Enhancing Teleoperated Manipulation with Visuotactile Sensors: Non-Intuitive but Informative Haptic Feedback

Maria Ramos Gonzalez*, Matthew Boyd*, Michael Hagenow, Julie Shah, Alberto Rodriguez
{mariarg, mcboyd, hagenow, arnoldj, albertor}@mit.edu
Massachusetts Institute of Technology

***Abstract*— Creating meaningful feedback for human operators of robotic teleoperation systems can be challenging when the robot end effector is not as sensorized as the human hand. Conveying touch, weight, fragility, and other characteristics to a human operator may be achieved with a combination of multimodal inputs and outputs. In this study, we explore the utility of deformable visuotactile sensors as input devices to relay manipulated object surface properties and local sensor-contact deformation as real-time haptic and image sensor output. Our teleoperation setup consists of a custom-designed parallel-jaw gripper equipped with visuotactile sensors attached to a 6-DoF robot arm and a custom-designed hand-held controller equipped with haptic motors. We designed two manipulation experiments where participants 1) determined the relative weights of objects and 2) repositioned a fragile object; participants received varying combinations of line-of-sight, haptic, and image sensor feedback. This extended abstract presents the results for 1) participants' sense of embodiment while using the system, 2) SUS scores, 3) PSSUQ scores, 4) feedback preferences and perception, and 5) NASA-TLX workload scores. The results of our study highlight the importance of integrating new sensory applications beyond what humans possess (we cannot "see" through our fingertips) in a thoughtful design to enhance human performance and understanding of feedback sensation during robotic teleoperation manipulation tasks.**

Keywords-Tactile and Haptic Sensing, Hardware Design for Contact

I. Past Works

During autonomous or human-in-the-loop robotic teleoperation tasks, there may occur an instance when the view of the object targeted for manipulation may be occluded, either by poor lighting conditions, limited camera angles in the system environment, or by the robotic gripper itself. Research studies have shown that including tactile sensors in robotic grippers enables autonomous robotic manipulation systems to perform tasks better than systems that lack tactile sensors [1-10]. Although the performance of these autonomous systems is enhanced by the use of tactile sensors, the time to complete a task is markedly slower than a human's ability to conduct the same task with their biological hands. The standard method for acquiring local gripper contact is by integrating force sensors on the grippers, which does not provide detailed information about the object coming in contact with the gripper. Few studies have integrated robotic teleoperation with visuotactile sensors as an input method for providing feedback to a human operator but none enabled the operator to see the deformation of the visuotactile sensor in real time as a mode of feedback [11-13]. We believe allowing operators to "see" the contact in real-time can enable richer modes of feedback.

II. Experiment Design

We designed two manipulation experiments where participants determined the relative weights of objects and placed them in order from least to greatest weight (Exp. 1) and manipulated a fragile object instrumented with a force sensitive resistor (FSR) (Exp. 2). Participants received four combinations of feedback:

1. **v -** line-of-sight
Participants are able to see the workspace with no obvious occlusions and may sit, stand, and move their heads as long as they do not cross into the workspace.
2. **vh -** line-of-sight with haptic vibration through the hand-held controller
Participants received haptic vibration through the hand-held controller which was mapped from the contact depth of the GelSight minis located on the fingers of the robot gripper.
3. **vg -** line-of-sight with live video of the image from the visuotactile sensor deformation displayed through a large screen
Participants were shown a live video feed of the GelSight mini contact deformation through a large screen placed within their line-of-sight with no obvious occlusions.
4. **vhg –** line-of-sight with haptic vibration and live video
Participants received all available forms of feedback.

A counterbalanced measures calculation determined the need for a minimum of 24 participants. In order to minimize learning effects, a Balanced Latin Squares calculator determined the order each participant would receive the feedback scenarios. The study was approved by MIT's Institutional Review Board. Participants were given 5 minutes to practice with the system prior to the start of each experiment but not for each feedback scenario. During the practice session, all output modalities were enabled to provide maximum information and participants were encouraged to handle the objects both with their hands and with the teleoperation system.

III. Results

The following figures and tables summarize the following: 1) participants' sense of embodiment while using the system with the gross robot movement mapping to their arm movement and the fine manipulation gripper movement mapping to their hand operating the controller, 2) SUS scores, 3) PSSUQ scores, 4) feedback preferences and perception, and

This work was supported in part by the MIT Postdoctoral Fellowship for Engineering Excellence.

5) NASA-TLX workload scores.

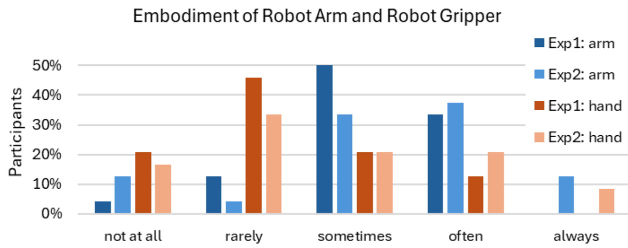

Fig. 1 Survey results of participants' embodiment with the teleop system.

Participants rated the System Usability Score overall as 66.6 for Exp. 1 and 75.4 for Exp. 2 indicating a higher-than-average usability compared to the standard benchmark score of 68 [14].

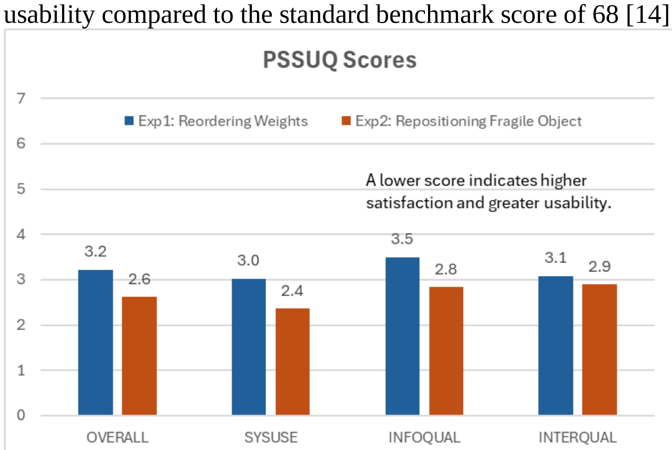

Fig. 2 Survey results of participants' PSSUQ survey on a scale of 1-7. A lower score indicates higher satisfaction and greater usability with the industry standard goal of 2.8 or lower [15].

TABLE I. PERCEPTION AND USE OF FEEDBACK

	A: Exp1	A: Exp2	B: Exp1	B: Exp2
Not at all	45.8%	54.2%	20.8%	37.5%
Rarely	33.3%	25.0%	33.3%	33.3%
Sometimes	8.3%	12.5%	16.7%	16.7%
Often	12.5%	8.3%	16.7%	8.3%
Always	0.0%	0.0%	12.5%	4.2%

Table 1 summarizes participants' perception of the various feedback scenarios for the following questions: **A:** When enabled, how well could you distinguish between the left haptic vibration and the right haptic vibration? **B:** When enabled how much did you rely on the GelSight images displayed on the television?

TABLE II. EXP 2 FEEDBACK SCENARIO PREFERENCES

Which scenario did you prefer?	Exp1	Exp2
visual only	4.2%	29.2%
visual + haptic	33.3%	37.5%
visual + gelsight	33.3%	8.3%
visual + haptic + gelsight	29.2%	25.0%

For Exp. 1, a repeated measures ANOVA revealed a significant difference in perceived Temporal Demand, **F**(3, 69) = 2.82, **p** = 0.045, $\eta^2 p$ = 0.11, indicating a medium effect size. A post-hoc analysis revealed a significant difference occurring between scenario **vg** and **vhg** with **p**<0.01 and a marginally significant difference occurring between **vh** and **vhg** with **p**=0.05. Participants perceived their ability to complete the task to take longer with **vhg**, potentially indicating a higher cognitive load as all feedback was enabled.

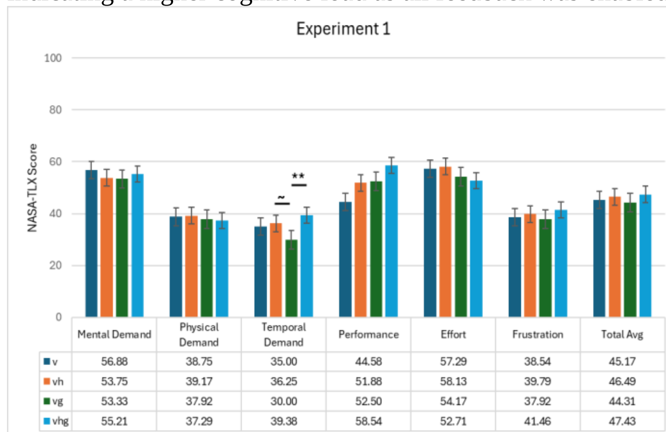

Fig. 3 Average NASA-TLX scores for Exp. 1. Standard error is shown for each score. **p<0.01, ~p=0.05 (marginally significant).

For Exp. 2, no significant difference was found in perceived workload, indicating one feedback scenario was not more demanding than another.

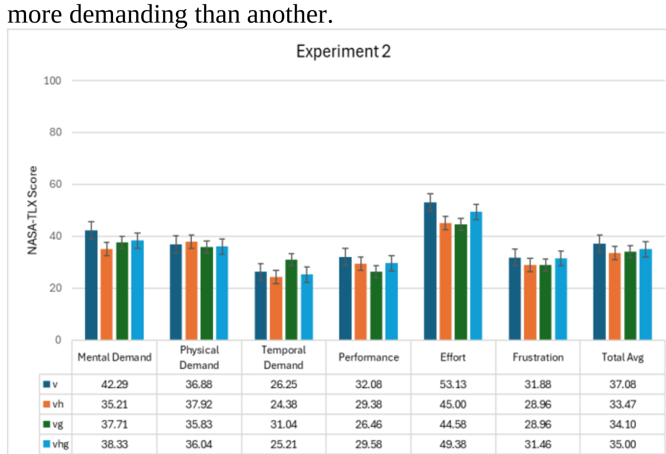

Fig. 4 Average NASA-TLX scores for Exp. 2. with standard error shown.

IV. CONCLUSION

Most striking is the difference in results depending on the task posed to participants. In Exp. 1, participants preferred and relied on the additional forms of feedback to order weights and felt a high sense of embodiment with the teleop system. In Exp. 2, participants least preferred the live GelSight video feed even though they were instructed to pick up the fragile object directly over the FSR, a task which occlusion from the gripper fingers would make difficult to achieve. Although we cannot "see" through our fingertips, the overall results indicate robot teleoperators are able to use visuotactile video feedback with little to no training to accomplish various manipulation tasks. Additionally, using visuotactile sensors to map to haptic feedback enhances manipulation tasks where accurate gripper placement is required.

References

[1] W. Yuan, R. Li, M. Srinivasan, and E. Adelson, "Measurement of Shear and Slip with a GelSight Tactile Sensor," in *IEEE International Conference on Robotics and Automation (ICRA)*, Seattle, May 2015.

[2] C. Wang, S. Wang, B. Romero, F. Veiga, and E. Adelson, "SwingBot: Learning physical features from in-hand tactile exploration for dynamic swing-up manipulation," *IEEE International Conference on Intelligent Robots and Systems*, pp. 5633–5640, 2020, doi: 10.1109/IROS45743.2020.9341006.

[3] E. Donlon, S. Dong, M. Liu, J. Li, E. Adelson, and A. Rodriguez, "Gel Slim : A High-Resolution , Compact , Robust , and Calibrated Tactile-sensing Finger," pp. 1927–1934, 2018.

[4] B. Romero, F. Veiga, and E. Adelson, "Soft, Round, High Resolution Tactile Fingertip Sensors for Dexterous Robotic Manipulation," *Proc IEEE Int Conf Robot Autom*, pp. 4796–4802, 2020, doi: 10.1109/ICRA40945.2020.9196909.

[5] R. Calandra *et al.*, "More than a feeling: Learning to grasp and regrasp using vision and touch," *IEEE Robot Autom Lett*, vol. 3, no. 4, pp. 3300–3307, 2018, doi: 10.1109/LRA.2018.2852779.

[6] J. Li, S. Dong, and E. Adelson, "Slip Detection with Combined Tactile and Visual Information," *Proc IEEE Int Conf Robot Autom*, pp. 7772–7777, 2018, doi: 10.1109/ICRA.2018.8460495.

[7] Y. She, S. Wang, S. Dong, N. Sunil, A. Rodriguez, and E. Adelson, "Cable manipulation with a tactile-reactive gripper," *International Journal of Robotics Research*, vol. 40, no. 12–14, pp. 1385–1401, 2021, doi: 10.1177/02783649211027233.

[8] S. Wang *et al.*, "3D Shape Perception from Monocular Vision, Touch, and Shape Priors," *IEEE International Conference on Intelligent Robots and Systems*, pp. 1606–1613, 2018, doi: 10.1109/IROS.2018.8593430.

[9] W. Yuan, Y. Mo, S. Wang, and E. Adelson, "Active clothing material perception using tactile sensing and deep learning," *Proc IEEE Int Conf Robot Autom*, vol. 1, pp. 4842–4849, 2018, doi: 10.1109/ICRA.2018.8461164.

[10] S. Wang, Y. She, B. Romero, and E. Adelson, "GelSight Wedge: Measuring High-Resolution 3D Contact Geometry with a Compact Robot Finger," *Proc IEEE Int Conf Robot Autom*, vol. 2021-May, no. Icra, pp. 6468–6475, 2021, doi: 10.1109/ICRA48506.2021.9560783.

[11] O. Jia and J. Yu, "Towards Improving Teletaction in Teleoperation Tasks Using Vision-Based Tactile Sensors," 2024.

[12] M. Lippi, M. C. Welle, M. K. Wozniak, A. Gasparri, and D. Kragic, "Low-Cost Teleoperation with Haptic Feedback through Vision-based Tactile Sensors for Rigid and Soft Object Manipulation," Mar. 2024, doi: 10.1109/RO-MAN60168.2024.10731383.

[13] Y. Zhu *et al.*, "Visual Tactile Sensor Based Force Estimation for Position-Force Teleoperation," in *2022 IEEE International Conference on Cyborg and Bionic Systems, CBS 2022*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 49–52. doi: 10.1109/CBS55922.2023.10115342.

[14] J. R. Lewis, "IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use," 1993.

[15] J. R. Lewis, "Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies," *Int J Hum Comput Interact*, vol. 14, no. 3–4, pp. 463–488, Sep. 2002, doi: 10.1080/10447318.2002.9669130.