AI Speaks with Hands: A Bidirectionally Actuated Pneumatic 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

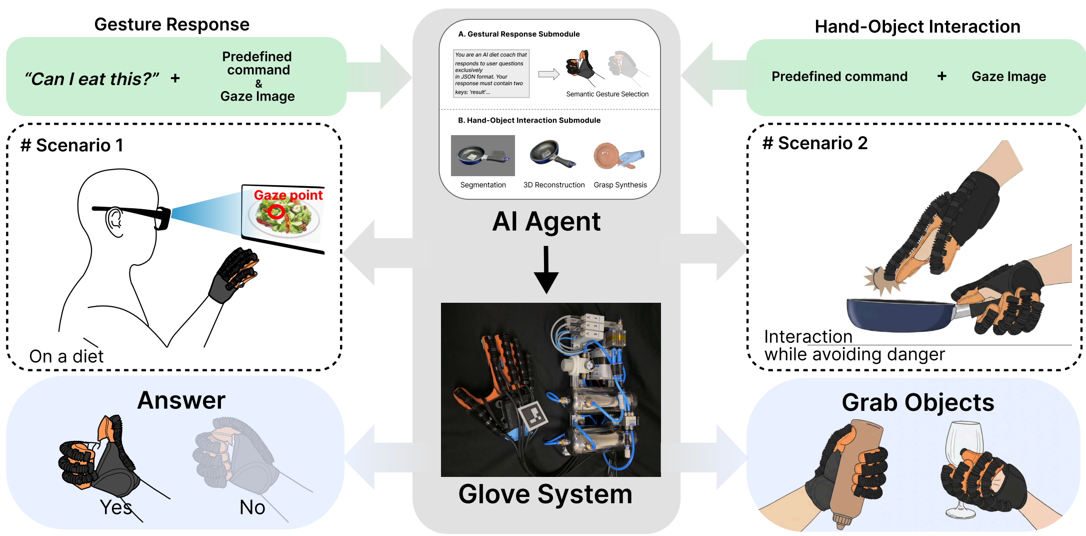

Fig. 1. Overview of AI Speaks with Hands. This illustrates interaction scenarios in our study and their technical implementation and workflow.

AI agents have transformed how people interact with computers, yet most still "speak" only through text or speech, missing an intuitive, embodied communication channel. 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 glove that delivers AI responses by constraining hand poses. The glove supports bi-directional actuation using both positive and negative pressure to enable responsive

Authors' addresses: 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.

Manuscript submitted to ACM

flexion and extension. A hand mocap device robustly recognizes poses even when the glove is worn. This I/O interface enables reciprocal gestural communication: users convey intent with gestures, and AI agents "replies" by guiding, holding, or inhibiting poses. The technical breakthrough of this study is a high-performance hand-pose constraint system, and we demonstrate its application to AI-human gestural communication. A user study examines the user experience with two AI-human communication scenarios.

CCS Concepts: • **Human-centered computing → Interaction paradigms**; *Gestural input*; *Haptic devices.*

Additional Key Words and Phrases: Hand Pose, AI Agent, Soft Pneumatic Actuation, Assistance

ACM Reference Format:

Minwoo Lee, Sungjoon Yoon, Seongmin Yun, Yunseo Do, and Seungjae Oh. 2018. AI Speaks with Hands: A Bidirectionally Actuated Pneumatic Glove for Pose-Based Human-AI Communication. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX).* ACM, New York, NY, USA, 17 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Recent advances in large language models (LLMs) have dramatically changed the way in which humans interact and communicate with computers. While early LLM-based interfaces were limited to text input and output, the development of vision-language models (VLMs) has expanded input modalities to include visual image data, enabling deeper contextual understanding. Nevertheless, even the most recent AI agents still primarily communicate via text or speech, which limits the intuitiveness and expressiveness of interaction. These constraints raise critical research questions in scenarios that demand nuanced, embodied, and context-aware communication.

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 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 [28], 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.

In this work, we propose a novel interaction paradigm with AI with a technically validated pneumatic pose constraining system. Our system employs a closed-loop control architecture with a dedicated pneumatic circuit that actively inflates and deflates pneumatic actuators for responsive control. We technically characterize the pneumatic actuation mechanism and evaluate the responsiveness of the system. 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 two key functionalities: (1) gesture-based communication with vision-enabled AI agents, allowing users to express intent and receive the agents' responses via hand poses; and (2) 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.

To validate the user experience of our approach, we conducted an empirical user study focusing on two interaction scenarios where AI agents respond and intervene through the pneumatic gestural communication system. Our results demonstrate that AI-mediated pose assistance can enhance the subjective experience of novelty and intuitiveness. In summary, this study makes the following contributions: (1) We propose a novel AI interaction paradigm that leverages hand pose recognition and assistance as a bi-directional communication channel between a user and an AI agent. (2) We validate the technical feasibility of a pneumatic hand pose constraining hardware platform with a software framework that supports AI-human gestural communication with vision-enabled AI agents. (3) We empirically evaluate the system, providing insights into user experience with AI-mediated physical intervention. By bridging the gap between intelligent agents and embodied interaction, our research opens new possibilities for intuitive, expressive, and collaborative human-AI interfaces.

2 RELATED WORK

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. We review recent technical advances in wearable actuation mechanisms for providing physical bodily assistance and their usages. We aim to discuss the implications of such assistance by examining its applications.

2.1 Pose Assistance Mechanism

Early research on inducing human bodily actions mostly focus on Electrical Muscle Stimulation (EMS). Tamaki et al. [33] introduced EMS to control finger movements with 28 stimulation points on the forearm. With an individual calibration feature, the system demonstrated direct control of 16 joints in the hand and wrist, but only three of them could be independently controlled. Lopes et al. [23] further explored the boundaries between voluntary user actions and involuntary system intervention. The study validated the technical feasibility of fine wrist flexion and extension control with a PID angle controller. These studies both reported that their subjects clearly perceived the loss of body ownership, yet responded positively to physical intervention, e.g., involuntary wrist and hand actuation. Takahashi et al. [32] achieved independent finger actuation by reflecting biomechanical knowledge in EMS actuation. DextrEMS [30] addressed fundamental limitations in motion accuracy. Mechanical braking mechanisms were combined with EMS to avoid inherent issues such as unwanted movements and oscillations. Nevertheless, EMS approaches inherently suffer from unavoidable electrotactile sensation or require per-user calibration to achieve pain-free stimulation [30, 32, 33]. These issues limit their practical usability for sustained use and their adoption in consumer-level devices.

Alternatively, mechanical structures can be used to confine hand and upper-body movements. Nishida et al. [27] designed a link-based passive exoskeleton to simulate the experience of having a smaller hand. Similarly, DigituSync [28] extended link-based passive exoskeletons to synchronize the hands of different users. The exoskeleton is worn by two different users, and the master user's movements are transmitted simultaneously to the other user. The variable-link mechanism of the exoskeleton makes it possible to control assistive forces, while still requiring human actuation. To achieve computer-mediated bodily assistance, active exoskeletons have been explored. In general, the actuation mechanisms of active exoskeletons include electrostatic clutches [12, 36], pneumatic actuators [7, 13, 20, 40], motorized mechanisms [1, 5, 10, 11, 14, 15]. Although these active wearable mechanisms can display large constraining forces, many systems rely on either pulling or pushing actuation directions, or sometimes support limited actuation bandwidths. We aim to improve the performance of pneumatic actuation for assisting hand movements in natural hand use.

2.2 Pneumatic Actuator and Glove

Soft pneumatic gloves have attracted significant attention due to their mechanical simplicity, intrinsic safety, and user comfort [31, 40]. However, they have not been regarded as suitable for fine and responsive tasks, as these advantages come at the cost of reduced motion precision and actuation speed due to the inherent mechanical compliance of pneumatic bellows [6]. To circumvent this, researchers have explored the use of multiple pneumatic actuators assigned to movements in different directions to increase the degree of freedom in motion and enhance control bandwidth for motion responsiveness [9, 19, 20, 38]. Introducing negative air pressure allows flexion and extension with a single actuator; however, only a few studies address this system due to technical difficulty in pneumatic system design and control [6, 7, 13]. Moreover, many studies still focus on pneumatic actuation mechanisms alone [6, 7], rather than holistic devices or systems.

Among the studies focusing on glove-type pneumatic devices, open-loop control can suffer from low motion precision and resolution [13, 20]. Some studies have introduced closed-loop control using sensors such as pressure sensors [31], QR codes [19], force sensors [38], and bending sensors [9, 38]; however, these presented low-to-moderate motion bandwidth (0.1-4 Hz) due to low sensor performance. Chen et al. [9] achieved the fastest motion bandwidth (up to 4 Hz), but it requires specialized optical-waveguide sensor for bending detection. In this study, we design a pneumatic circuit with positive and negative pressure reservoirs, which enables pose control responsiveness—especially for finger extesions by negative pressure actuation—and reliable sustained system operation with a closed loop control.

2.3 Involuntary Physical Feedback

Involuntary physical feedback focuses on system interventions that occur independently of user intention. The most common case is passive haptic feedback in immersive virtual environments. This mostly happens by environmental changes, such as virtual objects flying toward a user or wind blowing toward the user, delivered through wearable tactile devices (e.g., haptic exoskeleton [25], multimodal haptic controllers [16, 18], and full-body haptic suit [2, 3, 16]). Another underexplored area is system-initiated notification. There has been barely no noticeable improvements after the proliferation of smartphones. We revisit this type of physical feedback in light of recent technical advances in AI.

Affordance++ [24] sought to link affordances with involuntary interaction. This pioneering work envisioned the possibility of involuntary action assistance through subjective evaluations on compelling interaction scenarios. Although such involuntary system interventions can affect the body ownership and agency [29], recent studies have examined cases involving mixed-agency devices. For bi-manual tasks, SplitBody [29] automated one hand's actions by employing EMS stimulation, and they observed lower cognitive workload with improvements in the task performance. In Preemptive Action [26], the authors conducted an in-depth analysis of user experience with respect to the timing of system intervention. They found that delayed EMS interventions did not significantly degrade the sense of agency, which remained at a moderate level. Recent studies have started to focus on broader area of the body, such as full body [41] and the arm [37].

Building on these insights, our research aims to propose a novel form of physical intervention in communication with AI agents. With the advent of AI and LLMs, the possibility of combining assistance and guidance with involuntary interaction has emerged, as suggested by [15, 41]. In line with these theoretical foundations, we design, develop, and evaluate a gestural communication system in which AI agents can assist hand poses (e.g., grasp, release, hold, and refrain) and express attitudinal responses through gestures (e.g., a thumbs-up gesture). AI agents are enabled to deliver answers or responses by constraining hand poses based on inferred user intent or context. Our system integrates robust real-time hand tracking, LLM-driven intent inference, AI-driven hand-object interaction, and pneumatic 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.

3 SYSTEM: AI SPEAKS WITH HANDS

3.1 Pneumatic Hardware & Software

3.1.1 Pneumatic Actuator & Glove. For a wearable glove form factor, we modify a commercial pneumatic glove for rehabilitation (ML-115A). The glove is equipped with five bellow-type pneumatic actuators, one for each finger, providing 1 DoF flexion and extension. 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. This modification enables fast and robust acquisition of hand skeleton positions (>120 Hz), which is not feasible with typical vision-based approaches that could not result robust estimation due to the the distinct appearance of the glove and occlusion. We use a 3d-printed arUco marker onto the glove, enabling global hand position tracking useful for interaction design. Furthermore, we use a mobile computing system paired with the mocap device (Quantum Bodypack, Manus) to stream hand tracking data to the control system.

3.1.2 Electro-pneumatic Control & Regulation. The pneumatic glove is operated by pneumatic actuation and control systems. We use electro-pneumatic regulator valves (VEAB Proportional Regulator, Festo) that control target pressure values in the range of +100 to −100 kPa, directly supplying air to or removing air from the pneumatic bellows. Our early prototypes, which relied solely on conventional positive-pressure-only regulators, suffered from slow finger extension because the bellows could only vent passively. Even recent pneumatic gloves use more than one actuator to achieve bi-directional motion if only positive pressure actuation is available [9, 19, 20, 38]. In summary, incorporating negative pneumatic actuation enables rapid and responsive extension with a single bellow actuator (see Figure 4).

3.1.3 Pneumatic Circuit Design. We design a pneumatic circuit with a dual-reservoir system for stable system responses. For its compact pneumatic source, we employ a BLDC-driven rocking-piston compressor (30RNS-ED) capable of supplying both positive and negative pressures (+650 to −65 kPa). This single pressure source is electronically controlled to supply either the positive or negative pressure reservoir, which is dynamically switched by a 5-way/2-mode solenoid valve (KS105S, SMC). The positive pressure reservoir can store a sufficient amount of compressed air (over +600 kPa), which is then decompressed via a manual pressure regulator to match the maximum operating pressure requirement of the regulator (+200 kPa). Conversely for the negative pressure source, an amount of vacuum can't be charged in advance as much as the positive side does by the fact that vacuum is the state of pressure. The reservoir only functions as a buffer volume that smoothing pressure fluctuation due to sudden inflow of pressurized air exhaust when the pressure command is switched from positive to negative. With the same reason, air compressor must continuously depresses the influx of air from the negative pressure source to maintain vacuum pressure level. Therefore, aforementioned solenoid valve normally connects the air compressor to the negative pressure side and sparsely switches to the positive pressure reservoir when sensed pressure level drops below a pre-defined threshold, +200 kPa, which is the minimum level required by regulator valve. From this strategy, the maximum negative pressure (−65 kPa) is maintained in the reservoir for most time. Moreover, pneumatic check valves, functioning like diodes in electric circuits, are installed between the reservoirs and the pneumatic source to minimize air flow leakage while not charging. Overall, the system is compact enough for tabletop use but supports reliable operations for sustained uses due to the dual reservoir circuit design.

3.1.4 Micro Controller Unit. The MCU (Pico W, Raspberry Pi) controls the operation of all electro-pneumatic parts and coordinates communication between the pneumatic system and its closed-control loop controller. It receives digital set points from the control loop controller and generates logic-level signals, which are then amplified by a custom DAC-amp circuit to the analog voltage levels (0-10 V) required by the regulator valves.

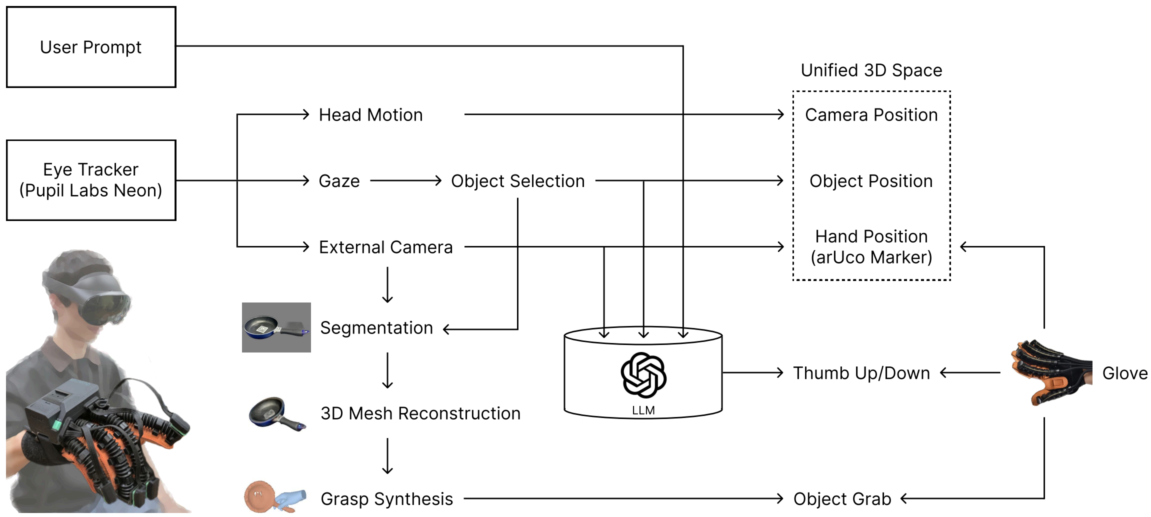

Fig. 2. Block diagram of our system workflow.

3.2 Control Software Framework

3.2.1 Middleware. Our middleware aggregates, processes, and transmits hand pose information, working between high-level applications and low-level embedded systems. We design it to support hand skeleton data from different APIs or sources. The middleware generates pressure commands with a closed-loop controller and facilitates communication with the microprocessor.

3.2.2 Hand Pose Tracking & Analysis. The pose tracking module received a 15-DOF finger joint position set (i.e., three joints per finger) from the mocap device via its proprietary C++ SDK. Joint composition and indexes are remain consistent, compatible across any combination of other hand tracking systems providing joint positions, opening rooms for future integration with computer vision-based systems like Google Mediapipe, or virtual reality frameworks such as OpenXR, and Meta Quest SDK. Since our glove is capable to impose only lumped flexion/extension for each finger, result of synergetic motion consist with multiple sub-finger joints. To simplifying the representation, values from the sum of flexion angles for all joints, the metric known as composite finger flexion(CFF), is used. This representation is interchangeably used in the system to compute either target poses from interaction context (any hand motion that are synthesized, recorded, synced with external hand tracking module being imposed to user) or measured poses (feedback data from the mocap device) to compute the 5 DoF per-finger pose errors.

3.2.3 2-Stage Cascaded Closed-Loop Control. Dynamics of pneumatic, soft robotic systems require a control scheme with robust feedback and error compensation for handling its nonlinearity. We also have the human user wearing the glove as the core of the system, timely and responsive operation is critical for ensuring pleasant interaction experience. With those requirement in mind, from given target finger flexion angle θ_{tar}, the pressure command p_{cmd} is derived from the error between θ_{tar} and acquired feedback θ using standard PID control law:

$$p_{\text{cmd}} = f^{-1}(\theta_{\text{tar}}) + K_p(\theta_{\text{tar}} - \theta) + K_i \int (\theta_{\text{tar}} - \theta)dt + K_d(\dot{\theta}_{\text{tar}} - \dot{\theta})$$

K_p, K_i and K_d are gains of the feedback controller while $f^{-1}(\theta)$ is the feedforward term which is calculated from the inverse of system response $f : p \rightarrow \theta$ which will be elaborated after. The output value of the positional closed-loop, p_{cmd} is feed into another closed loop PD control based law [8] :

$$u = \phi^{-1}(K_d \dot{p}_{cmd} + p_{cmd} + K_p(p - p_{\text{cmd}}))$$

In here, ϕ^{-1} is system-specific inverse relationship compensating of piezoelectric-introduced nonlinearity of pneumatic valve.

3.2.4 Performance Evaluation. Upon system characterization process that identifying aforementioned $f : p \rightarrow \theta$ from the sample pairs of pressure and flexion angle, we observed severe hysteresis depending on the direction of operation: actuator pressurization(finger flexion) and suction(extension). While feedback-only control suffers under-performance due to excessive residues from unexpected behavior, introducing feed-forward term with careful system identification can counter this. We model f to be a hyperbolic tangent function that fitted from the middlepoint of hysteresis, and with feeding this into control reduces residual burden that feedback control is handling.

Overall, with the optimal feedback gains of PID control determined empirically, the system exhibits an end-to-end finger flexion command to actual motion latency of 140 ms as from the Figure 4 measured from the initiation of target pose update to the point at which the tracking error stabilized within ±5%.

3.3 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.

3.3.1 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 point at the moment of activation, and it is segmented using SAM [17]. 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).

3.3.2 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 5). 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 [35]. The generated 3D mesh is fed into a offline hand-object interaction model G-HOP [39] 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 [4] 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 6). The selected pose is safely reproduced by the glove according to additional decision rules, including priority, conflict resolution, and timeout constraints.

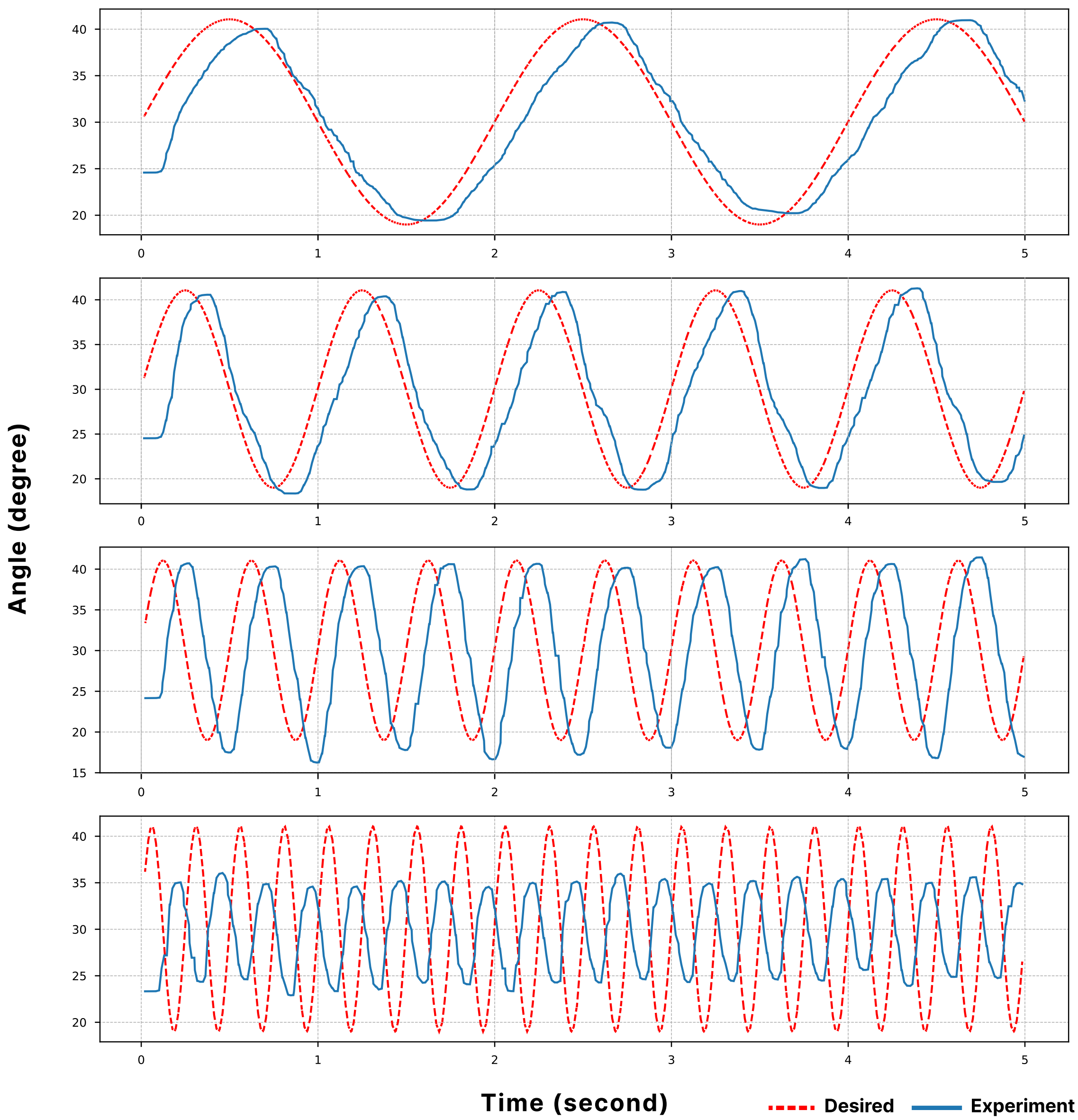

Fig. 3. Plots of index finger control responses for 0.5, 1, 2, and 4 Hz.

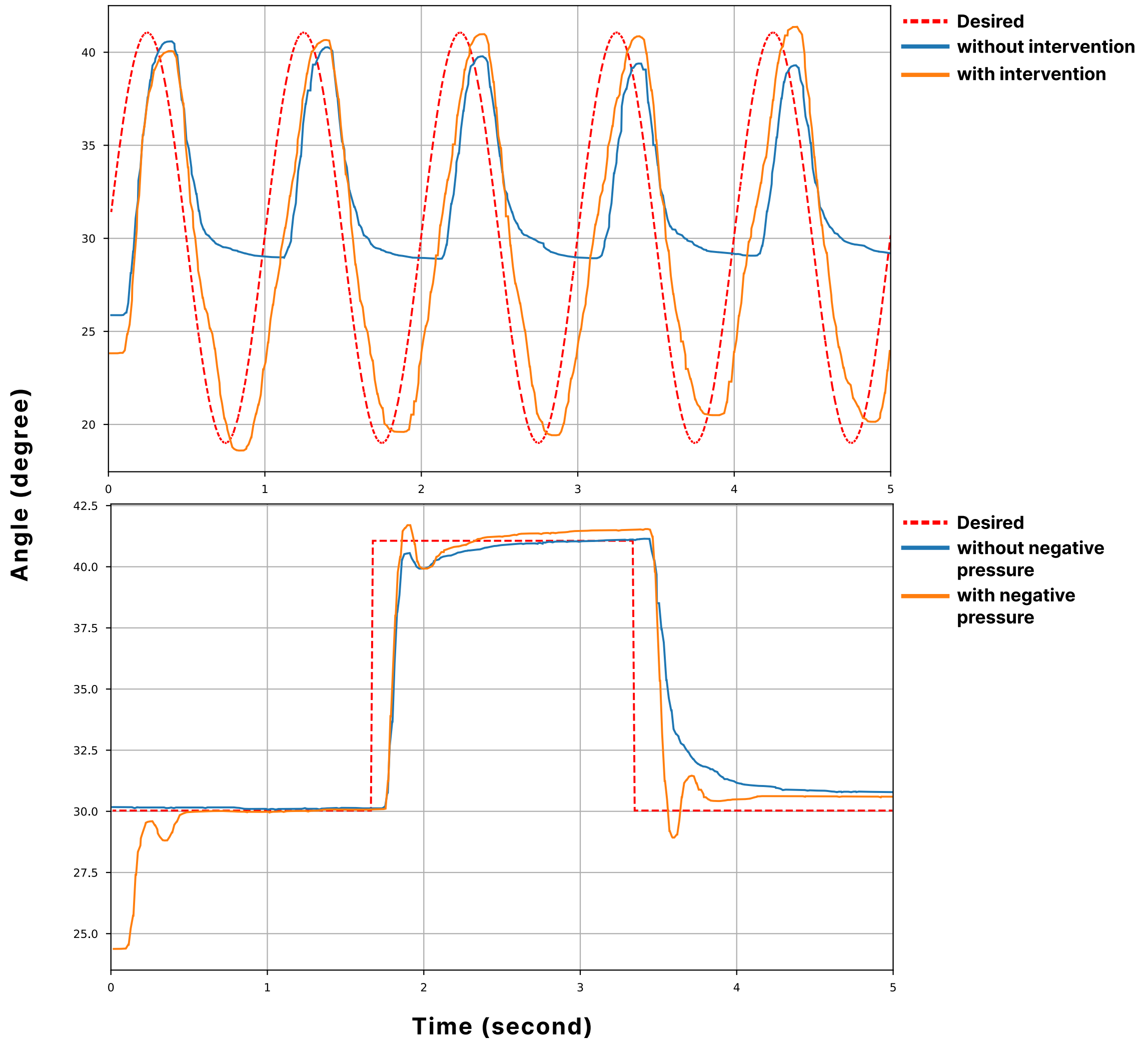

Fig. 4. Plots of index finger control responses, with and without the utilization of negative pressure on pneumatic control.

4 USER STUDY

In this study, we explore the user experience of our AI-mediated physical intervention and assistance in two potential use scenarios. The first scenario involves AI-Human gestural communication in a delivery application. The agent and the user communicate through our sensing and constraining glove, expressing positive and negative responses to each other via the user's gloved hand. The second scenario demonstrates agent intervention during ADL. When users interact with everyday objects, the agent intervenes to assist with grasping objects or inhibit touching certain objects. All features in the experimental scenarios were fully implemented and functional, but we deactivated the feedforward term to ensure stable and simplified operation. This user study was approved by the Institutional Review Board at the authors' institution.

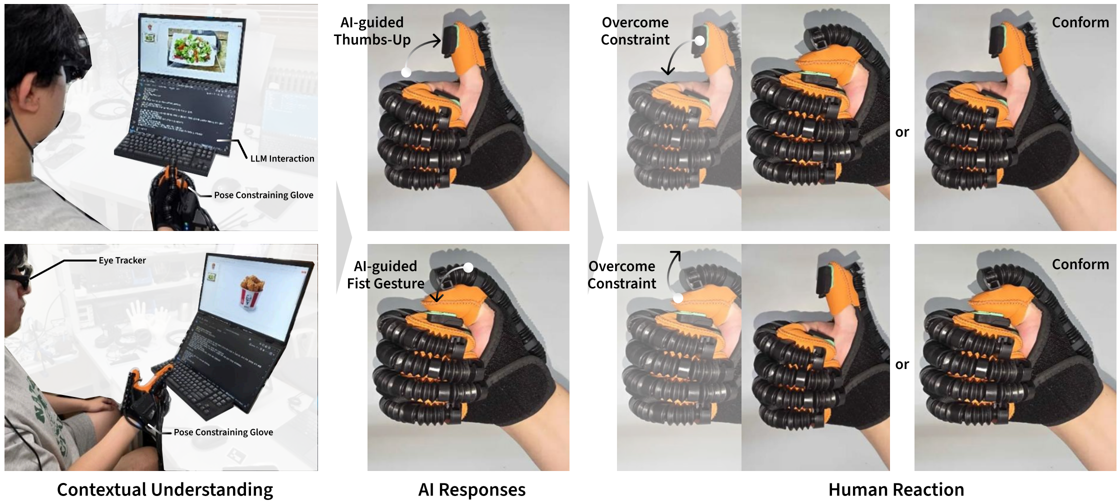

Fig. 5. Overview of Experiment 1. Poses transmitted from AI agents are illrustated. Then, there are basically two options for users, either overcome or conform to the contraint. The gradient-arrows indicate start (white) and end (black) of a motion.

Fig. 6. Overview of Experiment 2. Depending on motion trajectory toward the frying pan, different poses are assigned to assist or inhibit subsequent actions.

4.1 Experimental Design & Condition

For both scenarios, the independent variable was *Gesture Intervention*. The first scenario consisted of two levels that differed in the presence or absence of gesture responses (text-only and text+pose conditions). The second scenario also consisted of two conditions with and without physical interventions (no-intervention and with-intervention).

We designed a set of user experience questionnaire for each scenario, modifying questions from previous studies [21, 28, 37]. Participants completed three common questions: agency [21, 28], intuitiveness, and satisfaction. We additionally asked about novelty in the first scenario and matched expectation [34, 37] in the second scenario. All questions were rated on a 7-point Likert scale ranging from -3 (strongly disagree) to 3 (strongly agree).

The experiment used a one-factor within-subjects design for each scenario. We counterbalanced the presentation order in each scenario with a balanced Latin square. The entire process of completing the two interaction scenarios took approximately 50 minutes.

4.2 Participants

We recruited 16 right-handed participants (12 male, 4 female; age M = 23.1 years and SD = 1.8 years) from the authors' institution. Each participant was paid USD 10. Eleven out of sixteen participants reported using LLMs everyday, and the remaining five reported at least three to four times per week. Before the experiment, participants were informed of the procedure via a written document and signed a consent form.

4.3 Scenario 1: Positive-Negative Gesture Communication

4.3.1 Procedure. Participants performed a daily task of selecting diet-appropriate foods with the assistance of an AI agent (Figure 5). Participants wore the eye tracker and the pose glove. At the start of each trial, they were shown a photo of either a bucket of fried chicken or a salad, selected at random. For convenience, we pre-typed the text message, "I'm on a diet. Should I eat the food in this photo?" Participants pressed the enter key to confirm their food selection and submit the text query. The AI agent (GPT-4o) responded in real time via the ChatGPT response API, following our JSON query (in our supplementary materials). The AI agent responded with both textual answers and a gesture via the haptic glove. Following the agent's response, participants could also indicate their agreement or disagreement using either a hand gesture or text, depending on the condition. Specifically, participants could decide to comply with or resist the pose constraint imposed by the glove. They then repeated the trial with the other food and were asked to comply if they had previously resisted, and vice versa. After completing two trials, they answered the questionnaire and took a one-minute break. All participants completed both the text-only and text-and-pose conditions. A semi-structured interview was conducted after the completion of the first scenario twice.

4.3.2 Results and Discussion. We performed Wilcoxon signed rank tests with a significance level of 0.05. As expected, we observed a significant effect of physical intervention on agency ($W = 3.00, P = .022$). The median agency scores with and without gestural responses were 6 and 7, respectively. The results showed that the presence of intervention significantly increased perceived novelty ($W = 66.00, P = .003$), while no significant differences were found in intuitiveness ($W = 6.50, P = .461$) or satisfaction ($W = 20.50, P = .138$).

Participants reported that the system slightly limited their sense of free will but considered this AI-human communication modality to be novel (Figure 7). We analyzed interview responses to understand the reasons behind the lack of significance in intuitiveness and satisfaction. A few participants mentioned a preference for text outputs due to simplicity and clarity. Some participants noted that the system would be much more informative if it could perform more than two gestures. There were also comments that the gesture-constraining force was a bit stronger than expected and that they had to wear too many devices.

4.4 Scenario 2: Hand-Object Interaction

4.4.1 Procedure. Participants were equipped with the glove. We informed them that they would be interacting with everyday objects with the assistance of another AI agent (Figure 6). Participants were asked to reach their gloved hand toward a frying pan and a wine glass. 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. They experienced as much as they want and then answered the questionnaire. After finishing the trial, they had a one-minute break. All participants completed the trials with the two conditions of with and without intervention. We used pre-recorded hand poses in the user study to ensure stable and safe operation.

4.4.2 Results and Discussion. In line with the previous scenario, we observed a significant effect of physical intervention on agency ($W = 0.00, P = .002$), and the median agency scores with and without intervention were 3 and 7, respectively. The results showed that the presence of intervention significantly increased all other measures: matched expectation ($W = 98.00, P = .032$), intuitiveness ($W = 74.50, P = .006$), and satisfaction ($W = 74.50, P = .044$).

Interestingly, participants reported a greater loss of agency in the second scenario than the first, but they found the intervention to be intuitive and satisfactory (Figure 8). In the interview, all participants responded that the intervention was informative and beneficial, but some of them again reported that constraining forces in our experimental setting were too strong. Interestingly, some participants highlighted that it would be much more useful if these physical interventions took place based on the inference of safety contexts, e.g., for blind users, firefighters, or children.

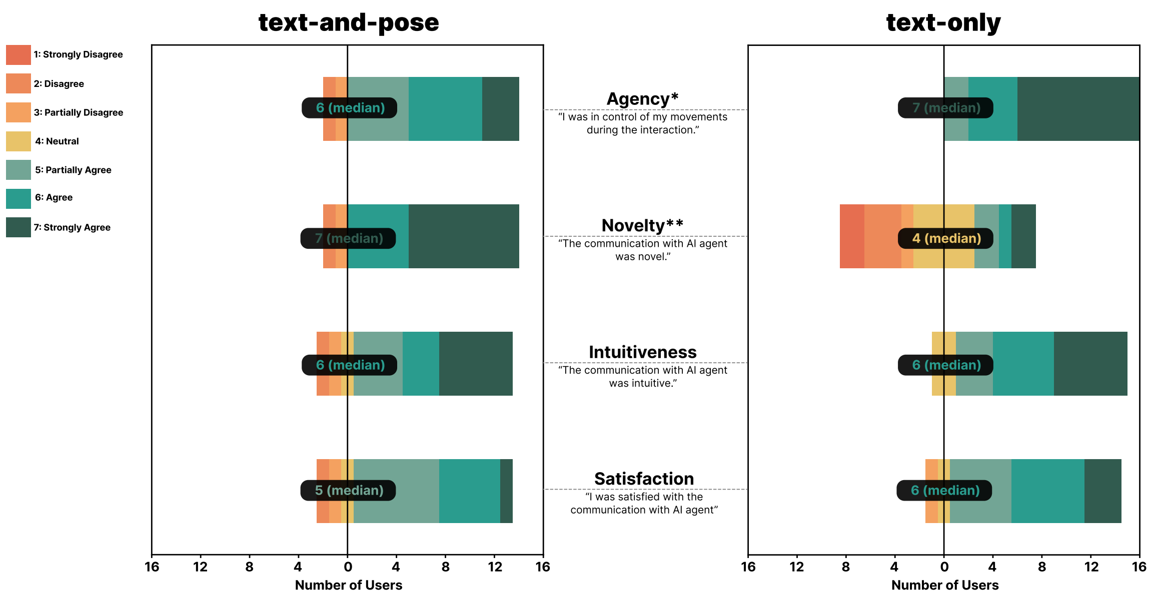

Fig. 7. Results of Experiment 1 with the distribution of participants' responses on the 7-point Likert scale for Agency, Novelty, Intuitiveness, and Satisfaction. Statistical significance is indicated by asterisks. (*: $p < .05$; **: $< .01$; ***: $p < .001$)

5 DISCUSSION

5.1 Designing AI-Human Gestural Interactions

Our findings highlight several important design considerations.

First, our results support the common understanding that gesture-based communication is essential, but it can be confusing when used alone. According to participants' subjective responses, they expect capabilities that are difficult to achieve with conventional methods, such as enhanced safety and agile responses. Similar to how gestures complement speech in human-human communication, AI gestures in our system are best received when used to augment—not replace—conventional text and voice outputs.

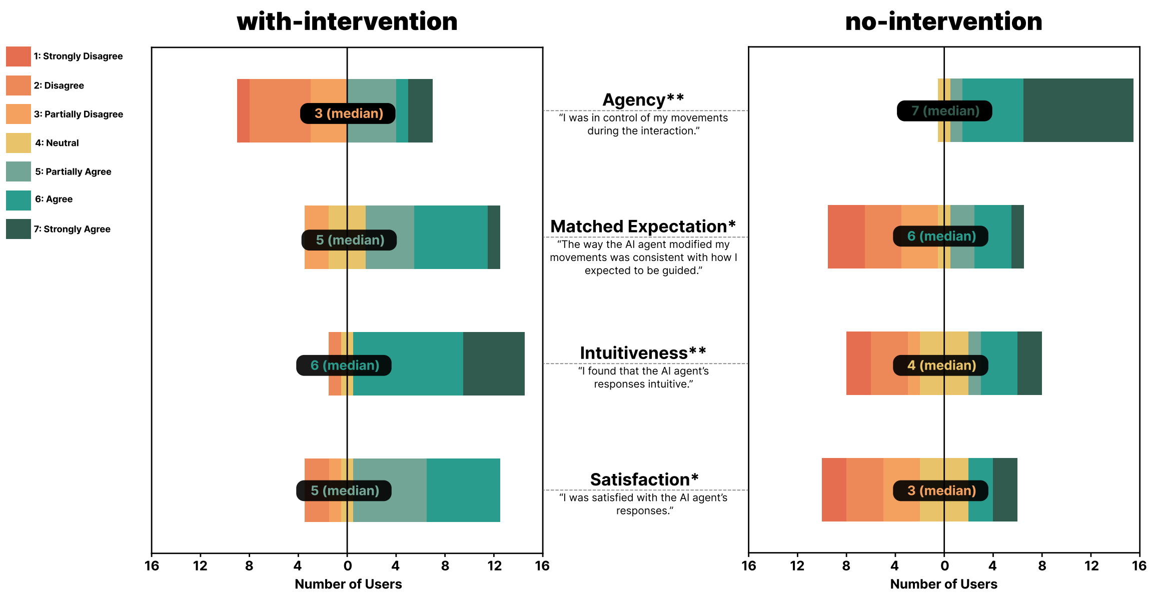

Fig. 8. Results of Experiment 2 with the distribution of participants' responses on the 7-point Likert scale for Agency, Matched Expectation, Intuitiveness, and Satisfaction. Statistical significance is indicated by asterisks.(*: $p < .05$; **: $< .01$; ***: $p < .001$)

Second, scenario-sensitive intervention emerged as a key requirement. While the system's physical feedback was perceived as even more forceful in the second scenario, our participants found such interventions satisfying. Moreover, they expressed a desire for more adaptive behaviors based on the inference of user intent and context. This points to the need for intelligent decision-making layers that determine *when* and *how* to intervene based on situational relevance.

Third, participants' qualitative feedback emphasized that the perceived strength of the constraint strongly influenced their experience. In future studies, we recommend to carefully calibrating the degree of intervention in conjunction with context and its importance.

5.2 Future Opportunities and Interaction Scenarios

Our system opens up a range of future applications by leveraging constrained hand poses as a communication modality. For example, in physical games like rock-paper-scissors, the glove can dynamically constrain gestures to compete with an AI agent (Figure 9 (a)). The glove can also assist transitions between static poses. In VR applications, snap-to-hand effects could be implemented, where a user's hand pose is automatically adapts to interacting objects (Figure 9 (b)). This approach could be further extended to incorporate material or functional properties of virtual objects; for example, magnetic objects could impose firmer constraints. Beyond static assistance, the glove system could support dynamic gesture sequences. An AI agent could signal index finger flexion as a deictic gesture or execute subtle movements such as finger tapping on a specific object to indicate selection (Figure 9 (c)). These temporally programmed gesture sequences expand the expressiveness of AI responses.

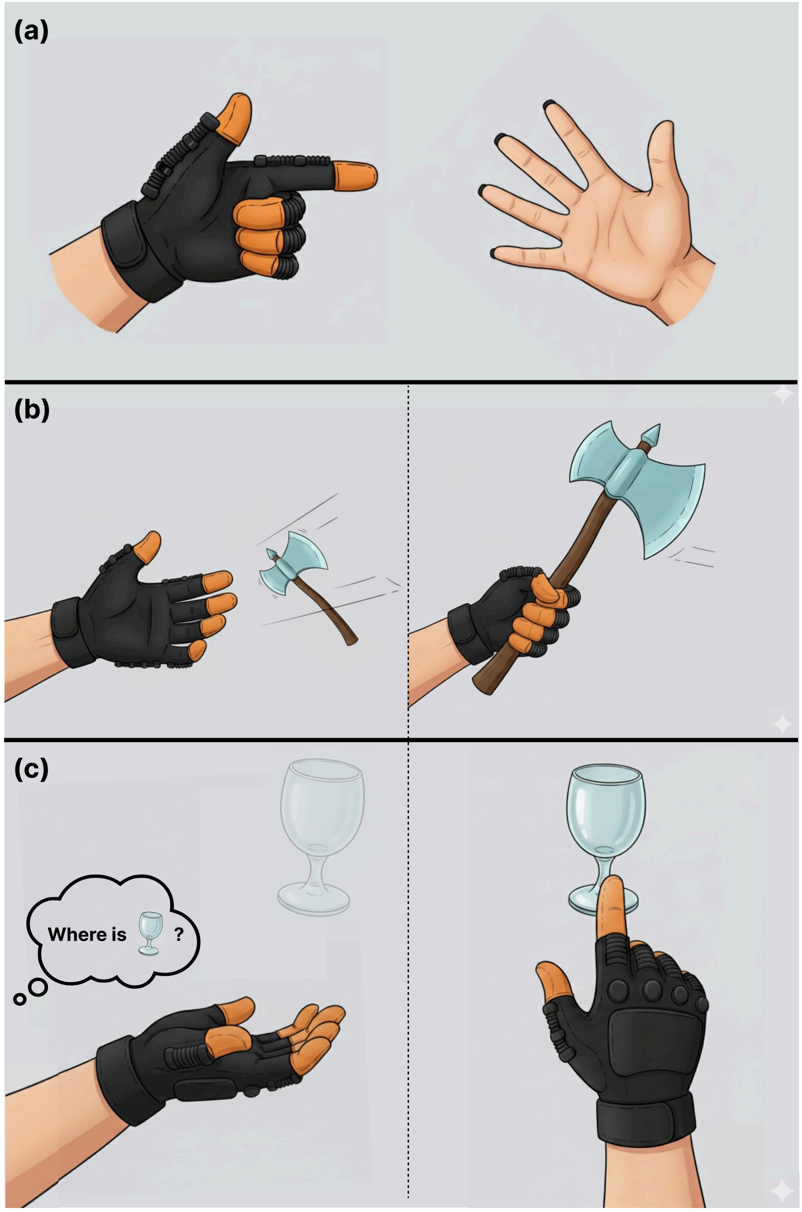

Fig. 9. Potential Scenarios in AI-human gestural communication.

5.3 Limitations and Future Work

Despite the promising results, several limitations remain. Although we implemented recent AI methods of hand-object interaction [22, 39], the estimated hand poses for unseen objects sometimes did not achieve satisfactory quality, and the methods proved infeasible for real-time inference. These issues can be addressed through future improvements in generalizable vision-to-grasp pipelines, which will be critical for enhancing user experience. From a hardware standpoint, the current bellow-type actuators are still bulky and mechanically rigid, which reduces user comfort during prolonged use. Although we demonstrated a desktop-scale, fully functional prototype, miniaturization and silent operation remain important challenges. In future iterations, replacing the bellows with softer actuators and adopting more compact compressed-air cartridges could further enhance both wearability and comfort during extended use.

6 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. Our user study finds that this modality influences the sense of agency; however, it opens up unique interaction capabilities with intuitive and satisfactory experiences, differently across intervention scenarios. Building on these insights, we discuss design guidelines, future opportunities, and limitations.

ACKNOWLEDGMENTS

University ethics review board approves human-subjects research and they approved this project.

REFERENCES

[1] [n. d.]. Dexta Robotics - Touch the Untouchable. https://www.dextarobotics.com/en-us
[2] [n. d.]. Full Body VR Haptic Suit with Motion Capture | TESLASUIT. https://teslasuit.io/products/teslasuit-4/
[3] [n. d.]. Next generation full body haptic suit - bHaptics TactSuit. https://www.bhaptics.com/en/
[4] [n. d.]. OpenCV: Detection of ArUco Markers. https://docs.opencv.org/4.x/d5/dae/tutorial{_}aruco{_}detection.html
[5] [n. d.]. The Future of VR: Interacting with SenseGlove Nova. https://www.senseglove.com/enhancing-virtual-reality-with-haptic-gloves-the-senseglove-nova/
[6] Brandon J. Caasenbrood, Alexander Y. Pogromsky, and Henk Nijmeijer. 2024. Sorotoki: A Matlab Toolkit for Design, Modeling, and Control of Soft Robots. *IEEE Access* 12, December 2023 (2024), 17604–17638. https://doi.org/10.1109/ACCESS.2024.3357351
[7] Brandon J. Caasenbrood, Femke E. Van Beek, Hoang Khanh Chu, and Irene A. Kuling. 2022. A Desktop-sized Platform for Real-time Control Applications of Pneumatic Soft Robots. In *IEEE International Conference on Soft Robotics*. IEEE, 217–223. https://doi.org/10.1109/RoboSoft54090.2022.9762137
[8] Brandon J. Caasenbrood, Femke E. van Beek, Hoang Khanh Chu, and Irene A. Kuling. 2022. A Desktop-sized Platform for Real-time Control Applications of Pneumatic Soft Robots. In *2022 IEEE 5th International Conference on Soft Robotics (RoboSoft)*. 217–223. https://doi.org/10.1109/RoboSoft54090.2022.9762137
[9] Wenbin Chen, Caihua Xiong, Chenlong Liu, Peimin Li, and Yonghua Chen. 2019. Fabrication and Dynamic Modeling of Bidirectional Bending Soft Actuator Integrated with Optical Waveguide Curvature Sensor. *Soft Robotics* 6, 4 (2019), 495–506. https://doi.org/10.1089/soro.2018.0061
[10] Cathy Fang and Matthew Dworman. 2020. Wireality : Enabling Complex Tangible Geometries in Virtual Reality with Worn Multi-String Haptics. In *Proceedings of ACM CHI Conference on Human Factors in Computing Systems*. 1–10.
[11] Xiaochi Gu, Yifei Zhang, Weize Sun, Yuanzhe Bian, Dao Zhou, and Per Ola Kristensson. 2016. Dexmo: An inexpensive and lightweight mechanical exoskeleton for motion capture and force feedback in VR. In *Conference on Human Factors in Computing Systems - Proceedings*. 1991–1995. https://doi.org/10.1145/2858036.2858487
[12] Ronan Hinchet, Velko Vechev, Herbert Shea, and Otmar Hilliges. 2018. Dextres: Wearable haptic feedback for grasping in VR via a thin form-factor electrostatic brake. *Proceedings of ACM Symposium on User Interface Software and Technology* (oct 2018), 901–912. https://doi.org/10.1145/3242587.3242657/SUPPL_FILE/P901-HINCHET.MP4
[13] Debin Hu, Jinhua Zhang, Yuhan Yang, Qiuyang Li, Dahai Li, and Jun Hong. 2020. A novel soft robotic glove with positive-negative pneumatic actuator for hand rehabilitation. In *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, Vol. 2020-July. 1840–1847. https://doi.org/10.1109/AIM43001.2020.9158826
[14] Brian Byunghyun Kang, Hyungmin Choi, Haemin Lee, and Kyu Jin Cho. 2019. Exo-Glove Poly II: A Polymer-Based Soft Wearable Robot for the Hand with a Tendon-Driven Actuation System. *Soft Robotics* 6, 2 (2019), 214–227. https://doi.org/10.1089/soro.2018.0006
[15] Daekyum Kim, Brian Byunghyun Kang, Kyu Bum Kim, Hyungmin Choi, Jeesoo 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. https://doi.org/10.1126/scirobotics.aav2949
[16] Jinsoo Kim, Seungjae Oh, Chaeyong Park, and Seungmoon Choi. 2020. Body-Penetrating Tactile Phantom Sensations. In *Proceedings of ACM CHI Conference on Human Factors in Computing Systems*. 1–13. https://doi.org/10.1145/3313831.3376619

[17] 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. https://doi.org/10.1109/ICCV51070.2023.00371 arXiv:2304.02643

[18] Robert Kovacs, Eyal Ofek, Mar Gonzalez Franco, Alexa Fay Siu, Sebastian Marwecki, Christian Holz, and Mike Sinclair. 2020. Haptic PIVOT: On-demand handhelds in VR. In *Proceedings of ACM Symposium on User Interface Software and Technology*. 1046–1059. https://doi.org/10.1145/3379337.3415854

[19] Jianwei Lai, Aiguo Song, Ke Shi, Qinjie Ji, Ye Lu, and Huijun Li. 2023. Design and Evaluation of a Bidirectional Soft Glove for Hand Rehabilitation-Assistance Tasks. *IEEE Transactions on Medical Robotics and Bionics* 5, 3 (2023), 730–740. https://doi.org/10.1109/TMRB.2023.3292414

[20] Jianwei Lai, Aiguo Song, Jiajin Wang, Ye Lu, Ting Wu, Huijun Li, Baoguo Xu, and Xiangshan Wei. 2023. A Novel Soft Glove Utilizing Honeycomb Pneumatic Actuators (HPAs) for Assisting Activities of Daily Living. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 31 (2023), 3223–3233. https://doi.org/10.1109/TNSRE.2023.3302612

[21] Kyungyeon Lee, Daniel S. Yang, Kriti Singh, and Jun Nishida. 2025. Hapticus: Exploring the Effects of Haptic Feedback and its Customization on Motor Skill Learning: Tactile, Haptic, and Somatosensory Approaches. *Proceedings of ACM CHI Conference on Human Factors in Computing Systems* (2025). https://doi.org/10.1145/3706598.3713821

[22] Shaowei Liu, Yang Zhou, Jimei Yang, Saurabh Gupta, and Shenlong Wang. 2023. ContactGen: Generative Contact Modeling for Grasp Generation. *Proceedings of the IEEE International Conference on Computer Vision* (2023), 20552–20563. https://doi.org/10.1109/ICCV51070.2023.01884 arXiv:2310.03740

[23] Pedro Lopes, Alexandra Ion, Willi Mueller, Daniel Hoffmann, Patrik Jonell, and Patrick Baudisch. 2015. Proprioceptive interaction. In *Proceedings of ACM CHI Conference on Human Factors in Computing Systems*, Vol. 2015-April. 939–948. https://doi.org/10.1145/2702123.2702461

[24] Pedro Lopes, Patrik Jonell, and Patrick Baudisch. 2015. Affordance++: Allowing objects to communicate dynamic use. In *Proceedings of ACM CHI Conference on Human Factors in Computing Systems*, Vol. 2015-April. 2515–2524. https://doi.org/10.1145/2702123.2702128

[25] Marie Muehlhaus, Alexander Liggesmeyer, and Jürgen Steimle. 2025. ExoKit: A Toolkit for Rapid Prototyping of Interactions for Arm-based Exoskeletons. *Conference on Human Factors in Computing Systems - Proceedings* (apr 2025). https://doi.org/10.1145/3706598.3713815/SUPPL_FILE/PN7402-TALK-VIDEO-CAPTION.VTT

[26] Jun Nishida, Shunichi Kasahara, and Pedro Lopes. 2019. Preemptive reaction: Accelerating human reaction using electrical muscle stimulation without compromising agency. In *Proceedings of ACM CHI Conference on Human Factors in Computing Systems*. 1–15. https://doi.org/10.1145/3305367.3327997

[27] Jun Nishida, Soichiro Matsuda, Hiroshi Matsui, Shan Yuan Teng, Ziwei Liu, Kenji Suzuki, and Pedro Lopes. 2020. HandMorph: A passive exoskeleton that miniaturizes grasp. In *Proceedings of ACM Symposium on User Interface Software and Technology*. 565–578. https://doi.org/10.1145/3379337.3415875

[28] Jun Nishida, Yudai Tanaka, Romain Nith, and Pedro Lopes. 2022. DigituSync: 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. https://doi.org/10.1145/3526113.3545630

[29] Romain Nith, Yun Ho, and Pedro Lopes. 2024. SplitBody: Reducing Mental Workload while Multitasking via Muscle Stimulation. *Conference on Human Factors in Computing Systems - Proceedings* (may 2024). https://doi.org/10.1145/3613904.3642629/SUPPL_FILE/PN6721-SUPPLEMENTAL-MATERIAL-1.ZIP

[30] Romain Nith, Shan Yuan Teng, Pengyu Li, Yujie Tao, and Pedro Lopes. 2021. DextrEMS: Increasing Dexterity in Electrical Muscle Stimulation by Combining it with Brakes. In *Proceedings of ACM Symposium on User Interface Software and Technology*. 414–430. https://doi.org/10.1145/3472749.3474759

[31] Panagiotis Polygerinos, Zheng Wang, Kevin C. Galloway, Robert J. Wood, and Conor J. Walsh. 2015. Soft robotic glove for combined assistance and at-home rehabilitation. *Robotics and Autonomous Systems* 73 (2015), 135–143. https://doi.org/10.1016/j.robot.2014.08.014

[32] Akifumi Takahashi, Jas Brooks, Hiroyuki Kajimoto, and Pedro Lopes. 2021. Increasing Electrical Muscle Stimulation's Dexterity by means of Back of the Hand Actuation. In *Proceedings of ACM CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3411764.3445761

[33] Emi Tamaki, Takashi Miyaki, and Jun Rekimoto. 2011. PossessedHand: Techniques for controlling human hands using electrical muscles stimuli. In *Proceedings of ACM CHI Conference on Human Factors in Computing Systems*. 543–552. https://doi.org/10.1145/1978942.1979018

[34] Adam Tapal, Ela Oren, Reuven Dar, and Baruch Eitam. 2017. The sense of agency scale: A measure of consciously perceived control over one's mind, body, and the immediate environment. *Frontiers in Psychology* 8, SEP (2017), 1–11. https://doi.org/10.3389/fpsyg.2017.01552

[35] 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

[36] Nicha Vanichvoranun, Hojeong Lee, Seoyeon Kim, and Sang Ho Yoon. 2024. EStatiG: Wearable Haptic Feedback with Multi-Phalanx Electrostatic Brake for Enhanced Object Perception in VR. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 3 (2024), 1–29. https://doi.org/10.1145/3678567

[37] David Wagmann and Marie Muehlhaus. 2025. Move with Style ! Enhancing Avatar Embodiment in Virtual Reality through Proprioceptive Motion Feedback. In *Proceedings of ACM Symposium on User Interface Software and Technology*. https://doi.org/10.1145/3746059.3747649

[38] Boran Wang, Andrew McDaid, Kean C. Aw, and Morteza Biglari-Abhari. 2017. Design and development of a skinny bidirectional soft glove for post-stroke hand rehabilitation. In *Intelligent Systems Conference*, Vol. 2018-Janua. 979–987. https://doi.org/10.1109/IntelliSys.2017.8324248

[39] 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. https://doi.org/10.1109/CVPR52733.2024.00187 arXiv:2404.12383

[40] Sung Sik Yun, Brian Byunghyun Kang, and Kyu Jin Cho. 2017. Exo-glove PM: An easily customizable modularized pneumatic assistive glove. *IEEE Robotics and Automation Letters* 2, 3 (2017), 1725–1732. https://doi.org/10.1109/LRA.2017.2678545

[41] 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*.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009