Integrating Visuo-tactile Sensing with Haptic Feedback for Teleoperated Robot Manipulation

Noah Becker[†1], Erik Gattung[†1], Kay Hansel[1], Tim Schneider[1], Yaonan Zhu[5], Yasuhisa Hasegawa[5], and Jan Peters[1−4]

***Abstract*—Telerobotics enables humans to overcome spatial constraints and allows them to physically interact with the environment in remote locations. However, the sensory feedback provided by the system to the operator is often purely visual, limiting the operator's dexterity in manipulation tasks. In this work, we address this issue by equipping the robot's end-effector with high-resolution visuotactile GelSight sensors. Using low-cost MANUS-Gloves, we provide the operator with haptic feedback about forces acting at the points of contact in the form of vibration signals. We propose two different methods for estimating these forces; one based on estimating the movement of markers on the sensor surface and one deep-learning approach. Additionally, we integrate our system into a virtual-reality teleoperation pipeline in which a human operator controls both arms of a Tiago robot while receiving visual and haptic feedback. We believe that integrating haptic feedback is a crucial step for dexterous manipulation in teleoperated robotic systems.**

***Index Terms*—Visuo-Tactile, Haptic-feedback, Teleoperation, Human-robot interaction**

I. INTRODUCTION

Teleoperation, the remote control of robots, enables humans to overcome spatial constraints and physically interact with the environment in remote locations [1, 2]. Recently, contributions such as the Nimbro system [3, 4], the ALOHA system [5], or the mobile ALOHA system [6] attracted a lot of attention in the field. A common challenge lies in providing feedback that is not easily conveyed through images, like temperature, surface structure, or current grip force. Haptic feedback has been looked to as a solution to this problem, providing additional information through touch instead of sight [7, 8]. This information is particularly relevant in our context, as we are aiming for an intuitive teleoperation system designed for people with limited technical knowledge.

In this work, we integrated vibrotactile feedback into a virtual reality (VR) teleoperation pipeline to improve performance in object manipulation tasks. This addition extends the visual feedback received from the VR system by providing tactile sensations. The user operates a dual-arm manipulation robot, benefiting from both visual and tactile cues to improve control and precision. Vibration feedback, also called vibrotactile feedback, is one of the most commonly used types of tactile feedback [9–12]; used due to its low profile, cost, and power consumption. To obtain haptic information about the objects to be manipulated, we attached visuotactile sensors to the robot's grippers. We use these sensors to detect the force exerted when interacting with objects. This information is crucial to ensure that objects are neither dropped nor crushed when manipulating them. The force data is then converted into a vibration intensity and transmitted to the user as vibrotactile feedback. In this way, the user receives tactile feedback in real time, providing a more intuitive sense of control when gripping or touching objects. While gripping force could also be measured with simple force sensors, visuotactile sensors can provide much richer information about the points of contact, such as object texture [13, 14], object shape [15], or the presence of slip [16]. In future work, we plan to feed back such information to the user in addition to force feedback.

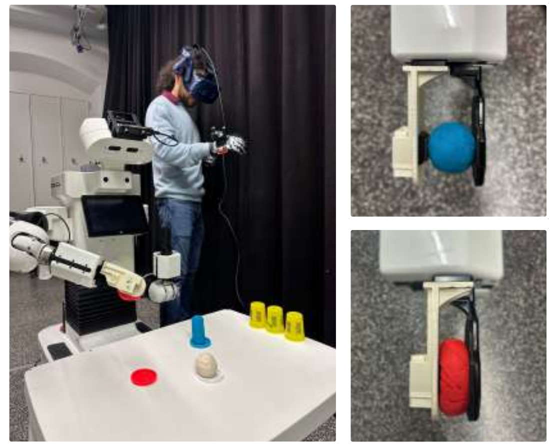


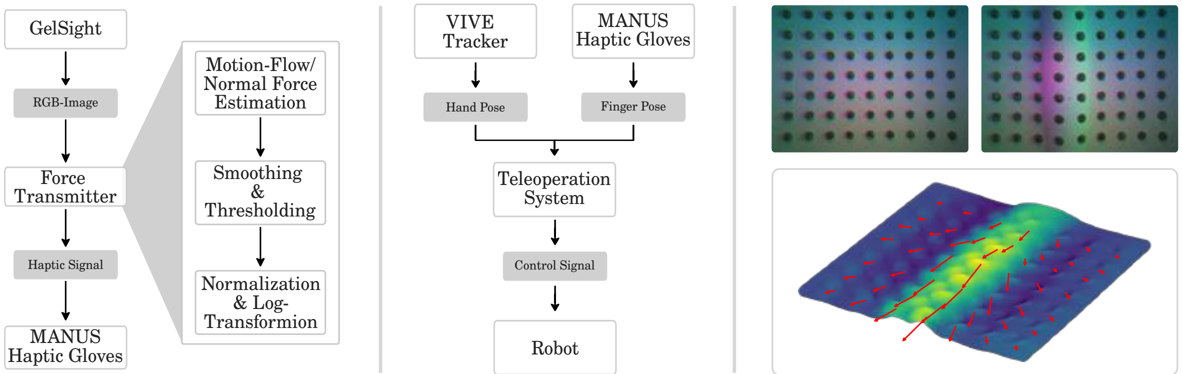
Fig. 1: The Teleoperation setup. **Left -** Tiago Robot equipped with tactile GelSight sensor. **Right -** Comparison of soft and hard grasps of modeling compound.

In the following, we provide an overview of the developed teleoperation system and detail methods for visuo-tactile force estimation. We present initial evaluations validating the proof of concept, followed by an outlook on future research directions and potential applications.

II. SYSTEM AND METHOD

We integrate haptic feedback into a VR teleoperation pipeline, as shown in Figure 2. The user wears a VR headset

This research received funding from CHIRON by DFG (PE 2315/8-1), the European Union's Horizon Europe programme under Grant Agreement No. 101135959 (project ARISE), BMWSB ZukunftBau (no. 10.08.18.7-21.34), and Aristotle by BMBF; [†]Equally contributing authors; [1]Computer Science Department, Technische Universität Darmstadt, Germany, [2]German Research Center for AI (DFKI), [3]Hessian Centre for Artificial Intelligence (Hessian.AI), [4]Centre for Cognitive Science (CogSci). {noah.becker, erik.gattung} @stud.tu-darmstadt.de, {kay.hansel, tim.schneider1, jan.peters} @tu-darmstadt.de, [5]Department of Micro-Nano Mechanical Science and Engineering, Nagoya University, Japan. zhu@robo.mein.nagoya-u.ac.jp, hasegawa@mein.nagoya-u.ac.jp;



(a) (b) (c)

Fig. 2: Overview of different components of the proposed system. **(a) -** The haptic feedback pipeline. **(b) -** The teleoperation pipeline. **(c) top -** The dot-matrix gel with/without an external force from a cylindrical object. **(c) bottom -** The corresponding optical motion flow visualised in 3D overlaid by upscaled vectors of the motion flow.

from [17] providing a first-person view through a stereo camera [18] mounted on the robot's head. This setup enables intuitive control of a dual-arm manipulation robot such as [19], akin to systems described in works [20, 21]. For haptic feedback, we use two Prime X Haptic gloves from MANUS [22], equipped with flex sensors and vibration motors for each finger. The flex sensors detect the user's finger pose, a gesture that the robot's grippers mimic. Additionally, a VR tracker from [17] on the glove tracks the user's hand pose and enables the robot to replicate arm motions.

The haptic feedback pipeline uses a GelSight Mini [23] with a dot-matrix-gel mounted to the robot's gripper to sense contact with an object. The GelSight streams the video of its gel to a transmitter node, which converts the images to a haptic signal. This haptic signal gets forwarded to the gloves as vibration feedback. We propose two different methods for generating the haptic signal: either the Lucas-Kanade optical motion flow method [24] or a neural network [25] trained to estimate the force acting on the gel.

For the first method, we track dots on the gel by analyzing small windows around each dot across consecutive frames. We then calculate spatial gradients of image intensity within these windows and relates them to intensity changes over time, solving for each dot's motion vector. This results in the estimated motion of each dot as a 2D vector. We sum the magnitude of all motion flow vectors to estimate the total force across the whole sensor. Previous research has shown that motion flow is an adequate approximation of force [23, 26].

The neural network we used was proposed by [25], using a similar process as described in [27]. It takes the image of a GelSight sensor with a dot matrix and outputs shear force and normal force as a single vector. The magnitude of this vector is used as the total force.

We apply a threshold to the total force calculated by either algorithm to prevent the glove from vibrating due to noise or arm movements. In addition, the values are normalized, and a log scale is applied so that the haptic feedback range for smaller forces is larger, assisting with fine control. These values are sent to the Manus glove, which controls the haptic motors to vibrate proportional to the total force. The maximum vibration is reached when the robot presses its fingers together with the maximum possible force. Communication between system components is done using the Robot Operating System (ROS) [28] and shared memory to minimize latency.

III. EVALUATION

To evaluate our system's strengths and weaknesses, we conducted a small preliminary user study with 7 participants. In this study, participants were tasked with using our teleoperation system to pick up a plasticine ball as gently as possible without dropping it. We conducted trials both with and without haptic feedback and using balls of various sizes. The extent of the ball's deformation served as a quantitative metric for the users' ability to execute tasks requiring precise force control. Preliminary results show that, on average, ball deformation was reduced by 48%

For the qualitative analysis, we use the NASA Task Load Index (TLX) [29], evaluating the subjective workload of the task described above. Preliminary results of the test trial show that users perceive the haptic feedback as enhancing performance, although the increased feedback while performing fine control using haptic feedback also requires more effort to handle. Further studies, both on quantitative and qualitative measurements, are still ongoing.

IV. CONCLUSION

In this work, we developed a teleoperation system in which users get haptic force feedback through vibration. Our preliminary user study indicates that this haptic feedback improves dexterity and perceived performance.

In future work, we plan to utilize the rich data of visuotactile sensors to provide additional haptic feedback to the user, such as shear forces, slip, and texture, further reducing the reliance on visual cues. Another exciting avenue for further research is to utilize these sensors in shared control, e.g., by automatically adjusting gripping strength based on tactile feedback.

References

[1] T. B. Sheridan, *Telerobotics, automation, and human supervisory control*. MIT press, 1992.

[2] ——, "Teleoperation, telerobotics and telepresence: A progress report," *Control Engineering Practice*, vol. 3, no. 2, pp. 205–214, 1995.

[3] M. Schwarz, C. Lenz, A. Rochow, M. Schreiber, and S. Behnke, "Nimbro avatar: Interactive immersive telepresence with force-feedback telemanipulation," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 5312–5319.

[4] C. Lenz, M. Schwarz, A. Rochow, B. Pätzold, R. Memmesheimer, M. Schreiber, and S. Behnke, "Nimbro wins ana avatar xprize immersive telepresence competition: Human-centric evaluation and lessons learned," *International Journal of Social Robotics*, pp. 1–25, 2023.

[5] T. Z. Zhao, V. Kumar, S. Levine, and C. Finn, "Learning fine-grained bimanual manipulation with low-cost hardware," *arXiv preprint arXiv:2304.13705*, 2023.

[6] Z. Fu, T. Z. Zhao, and C. Finn, "Mobile aloha: Learning bimanual mobile manipulation with low-cost whole-body teleoperation," *arXiv preprint arXiv:2401.02117*, 2024.

[7] R. Muradore, P. Fiorini *et al.*, "A review of bilateral teleoperation algorithms," *Acta Polytechnica Hungarica*, vol. 13, no. 1, pp. 191–208, 2016.

[8] G. Niemeyer, C. Preusche, S. Stramigioli, and D. Lee, "Telerobotics," *Springer handbook of robotics*, pp. 1085–1108, 2016.

[9] J. Martínez, A. García, M. Oliver, J. P. Molina, and P. González, "Identifying virtual 3d geometric shapes with a vibrotactile glove," *IEEE computer graphics and applications*, vol. 36, no. 1, pp. 42–51, 2014.

[10] L. Fritsche, F. Unverzag, J. Peters, and R. Calandra, "First-person tele-operation of a humanoid robot," in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2015, pp. 997–1002.

[11] P. Weber, E. Rueckert, R. Calandra, J. Peters, and P. Beckerle, "A low-cost sensor glove with vibrotactile feedback and multiple finger joint and hand motion sensing for human-robot interaction," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2016, pp. 99–104.

[12] S. Tachi, "Telexistence: Enabling humans to be virtually ubiquitous," *IEEE Computer Graphics and Applications*, vol. 36, no. 1, pp. 8–14, 2016.

[13] A. Böhm, T. Schneider, B. Belousov, A. Kshirsagar, L. Lin, K. Doerschner, K. Drewing, C. A. Rothkopf, and J. Peters, "What matters for active texture recognition with vision-based tactile sensors," in *International conference on robotics and automation*, 2024.

[14] W. Yuan, Y. Mo, S. Wang, and E. H. Adelson, "Active Clothing Material Perception Using Tactile Sensing and Deep Learning," in *International conference on robotics and automation*, 2018.

[15] S. Suresh, Z. Si, J. G. Mangelson, W. Yuan, and M. Kaess, "Shapemap 3-d: Efficient shape mapping through dense touch and vision," in *International conference on robotics and automation*, 2022.

[16] N. Funk, E. Helmut, G. Chalvatzaki, R. Calandra, and J. Peters, "Evetac: An event-based optical tactile sensor for robotic manipulation," 2023.

[17] H. Corporation, "Vive deutschland," 2024, accessed Mar. 20, 2024. [Online]. Available: https://www.vive.com/de/

[18] Stereolabs, "Zed mini stereo camera," 2024, accessed Mar. 20, 2024. [Online]. Available: https://store.stereolabs.com/en-de/products/zed-mini

[19] P. Robotics, "Tiago++, the robot you need for bi-manual tasks," 2019, accessed Mar. 14, 2024. [Online]. Available: https://blog.pal-robotics.com/tiago-bi-manual-robot-research/

[20] 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)*. IEEE, 2023, pp. 49–52.

[21] Y. Zhu, B. Jiang, Q. Chen, T. Aoyama, and Y. Hasegawa, "A shared control framework for enhanced grasping performance in teleoperation," *IEEE Access*, 2023.

[22] M. Meta, "Prime x haptic vr," 2023, accessed Mar. 15, 2024. [Online]. Available: https://www.manus-meta.com/products/prime-x-haptic

[23] W. Yuan, S. Dong, and E. H. Adelson, "Gelsight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors*, vol. 17, no. 12, p. 2762, 2017.

[24] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *IJCAI'81: 7th international joint conference on Artificial intelligence*, vol. 2, 1981, pp. 674–679.

[25] E. Helmut, L. Dziarski, N. Funk, B. Belousov, and J. Peters, "Visuotactile shear force estimation," 2023, unpublished manuscript.

[26] W. Yuan, R. Li, M. A. Srinivasan, and E. H. Adelson, "Measurement of shear and slip with a gelsight tactile sensor," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 304–311.

[27] k. Funk, P. O. Müller, B. Belousov, A. Savchenko, R. Findeisen, and J. Peters, "High-resolution pixelwise contact area and normal force estimation for the gelsight mini visuotactile sensor using neural networks," in *Embracing Contacts-Workshop at ICRA 2023*, 2023.

[28] Stanford Artificial Intelligence Laboratory et al., "Robotic operating system." [Online]. Available: https://www.ros.org

[29] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Advances in psychology*, vol. 52. North Holland, 1988, pp. 139–183.