2025 IEEE/SICE International Symposium on System Integration (SII)
January 21-24, 2025, Munich, Germany

Low-Cost Robot Operation Interface for Simultaneous Hand Position Input and Force Fine-Tuning Using Visual-Based Tactile Sensor

Tomoya Morita[1], Yuki Yamashita[1], Yaonan Zhu[2] and Yasuhisa Hasegawa[1]

***Abstract*— We propose a low-cost robot operation interface that combines a mixed-reality device and vision-based tactile sensor for simultaneous grip force fine-tuning and position control. In a paper creasing demonstration, we confirm that the proposed system dynamically reproduces a small force input by the operator to the controller as the grasping force of a robot gripper.**

I. INTRODUCTION

Recent advancements in virtual reality have promoted the research and development of intuitive interfaces for robots to replicate human head and hand movements. However, when handling deformable objects, such as fabric or paper, synchronizing the pose alone is insufficient for robots or operators to accurately recognize the grasping or contact states, making task execution extremely challenging.

To address this manipulation problem, robots must be equipped with tactile sensors to provide human-like sensing capabilities. Tactile sensors provide robots with sensing information, such as object contact and hardness, thereby enhancing robot perception and enabling sophisticated object interactions. In particular, vision-based tactile sensors (VBTSs) can measure various types of sensing information and are compact and cost-effective, facilitating their integration into robots. Thus, VBTSs have been widely used in recent studies.

Zhu et al. [1] created depth maps from tactile images acquired by a DIGIT VBTS attached to a robot gripper and dynamically estimated normal forces using a relation derived from the forces exerted on the sensor. The estimated force was fed back to the operator via a force feedback device. In this study, we used the GelSight sensor [2] that includes a soft sensor surface for high-resolution geometric sensing. This VBTS can acquire detailed sensing information, including the surface shape of the contact objects, contact force, torque, and shear force, from images acquired by its optical system and internal camera.

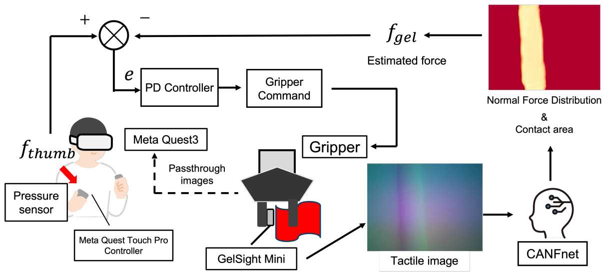

Fig. 1: Force estimation and fine-tuning using proposed interface.

Although force feedback delivers tactile information to operators, it has a limited operating range and very high implementation costs. Instead, we adopted a commercially available mixed-reality (MR) device as a low-cost robot operation interface without range limitations. Using a controller that measures the pushing force of the thumb, the gripper could be controlled such that the normal force estimated from the tactile image was the pushing force (Fig. 1). This allowed the reproduction of subtle forces from the operator to control the robot. In addition, the force input surface was clear to the operator, allowing for an intuitive operation as they gripped the hand while feeling the reaction force. The proposed robot operation interface enabled the fine-tuning of both force and position inputs.

II. PROPOSED SYSTEM

A. Hardware Setup

Fig. 2 shows the configuration of the proposed system. In this study, we used a commercially available MR head-mounted display (HMD) (Meta MetaQuest 3) and controllers (MetaQuest Touch Pro Controllers) as the operational input interfaces. Images captured by multiple cameras built into the HMD were presented to the operator as passthrough images, and pressure sensors in the controllers measured the thumb pushing force up to 6 N. A two-finger robot gripper (Hand-E Robotiq Adaptive Gripper) was attached to the end effector of a six-axis collaborative robot (COBOTTA PRO DENSO WAVE). One of the gripper fingers was replaced with a 3D-printed attachment for the GelSight Mini VBTS. To prevent damage to the tactile sensor surface due to excessive shear force, the sensor surface was covered with a thin transparent film. The other fingertip was replaced with a 3D-printed convex fingertip. The position of the MR controller was sent as the target position for the gripper tip to a computer with ROS (Robot Operating System) installed via Unity, which is a 3D game engine running on a computer that can be

---

*This work was supported in part by JST SICORP, Japan, under Grant JPMJSC2305. This work was partly supported by the DII Collaborative Graduate Program for Accelerating Innovation in Future Electronics, Nagoya University. This work was financially supported by JST SPRING, Grant Number JPMJSP2125.The author T. Morita would like to take this opportunity to thank the “THERS Make New Standards Program for the Next Generation Researchers.”

[1] Tomoya Morita, Yuki Yamashita and Yasuhisa Hasegawa are with the Department of Micro-Nano Mechanical Science and Engineering, Nagoya University, Nagoya, 464-8603 Japan (email: morita@robo.mein.nagoya-u.ac.jp; yamashita@robo.mein.nagoya-u.ac.jp; hasegawa@mein.nagoya-u.ac.jp).

[2] Yaonan Zhu is with the School of Engineering, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo, 113-8656, Japan (email: yaonan.zhu@weblab.t.u-tokyo.ac.jp).

979-8-3315-3161-4/25/$31.00 ©2025 IEEE

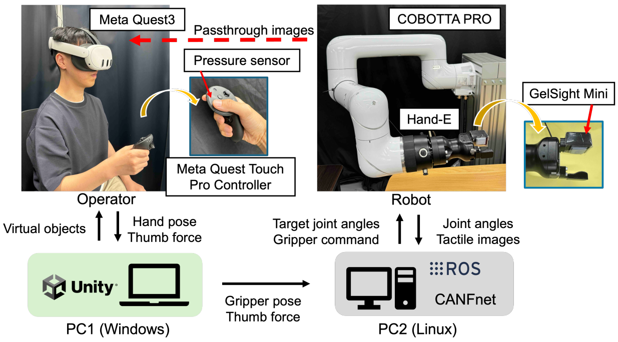

Fig. 2: Configuration of proposed system

connected to the HMD via a cable. The target position in a Cartesian coordinate system was converted into target joint angles of the robot arm by solving the inverse kinematics.

B. Force Estimation and Fine-Tuning

We used CANFnet [3], a neural network trained on a labeled dataset collected using a GelSight Mini VBTS and a force/torque sensor. CANFnet can estimate the contact area and normal force per pixel with high accuracy from tactile images captured by a VBTS, even in dynamic scenarios. The gripper open/close command (g) obeyed PD control such that the estimated normal force (f_{gel}) approached the operator's thumb pushing force (f_{thumb}) as follows:

$$\begin{cases} e(t) = f_{thumb}(t) - f_{gel}(t) \\ g(t) = K_p e(t) + K_d \dfrac{d}{dt} e(t), \end{cases} \tag{1}$$

where K_p and K_d are experimentally determined control gains. The gripper open/close command was converted into an integer between 0 and 255.

III. EXPERIMENT

A. Experimental Procedure

We conducted an illustrative experiment to create creases in paper using a manipulator. Fig. 3 shows the experimental setup during task execution. The operator wore an HMD and held the MR controller in the right hand while sitting. The operator was presented with pass-through images, allowing him to see the workspace while wearing the HMD (Fig. 4). The position and orientation of the right hand were input to the gripper only if the grip trigger on the controller was pulled. The index-finger trigger was linked to the rough opening and closing of the gripper, which performed a closing motion until the two fingers made contact. After the gripper grasped the paper, the pushing force was input to the PD controller when the operator' s thumb pressed the controller surface. The strength of the operator's pushing force was visualized as the size of a virtual sphere appearing near the controller and reflected as the intensity of the controller vibration. The operator grasped the curved part of the paper, which was folded in half without a crease, with the gripper and moved the arm to create a crease under fine-tuning of the gripping force.

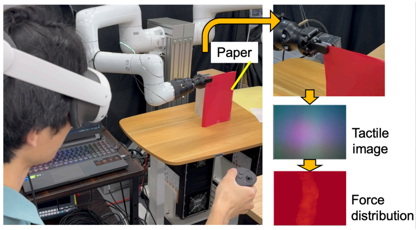

Fig. 3: Paper creasing task

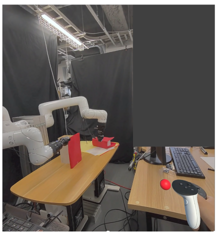

Fig. 4: Operator's view

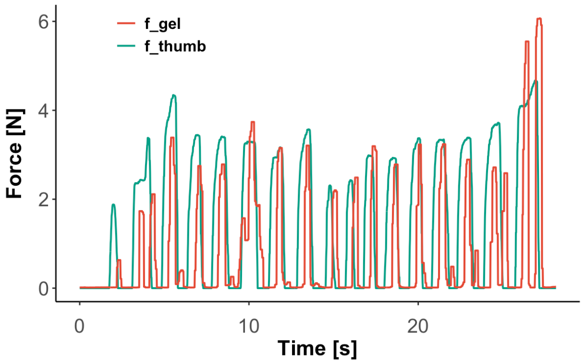

Fig. 5: Operator's thumb force and estimated force while creasing paper.

B. Experimental Results

The operator successfully created a crease while fine-tuning the grasping force and moving the arm. Fig. 5 shows forces f_{thumb} and f_{gel} during task execution. In the experimental environment, CANFnet returned the normal force value 340 ms after the object contacted the tactile sensor, resulting in a corresponding delay of f_{gel} with respect to f_{thumb}. The mean absolute error calculated after correcting for this delay was 0.83 ± 0.027 N, indicating that the opening and closing of the gripper were properly controlled to achieve a grasping force close to that delivered by the operator.

IV. CONCLUSION

We propose a robot operation interface to simultaneously deliver fine-tuned force and position inputs using a VBTS and commercially available MR device. A paper creasing task was performed using the proposed system, confirming that the grasping force of the gripper can be controlled by a fine-tuned small force input delivered by the operator and the estimated normal force from tactile images.

REFERENCES

[1] Y. Zhu, S. Nazirjonov, B. Jiang, J. Colan, T. Aoyama, Y. Hasegawa, B. Belousov, K. Hansel, and J. Peters, "Visual Tactile Sensor Based Force Estimation for Position-Force Teleoperation," in *2022 IEEE International Conference on Cyborg and Bionic Systems (CBS)*, 2023, pp. 49–52.

[2] W. Yuan, S. Dong, and E. H. Adelson, "GelSight: High-Resolution Robot Tactile Sensors for Estimating Geometry and Force," *Sensors*, vol. 17, no. 12, 2017.

[3] N. Funk, P.-O. Müller, B. Belousov, A. Savchenko, R. Findeisen, and J. Peters, "CANFnet: High-Resolution Pixelwise Contact Area and Normal Force Estimation for Visuotactile Sensors Using Neural Networks," in *Embracing Contacts - Workshop at ICRA 2023*, 2023.