

Body Sublimation: Designing State-Changing Extended Hand for Embodied Out-of-Reach Interaction

Masahiro Ohta*
The University of Tokyo
Tokyo, Japan
Keio University
Yokohama, Kanagawa, Japan
ohta-masa@g.ecc.u-tokyo.ac.jp

Reigo Ban*
The University of Tokyo
Tokyo, Japan
reigo@cyber.t.u-tokyo.ac.jp

Hideaki Nii
Keio University
Yokohama, Kanagawa, Japan
nii@kmd.keio.ac.jp

Guillaume Gourmelen
Keio University
Yokohama, Kanagawa, Japan
guillaume@kmd.keio.ac.jp

Takuji Narumi
The University of Tokyo
Tokyo, Japan
narumi@cyber.t.u-tokyo.ac.jp

Mauricio Sousa
Keio University
Yokohama, Kanagawa, Japan
mauricio@keio.jp

Takatoshi Yoshida
Keio University
Yokohama, Kanagawa, Japan
yoshida@kmd.keio.ac.jp

Kouta Minamizawa
Keio University
Yokohama, Kanagawa, Japan
kouta@kmd.keio.ac.jp

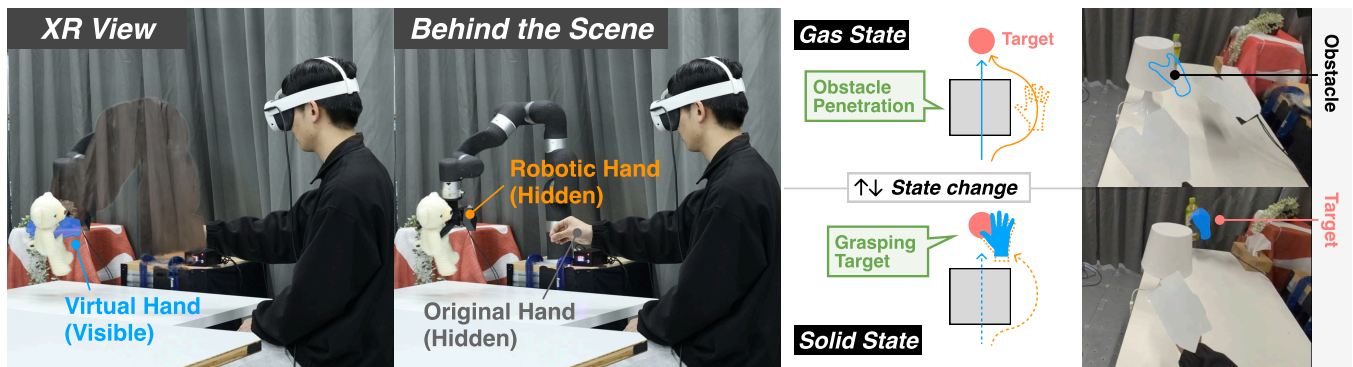


Figure 1: Overview of *Body Sublimation*. In the XR view (left), the user interacts with a virtual hand while both the user’s original hand and the robotic hand are visually removed. The behind-the-scenes view shows the hidden robotic manipulator that follows the virtual hand to perform physical interaction. The extended hand operates through two states (right): in the gas state, it can pass through obstacles to reach distant targets, while in the solid state, it becomes physically grounded to grasp and manipulate real-world objects.

Abstract

Recent research in extended reality (XR) and robotics has explored approaches for extending human reach beyond the inert physical body. XR enables embodied interaction with distant targets but lacks physical manipulation. In contrast, robot-mediated systems enable real-world interaction but often disrupt embodiment due to physical constraints such as latency, limited workspace, kinematic mismatch, and collision avoidance. Therefore, reconciling

physical manipulation with perceptual embodiment remains an open challenge. In this paper, we present *Body Sublimation*, an interaction concept that integrates the perceptual flexibility of XR with the physical capabilities of robotic systems. A virtual hand in XR serves as the user’s extended body while both the original hand and robotic manipulator are hidden using diminished reality. A hidden robotic hand then performs physical manipulation, offloading the physical constraints from the user’s perception, thereby maintaining embodiment continuity. We validate the *Body Sublimation* approach through a user study against a pass-through baseline, where the robot body is always visible, and application scenarios. Compared with a pass-through baseline in which the robot remains visible, our approach maintains task performance while improving trajectory efficiency in obstacle-reaching scenarios, and contributes to a stronger sense of embodiment.

*Both authors contributed equally to this work.



This work is licensed under a Creative Commons Attribution 4.0 International License.
UIST '26, Detroit, MI, USA

© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2856-3/2026/11
<https://doi.org/10.1145/3830398.3830583>

CCS Concepts

- **Human-centered computing** → **Mixed / augmented reality**;
- **Computer systems organization** → *Robotic control*.

Keywords

Human Augmentation, Embodiment, Diminished Reality

ACM Reference Format:

Masahiro Ohta, Reigo Ban, Hideaki Nii, Guillaume Gourmelen, Takuji Narumi, Mauricio Sousa, Takatoshi Yoshida, and Kouta Minamizawa. 2026. Body Sublimation: Designing State-Changing Extended Hand for Embodied Out-of-Reach Interaction. In *The 39th Annual ACM Symposium on User Interface Software and Technology (UIST '26)*, November 02–05, 2026, Detroit, MI, USA. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3830398.3830583>

1 Introduction

Human beings interact with their environment through their bodies, particularly through the hand, which enables reaching, grasping, and manipulating physical objects. However, this capability is inherently limited to the space surrounding the body. Extending human reach beyond the inert physical body to directly interact with distant objects has therefore attracted increasing attention [28, 42].

Recent research in XR and robotics has explored approaches toward this goal. XR enables interaction with distant targets by extending the user's body into virtual space, while preserving the sense of embodiment, which is critical for intuitive and low-effort interaction [6, 8, 54]. However, such virtual bodies cannot physically interact with real-world objects. In contrast, robot-mediated systems enable real-world interaction by physically manipulating objects. By directly mapping human motion onto robots, prior works enabled users to acquire additional arms [50, 51, 58, 60], operate at larger scales [11, 29], or interact with objects beyond natural reach [23, 47]. Yet, physical constraints such as latency, limited workspace, kinematic mismatch, and collision avoidance often introduce spatial and temporal discrepancies between user intention and system behavior, disrupting the sense of embodiment [7, 41, 52]. Therefore, reconciling physical manipulation with perceptual embodiment remains an open challenge.

In this paper, we propose a novel extended hand design that integrates the ability of XR to support embodiment through accurate body representation with the physical capabilities of robotic systems. In our approach, the user interacts with real-world objects through a virtual hand in XR, which provides high-fidelity and low-latency control, while a robotic hand follows the virtual hand to perform physical manipulation. To enhance the user's sense of embodiment toward the virtual hand, we employ diminished reality (DR) to visually remove both the user's original hand and the robotic hand, allowing the virtual hand to be perceived as the sole embodiment of the user's action. Through this mechanism, the user can interact with objects beyond their natural reach while maintaining user's sense of embodiment toward the extended hand. This extended hand can operate in two modes, depending on whether the hidden robotic hand matches the virtual hand. When the robotic hand matches the virtual hand, users can physically grasp and manipulate objects beyond their natural reach. When it does not, the

virtual hand behaves like a conventional XR hand and can move freely without physical constraints.

We refer to this design concept as *Body Sublimation*, inspired by material phase transitions, and in particular by *Sublimate* by Leithinger et al. [36]. In this concept, the extended hand dynamically transitions between two states: a *solid state*, in which the virtual hand is physically grounded by the hidden robotic hand and enables physical manipulation, and a *gas state*, in which the hand remains purely virtual and unconstrained by physical contact. We further characterize the interaction enabled by Body Sublimation through representative application scenarios. To evaluate the efficiency of our approach, we compared it with a pass-through baseline, in which both the robotic arm and the user's original hand remain visible. A user study across multiple tasks showed that, compared with a pass-through baseline, Body Sublimation enhances users' sense of embodiment while maintaining task performance.

The contributions of this paper are as follows:

- Exploration of transitions between robot-mediated physical embodiment and graphically rendered virtual embodiment, and interaction techniques leveraging such state changes.
- Practical implementations to prototype interactions using a hidden robotic hand and a virtual hand in XR.
- We evaluate our approach through a user study against a pass-through baseline, demonstrating maintained task performance and improvements in embodiment-related subjective measures, including ownership and agency.

2 Related Work

2.1 Expanding the Reach in XR

In the field of XR, techniques have been developed that expand bodily reach by using the hand as a controller and transforming spatial mappings. By amplifying hand movements [54] or remapping them [6, 8], these approaches enable intuitive manipulation of distant virtual objects. They preserve natural bodily motion while extending reach within virtual space. Furthermore, such interaction methods have been applied to directly control remotely located IoT devices through bodily movements [45, 69]. However, these approaches do not involve dynamic physical interaction with physical objects, and they do not allow users to physically grasp or move objects in the physical world. As an attempt to extend reach toward real-world objects, Feuchtner et al. proposed a method that uses an extended virtual hand in XR to manipulate physical objects [18]. However, their approach required instrumentation or mechanisms on the object side, resulting in strong environmental dependency.

2.2 Extending the Body with Robotics

Robotic systems have also been explored as a means of extending human capabilities. By mapping human motion to robotic manipulators, these systems enable users to physically interact with distant objects, operate at different scales, and augment their bodies with robotic embodiments. Such approaches have been widely studied in the context of robot teleoperation [12, 19, 23, 32, 38, 55]. When deployed in the same physical space as the user, these systems have also been explored for assistive applications, enabling users to interact with the environment beyond their physical limitations. [47, 78]

Furthermore, some researchers have explored extending human capabilities through Supernumerary Robotic Limbs (SRLs) [46, 80]. SRLs introduce additional body parts, such as extra arms or fingers that humans do not originally possess, and augment human capabilities by controlling them through users' bodily movements or physiological signals [58, 70]. Several studies have demonstrated that humans can embody body structures different from their natural body, such as six-fingered hands and third or fourth arms, in virtual environments [4, 24].

However, robot-mediated bodily interaction in real environments entails inherent constraints. Temporal limitations such as control latency and differences in response speed, as well as morphological differences including joint structure, range of motion, and dynamics, introduce mismatches between the user's natural movements and the robot's behavior [7, 41, 52]. In addition, autonomous robot behaviors such as obstacle avoidance can create inconsistencies between user input and robot behavior, thereby reducing the sense of embodiment. Although embodiment is critical for intuitive and low-effort interaction [3, 15], such inconsistencies disrupt the alignment between users' actions and system behavior, thereby undermining embodiment and degrading the quality of interaction.

2.3 Supporting Embodiment in Human-Robot Interaction

To mitigate the degradation of embodiment caused by physical constraints in real-world robot operation, prior work has explored various approaches to align system behavior with user intention. For example, latency introduced by communication and control has been addressed through predictive control methods that estimate future states and compensate for delays [34, 66]. Differences in embodiment, such as mismatches in kinematics or orientation between the human body and robotic system, have also been addressed through coordinate transformations and motion retargeting techniques [52].

Alternatively, when delays or other system limitations are unavoidable, several approaches explicitly communicate the internal state of the robot to the user in order to reduce ambiguity in interaction. For instance, feedback about robot constraints, planned motion, or safety interventions has been provided through haptic feedback [48, 49, 79] or visual feedback mechanisms [61, 67, 74]. Among these approaches, visual feedback by XR has been widely investigated due to its ability to convey rich information intuitively [65, 71, 75]. For example, previous studies have visualized robot states through augmented reality overlays [20, 73, 81], projection-based feedback [21, 37], or digital twin representations [52, 72]. By exposing robot intentions, constraints, or predicted motion, such systems help users better understand the robot's behavior and improve interaction performance.

However, in direct manipulation scenarios where users control a robotic hand as their own, explicitly overlaying the robot can interfere with embodiment by introducing additional representations [18]. Our work therefore explores an alternative design in which robot-side constraints are conveyed through changes in the embodied virtual hand rather than through explicit visualization of the robot.

2.4 Virtual-Physical Decoupling

In conventional XR systems, interaction is often grounded in a tight spatiotemporal alignment between virtual and physical elements. When such synchronization is maintained, users' actions are consistently reflected in both virtual representations and physical outcomes, resulting in interaction without perceptual discrepancy between bodily motion and environmental response. Such strong coupling enables users to directly manipulate physical objects through virtual representations [18, 64], as well as to realize interactions in which virtual and physical entities mutually influence each other [26, 36].

In contrast, intentionally relaxing this synchronization enables interaction beyond physical constraints. For example, some researches manipulate the temporal structure of reality by presenting past states or alternative scenes within the pass-through view [17, 27, 63]. Reality Promises [33] compensates for physical latency by presenting predicted outcomes in advance, thereby maintaining perceived immediacy. Remixed Reality [40] further reconfigures the environment by combining real and virtual elements at spatial and semantic levels. These approaches introduce decoupling between virtual and physical elements, allowing interaction to transcend the limitations imposed by the physical world.

A key technical approach for enabling such virtual-physical decoupling is diminished reality (DR) [13, 43]. DR is a technique that visually removes or replaces physical objects and bodies from the user's perception. Asynchronous Reality [17], for example, uses DR to hide physical events from the user while they are immersed in another task. Instead of presenting the physical events as they occur, the system later presents virtual replays triggered by interactions with the environment, effectively decoupling the perceived experience from the underlying physical timeline.

Various approaches have been proposed, including real-time reconstruction methods that estimate and complete the background from the current environment using RGB-D sensors or multiple cameras [16, 30], and methods that replace target regions using pre-captured scene models [2, 25]. While real-time reconstruction approaches can adapt to dynamic environmental changes, they require additional sensing setups and computational resources for online processing [10]. In contrast, pre-captured approaches are limited in their ability to handle environmental changes but can provide high-quality visual representations with lower runtime costs.

The visual removal and replacement enabled by DR have been widely used to improve users' perception and interaction with physical environments. For example, removing unnecessary objects or visual occlusions has been shown to reduce cognitive load, improve visibility, and enhance task performance [9, 44]. In robotic applications, DR has also been used to visually remove robots themselves or surrounding occluders from the user's view, improving visibility of the workspace and supporting more efficient operation [53, 62, 68].

In this work, we also visually remove the robot controlled by the user through DR, however, our goal is not to organize visual information or improve visibility of the workspace, but to maintain embodiment during robot-mediated interaction. In addition, we remove the user's original hand and present only the extended virtual hand as the user's body representation [18, 39], allowing

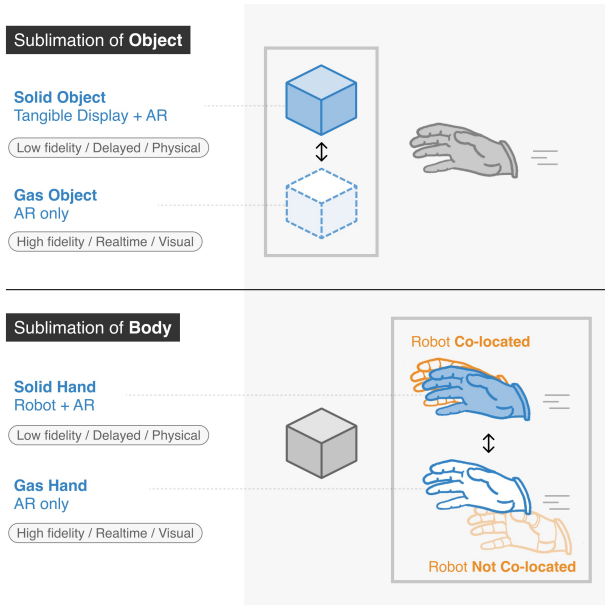


Figure 2: Comparison between Sublimate [36] and our body-oriented extension of its concept.

users to embody the extended hand. This allows users to leverage the robot’s physical interaction capability during contact and manipulation, while separating the virtual representation from the robot’s physical constraints when necessary, enabling remote object manipulation with a consistent sense of embodiment.

3 Body Sublimation

3.1 Vision and Concept

Our vision is to enable users to perform physical interaction beyond the spatial limits of their biological body. To achieve this, robotic arms are necessary as a means to reach and manipulate out-of-reach objects. However, in order to support such interaction in a natural and intuitive manner, the robotic arm must be embodied as part of the user’s own body.

In considering embodiment for robotic arms, we draw inspiration from Sublimate [36] by Leithinger et al. Sublimate explored a way to integrate two complementary representations: tangible display objects, which enable physical interaction but suffer from latency and limited resolution, and augmented reality objects, which provide high visual fidelity and responsiveness but lack physical interaction. It introduced a metaphor for transitioning between these representations to leverage the strengths of both; in Sublimate, objects can dynamically switch between a solid physical state, where they are rendered as tangible objects through a shape display with spatially co-located 3D graphics, and a gas-like state, where they are represented by 3D graphics only (Fig. 2 top).

While Sublimate focused on the design of objects being touched, our work shifts the focus toward the body that performs the interaction. With recent advances in embodied interfaces, a similar trade-off to that raised by Sublimate now appears in extended body systems. Robotic arms and hands enable physical interaction but

suffer from physical constraints such as latency, limited workspace, and collision avoidance. In contrast, virtual hands in XR provide fast and high-fidelity control but cannot directly interact with physical objects (Fig. 2 bottom). To address this trade-off, we propose Body Sublimation. In our approach, the user embodies a virtual hand as an extended hand, while physical interaction is performed by a hidden robotic manipulator. The extended hand can take two states depending on the correspondence between the virtual hand and the robot hand. When the robot matches the virtual hand, it enters a *solid state*, in which it becomes physically grounded and can interact with real-world objects. When they do not match, it enters a *gas state*, in which it becomes a purely virtual entity without physical constraints (Fig. 2 bottom).

To enable users to embody the virtual hand and operate it as an extended hand, we employ DR to visually remove both the robotic hand and the user’s original hand, leaving only the state-changing virtual hand visible (Fig. 1).

3.2 Benefit of Our Interaction Approach

3.2.1 Maintaining Embodiment despite Human–Robot Kinematic Differences. Although robotic arms and human arms differ in morphology and kinematics, Body Sublimation enables users to maintain embodiment during physical interaction by visually removing both the robotic hand and the user’s original hand. As a result, the virtual hand becomes the sole embodied representation, allowing users to interact with physical objects without being affected by discrepancies between human and robotic bodies (Fig. 3 left).

3.2.2 Maintaining Embodiment under Collision Avoidance. When operating robots in real-world environments, obstacles are often present, making autonomous collision avoidance necessary for safety and performance; however, such automation can disrupt the user’s sense of embodiment. In Body Sublimation, the robot is not directly exposed to the user, and collision avoidance is handled autonomously by the invisible robot. When a potential collision is detected, the system transitions the extended hand into the gas state while the robot adjusts its motion to avoid obstacles. This allows the user to perceive a continuous embodied interaction, while the virtual hand can pass through obstacles and reach target objects beyond them.

3.2.3 Maintaining Embodiment beyond the Robot Workspace. When the extended hand moves outside the robot’s workspace, the robotic hand cannot follow it. In such cases, the extended hand transitions from solid to gas. This allows the system to convey the robot’s state through changes in the state of the hand, preserving embodiment while avoiding explicit exposure of the robot itself.

3.2.4 Interaction Across Multiple Robotic Platforms. Since the robotic system is not directly exposed to the user, the interaction is not tied to a specific robotic embodiment and can seamlessly adapt to different robotic systems. Various mechanisms, such as robotic arms, mobile manipulators [78], or parallel-link systems [47], can be used to realize physical interaction of the extended hand. Furthermore, the system enables continuous interaction across multiple robots by dynamically switching which robot follows the virtual hand based on its spatial position. This allows users to maintain continuous

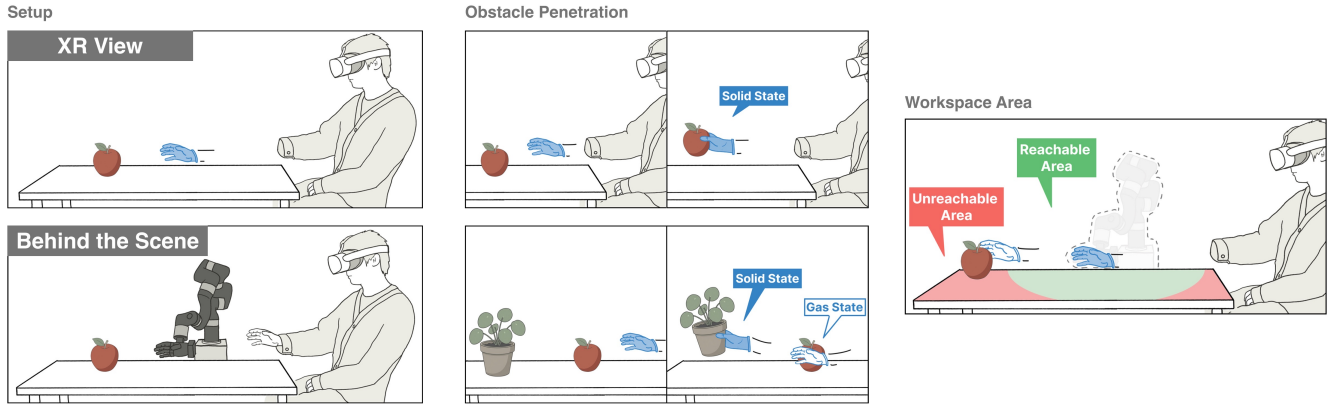


Figure 3: Interactions enabled by Body Sublimation. The setup (left) enables two states, gas and solid (center), allowing the extended hand to penetrate obstacles and implicitly convey robot constraints, such as workspace limitations, to the user through state changes (right).

embodied interaction across a workspace larger than what a single robot can cover.

3.3 Design Considerations

To ensure safety and usability in interactions enabled by Body Sublimation, which involves an invisible robotic hand and manipulator, the system must satisfy several key requirements.

3.3.1 Communicating the state of the extended hand. While Body Sublimation maintains embodiment by visually removing the robotic system, users still need to understand whether the extended hand can physically interact with the environment to ensure usability. Therefore, the system should convey the state of the extended hand to the user through visual or other forms of feedback. In particular, it is important to inform the user whether the extended hand is physically grounded or not, i.e., whether it is in the solid or gas state. This feedback allows users to understand whether the extended hand can physically interact with the environment through changes in its state, without directly perceiving the physical constraints of the underlying robot, thereby maintaining a sense of embodiment.

3.3.2 Ensuring safety through autonomous collision avoidance. Since the robotic hand is invisible to the user, the system must autonomously ensure safe robot behavior. The robot should avoid collisions with environmental obstacles, objects, and people including the user. When a potential collision is detected, the robot should temporarily decouple from the virtual hand and adjust its motion to avoid the obstacle while maintaining perceptual continuity.

3.3.3 Supporting attachment and detachment function between the user and robot. When the robot workspace overlaps with the user’s reachable space, users may want to switch between using the robot for remote interaction and using their original hand for direct interaction. Therefore, the system should allow users to intentionally attach and detach the extended hand from the robot depending on their interaction goals.

4 Implementation

4.1 User Experience

We first describe the end-to-end user experience before detailing its technical implementation. Please also refer to the supplemental video for a demonstration of the interaction.

The user wears a VR headset (Meta Quest 3) and sits in front of a desk. The passthrough view of the physical environment is displayed, but the user’s original hand and the robotic arm/hand are visually hidden through Gaussian splat-based diminished reality. In their place, a virtual hand is displayed.

The virtual hand employs a Go-Go-style reach amplification: the farther the user extends their hand, the farther the virtual hand travels. A blue sphere is displayed at the position of the (hidden) robotic hand. When the user reaches out with the virtual hand and touches this sphere, the virtual hand turns into a blue filled hand, indicating attachment. In this attached state, the robotic arm and hand track the virtual hand’s motion, enabling physical interaction such as grasping objects (solid state). If the robotic arm encounters an obstacle, reaches the boundary of its workspace, or cannot match the virtual hand’s position for other reasons, meaning that the physical interaction afforded by the robot does not correspond to what the virtual hand visually suggests, the virtual hand transitions to a blue outline, representing the gas state.

To detach from the robot, the user performs a stop sign gesture (palm facing away). Upon detachment, the blue sphere remains at the robot’s current position, and the virtual hand returns to white, signifying the detached state.

4.2 System Architecture

Our system consists of two main components: a Meta Quest 3 head-mounted display (HMD) for hand tracking, diminished reality rendering, and virtual hand visualization, and an xArm7 (uFactory) robotic arm equipped with an Ability Hand (PSYONIC) as the end effector for realizing physical interaction. The user application was implemented in Unity and deployed to a PC connected to the Meta

Quest 3 via USB (Quest Link). The robotic control system was implemented using Python and TouchDesigner. The two components communicate bidirectionally via UDP using Protocol Buffers for serialization. To establish a shared coordinate system between the Quest and the robot, we manually placed a spatial anchor aligned with the robot's base origin.

4.3 Extended Virtual Hand

To enable the user to reach objects beyond their physical arm's length, we employ a Go-Go-like [54] hand extension technique.

We compute the distance r_p between the user's original hand and the estimated center of the head on the horizontal (XZ) plane only, excluding the vertical axis to avoid unintended amplification from raising or lowering the hand. Given a distance threshold $D = 0.1$ m, the virtual hand position is computed as follows: when $r_p \leq D$, the virtual hand maps one-to-one with the original hand. When $r_p > D$, the virtual hand distance r_v from the head center is computed as:

$$r_v = D + g \cdot (r_p - D)^2 \quad (1)$$

where $g = 10$ is the amplification gain. The direction of the virtual hand on the XZ plane remains identical to that of the original hand, and the Y coordinate is passed through without modification.

The visual appearance of the virtual hand encodes two independent dimensions of system state—attachment and the solid/gas state—through color and fill style. When the virtual hand is *not* attached to the robot, it is rendered in white: as a filled hand in the solid state (one-to-one correspondence with the original hand, $r_p \leq D$), and as an outline in the gas state (reach is amplified, $r_p > D$). When the virtual hand is *attached* to the robot, it is rendered in blue: as a filled hand in the solid state (the robotic arm successfully tracks the virtual hand position), and as an outline in the gas state (the robotic arm cannot match the virtual hand's position due to obstacles, workspace limits, or other constraints).

4.4 Robotic Hand Control

The robotic hand was controlled to follow the virtual hand position received from Unity as the target position. During tracking, the robot's target motion was computed while considering collision avoidance and speed constraints.

The robot was controlled by sending the end-effector pose (6DoF), and inverse kinematics was automatically solved by the robot controller. Collision avoidance was implemented as a retargeting of the target position. Let $\mathbf{x}_{target}$ denote the desired target position of the robot and $\mathbf{x}_{obs}$ denote the position of an obstacle. The distance between them is defined as $d = \|\mathbf{x}_{target} - \mathbf{x}_{obs}\|$. When this distance becomes smaller than a predefined safety threshold δ , the target position is pushed outward along the radial direction from the obstacle to maintain a safe clearance, as defined in Eq. (2).

$$\mathbf{x}'_{target} = \begin{cases} \mathbf{x}_{target} + (\delta - d) \frac{\mathbf{x}_{target} - \mathbf{x}_{obs}}{\|\mathbf{x}_{target} - \mathbf{x}_{obs}\|} & (d < \delta) \\ \mathbf{x}_{target} & (d \geq \delta) \end{cases} \quad (2)$$

Here, $\mathbf{x}'_{target}$ represents the adjusted target position used for robot control. Potential collision targets included the user's original hand, a base position defined to prevent self-collision of the robot,

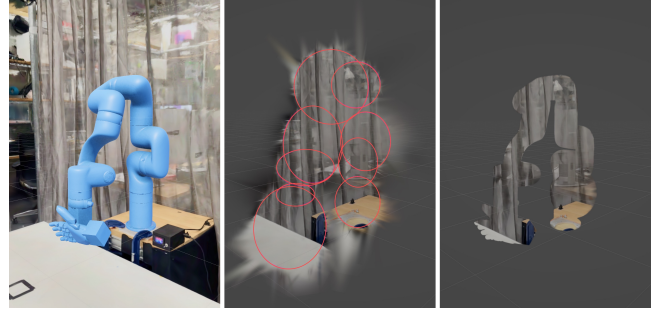


Figure 4: (Left) Gaussian splat of the environment and the 3D-reconstructed robot arm. (Center) Splat clipped using cone cutouts. The red ellipses represent the base of each cone. (Right) Final splat rendering clipped and edge-feathered using alpha masking. To account for latency and positional misalignment of the robot arm, the mask boundaries in the actual scene are slightly expanded compared to those shown in this image.

and obstacles tracked by OptiTrack, as well as obstacles fixed at predetermined positions in the environment.

To ensure safe motion, the robot speed was constrained such that the speed of end-effector remained below 1000 mm/s at all times. In addition, when abnormal torque was detected, the robot was immediately stopped according to the robot controller's safety mechanism. The robot arm has a reach of 750 mm; therefore, the target positions sent to the robot were constrained within a sphere of 750 mm radius from the robot base. In addition, to avoid overly complex robot configurations, the orientation of the end effector (roll, pitch, yaw) was constrained within predefined limits of $\pm 90^\circ$. When the target position moved outside the predefined workspace, it was retargeted to remain within the workspace boundaries.

When the target position was modified due to these operational constraints, such as velocity limitation, workspace limitation, and collision avoidance, the system considered that the robotic hand could no longer accurately follow the extended virtual hand. This mismatch was sent back to the Unity system as a state flag and used to control the transition from the solid state to the gas state.

4.5 Diminished Reality Rendering

Here, we describe our DR implementation. As discussed in the Related Work section, we adopt a pre-captured scene representation to prioritize visual fidelity over adaptability to dynamic environments. To visually remove both the robotic arm/hand and the user's original hand from the passthrough view, we employ a 3D Gaussian splatting [35] system that renders a previously captured splat of the environment over the entities to be hidden. We use an open-source Gaussian splatting renderer [1], extended with additional functionalities described below.

4.5.1 Splat Preparation and Masked Rendering. The environment was captured using an iPhone with a Gaussian splatting application (Luma 3D) and imported into Unity. We color-corrected the splat to match the Quest's passthrough view by enabling a tone curve adjustment in the renderer.

To clip the splat to the regions to be hidden, we adopt a two-stage masking similar to that of Reality Promises [33]: cone cutouts at the compute stage for coarse region selection, followed by screen-space alpha masking with edge feathering in the fragment shader for pixel-accurate silhouette clipping. Cone cutouts are attached to the tracked positions of the robotic arm and the user’s original hand, so that only the splats necessary to cover these entities are rendered. Whereas Reality Promises masked a low-DoF robot by combining two elliptical masks, our system must cover the articulated 7-DoF arm and multi-fingered hand, removing only the targeted regions with high fidelity so that the surrounding view remains intact and real-time embodied interaction is not disrupted. We therefore reconstruct the full robot within Unity, using publicly available 3D models of the xArm and Ability Hand driven by real-time joint angles received via UDP, and render it off-screen to a mask texture, to which a Gaussian blur is applied to feather the boundary. During splat rendering, the fragment shader samples this mask at each pixel and multiplies the splat’s output by the mask value, ensuring that the splat patch covers the robotic arm accurately without bleeding into the surrounding passthrough view (Fig. 4).

4.5.2 Diminished Reality for the Original Hand. To prevent conflicting body ownership cues between the original hand and the virtual hand, the original hand must be hidden from the passthrough view. However, our initial attempts revealed a critical constraint: using precise hand-tracking data to cutout the original hand creates a distinct, hand-shaped outline due to subtle color and alignment discrepancies. This silhouette inadvertently draws the user’s attention back to the original hand’s location, undermining the illusion of hand extension. To address this problem, we employed a low-fidelity masking that deliberately avoids a precise silhouette. We approximate the original hand using three cuboids, representing the palm, thumb, and the remaining fingers, resulting in a mitten-like shape. The resulting lower-fidelity cutout at the original hand’s location reduces the visual salience of the boundary, encouraging the user to attend to the higher-fidelity virtual hand displayed at the extended position.

4.6 Technical Analysis

To evaluate the system performance of Body Sublimation, we measured the trajectories of the user’s hand, the virtual hand, and the robotic hand during interaction. The position of the user’s original hand was obtained from the hand tracking data of the Meta Quest 3. The position of the robotic hand was reconstructed from the robot’s joint states using forward kinematics to obtain the end-effector XYZ position. The position of the virtual hand was computed by applying a hand extension function to the user’s hand motion.

Fig. 5 shows the trajectories of the original hand, virtual hand, and robotic hand during an interaction sequence, together with the system state (gas / solid). The results illustrate how the virtual hand extends the user’s reach, while the robotic hand follows the motion when physical interaction is required.

The mechanical time lag was measured through cross-correlation between the control position (user’s original hand) and the measured position (robot). The measured delays were 0.025 s in the X direction, 0.30 s in the Y direction, and 0.044 s in the Z direction.

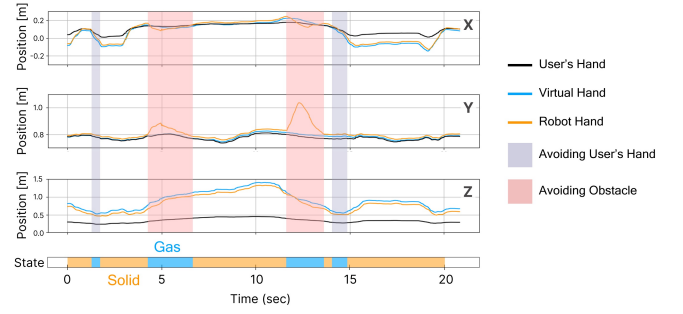


Figure 5: Time-series position data along each axis (X, Y, Z) for the user’s hand, virtual hand, and robot hand during a sequence of motion. Periods in the gas state are highlighted with color.

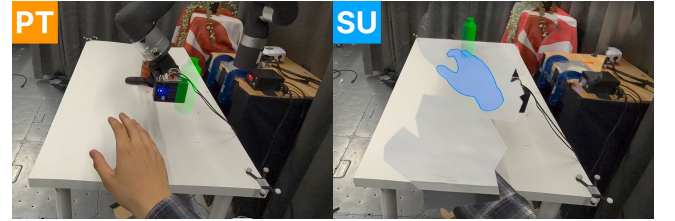


Figure 6: Two conditions for user study in participants’ view: PT, a pass-through baseline (left), and SU, the Body Sublimation approach (right).

5 User Study

We conducted a user study to assess the potential benefits of Body Sublimation relative to the baseline pass-through system.

5.1 Experimental Design

We conducted a within-subject study to evaluate whether Body Sublimation enables users to experience remote physical manipulation as part of their own body while maintaining task performance. Participants performed three types of tasks: reaching without obstacles, reaching with obstacle, and a grasping (pick-and-place) task.

Each task was performed under two conditions: (PT) a pass-through baseline, in which both the user’s real hand and the robot are visible (Fig. 6 left), and (SU) Sublimation approach, in which both are visually removed and replaced with a virtual hand representation (Fig. 6 right). Note that the underlying robotic system and control pipeline were identical across the two conditions; only the visual representation differed. This allows us to isolate the effect of perceptual embodiment introduced by Body Sublimation.

5.2 Participants

12 participants (10 male, 2 female) voluntarily joined the study. Participants’ age ranged from 19 to 29 years (Mean = 23.25, SD = 2.87), and all participants were right-handed. All participants provided informed consent prior to the experiment and were informed about the procedure and the approximate duration of the study.

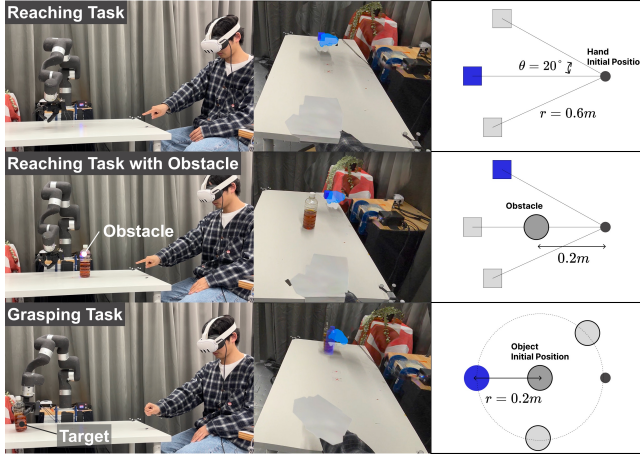


Figure 7: Spatial configuration of the user study. The initial hand position was fixed across all tasks. In the reaching task, three target candidates were placed at equal distances, and one was randomly selected as the target. An obstacle was positioned 0.2 m from the initial hand position. In the grasping task, multiple target candidates were placed at equal distances from the initial object position.

5.3 Procedure

Each experimental session lasted approximately 80 minutes and was conducted individually. The order of the PT and SU conditions was counterbalanced across participants, while the other factors, including task order, were fixed. Participants first completed a short practice session to familiarize themselves with the system and interaction method. Participants then performed three tasks in the following order: reaching without obstacles, reaching with obstacle, and grasping. After completing all three tasks under one condition, participants repeated the same procedure under the other condition. After completing each task under both conditions, participants reported their subjective experience via questionnaires. At the end of the session, a semi-structured interview was conducted to collect qualitative feedback.

5.4 Tasks

The spatial configuration of the tasks is shown in Fig. 7. The initial hand position was fixed across all tasks.

Reaching task. Participants moved their extended hand from the initial position toward a visually presented target. Three target candidates were placed at equal distances, and one was randomly selected in each trial. A trial was considered successful when the robot’s index finger touched the target and remained within the target region for 2 seconds.

Reaching task with obstacle. A physical object was placed 0.2 m from the initial hand position. The robot performed obstacle avoidance during motion. Participants were informed in advance that the robot would automatically avoid obstacles.

Grasping task. Participants reached for and grasped a physical object (a water bottle), then moved it to a goal region indicated in AR. Multiple target candidates were placed at equal distances

from the initial object position. A trial was considered successful when the object was placed within the goal region and held for 2 seconds. Trials were terminated as failures if the object was dropped or if the task exceeded 60 seconds. In each condition, participants performed three trials per task, and target positions were presented in randomized order.

5.5 Metrics

We evaluated both behavioral performance and subjective experience. In all trials, we recorded the positions of the virtual hand, the robot end-effector, and the user’s original hand. The virtual hand represents the target position generated from the user’s original hand using the Go-Go interaction mapping. Although it is not visually presented in the baseline condition, it is still computed and used as the control target for the robot, ensuring identical control inputs across conditions.

For the reaching task, we measured task completion time and trajectory straightness of the virtual hand in order to measure the effectiveness of reaching. For the grasping task, we measured success rate. Following prior work [14, 76, 77], trajectory straightness was computed as the ratio between the Euclidean distance from the start position to the target and the actual path length. A value closer to 1 indicates a more efficient trajectory.

To assess subjective experience, we used the Virtual Embodiment Questionnaire (VEQ)[57] to evaluate embodiment and the UEQ-S[22] to assess perceived usability. We adapted the VEQ by replacing the term “virtual body” in the questionnaire items with “extended hand”. Additionally, we conducted semi-structured interviews to analyze how participants interpreted discrepancies between intended motion and robot behavior, and whether the system was perceived as an external tool or as an extension of the body.

5.6 Results

We analyzed task performance and subjective responses to compare the two conditions. Statistical comparisons were conducted independently for each task between PT and SU conditions, and no correction for multiple comparisons was applied. For task performance, we first assessed normality using the Shapiro–Wilk test and selected either paired t-tests or Wilcoxon signed-rank tests accordingly. For subjective ratings obtained from Likert-scale questionnaires, we treated the data as ordinal and applied Wilcoxon signed-rank tests. In addition, qualitative preferences were collected through semi-structured interviews.

5.6.1 Task Performance. Fig. 8 left shows the task performance results. For the reaching task without obstacles, paired t-tests revealed no significant differences between conditions in either completion time or trajectory straightness. For the reaching task with obstacle, the Wilcoxon signed-rank test showed no significant difference in completion time. In contrast, trajectory straightness was significantly higher in SU than in PT (paired t-test, $p = 0.04$). For the grasping task, the Wilcoxon signed-rank test revealed no significant differences between conditions.

5.6.2 User Preferences. Subjective ratings were collected using 7-point Likert-scale questionnaires. Ownership, Agency, and Change

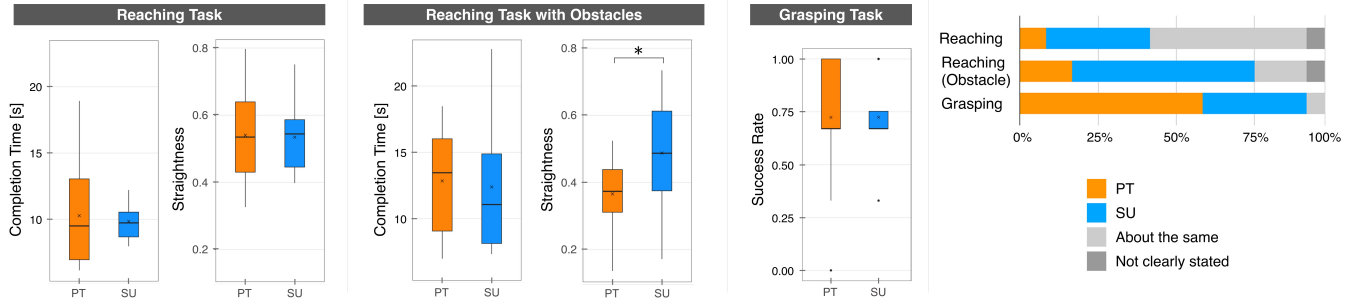


Figure 8: Task performance across conditions (left) and user preferences in each task (right). Asterisks indicate statistical significance: $*p < .05$

Table 1: Questionnaire results (median, IQR). $*$: $p < .05$, $**$: $p < .01$.

Measure	Reaching		Reaching (Obstacle)		Grasping	
	PT	SU	PT	SU	PT	SU
Ownership	4.1 (1.4)	5.4 (0.9)*	4.3 (0.8)	5.5 (0.9)*	4.9 (1.7)	5.5 (1.6)*
Agency	5.1 (1.8)	6.3 (1.0)*	4.8 (1.8)	6.0 (0.6)*	5.8 (1.4)	5.9 (1.1)
Change	1.3 (0.6)	1.8 (0.9)*	1.3 (1.2)	2.5 (1.9)	1.3 (1.9)	2.4 (1.7)
Pragmatic	4.9 (0.6)	5.9 (1.6)*	5.3 (1.4)	5.9 (0.6)	6.0 (1.1)	6.0 (1.9)
Hedonic	5.5 (1.4)	6.9 (0.6)**	5.9 (1.4)	6.6 (0.8)*	6.1 (1.1)	6.9 (0.3)**

subscales were derived from the VEQ, while pragmatic and hedonic quality were obtained from the UEQ-S. All comparisons were conducted using the Wilcoxon signed-rank test.

Table 1 summarizes the results for each task. For the reaching task without obstacles, SU showed significantly higher scores than PT in ownership ($p = 0.006$), agency ($p = 0.02$), change ($p = 0.03$), pragmatic quality ($p = 0.04$), and hedonic quality ($p = 0.009$). For the reaching task with obstacle, SU was significantly higher than PT in ownership ($p = 0.02$), agency ($p = 0.02$), and hedonic quality ($p = 0.02$). For the grasping task, SU was significantly higher than PT in ownership ($p = 0.04$) and hedonic quality ($p = 0.009$).

Finally, in the semi-structured interview, participants were asked which condition they found easier to perform for each task. The results are summarized in Fig. 8 right.

6 Discussion

The aim of this study was to examine whether the Body Sublimation approach, where both the user’s original hand and the robotic hand are visually removed and only a virtual hand representing the user’s intended position is presented, can improve users’ sense of embodiment while maintaining task performance.

Task Performance. Overall, the results suggest that Body Sublimation enhances embodiment while maintaining task performance.

Furthermore, in the reaching task with obstacle, trajectory straightness was significantly improved in SU. Although participants were informed that the robot would automatically avoid obstacles in both conditions, some participants noted that in PT, the visibility of the avoidance behavior led them to adjust their own movements (P2), or made them feel inclined to avoid obstacles themselves even when they knew the robot would do so (P8, P10). On the other hand, some participants reported that they avoided obstacles on their own regardless of the condition (P7), while others felt increased uncertainty when the robot was not visible (P4).

User Preferences. The results of the subjective questionnaire indicated that Body Sublimation improves both embodiment and usability. In the reaching task, more participants preferred SU (Fig. 8 right). Many participants reported that it was easier to understand the spatial relationship between the hand and the environment (P3, P4, P7, P10, P12). P4 noted, “PT was more difficult, as the robot arm and the control box obstructed the view and made it harder to perceive depth and spatial relationships.” On the other hand, one participant noted, “In SU, the mismatch between the robotic hand and the virtual hand was noticeable” (P3). In the reaching task with obstacle, several participants reported that SU felt less stressful (P3, P5, P9). Participants tended to avoid obstacles themselves in PT despite knowing it was unnecessary, and the visibility of the robot made them anxious about potential collisions (P9). Some also reported a sense of inconsistency between their own movements and the robot’s avoidance behavior in PT (P12). However, one participant perceived both conditions as using a tool rather than an extension of the body and therefore found PT easier to use (P10). Taken together, these findings suggest that Body Sublimation reduces visual interference from the robot, enabling more direct spatial understanding and interaction. At the same time, removing the robot’s visual representation can introduce perceptual discrepancies and reduce transparency for some users, indicating a trade-off between embodiment and system interpretability in Body Sublimation.

In contrast, a different tendency was observed in the grasping task. Although ownership and hedonic usability were significantly higher in SU, more participants preferred PT. A major reason was that the Gaussian splatting mask occasionally occluded parts of the object being grasped, and the mismatch between the virtual hand

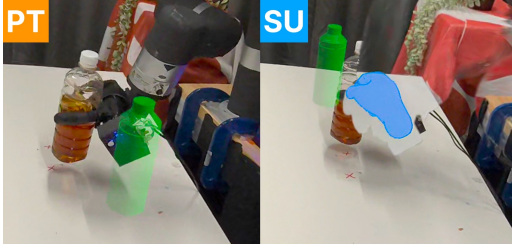


Figure 9: Comparison between PT and SU during grasping a bottle. In PT, the robot grasping the bottle is visible, whereas in SU, parts of the bottle are also occluded due to the Gaussian splatting mask.

and the robotic hand made it difficult to judge whether the object was successfully grasped (P2, P3, P6, P8, P9, P10, P11, P12). P11 also noted that haptic feedback would be necessary to confirm successful grasping. Fig. 9 shows a comparison between PT and SU during grasping. Precisely masking only the robot remains technically challenging, and the current approach can obscure relevant parts of the scene, which is a limitation. On the other hand, several participants reported that SU was easier to use. They noted that the virtual hand, being closer to the shape of a human hand, made it easier to determine the timing of finger and hand movements (P1), and that removing the visible robot reduced stress (P4, P5).

Findings from the Interview. Finally, the semi-structured interviews provided further insights into how Body Sublimation addresses challenges inherent to physical interaction in real environments. First, regarding environmental obstacles, as reflected in the feedback from the reaching task with obstacle, many participants reported that hiding the robot’s avoidance behavior reduced stress and prevented a decrease in their sense of agency. Second, regarding differences in embodiment, as observed in the grasping task feedback, several participants reported that SU made grasping easier by masking the robot and reducing the discrepancy between their own body and the robotic system. In addition, many participants noted that removing the visible robot arm reduced a sense of intimidation, resulting in a less stressful interaction. Regarding latency, most participants reported that it was not a major concern in either condition (7/12 participants). However, in the reaching with obstacle task, some participants perceived additional delay due to the robot’s avoidance behavior, particularly in PT (P2). In PT, the visible delay of the robot arm also influenced users’ movements, causing them to slow down or adapt their actions accordingly (P8, P11). In contrast, some participants reported that latency was more noticeable in SU, as the virtual hand made the intended position of the robot more explicit, thereby highlighting temporal discrepancies between the virtual and physical motion (P1, P3).

7 Applications

Body Sublimation enables embodied out-of-reach interaction through robotic mediation, even in cluttered daily environments or across spatially separated spaces. Here we describe example scenarios.

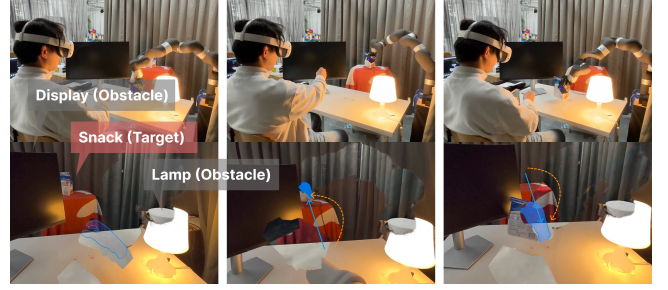


Figure 10: The system assists the user in fetching a distant snack without moving by extending the user’s reachable space, even in a desktop environment where many obstacles exist for the robot.

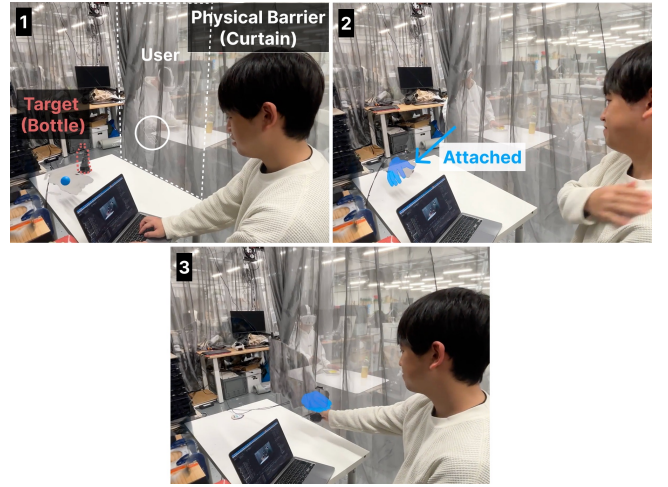


Figure 11: The user attaches to a robot beyond the curtain to hand over a bottle to a person on the other side (1, 2), and directly transfers the object (3).

7.1 Enhancing Fetching Action for Reachability

In everyday living environments, physical spaces often contain many obstacles such as monitors, furniture, and shelves, which limit direct access to objects. Body Sublimation enables users to interact with objects beyond such obstacles without requiring physical repositioning. For example, a user can fetch a snack placed on a shelf behind monitors without standing up. By extending the virtual hand beyond the obstructed region, the system allows the user to reach and retrieve objects that are otherwise physically inaccessible from their current position (Fig. 10).

7.2 Social Interaction Beyond Physical Barriers

Our system also enables embodied interaction across physical barriers by perceptually bridging the gap between the user’s reachable space and the robot’s operational space. When the user extends their arm toward a barrier, the extended hand transitions into a gas state and mediates the spatial discontinuity, visually and perceptually connecting the user’s body to the manipulation point

beyond the physical obstruction. Rather than switching to a separate robotic agent, the extended hand preserves a continuous bodily trajectory across the barrier. For example, a user can reach through a curtain and grasp a bottle located on the opposite side, then hand it to a friend standing beyond the barrier (Fig. 11).

8 Limitations and Future Work

This work has several possible directions for future improvement.

First, Body Sublimation can be extended beyond a single robotic arm to room-scale systems [59], such as cable-driven parallel robots [47]. Although such systems provide a larger workspace, they often suffer from dead zones in cluttered environments. Body Sublimation can mitigate this limitation by enabling seamless transitions between gas and solid states, maintaining interaction continuity even when the robot cannot perfectly follow the path. Body Sublimation is also applicable to telepresence scenarios. In teleoperation, limited viewpoint control and reduced situational awareness make it difficult to perceive obstacles and plan safe motions [56]. By allowing users to maintain a sense of agency while the system autonomously handles constraints such as obstacle avoidance, Body Sublimation can improve both usability and robustness in remote interaction.

Second, there are several limitations in our current implementation. In our implementation, DR masking relies on a digital twin reconstructed from joint angles, which can lead to misalignment and unintended occlusion of target objects. More visually accurate approaches, such as directly tracking occluding geometry, could improve visual consistency. Additionally, our method relies on a pre-captured scene representation using Gaussian splatting. While this enables high-fidelity rendering with moderate computational cost, it requires a time-consuming preprocessing step and cannot handle dynamic changes in the environment. Future work could explore multi-camera or real-time reconstruction methods to address these limitations. Furthermore, the current system provides only visual feedback, but haptic sensations during contact and grasping would improve the interaction experience. Integrating haptic feedback may further enhance embodiment and usability, and remains an important direction for future implementation.

Third, there are also limitations in our evaluation. The primary objective of the user study was to assess the overall Body Sublimation experience, particularly in terms of embodiment and agency, rather than to isolate the effects of individual design components or provide a comprehensive evaluation of task performance. Therefore, the individual contributions of diminished reality, robot masking, original-hand masking, and the gas/solid state transition remain unclear. Future work should clarify these contributions through controlled ablation studies. In addition, our evaluation was limited to relatively simple reaching and grasping tasks. Although grasping was important as a representative physical interaction task, our analysis primarily focused on subjective embodiment and usability rather than detailed manipulation performance. A more comprehensive evaluation of manipulation performance, including grasp accuracy, grasp stability, and failure modes, together with validation in more complex real-world manipulation scenarios, remains important future work.

Finally, participants' interview comments suggest a trade-off between embodiment and interpretability. While implicitly handling

processes such as obstacle avoidance can improve efficiency and enhance embodiment, the current binary gas/solid representation may limit users' understanding of system behavior. For example, some participants appeared not to understand why the robot failed to reach targets near the boundary of its workspace. These observations indicate that explicitly communicating robot states may improve task performance, but could also reduce the sense of embodiment. This suggests the need to convey rich information about robot-side constraints without exposing the robot itself. Prior work in object-oriented HRI has explored encoding robot reachability on the object side [74]. Similarly, a body-oriented approach may be required, in which such constraints are communicated through the user's embodied representation rather than through explicit visualization. This trade-off between embodiment and interpretability constitutes an important design consideration for future robot-mediated embodied interaction. A key characteristic of Body Sublimation is its tolerance to mismatches between user actions and the robot's actual behavior. Leveraging this property, future work could incorporate predictive control for latency compensation [34] or introduce autonomous assistance for precise manipulation [5, 31], enabling more accurate and robust interaction.

9 Conclusions

In this paper, we presented Body Sublimation, an interaction approach that enables embodied physical interaction beyond the limits of the human body by integrating a virtual hand representation with a hidden robotic manipulator. By decoupling the user's perceived interaction space from the physical constraints of the robot, our approach allows users to seamlessly transition between unconstrained virtual interaction and physically grounded manipulation.

Through a user study comparing our method with a pass-through baseline, we demonstrate that Body Sublimation enhances users' sense of agency and ownership, while maintaining task performance and improving trajectory straightness during reaching in the presence of obstacles. These results suggest that removing explicit robot representations and instead presenting a unified virtual hand can support more coherent embodiment, even in the presence of physical constraints such as reachability and obstacle avoidance.

Our findings highlight a new design direction for integrating XR and robotics, where physical constraints are handled implicitly while preserving a continuous embodied experience. We believe Body Sublimation provides a foundation for future systems that extend human interaction beyond natural reach, bridging the gap between virtual flexibility and physical capability.

Acknowledgments

This work is supported by JST Moonshot R&D Program (JPMJMS2013), JST SPRING (JPMJSP2108) and JSPS KAKENHI (25K22814, 25K03162). We also appreciated Mark Armstrong and Takayoshi Hagiwara for insightful comments.

References

- [1] [n. d.]. aras-p/UnityGaussianSplatting: Toy Gaussian Splatting visualization in Unity. <https://github.com/aras-p/UnityGaussianSplatting> [Online; accessed 2026-03-30].
- [2] Erwin Andre and Helmut Hlavacs. 2019. Diminished reality based on 3d-scanning. In *Joint international conference on entertainment computing and serious games*. Springer, 3–14.

- [3] Alissa N Antle, Paul Marshall, and Elise Van Den Hoven. 2011. Workshop on embodied interaction: theory and practice in HCI. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*. 5–8.
- [4] Ken Arai, Hiroto Saito, Masaaki Fukuoaka, Sachiyo Ueda, Maki Sugimoto, Michiteru Kitazaki, and Masahiko Inami. 2022. Embodiment of supernumerary robotic limbs in virtual reality. *Scientific reports* 12, 1 (2022), 9769.
- [5] Umur Atan, Varun R Bharadwaj, and Chao Jiang. 2024. Assistive control of robot arms via adaptive shared autonomy. In *2024 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*. IEEE, 1096–1102.
- [6] Reigo Ban, Keigo Matsumoto, and Takuji Narumi. 2026. Hitchhiking Hands: Enabling “Virtually Direct” VR Manipulation by Switching Multiple Hand Avatars with Gaze. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [7] Claire Blackett, Alexandra Fernandes, Espen Teigen, and Thomas Thoresen. 2021. Effects of signal latency on human performance in Teleoperations. In *International Conference on Human Interaction and Emerging Technologies*. Springer, 386–393.
- [8] Doug A Bowman and Larry F Hodges. 1997. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proceedings of the 1997 symposium on Interactive 3D graphics*. 35–ff.
- [9] Samantha WT Chan, Bektur Ryskeldiev, and Suranga Nanayakkara. 2022. DeclutterAR: mobile diminished reality and augmented reality to address hoarding by motivating decluttering and selling on online marketplace. In *2022 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-Adjunct)*. IEEE, 870–874.
- [10] Siru Chen, Lingxin Yu, Yuxuan Liu, Zhifei Ding, Jiacheng Zhang, Xinyue Wang, Jiahao Han, and Richen Liu. 2024. Diminished reality techniques for metaverse applications: A perspective from evaluation. *IEEE internet of things journal* 11, 21 (2024), 34734–34748.
- [11] Tongqing Chen, Hang Wu, Jiasen Wang, Xiaotao Li, Zhu Jin, and Lu Fang. 2026. SuperSuit: An Isomorphic Bimodal Interface for Scalable Mobile Manipulation. *arXiv preprint arXiv:2603.06280* (2026).
- [12] Hung Hon Cheng and Josie Hughes. 2025. Enhancing bimanual teleoperation with variable shoulder distance: manipulation in varying-scale of applications. *npj Robotics* 3, 1 (2025), 38.
- [13] Yi Fei Cheng, Hang Yin, Yukang Yan, Jan Gugenheimer, and David Lindlbauer. 2022. Towards understanding diminished reality. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–16.
- [14] Ki-Hun Cho, Mi-Ran Hong, and Won-Kyung Song. 2022. Upper-limb robot-assisted therapy based on visual error augmentation in virtual reality for motor recovery and kinematics after chronic hemiparetic stroke: a feasibility study. In *Healthcare*, Vol. 10. MDPI, 1186.
- [15] Paul Dourish. 2001. *Where the action is: the foundations of embodied interaction*. MIT press.
- [16] Akihito Enomoto and Hideo Saito. 2007. Diminished reality using multiple handheld cameras. In *Proc. ACCV*, Vol. 7. 130–135.
- [17] Andreas Rene Fender and Christian Holz. 2022. Causality-preserving asynchronous reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15.
- [18] Tiare Feuchtner and Jörg Müller. 2017. Extending the body for interaction with reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 5145–5157.
- [19] Lars Fritsche, Felix Unverzag, Jan Peters, and Roberto Calandra. 2015. First-person tele-operation of a humanoid robot. In *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*. IEEE, 997–1002.
- [20] Uwe Gruenefeld, Lars Prädell, Jannike Illing, Tim Stratmann, Sandra Drolshagen, and Max Pfingsthorn. 2020. Mind the arm: realtime visualization of robot motion intent in head-mounted augmented reality. In *Proceedings of Mensch und Computer 2020*. 259–266.
- [21] Zhao Han, Jenna Parrillo, Alexander Wilkinson, Holly A Yanco, and Tom Williams. 2022. Projecting robot navigation paths: Hardware and software for projected ar. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 623–628.
- [22] Andreas Hinderks. 2017. Design and evaluation of a short version of the user experience questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence* (2017).
- [23] Yusuke Hirao, Weiwei Wan, Dimitrios Kanoulas, and Kensuke Harada. 2023. Body extension by using two mobile manipulators. *Cyborg and Bionic Systems* 4 (2023), 0014.
- [24] Ludovic Hoyet, Ferran Argelaguet, Corentin Nicole, and Anatole Lécuyer. 2016. “Wow! i have six Fingers!”: Would You accept structural changes of Your hand in Vr? *Frontiers in Robotics and AI* 3 (2016), 27.
- [25] Masahiko Inami, Naoki Kawakami, and Susumu Tachi. 2003. Optical camouflage using retro-reflective projection technology. In *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings*. IEEE, 348–349.
- [26] Hiroshi Ishii, Dávid Lakatos, Leonardo Bonanni, and Jean-Baptiste Labrune. 2012. Radical atoms: beyond tangible bits, toward transformable materials. *interactions* 19, 1 (2012), 38–51.
- [27] Daichi Ito, Sohei Wakisaka, Atsushi Izumihara, Tomoya Yamaguchi, Atsushi Hiyama, and Masahiko Inami. 2019. EyeHacker: gaze-based automatic reality manipulation. In *ACM SIGGRAPH 2019 Emerging Technologies*. 1–2.
- [28] Yukiko Iwasaki, Kozo Ando, Shuhei Iizuka, Michiteru Kitazaki, and Hiroyasu Iwata. 2020. Detachable body: The impact of binocular disparity and vibrotactile feedback in co-presence tasks. *IEEE Robotics and automation Letters* 5, 2 (2020), 3477–3484.
- [29] Atsushi Izumihara, Tomoya Sasaki, Masahiro Ogino, Reona Takamura, and Masahiko Inami. 2019. Transfantom: transformation into bodies of various scale and structure in multiple spaces. In *ACM SIGGRAPH 2019 Emerging Technologies*. 1–2.
- [30] Songkran Jarusrisawad, Takahide Hosokawa, and Hideo Saito. 2010. Diminished reality using plane-sweep algorithm with weakly-calibrated cameras. *Progress in Informatics* 7, 7 (2010), 11–20.
- [31] Shervin Javdani, Siddhartha S Srinivasa, and J Andrew Bagnell. 2015. Shared autonomy via hindsight optimization. *Robotics science and systems: online proceedings* 2015 (2015), 10–15607.
- [32] Kai Junge and Josie Hughes. 2025. ADAPT-Teleop: robotic hand with human matched embodiment enables dexterous teleoperated manipulation. *npj Robotics* 3, 1 (2025), 31.
- [33] Mohamed Kari and Parastoo Abtahi. 2025. Reality Promises: Virtual-Physical Decoupling Illusions in Mixed Reality via Invisible Mobile Robots. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology*. 1–17.
- [34] Yusuke Kawashima, Yoshiaki Abe, Takeru Hashimoto, Shunichi Kasahara, and Ryuma Niiyama. 2024. AgiLimb: Embodied Agile Robotic Arm by Integrating Digital Reflex and Seamless Action Takeover. In *SIGGRAPH Asia 2024 Emerging Technologies*. 1–2.
- [35] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, George Drettakis, et al. 2023. 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.* 42, 4 (2023), 139–1.
- [36] Daniel Leithinger, Sean Follmer, Alex Olwal, Samuel Luescher, Akimitsu Hogge, Jinha Lee, and Hiroshi Ishii. 2013. Sublimate: state-changing virtual and physical rendering to augment interaction with shape displays. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1441–1450.
- [37] Florian Leutert, Christian Herrmann, and Klaus Schilling. 2013. A spatial augmented reality system for intuitive display of robotic data. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 179–180.
- [38] Shuang Li, Jiaxi Jiang, Philipp Ruppel, Hongzhuo Liang, Xiaojian Ma, Norman Hendrich, Fuchun Sun, and Jianwei Zhang. 2020. A mobile robot hand-arm teleoperation system by vision and imu. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 10900–10906.
- [39] Shuqi Liao, Yuqi Zhou, and Voicu Popescu. 2025. The Effect of Realism on Hand Redirection in Immersive Environments. *IEEE Transactions on Visualization and Computer Graphics* (2025).
- [40] David Lindlbauer and Andy D Wilson. 2018. Remixed reality: Manipulating space and time in augmented reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [41] Mitchell JH Lum, Jacob Rosen, Hawkeye King, Diana CW Friedman, Thomas S Lendvay, Andrew S Wright, Mika N Sinanan, and Blake Hannaford. 2009. Teleoperation in surgical robotics—network latency effects on surgical performance. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 6860–6863.
- [42] Mark R Mine, Frederick P Brooks Jr, and Carlo H Sequin. 1997. Moving objects in space: exploiting proprioception in virtual-environment interaction. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. 19–26.
- [43] Shohei Mori, Sei Ikeda, and Hideo Saito. 2017. A survey of diminished reality: Techniques for visually concealing, eliminating, and seeing through real objects. *IPSP Transactions on Computer Vision and Applications* 9, 1 (2017), 17.
- [44] Shohei Mori, Momoko Maezawa, and Hideo Saito. 2017. A work area visualization by multi-view camera-based diminished reality. *Multimodal Technologies and Interaction* 1, 3 (2017), 18.
- [45] Kohei Morita, Takefumi Hiraki, Haruka Matsukura, Daisuke Iwai, and Kosuke Sato. 2020. Extension of projection area using head orientation in projected virtual hand interface for wheelchair users. In *2020 59th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*. IEEE, 421–426.
- [46] Marie Muehlhaus, Marion Koelle, Artin Saberpour, and Jürgen Steimle. 2023. I need a third arm! eliciting body-based interactions with a wearable robotic arm. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–15.
- [47] Masahiro Ohta, Takatoshi Yoshida, Tomoya Sasaki, Jinsuke Morita, Hideaki Nii, and Kouta Minamizawa. 2026. ROomBOT: A Room as an Assistive Robotic Interface based on Cable-Driven Parallel Robotics. In *Proceedings of the twentieth international conference on tangible, embedded, and embodied interaction*.
- [48] Allison M Okamura. 2009. Haptic feedback in robot-assisted minimally invasive surgery. *Current opinion in urology* 19, 1 (2009), 102–107.

- [49] Claudio Pacchierotti et al. 2015. *Cutaneous haptic feedback in robotic teleoperation*. Springer.
- [50] Federico Parietti and H Harry Asada. 2014. Supernumerary robotic limbs for aircraft fuselage assembly: body stabilization and guidance by bracing. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 1176–1183.
- [51] Federico Parietti, Kameron Chan, and H Harry Asada. 2014. Bracing the human body with supernumerary robotic limbs for physical assistance and load reduction. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 141–148.
- [52] Siyou Pei, Alexander Chen, Ronak Kaoshik, Ruofei Du, and Yang Zhang. 2024. Arm Robot: AR-Enhanced Embodied Control and Visualization for Intuitive Robot Arm Manipulation. *arXiv preprint arXiv:2411.13851* (2024).
- [53] Alexander Plopski, Ada Virginia Taylor, Elizabeth Jeanne Carter, and Henny Admoni. 2019. InvisibleRobot: facilitating robot manipulation through diminished reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 165–166.
- [54] Ivan Poupyrev, Mark Billinghurst, Suzanne Weghorst, and Tadao Ichikawa. 1996. The go-go interaction technique: non-linear mapping for direct manipulation in VR. In *Proceedings of the 9th annual ACM symposium on User interface software and technology*. 79–80.
- [55] Yuzhe Qin, Wei Yang, Binghao Huang, Karl Van Wyk, Hao Su, Xiaolong Wang, Yu-Wei Chao, and Dieter Fox. 2023. Anyteleop: A general vision-based dexterous robot arm-hand teleoperation system. *arXiv preprint arXiv:2307.04577* (2023).
- [56] Daniel Rakita, Bilge Mutlu, and Michael Gleicher. 2019. Remote telemanipulation with adapting viewpoints in visually complex environments. *Robotics: Science and Systems XV* (2019).
- [57] Daniel Roth and Marc Erich Latoschik. 2020. Construction of the virtual embodiment questionnaire (veg). *IEEE transactions on visualization and computer graphics* 26, 12 (2020), 3546–3556.
- [58] MHD Yamen Saraiji, Tomoya Sasaki, Kai Kunze, Kouta Minamizawa, and Masahiko Inami. 2018. Metaarms: Body remapping using feet-controlled artificial arms. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. 65–74.
- [59] Tomomasa Sato, Yoshifumi Nishida, and Hiroshi Mizoguchi. 1996. Robotic room: Symbiosis with human through behavior media. *Robotics and autonomous systems* 18, 1-2 (1996), 185–194.
- [60] Hideki Shimobayashi, Tomoya Sasaki, Arata Horie, Riku Arakawa, Zendai Kashino, and Masahiko Inami. 2021. Independent control of supernumerary appendages exploiting upper limb redundancy. In *Proceedings of the Augmented Humans International Conference 2021*. 19–30.
- [61] Michael Stilman, Philipp Michel, Joel Chestnutt, Koichi Nishiwaki, Satoshi Kagami, and James Kuffner. 2005. Augmented reality for robot development and experimentation. *Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-05-55 2, 3* (2005).
- [62] Kazuya Sugimoto, Hiromitsu Fujii, Atsushi Yamashita, and Hajime Asama. 2014. Half-diminished reality image using three rgb-d sensors for remote control robots. In *2014 IEEE International Symposium on Safety, Security, and Rescue Robotics (2014)*. IEEE, 1–6.
- [63] Keisuke Suzuki, Sohei Wakisaka, and Naotaka Fujii. 2012. Substitutional reality system: a novel experimental platform for experiencing alternative reality. *Scientific reports* 2, 1 (2012), 459.
- [64] Ryo Suzuki, Hooman Hedayati, Clement Zheng, James L Bohn, Daniel Szafir, Ellen Yi-Luen Do, Mark D Gross, and Daniel Leithinger. 2020. Roomshift: Room-scale dynamic haptics for vr with furniture-moving swarm robots. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–11.
- [65] Ryo Suzuki, Adnan Karim, Tian Xia, Hooman Hedayati, and Nicolai Marquardt. 2022. Augmented reality and robotics: A survey and taxonomy for ar-enhanced human-robot interaction and robotic interfaces. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–33.
- [66] Kazuma Takada, Midori Kawaguchi, Akira Uehara, Yukiya Nakanishi, Mark Armstrong, Adrien Verhulst, Kouta Minamizawa, and Shunichi Kasahara. 2022. Parallel ping-pong: Exploring parallel embodiment through multiple bodies by a single user. In *Proceedings of the augmented humans international conference 2022*. 121–130.
- [67] Takumi Taki, Masato Kobayashi, Eduardo Iglesias, Naoya Chiba, Shizuka Shirai, and Yuki Uranishi. 2025. MRHaD: Mixed Reality-based Hand-Drawn Map Editing Interface for Mobile Robot Navigation. In *2025 34th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1888–1895.
- [68] Ada V Taylor, Ayaka Matsumoto, Elizabeth J Carter, Alexander Plopski, and Henny Admoni. 2020. Diminished reality for close quarters robotic telemanipulation. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 11531–11538.
- [69] Yuta Ueda, Yuki Asai, Ryuichi Enomoto, Kai Wang, Daisuke Iwai, and Kosuke Sato. 2017. Body cyberization by spatial augmented reality for reaching unreachable world. In *Proceedings of the 8th Augmented Human International Conference*. 1–9.
- [70] Kohei Umezawa, Yuta Suzuki, Gowrishankar Ganesh, and Yoichi Miyawaki. 2022. Bodily ownership of an independent supernumerary limb: an exploratory study. *Scientific reports* 12, 1 (2022), 2339.
- [71] Michael Walker, Thao Phung, Tathagata Chakraborti, Tom Williams, and Daniel Szafir. 2023. Virtual, augmented, and mixed reality for human-robot interaction: A survey and virtual design element taxonomy. *ACM Transactions on Human-Robot Interaction* 12, 4 (2023), 1–39.
- [72] Chao Wang, Anna Belardinelli, Stephan Hasler, Theodoros Stouraitis, Daniel Tanneberg, and Michael Gienger. 2023. Explainable human-robot training and co-operation with augmented reality. In *Extended abstracts of the 2023 CHI conference on human factors in computing systems*. 1–5.
- [73] Hong Wang, Ridhima Phatak, James Ocampo, and Zhao Han. 2026. Indicating Robot Vision Capabilities with Augmented Reality. *International Journal of Social Robotics* 18, 3 (2026), 42.
- [74] Lauren W Wang, Mohamed Kari, and Parastoo Abtahi. 2026. Explainable OOHRI: Communicating Robot Capabilities and Limitations as Augmented Reality Affordances. *arXiv preprint arXiv:2601.14587* (2026).
- [75] Xian Wang, Luyao Shen, and Lik-Hang Lee. 2024. Towards massive interaction with generalist robotics: a systematic review of XR-enabled remote human-robot interaction systems. *arXiv preprint arXiv:2403.11384* (2024).
- [76] Ziheng Wang and Ann Majewicz Fey. 2018. Human-centric predictive model of task difficulty for human-in-the-loop control tasks. *PloS one* 13, 4 (2018), e0195053.
- [77] Nicolas Wenk, Joaquin Penalver-Andres, Rara Palma, Karin A Buetler, René Müri, Tobias Nef, and Laura Marchal-Crespo. 2019. Reaching in several realities: motor and cognitive benefits of different visualization technologies. In *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*. IEEE, 1037–1042.
- [78] Jehan Yang, Kent Shibata, Douglas Weber, and Zackory Erickson. 2025. High-density electromyography for effective gesture-based control of physically assistive mobile manipulators. *npj Robotics* 3, 1 (2025), 2.
- [79] Dawei Zhang, Roberto Tron, and Rebecca P Khurshid. 2021. Haptic feedback improves human-robot agreement and user satisfaction in shared-autonomy teleoperation. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 3306–3312.
- [80] Hongyu Zhou, Tom Kip, Yihao Dong, Andrea Bianchi, Zhanna Sarsenbayeva, and Anusha Withana. 2025. Juggling extra limbs: Identifying control strategies for supernumerary multi-arms in virtual reality. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–16.
- [81] Qiushi Zhou, Antony Chacon, Jiahe Pan, and Wafa Johal. 2025. Assisting MoCap-Based Teleoperation of Robot Arm using Augmented Reality Visualisations. In *2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 1765–1769.