pulley A*. Since the finger linkage is fixed to the *Finger Pulley*, it also rotates clockwise, resulting in the extension movement of the MCP_B joint. Similarly, when the *Servo Pulley* rotates counterclockwise, the tension shifts to *Cable A*, producing a flexion movement of the MCP_B joint.

*3) **Fingertip Haptic Feedback**:*

To further enhance the operator's tactile experience, each fingertip in DOGlove is equipped with a tactile actuator.

In DOGlove, we use LRAs with a diameter of 8 mm and a height of 2.5 mm, installed close to the fingertips. These LRAs provide high-quality vibration stimuli by resonating at approximately 240 Hz along Z axis, which is orthogonal to the fingertip surface. To fully leverage the potential of the LRA, we employ the `TI DRV2605L` motor driver, which includes the licensed `Immersion TouchSense© 2200` haptic library. This driver supports over 100 pre-programmed waveforms, allowing DOGlove to deliver realistic and refined haptic feedback.

C. Wrist Localization

In DOGlove, we design a shell with a 1/4 inch screw connector to accommodate external wrist localization devices. For our experiments, we use the `HTC Vive Tracker` for real-time wrist position tracking. However, the design is compatible with other solutions, depending on the user's requirement.

III. Retargeting

A. Action Retargeting

Fig. 4: **Action retargeting results**: Teleoperating the LEAP Hand to grasp a toy in the real world and teleoperating the Shadow Hand, Inspire Hand, and Allegro Hand in simulation.

Our approach combines `Forward Kinematics (FK)` to compute human fingertip positions and `Inverse Kinematics (IK)` through Mink [39] to calculate the corresponding robotic hand positions.

The rigid-body glove design with anthropomorphic kinematics enables precise fingertip tracking via FK calculation. To bridge size disparities between human hand and target robotic hand, we we introduce a scaling factor in IK calculation. This ensures an intuitive teleoperation experience where the robotic hand naturally mirrors the human hand's gestures.

In our experiments, we deploy the system on the LEAP hand [26] in real-world scenarios and test it with various robotic hands in the MuJoCo simulator. Figure 4 presents the retargeting results of DOGlove in both simulation and real-world environments.

B. Haptic+Force Retargeting

In our experimental setup, we install a `1-D force sensor` on each fingertip of the LEAP Hand, with a measurement range of 3 kg and a precision of 1 g. A combination strategy for integrating haptic and force feedback is designed to optimize performance. This strategy along with the corresponding thresholds and feedback patterns is summarized in Table I. For force feedback, the Dynamixel servos operate in `current-based position control mode`. The force readings from the LEAP Hand fingertips are clamped to the range [0g, 3000g], and mapped linearly to the K_P gain of the Dynamixel servos. For haptic feedback, we use `waveform ID 56` from the haptic engine library, corresponding to `Pulsing Sharp 1-100%`.

Force Sensor Readings (g)	Haptic Feedback	Force Feedback
<10	✗	✗
10–50	✓	✗
50–100	✓	✓
>100	✗	✓

Table I: **The combination strategy** for haptic+force feedback in DOGlove.

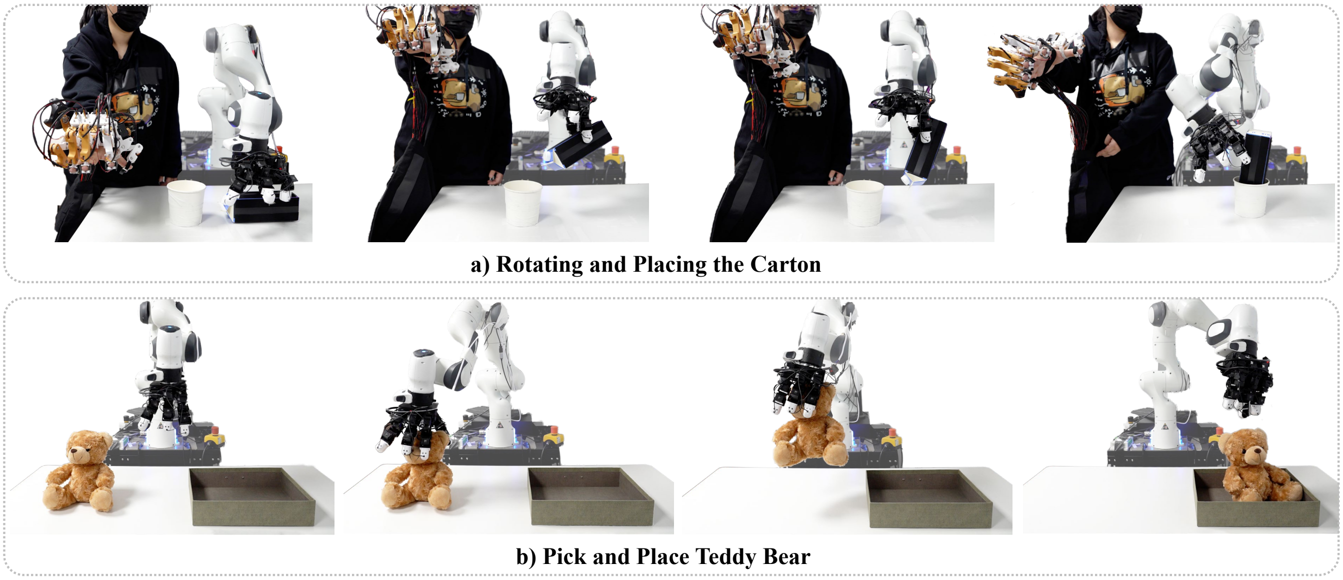


Fig. 5: (a) **Teleoperation experiment**: In in-hand rotation, the challenge is to slightly release the fingers, allowing the carton to rotate without slipping out (as shown in the middle two images). (b) **Imitation learning evaluation**: The robot must first locate the bear, then open its hand to grasp it. Due to the bear's size, precise grasping control is required. An inaccurate grasp deforms the bear and causes it to slip out of the fingertips.

IV. Experiments

In this section, we use DOGlove to teleoperate the `LEAP Hand` [26] mounted on the `Franka Robot Arm` to evaluate its effectiveness.

A. Rotating and Placing the Carton

	Success Rate	Average Completion Time (s)
Only Force	9/10	**18.92**
Only Haptic	9/10	21.16
Haptic+Force	**10/10**	**19.89**
No Haptic/Force	4/10	24.76
AnyTeleop	1/10	54.85

Table II: **Quantitative experiment results.** Haptic+force feedback enables operators to achieve a higher success rate and a faster average completion time, as haptic feedback provides contact information, while force feedback indicates the proper timing for in-hand rotation.

Task: This is a long-horizon contact-rich task. As shown in Fig 5a, the operator must first pick up the carton horizontally, then perform an in-hand rotation, orienting the carton vertically before placing it into a small bucket.

Metrics: Performance is evaluated using two metrics:

- Success Rate: A trial is denoted as successful if the carton rotates more than 45 degrees and is successfully placed into the bucket.
- Completion Time: The total time taken to complete the entire process.

Challenges:

- Precise Manipulation: The operator must accurately teleoperate to rotate the carton while preventing it from falling.
- Visual Obstacle: Grasping the carton is hindered by visual obstacles, as the operator cannot see the contact points between the robotic hand's fingers and the carton.

Performance: Table II shows that both haptic and force feedback significantly improve the teleoperation success rate and reduce completion time. While force feedback alone results in a comparable average completion time, haptic+force feedback achieves a higher success rate. The vision-based MoCap method AnyTeleop [21] struggles with in-hand rotation in this task.

B. Imitation Learning

We show DOGlove is capable of collecting high-quality demonstrations. **3D Diffusion Policy (DP3)** [40] is selected as our imitation learning algorithm, and we use `Realsense L515` to acquire the point cloud inputs, which are then downsampled to 1024 points using farthest point sampling [20]. The data collected by DOGlove is used to train policies for various downstream tasks.

We evaluate imitation learning performance on basic contact-rich task. As shown in Fig 5b, the robot must grasp a teddy bear and place it into a designated box. During data collection, the teddy bear's initial position is randomized within a 30$\times$20 cm area, and DOGlove collects 40 demonstrations to train the policy. In evaluation, the bear is again randomly placed in the same area. Across 20 trials, the success rate is 70% (14/20), with failures primarily due to the teddy bear slipping out of the robotic hand when not grasped firmly.

V. Conclusion

In this paper, we present DOGlove, a low-cost, open-source haptic+force feedback glove designed for dexterous manipulation. DOGlove enables precise and efficient execution of long-horizon, contact-rich tasks. Experimental results show that DOGlove enhances the operator's immersive teleoperation experience while also serving as an effective tool for training imitation learning policies. To support further research and contribute to the community, all hardware designs and code are open-sourced.

References

[1] Jorge Aldaco, Travis Armstrong, Robert Baruch, Jeff Bingham, Sanky Chan, Kenneth Draper, Debidatta Dwibedi, Chelsea Finn, Pete Florence, Spencer Goodrich, et al. Aloha 2: An enhanced low-cost hardware for bimanual teleoperation. *arXiv preprint arXiv:2405.02292*, 2024.

[2] Homanga Bharadhwaj, Abhinav Gupta, Vikash Kumar, and Shubham Tulsiani. Towards generalizable zero-shot manipulation via translating human interaction plans. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6904–6911. IEEE, 2024.

[3] Ilaria Cerulo, Fanny Ficuciello, Vincenzo Lippiello, and Bruno Siciliano. Teleoperation of the schunk s5fh underactuated anthropomorphic hand using human hand motion tracking. *Robotics and Autonomous Systems*, 89: 75–84, 2017.

[4] Pietro Cerveri, Elena De Momi, N Lopomo, Gabriel Baud-Bovy, RML Barros, and Giancarlo Ferrigno. Finger kinematic modeling and real-time hand motion estimation. *Annals of biomedical engineering*, 35:1989–2002, 2007.

[5] Xuxin Cheng, Jialong Li, Shiqi Yang, Ge Yang, and Xiaolong Wang. Open-television: Teleoperation with immersive active visual feedback. In *8th Annual Conference on Robot Learning*, 2024.

[6] Cheng Chi, Zhenjia Xu, Siyuan Feng, Eric Cousineau, Yilun Du, Benjamin Burchfiel, Russ Tedrake, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion. *The International Journal of Robotics Research*, page 02783649241273668, 2023.

[7] Cheng Chi, Zhenjia Xu, Chuer Pan, Eric Cousineau, Benjamin Burchfiel, Siyuan Feng, Russ Tedrake, and Shuran Song. Universal manipulation interface: In-the-wild robot teaching without in-the-wild robots. In *Proceedings of Robotics: Science and Systems (RSS)*, 2024.

[8] Dexta Robotics. Dexta robotics official website. https://www.dextarobotics.com/, 2025.

[9] Runyu Ding, Yuzhe Qin, Jiyue Zhu, Chengzhe Jia, Shiqi Yang, Ruihan Yang, Xiaojuan Qi, and Xiaolong Wang. Bunny-visionpro: Real-time bimanual dexterous teleoperation for imitation learning. *arXiv preprint arXiv:2407.03162*, 2024.

[10] Zicong Fan, Omid Taheri, Dimitrios Tzionas, Muhammed Kocabas, Manuel Kaufmann, Michael J Black, and Otmar Hilliges. Arctic: A dataset for dexterous bimanual hand-object manipulation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12943–12954, 2023.

[11] Hongjie Fang, Hao-Shu Fang, Yiming Wang, Jieji Ren, Jingjing Chen, Ruo Zhang, Weiming Wang, and Cewu Lu. Airexo: Low-cost exoskeletons for learning whole-arm manipulation in the wild. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 15031–15038. IEEE, 2024.

[12] Zipeng Fu, Tony Z. Zhao, and Chelsea Finn. Mobile aloha: Learning bimanual mobile manipulation with low-cost whole-body teleoperation. In *Conference on Robot Learning (CoRL)*, 2024.

[13] HaptX. Haptx official website. https://haptx.com/, 2025.

[14] Aadhithya Iyer, Zhuoran Peng, Yinlong Dai, Irmak Guzey, Siddhant Haldar, Soumith Chintala, and Lerrel Pinto. OPEN TEACH: A versatile teleoperation system for robotic manipulation. In *8th Annual Conference on Robot Learning*, 2024.

[15] Hangxin Liu, Xu Xie, Matt Millar, Mark Edmonds, Feng Gao, Yixin Zhu, Veronica J Santos, Brandon Rothrock, and Song-Chun Zhu. A glove-based system for studying hand-object manipulation via joint pose and force sensing. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6617–6624. IEEE, 2017.

[16] Hangxin Liu, Zhenliang Zhang, Xu Xie, Yixin Zhu, Yue Liu, Yongtian Wang, and Song-Chun Zhu. High-fidelity grasping in virtual reality using a glove-based system. In *2019 international conference on robotics and automation (icra)*, pages 5180–5186. IEEE, 2019.

[17] Manus Meta. Manus meta robotics official website. https://www.manus-meta.com/robotics, 2025.

[18] Malte Mosbach, Kara Moraw, and Sven Behnke. Accelerating interactive human-like manipulation learning with gpu-based simulation and high-quality demonstrations. In *2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids)*, pages 435–441. IEEE, 2022.

[19] OptiTrack. OptiTrack Motion Capture Systems. https://www.optitrack.com/, 2025.

[20] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 652–660, 2017.

[21] Yuzhe Qin, Wei Yang, Binghao Huang, Karl Van Wyk, Hao Su, Xiaolong Wang, Yu-Wei Chao, and Dieter Fox. Anyteleop: A general vision-based dexterous robot arm-hand teleoperation system. In *Proceedings of Robotics: Science and Systems (RSS)*, 2023.

[22] Felipe Sanches, Geng Gao, Nathan Elangovan, Ricardo V Godoy, Jayden Chapman, Ke Wang, Patrick Jarvis, and Minas Liarokapis. Scalable. intuitive human to robot skill transfer with wearable human machine interfaces: On complex, dexterous tasks. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6318–6325. IEEE, 2023.

[23] Gaspare Santaera, Emanuele Luberto, Alessandro Serio, Marco Gabiccini, and Antonio Bicchi. Low-cost, fast and accurate reconstruction of robotic and human postures via imu measurements. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages

2728–2735. IEEE, 2015.

[24] Max Schwarz, Christian Lenz, Andre Rochow, Michael Schreiber, and Sven Behnke. Nimbro avatar: Interactive immersive telepresence with force-feedback telemanipulation. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5312–5319. IEEE, 2021.

[25] SenseGlove. Senseglove official website. https://www.senseglove.com/, 2025.

[26] Kenneth Shaw, Ananye Agarwal, and Deepak Pathak. Leap hand: Low-cost, efficient, and anthropomorphic hand for robot learning. *Robotics: Science and Systems (RSS)*, 2023.

[27] Kenneth Shaw, Yulong Li, Jiahui Yang, Mohan Kumar Srirama, Ray Liu, Haoyu Xiong, Russell Mendonca, and Deepak Pathak. Bimanual dexterity for complex tasks. In *8th Annual Conference on Robot Learning*, 2024.

[28] Aravind Sivakumar, Kenneth Shaw, and Deepak Pathak. Robotic telekinesis: Learning a robotic hand imitator by watching humans on youtube. In *Robotics: Science and Systems*, 2022.

[29] Shuran Song, Andy Zeng, Johnny Lee, and Thomas Funkhouser. Grasping in the wild: Learning 6dof closed-loop grasping from low-cost demonstrations. *IEEE Robotics and Automation Letters*, 5(3):4978–4985, 2020.

[30] Omid Taheri, Nima Ghorbani, Michael J Black, and Dimitrios Tzionas. Grab: A dataset of whole-body human grasping of objects. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IV 16*, pages 581–600. Springer, 2020.

[31] Ultraleap. Ultraleap official website. https://www.ultraleap.com/, 2025.

[32] Vicon Motion Systems. Vicon Motion Capture Systems. https://www.vicon.com/, 2025.

[33] Chen Wang, Linxi Fan, Jiankai Sun, Ruohan Zhang, Li Fei-Fei, Danfei Xu, Yuke Zhu, and Anima Anandkumar. Mimicplay: Long-horizon imitation learning by watching human play. In *7th Annual Conference on Robot Learning*, 2023.

[34] Chen Wang, Haochen Shi, Weizhuo Wang, Ruohan Zhang, Li Fei-Fei, and C Karen Liu. Dexcap: Scalable and portable mocap data collection system for dexterous manipulation. In *Proceedings of Robotics: Science and Systems (RSS)*, 2024.

[35] Han Xu, Mingqi Chen, Gaofeng Li, Lei Wei, Shichi Peng, Haoliang Xu, and Qiang Li. An immersive virtual reality bimanual telerobotic system with haptic feedback. *arXiv preprint arXiv:2501.00822*, 2025.

[36] Mengda Xu, Zhenjia Xu, Cheng Chi, Manuela Veloso, and Shuran Song. XSkill: Cross embodiment skill discovery. In *7th Annual Conference on Robot Learning*, 2023.

[37] Mengda Xu, Zhenjia Xu, Yinghao Xu, Cheng Chi, Gordon Wetzstein, Manuela Veloso, and Shuran Song. Flow as the cross-domain manipulation interface. In *8th Annual Conference on Robot Learning*, 2024.

[38] Jingyun Yang, Junwu Zhang, Connor Settle, Akshara Rai, Rika Antonova, and Jeannette Bohg. Learning periodic tasks from human demonstrations. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 8658–8665. IEEE, 2022.

[39] Kevin Zakka. Mink: Python inverse kinematics based on MuJoCo, July 2024. URL https://github.com/kevinzakka/mink. Software.

[40] Yanjie Ze, Gu Zhang, Kangning Zhang, Chenyuan Hu, Muhan Wang, and Huazhe Xu. 3d diffusion policy: Generalizable visuomotor policy learning via simple 3d representations. In *Proceedings of Robotics: Science and Systems (RSS)*, 2024.

[41] Tony Z Zhao, Vikash Kumar, Sergey Levine, and Chelsea Finn. Learning fine-grained bimanual manipulation with low-cost hardware. In *Proceedings of Robotics: Science and Systems (RSS)*, 2023.

[42] Tony Z Zhao, Jonathan Tompson, Danny Driess, Pete Florence, Seyed Kamyar Seyed Ghasemipour, Chelsea Finn, and Ayzaan Wahid. Aloha unleashed: A simple recipe for robot dexterity. In *8th Annual Conference on Robot Learning*, 2024.

[43] Wenping Zhao, Jinxiang Chai, and Ying-Qing Xu. Combining marker-based mocap and rgb-d camera for acquiring high-fidelity hand motion data. In *Proceedings of the ACM SIGGRAPH/eurographics symposium on computer animation*, pages 33–42, 2012.

[44] Christian Zimmermann, Duygu Ceylan, Jimei Yang, Bryan Russell, Max Argus, and Thomas Brox. Freihand: A dataset for markerless capture of hand pose and shape from single rgb images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 813–822, 2019.

APPENDIX

A. Related Work

1) Data Collection from Human Demonstrations:

A substantial amount of task-specific data is essential for imitation learning. In dexterous manipulation, obtaining high-quality hand motion data is critical for training effective policies. Prior work includes extracting demonstration data from human videos [28, 36, 37, 2] and hand trajectories [33, 38]. While these approaches are accessible and have shown promising results, the significant visual gap between recorded human demonstrations and the robot's perception often makes real-world transfer challenging. An alternative is using dedicated hardware for data collection to bridge this gap. Hand-held grippers [29, 7, 22] have proven effective in capturing robot manipulation data. However, these systems are primarily designed for parallel grippers. Another widely used approach is MoCap systems, which record human demonstrations and extract hand motion data. These systems include camera-based methods [21, 44], glove-based tracking systems [34, 15, 16], marker-based tracking [43], and commercial MoCap solutions [30, 10]. While MoCap offers high-precision tracking, bridging the embodiment gap between human and robotic hands remains a persistent challenge.

2) Dexterous Hand Teleoperation:

Collecting high-quality human demonstrations through robotic teleoperation systems [11, 5, 9, 14] also plays a critical role for advancing dexterous manipulation. Existing research has explored teleoperation from various perspectives, including leader-follower setups such as ALOHA [41, 42, 12, 1]. However, teleoperating dexterous hands remains a significant challenge. OpenTelevision [5] leverages VR devices to capture hand poses and streams the pose information for retargeting to robotic hands. BiDex [27], on the other hand, implements a teleoperation system based on commercial motion capture gloves [17] and leader arms. Compared to these frameworks and other glove-based systems [15, 16], DOGlove offers distinct advantages. It eliminates the need for expensive equipment while precisely capturing fingertip positions and delivering richer haptic force feedback to the operator. This system achieves accurate dexterous hand teleoperation with a low-cost setup, making it an efficient alternative.

3) Teleoperation with Haptic Force Feedback:

While recent studies rely on visual information to capture environmental characteristics, vision alone inherently limits the richness of available sensory data. In contrast, haptic force feedback enhances the teleoperation experience by providing greater immersion and improving perception of the robot's status and movement compared to vision-based methods. Bunny-VisionPro [9] and Liu et al. [16] apply real-time haptic feedback to enable more accurate manipulation. Xu et al. [35] build a bilateral isomorphic bimanual telerobotic system using a commercial force feedback glove [8] to enhance perception and improve performance in complex tasks. NimbRo-Avatar [24] and Mosbach et al. [18] integrate commercial force feedback glove [25] into dexterous teleoperation systems. However, these approaches rely on specialized or expensive equipment. In contrast, DOGlove provides a highly accurate teleoperation system with integrated haptic force feedback at a significantly lower cost and can be widely used in dexterous manipulation.

B. Hardware Design

1) The linkage length:

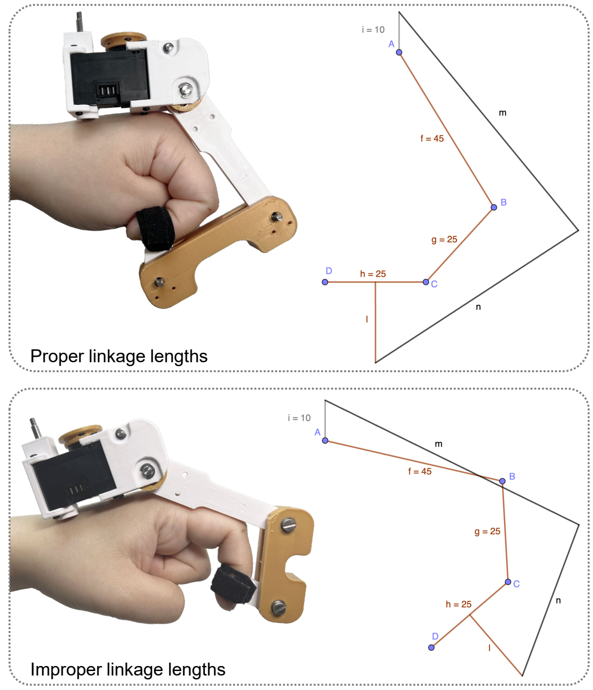


Fig. 6: **Improper linkage lengths can cause collisions** between the human finger (**link** $\boldsymbol{f}, \boldsymbol{g}$) and the glove (**link** $\boldsymbol{m}$), restricting finger movements and leading to discomfort and poor MoCap performance.

The linkage lengths in DOGlove are designed to accommodate the majority of adult human sizes. To achieve this, a standard human finger length was first modeled, and the glove's linkage parameters were simulated to ensure an optimal range of motion. As shown in Figure 6, improper linkage lengths can obstruct the natural flexion-extension of the fingers, leading to discomfort and reduced MoCap performance. Furthermore, DOGlove features a modular design where all fingers share a common structural framework. This modularity enables users to replace linkages with customized sizes as needed, enhancing both adaptability and usability.

2) The exploded view of a single finger:

The exploded view of a single finger is illustrated in Figure 7. The highlighted area indicates the basic components of a rotary joint. Each rotary joint is constructed using an M4×15 shoulder screw to connect the finger linkages, ball bearing, and joint encoder, secured with an M3 locknut. This design ensures smooth and reliable joint rotation. The main body of the finger, colored white and gold, is 3D printed using PETG material for ease of fabrication and durability.

Given the limited space on the back of the human hand, the finger assembly's width is constrained to less than 26 mm. Simultaneously, to provide effective force feedback, the

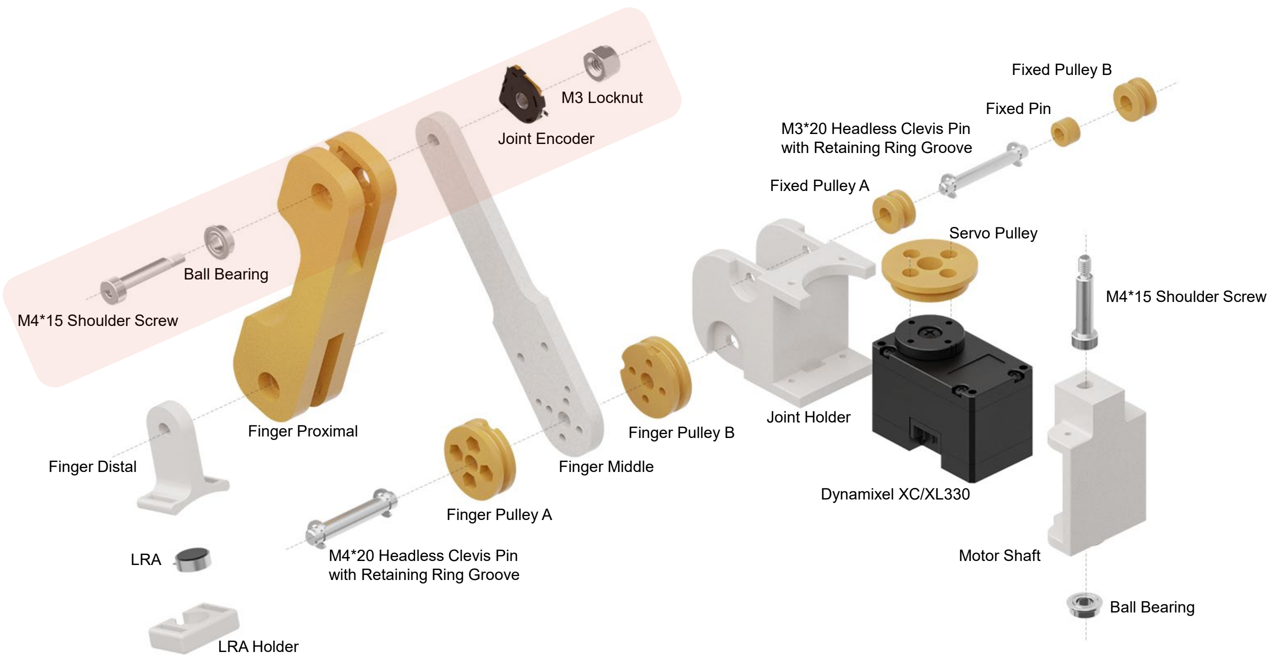


Fig. 7: **Exploded view of the finger assembly**, with the highlighted area indicating the basic components of a rotary joint.

actuator must deliver a stall torque of at least 0.5 N·m. Additionally, adjustable stiffness requires the actuator's current to be regulated. Since the actuator is directly connected to the pulley system as a rotary joint for MCP_B, it is essential to measure its rotary position in real time to achieve precise joint angle control. Considering these design requirements, the `Dynamixel XC/XL330` servo motors were selected as the actuators for force feedback. It fulfills the torque, size, and real-time position measurement needs, making it a suitable choice for DOGlove.

*C. User Study: Object Perception **w/o Visual Input***

Fig. 8: **User Study.** (a) Experiment setup: Users wear an eyemask and headphones to eliminate visual and auditory feedback. (b)–(f) Object pairs tested in the study.

Task: Five untrained human operators participate in this user study. During the experiment, they are required to distinguish between five pairs of objects solely through feedback from DOGlove, without any visual or auditory input (achieved by wearing an eyemask and headphones). In each trial, a pair of objects is randomly selected, and users provide their answers immediately after experiencing feedback from DOGlove for both objects. Figure 8 illustrates the experiment setup and the five object pairs, selected based on factors such as shape, size, and softness.

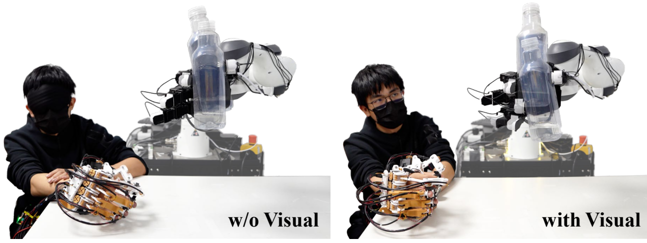
	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5
Only Force	5/5	5/5	5/5	**4/5**	0/5
Only Haptic	5/5	5/5	5/5	3/5	**3/5**
Haptic+Force	5/5	5/5	5/5	3/5	2/5

Table III: **Success rates in the user study.** All feedback modes perform well for the basic pairs. For the challenging pairs, force feedback is more sensitive to softness, while haptic feedback is more sensitive to shape.

Metrics: Users' ability to distinguish object pairs is evaluated based on their success rate.

Challenges: The five object pairs are intentionally chosen based on the following considerations:

- Pair 1: <u>Basic Pair</u>, different shape. The ball and the box have distinctly different shapes (Fig 8b).
- Pair 2: <u>Basic Pair</u>, similar shape, different size. The peanut bottle and the coffee paper cup share a similar cylindrical shape, but their diameters differ slightly (Fig 8c).
- Pair 3: <u>Basic Pair</u>, similar softness, different size. The two toys have similar softness and shapes but vary in size (Fig 8d).
- Pair 4: <u>Challenging Pair</u>, similar size and shape, different softness. Two identical bottles are used, one filled with pure water (soft) and the other filled with carbonated cola, shaken to increase its hardness (Fig 8e).
- Pair 5: <u>Challenging Pair</u>, similar shape, different size and softness. A toy cabbage (softer, larger) and a real cab-



	w/o Visual	with Visual	
	Success Rate	Success Rate	Avg. Slip Dev. (cm)
Only Force	9/10	**9/10**	**1.1**
Only Haptic	5/10	7/10	2.4
Haptic+Force	**10/10**	6/10	1.5
No Haptic/Force	5/10	7/10	2.1
AnyTeleop	0/10	4/10	2.4

Bottle-Slipping Experiment

Fig. 9: **Quantitative results of the bottle-slipping experiment**: Without visual feedback, force feedback significantly improves the task success rate. With visual feedback, it enhances precise control.

bage (Fig 8f).

Performance: As shown in Table III, even without visual and auditory feedback, all participants effortlessly distinguish basic pairs 1-3. For challenging pair 4, most participants can perceive softness using only force feedback. Some also discern softness using only haptic feedback by evaluating the duration of contact during deformation.

For challenge pair 5, when the robotic hand grasps the softer toy cabbage, it deforms to resemble the size of the real cabbage. This deformation increases its perceived softness, making it difficult for participants to distinguish using force feedback alone.

For both challenge pairs, combining haptic and force feedback slightly reduces user sensitivity, leading to a marginally lower accuracy.

D. Bottle-Slipping Experiment

1) Teleoperation w/o Visual Input

Task: In this experiment, the human operator must perform a bottle-slipping action relying solely on feedback from DOGlove. A 15-second countdown timer is set for each trial. If the bottle successfully slips without falling within the 15 seconds, the trial is denoted as successful.

Metrics: The success rate.

Challenges: Without any visual or auditory input (achieved by wearing an eyemask and headphones), the operator must determine if the bottle is slipping at the right speed or too quickly, risking a fall.

Performance: As shown in Fig 9a, force feedback significantly improves the success rate of this task. Additionally, incorporating haptic feedback further enhances overall performance. However, since the fingers of the LEAP Hand maintain continuous contact with the bottle during the task, haptic feedback does not provide additional information beyond using the glove solely as a MoCap device, resulting in the same success rate for both conditions.

Due to differences in retargeting strategies, even a slight change in human finger position can lead to a significant deviation in the LEAP Hand's movements. As a result, AnyTeleop [21] struggles to perform the slipping task effectively.

2) Teleoperation w/ Visual Input

Task: Unlike the previous blindfolded experiment, this experiment allows operators to have visual feedback. To further evaluate the operator's control ability, they are required to slip the bottle to a specified distance (9 cm). A trial is denoted as successful if the bottle slips without falling. Additionally, We measure the deviation between the actual slipping distance and the target distance (9 cm).

Metrics: Performance is evaluated using two metrics:

- Success Rate: A trial is denoted as successful if the bottle slips without falling.
- Slipping Deviation: This measures the difference between the target sliding distance (9 cm) and the actual slipping distance, with a smaller deviation indicating greater operational accuracy.

Challenges: Operators must precisely control the bottle to achieve the desired distance. While a greedy approach often causes the bottle to fall and results in failure, a conservative approach leads to an unsatisfactory distance deviation.

Performance: This task evaluates not only success rate but also teleoperation precision. To minimize slipping deviation, operators are instructed to control the LEAP Hand carefully and optimally. As shown in Fig 9a, similar to previous results, haptic feedback does not provide additional information and may even interfere with task precision. However, force feedback enables operators to minimize slipping deviation more effectively. While using DOGlove solely as a MoCap device achieves the same success rate as with haptic force feedback, it results in a larger average slipping deviation.

E. Limitations and Future Work

DOGlove is a powerful haptic force feedback glove for dexterous manipulation, but several limitations remain. First, the weight of DOGlove is inevitably high, as it utilizes 5 commercial servos, bringing the total weight to 450g. Additionally, in agile teleoperation scenarios, performance is constrained by the servos' maximum speed and torque output. To address these issues, we are investigating the use of lighter servos with smaller reduction ratios and designing a customized reduction mechanism to balance speed and torque more effectively. Second, although DOGlove is designed to accommodate most hand sizes, it may be uncomfortable for some users. To enhance adaptability and wearability, we are developing CAD files for linkages in multiple sizes, enabling customization for various hand dimensions.