

A Pneumatic Haptic Glove for Pose-Based Human-AI Communication

MINWOO LEE, Kyung Hee University, Republic of Korea

SUNGJOON YOON, Kyung Hee University, Republic of Korea

SEONGMIN YUN, Kyung Hee University, Republic of Korea

YUNSEO DO, Kyung Hee University, Republic of Korea

SEUNGJAE OH, Kyung Hee University, Republic of Korea

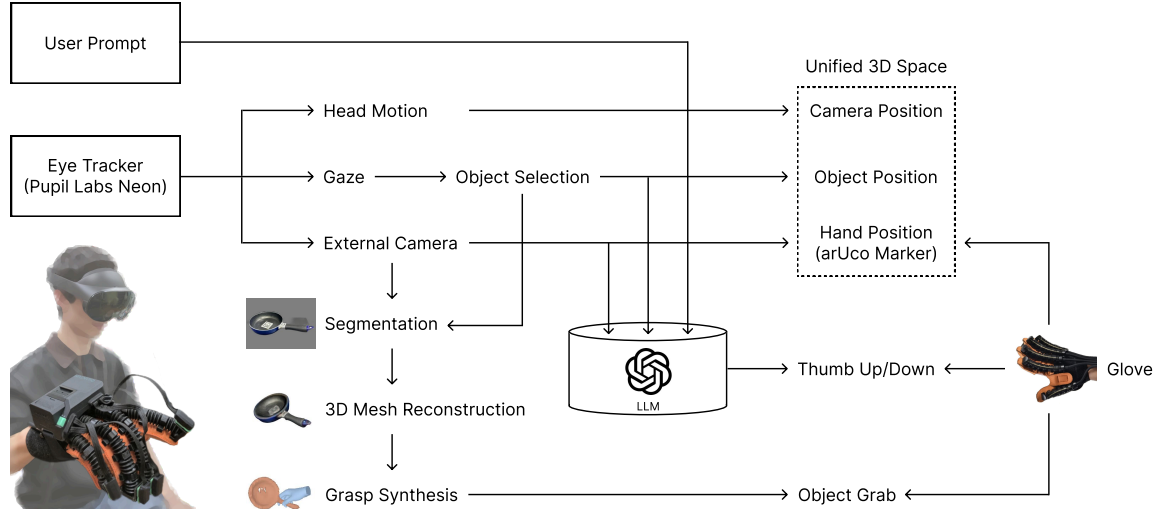


Fig. 1. Overview of pose-based Human-AI communication

We propose hand pose as a novel output modality, treating the user's hand as a dynamic medium through which AI agents can express their responses. To realize this, we introduce a pneumatic haptic glove that delivers AI responses by constraining hand poses. This I/O interface enables reciprocal gestural communication: users convey intent with gestures, and AI agents "replies" by guiding, holding, or inhibiting poses.

Additional Key Words and Phrases: Pneumatic Glove, Haptic Interaction, Pose-based Human-AI Communication

Reference Format:

Minwoo Lee, Sungjoon Yoon, Seongmin Yun, Yunseo Do, and Seungjae Oh. 2026. A Pneumatic Haptic Glove for Pose-Based Human-AI Communication. Proceedings of the *AI for Haptics and Haptics for AI: Challenges and Opportunities* Workshop at the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26), April 13–17, 2026, Barcelona, Spain. 4 pages.

1 Introduction

In this study, we propose hand gestures as a novel modality for enabling direct, bodily, context-rich communication. Our approach is grounded in the common sense that hand gestures are a natural and expressive medium for conveying one's

Authors' Contact Information: Minwoo Lee, minnov13@khu.ac.kr, Kyung Hee University, Yongin-si, Gyunggi-do, Republic of Korea; Sungjoon Yoon, sungjoon2022@khu.ac.kr, Kyung Hee University, Yongin-si, Gyunggi-do, Republic of Korea; Seongmin Yun, dbstjdals961@khu.ac.kr, Kyung Hee University, Yongin-si, Gyunggi-do, Republic of Korea; Yunseo Do, ysdoh0209@khu.ac.kr, Kyung Hee University, Yongin-si, Gyunggi-do, Republic of Korea; Seungjae Oh, oreo329@khu.ac.kr, Kyung Hee University, Yongin-si, Gyunggi-do, Republic of Korea.

intent and sentiment. Gesture recognition is widely adopted in HCI studies and practices as a means of communication, e.g., using a thumbs-up gesture to indicate a positive response. Conversely, movement assistance systems can be used to induce hand movements in the context of teaching and learning manual tasks [4], such as helping users coordinate timing and finger placement while learning to play the piano. Through these bi-directional gestural communications, AI agents can express their subtle intentions via a user’s hand and understand the user’s intentions by observing their hand poses. A key technical challenge addressed in our system is how to assist—or occasionally inhibit—fine-grained human hand movements in a fast and reliable manner.

2 Problem Framing and Position Statement

Our overarching goal is to enable seamless, reciprocal gestural communication between users and AI agents. For this purpose, we demonstrate an AI software framework supporting fine-grained object-hand interaction, where AI agents can assist or even constrain hand poses during manual tasks. By integrating real-time hand tracking and offline pose estimation, our system allows for dynamic, context-aware physical assistance that adapts to user context.

3 Pose-based Human-AI Communication

In this paper, we define bodily assistance as any form of involuntary physical actuation that either intervenes in or supports the user’s bodily movements. This includes mechanisms that cue or prompt the user’s actions to facilitate or enhance physical interactions, as well as mechanisms that constrain movements to modify actions.

Our research aims to propose a novel form of physical intervention in communication with AI agents. With the advent of VLMs, the possibility of combining assistance and guidance with involuntary interaction has emerged, as suggested by [2, 7]. In line with these theoretical foundations, we design and develop a gestural communication system in which AI agents can assist hand poses (e.g., grasp, release, hold, and refrain) based on inferred user intent or context.

Our system integrates robust real-time hand/eye tracking, AI-driven hand-object interaction, and haptic actuation system, forming a complete and functional pipeline for AI-human gestural communication. Through this implementation, we demonstrate how physical intervention can support expressive, intuitive, and agent-driven interaction beyond conventional input-output modalities.

4 System and Future Directions

4.0.1 Pneumatic Actuator & Glove. For a wearable glove form factor, we modify a commercial pneumatic glove for rehabilitation (ML-115A), equipped with five bellow-type pneumatic actuators and operated by our pneumatic control systems. Moreover, we integrate a hand motion capture device (Quantum Metaglove, Manus) to the glove by installing the mount structure of its sensing components for each digit on the pad side of the glove. We use a 3d-printed arUco marker onto the glove, enabling global hand position tracking useful for interaction design.

4.1 AI Software Framework

Starting from pose and eye tracking, we create an end-to-end pipeline that generates output poses for the pneumatic glove according to user context. Our AI framework has two major functionalities: scene understanding and interaction.

4.1.1 Scene Understanding. The scene understanding module utilizes images captured by the eye-tracking device (Neon, Pupil Labs) and gaze points. The overall pipeline is activated by a predefined command, such as a text prompt input or a wake-up word like "Hey, Siri." Upon activation, a target object is specified as the object nearest to the gaze

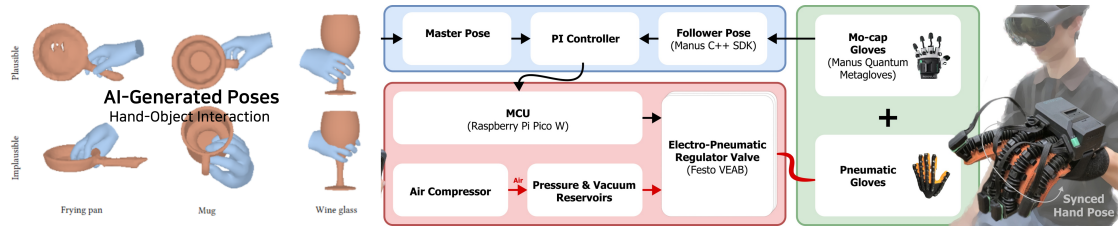


Fig. 2. This illustrates the workflow of our system.

point at the moment of activation, and it is segmented using SAM [3]. This segmented scene image can then be used as visual input for AI agents (or LLMs) to support contextual queries with ChatGPT API (see our supplementary materials).

4.1.2 Interaction. Our interaction module selectively executes either the gestural response submodule or the hand-object interaction submodule.

First, the gestural response submodule is responsible for answering to user queries with affirmative or negative gestures and their corresponding intensity. Depending on the polarity of the agent’s response, the appropriate gesture is commanded; for example, a thumbs-up gesture for a positive response and a fist gesture for a negative response. Here, users may choose to comply with or reject the given constraint, whose intensity is naturally conveyed as haptic feedback (see Figure 2). Their decision can be explicitly confirmed by, for instance, pressing a key, and the final pose is recorded by the mocap system.

Second, the hand-object interaction submodule is designed to output a suitable grasp pose when users attempt to manipulate an object. This component first reconstructs a 3D mesh from the segmented 2D object image using TripoSR [5]. The generated 3D mesh is fed into an offline hand-object interaction model G-HOP [6] to infer potential hand poses for interacting with the segmented object. We select a set of hand-object configurations and then encode these configurations with ArUco markers [1] and their positions. A grasp pose is selected by checking the proximity between the markers attached to the real object and one on the hand (see Figure 3). The selected pose is safely reproduced by the glove according to additional decision rules, including priority, conflict resolution, and timeout constraints.



Fig. 3. Scenarios of Pose-based Human-AI Communication. Depending on motion trajectory toward the frying pan, different poses are assigned to assist or inhibit subsequent actions.

4.2 Scenario: Hand-Object Interaction

We prepared two objects (pan and wine glass) for demonstration. We attached three fiducial markers to the pan: one on the handle, one on the base, and one representing hot oily foods. They repeated with or without the marker representing hot oily foods in the pan. For example, all fingers are extended to signal caution when they approaches a pan with a hot food marker. For the wine glass, a marker was place on the top to assist a grasping pose suitable for holding the stem. We used pre-recorded hand poses to ensure stable and safe operation.

5 Conclusion

This work advances a new interaction paradigm—to our knowledge—that lets AI agents convey semantically grounded, context-aware information through physical hand-pose intervention. This enables bi-directional AI-human gestural communication that goes beyond text or speech. This modality can influence the sense of agency; however, it opens up unique interaction capabilities with intuitive and satisfactory experiences, differently across intervention scenarios.

Acknowledgments

References

- [1] [n. d.]. OpenCV: Detection of ArUco Markers. https://docs.opencv.org/4.x/d5/dae/tutorial_aruco_detection.html
- [2] Daekyum Kim, Brian Byunghyun Kang, Kyu Bum Kim, Hyungmin Choi, Jeessoo Ha, Kyu Jin Cho, and Sungho Jo. 2019. Eyes are faster than hands: A soft wearable robot learns user intention from the egocentric view. *Science Robotics* 4, 26 (2019), 4–6. doi:10.1126/scirobotics.aav2949
- [3] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan Yen Lo, Piotr Dollár, and Ross Girshick. 2023. Segment Anything. *Proceedings of the IEEE International Conference on Computer Vision* (2023), 3992–4003. arXiv:2304.02643 doi:10.1109/ICCV51070.2023.00371
- [4] Jun Nishida, Yudai Tanaka, Romain Nith, and Pedro Lopes. 2022. Digitsync: A Dual-User Passive Exoskeleton Glove That Adaptively Shares Hand Gestures. In *Proceedings of ACM Symposium on User Interface Software and Technology*, Vol. 1. Association for Computing Machinery, doi:10.1145/3526113.3545630
- [5] Dmitry Tochilkin, David Pankratz, Zexiang Liu, Zixuan Huang, Adam Letts, Yangguang Li, Ding Liang, Christian Laforte, Varun Jampani, and Yan-Pei Cao. 2024. TripoSR: Fast 3D Object Reconstruction from a Single Image. (2024), 33–35. arXiv:2403.02151 <http://arxiv.org/abs/2403.02151>
- [6] Yufei Ye, Abhinav Gupta, Kris Kitani, and Shubham Tulsiani. 2024. G-HOP: Generative Hand-Object Prior for Interaction Reconstruction and Grasp Synthesis. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2024), 1911–1920. arXiv:2404.12383 doi:10.1109/CVPR52733.2024.00187
- [7] Pedro Lopes Yun Ho, Romain Nith, Peili Jiang, Shan-Yuan Teng. 2025. Generative Muscle Stimulation : Physical Assistance by Constraining Multimodal-AI with Biomechanical Knowledge. In *arXiv*.

The Author

Minwoo Lee received the B.S. degree in Mechanical Engineering and the Master degree in Software Convergence from Kyung Hee University, South Korea, in 2024 and 2026, respectively. Sungjoon Yoon, Seongmin Yun, and Yunseo Do are undergraduate interns of Interactive TEchnology and Methodology (ITEM) laboratory at Kyung Hee University. Seungjae Oh received the B.S. degree in Electrical Engineering and the Ph.D. degree in Creative IT Engineering from Pohang University of Science and Technology (POSTECH), Pohang, South Korea, in 2012 and 2020, respectively. From 2020 to 2022, he was a Postdoctoral Researcher in the Department of Computer Science and Engineering at POSTECH. He is currently an Assistant Professor in the Department of Software Convergence at Kyung Hee University. He leads the Interactive TEchnology and Methodology (ITEM) laboratory at Kyung Hee University. His research interests include designing novel input and output devices for haptic applications.