Ronak R. Mohanty
J.Mike Walker '66 Department of Mechanical Engineering, Texas A&M University, College Station, TX 77840
e-mail: ronakmohanty@aggienetwork.com

Abhijeet S. Raina
J.Mike Walker '66 Department of Mechanical Engineering, Texas A&M University, College Station, TX 77840
e-mail: abhijeetsinghraina@tamu.edu

Subhrajyoti Chaudhuri
J.Mike Walker '66 Department of Mechanical Engineering, Texas A&M University, College Station, TX 77840
e-mail: schaud_94@tamu.edu

Francis Quek
Department of Visualization, Texas A&M University, College Station, TX 77840
e-mail: quek@tamu.edu

Shinjiro Sueda
Department of Computer Science & Engineering, Texas A&M University, College Station, TX 77840
e-mail: sueda@tamu.edu

Vinayak R. Krishnamurthy
J.Mike Walker '66 Department of Mechanical Engineering, Texas A&M University, College Station, TX 77840;
Department of Computer Science and Engineering (by Affiliation), Texas A&M University, College Station, TX 77840
e-mail: vinayak@tamu.edu

Spatial Manipulation in Virtual Peripersonal Space: A Study of Motor Strategies

This article studies fine motor strategies for precise spatial manipulation in close-to-body interactions. Our innate ability for precise work is the result of the confluence of visuo-tactile perception, proprioception, and bi-manual motor control. Contrary to this, most mixed-reality (MR) systems are designed for interactions at arms length. To develop guidelines for precise manipulations in MR systems, there is a need for a systematic study of motor strategies including physical indexing, bi-manual coordination, and the relationship between visual and tactile feedback. To address this need, we present a series of experiments using three variations of a tablet-based MR interface using a close-range motion capture system and motion-tracked shape proxies. We investigate an elaborate version of the classic peg-and-hole task that our results strongly suggests the critical need for high precision tracking to enable precise manipulation. [DOI: 10.1115/1.4054277]

Keywords: human computer interfaces/interactions, virtual and augmented reality environments

1 Introduction

1.1 Context and Motivation. Mixed reality (MR) systems have a rich and extensive history in the context of spatial manipulation [1–3]. The commodification of these systems has helped create a large body of work on spatial interactions, tangible user interfaces, and immersive displays [4–6]. As a result, there is also significant interest in employing spatial user interactions in the engineering domain in the context of VR for training assembly personnel [7,8]. Case in point, human assembly strategies have inspired human–robot interaction research by using peg-in-the-hole micro-assembly as their primary evaluation task [9–12].

At least in principle, the developments in MR systems and interaction techniques [13–15] are generally aligned with the embodied interactions viewpoint, wherein the intention is to incorporate bodily practice into interactions with virtual artifacts such that users *perceive the artifact as an extension of themselves; they act through it rather than on it* [16,17]. Despite this, spatial manipulation is still, by and large, achieved through *arms–length interactions*.

History, limitations of technology, and inertia of development have combined to favor interaction in the space farther away from the body in HCI [18] (Fig. 1). We must point out here that precision interfaces are not new as such. They have existed in laparoscopic surgery for some time wherein highly specialized tooling and high-precision motion tracking is used to enable precise manipulation. However, the same cannot be said for spatial user interfaces (SUIs) in general. For example, the distal display monitor and the ubiquity of mouse-based interaction have pushed spatial object manipulation to arms–length interaction [19–21]. Early limitations of stereoscopic displays and the interest in immersive experience created a bias toward deploying display resources to the visual periphery at the expense of the visual fovea [22,23] with which fine motor ability is paired. As a case in point, recent studies in HCI literature [24–26] discuss a decrease in visual perception of a virtual object when placed at distance (beyond arm's length), as well as, its influence on the manipulation action performed by the user.

We argue that existing design methodologies for spatial interaction techniques need to reconcile with our mental representations of the actions we perform in physical spaces with physical objects — *the proximity of the action, the size and shape of the manipulated object, visual feedback, and the corresponding tactile perception* should be in synergy. The inspiration for our work comes from some seminal systems such as *HoloDesk* [27], *SpaceTop* [28], and *MixFab* [29] that underscore the importance of spatial

Contributed by the Computers and Information Division of ASME for publication in the Journal of Computing and Information Science in Engineering. Manuscript received November 29, 2021; final manuscript received March 28, 2022; published online May 20, 2022. Assoc. Editor: Chih-Hsing Chu.

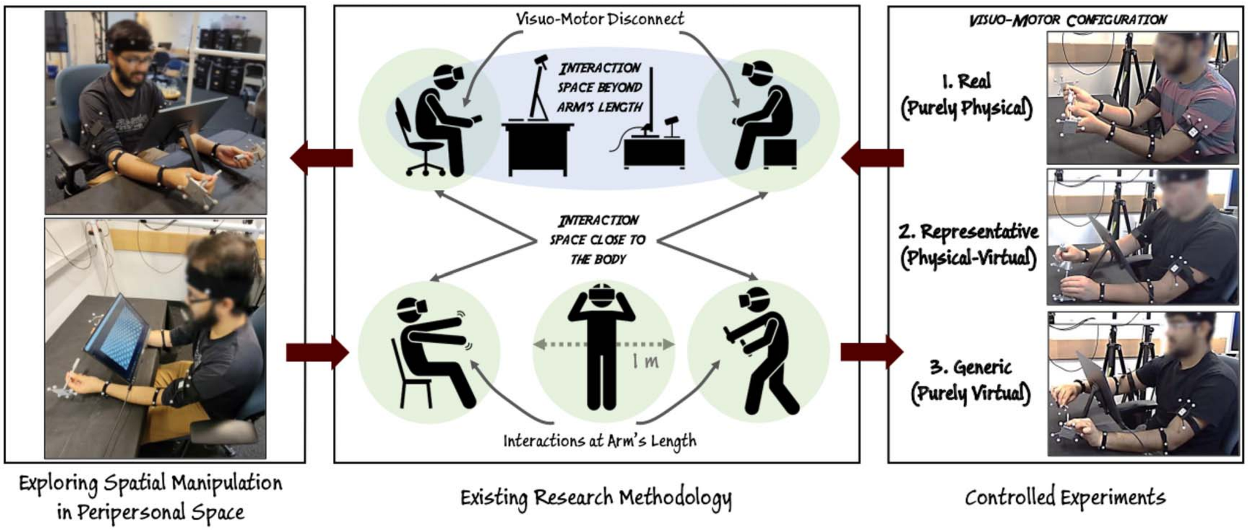

Fig. 1 We explore the interaction design space for precise manipulations (left) in the interaction space close to the body (middle) for which we conduct three controlled lab experiments comparing physical precise tasks with virtual manipulation (right)

interactions in close to the body. While these seminal works show us the power of precise manipulation close to the body, a systematic capturing, characterization, and analysis of actions in this space is missing from the literature. What is more, such a systematic study is critical to tap the potential of precise manipulation of virtual objects in the user's peripersonal space. *Therefore, our goal in this article in to explore the interaction space that merges physical actions (as we perform in daily physical tasks) with virtual interactions in a way that it supports our internal mental representation of the physical world.*

1.2 Knowledge Gaps. Our research draws from Gibson's seminal adage: "perception is for action'" [30]. The corollary of this concept for interaction design is that action should be designed to match the powers of human perception. While underexplored, HCI literature discuss similar thoughts on co-located action–perception space from the perspective of proximity to the action space [18]. While prior works [27–29,31] have implemented similar co-located action–perception interaction spaces, there is still much to understand about precise manipulations in terms of the types of actions and motor strategies that are active in the peripersonal space. Few works have highlighted spatial interactions from the point of natural user interfaces [32,33]; however, we believe the broader sense of this terminology shifts focus from the fundamental need to understand and advocate the design of interaction spaces that facilitate fine motor control. More importantly, there are also methodological gaps regarding how such interactions should be measured and analyzed. The metrics [34] that typically work for coarse-grained actions such as object docking are likely to miss the fine-grained finger movements, non-Euclidean kinematics of the arms, and concurrent head movements among other things. With this in view, we focus on investigating the following research questions from the point of exploring spatial manipulation in the peripersonal space:

- Q1 How do we appropriately measure precise spatial actions performed close to the body in virtual environments ?
- Q2 What is the quantitative extent of technological factors that affect the blending of the visuo-motor space close to body for precise motor control ?
- Q3 What are the implications of this action–perception blending on design of interaction spaces in the user's peripersonal space ?

Hence, we conduct a systematic and fundamental investigation of motor strategies including bio-mechanical stability, physical indexing, bi-manual coordination, and the visuo-tactile perception.

1.3 Approach and Contribution. Taking the aforementioned broader questions into consideration, we systematically investigate and present our findings through a set of three controlled lab experiments to understand how spatial interactions designed close to the body (peripersonal space) affect fine motor control and their influence on action-specific perception. In this article, we make three primary contributions: (1) we design and prototype a spatial interface to analyze user behavior for precise spatial task in a *real-world* scenario having a co-located action–perception space. We use this experiment as a reference to further evaluate user behavior across two experiments having visuo-tactile and visual-only feedback, respectively, for precise virtual tasks; (2) we evaluate user performance through an elaborate version of the classic peg-and-hole bi-manual task across different shapes with the increasing complexity in peg–hole geometry and quantity; (3) we propose a new set of user evaluation metrics in terms of cumulative kinetic energy, energetic transitions for *fine* and *coarse* spatial actions. We also evaluate using traditional metrics of the task completion time and peg–hole insertion accuracy to corroborate our findings.

Our analysis highlights the key motor strategies that are followed to perform precise bi-manual tasks in the interaction space close to the body. We quantitatively demonstrate how sensitive the overall user performance is to the blending of the visual and proprioceptive cues in the peripersonal space. The kinetic energy analysis provides fundamental insights into user performance across the three interfaces as it helped capture nuances that are typically lost in arm's length interactions.

2 Background

The background of our work stems from the broader areas of action–perception and proprioceptive feedback in spatial interactions focused on precise motor control.

2.1 Proximity to Action in Spatial User Interfaces. Looking across the mixed-reality continuum [1,2,6], we find that interfaces with head mounted displays, desktop-VR, tablet-AR, and augmented virtuality displays work within medium to large interaction volume for 3D object manipulation tasks [13–15,35]. While there are several works that study SUIs in the peripersonal space, they primarily focus on social behavior of virtual avatar of humans in a VR environment toward enhancing user engagement and immersion [36–39] or *social interaction through cross-device interaction* [40]. Similarly in *ubiquitous computing*, the focus is on developing spatially aware systems such as smart display setups [41–48], and input-control mechanisms that control hardware and their software elements using the portable hand-held devices in *close proximity* to their body [49–51]. However, little is currently known about 3D design and shape manipulation-type interactions in close-to-the-body scenarios where precise and fine motor actions are actually carried out. If anything, these few recent studies show that the role of proximity to action for spatial tasks is quite rich and still largely remains underexplored. Our work seeks to systematically explore this interaction space through three controlled lab experiments by building on the current and growing literature on spatial interactions performed in the peripersonal space.

2.2 Proprioceptive Feedback. Another aspect that is of great importance is proprioception. Proprioceptive and kinesthetic control play a key role in spatial manipulation, especially close to the body [52]. The lack of force (kinesthetic control) or tactile (touch) feedback as observed commonly across these systems severely impedes the ability to make fine spatial control; as a result, object manipulation becomes a difficult and *high effort task* [53]. This is mainly experienced with spatial actions performed close to the body such as MR systems for mid-air pointing and selection actions [54,55]. It has been found that proprioception at an egocentric distance around the periphery of human body helps reduce dependency on the visual feedback for manipulation actions in local and distant mid-air interaction spaces. Recent work by Plaumann et al. discusses a formal study focused on studying the visuo-motor relationship in spatial actions [56]. While the study primarily focuses on *macro*(coarse), as well as, *micro*(fine-grain) interactions, the key finding here is that the users experience a *visuo-motor mismatch* for spatial pointing tasks at distances away from the body, and this discrepancy reduces in pointing actions performed close to the body. On the other hand, Argelaguet and Andujar note visual dependency in virtual environments attributing to distinct motor and visual spaces for spatial interactions [14]. Further, De Boeck et al. [57] also demonstrate that interactions performed proximal to the body improves kinesthetic control by exploiting proprioception. The key finding relevant to our work is that while proprioception enhances manipulative precision close to the body, the same is not true for distal interactions, and these can be further influenced by other sensory perceptions such as visual, audio, tactile, and kinesthetics [58]. To further explore this, we study the differences between users' motor indexing and movement strategies for peg-and-hole tasks.

2.3 Bi-manual Action in Spatial Interactions. Work by Hinckley et al. [59] studied cooperative bi-manual interactions for virtual manipulation and provides a strong evidence for augmenting hand–eye coordination through the use of two hands in conjunction with haptics feedback. In general, two-handed interactions in coordinated tasks have been shown to increase cognitive engagement of the user [60–62] and efficiency of 3D object assembly [63]. Several works explore the advantages of bi-manual spatial interactions [20,64–70] for object selection [71] and manipulation (rotation, translation, scaling) of 3D objects. Alternatively, few recent works [72,73] showcase a *hybrid 2D–3D* input mode using a tablet surface and a 6DoF controller for bi-manual interactions in a VR setup. Similarly, Brandl et al. [74] explore the combination of two-handed interactions with pen and multi-touch inputs on a surface. Regardless of the wealth of literature, we believe that much is to be discovered regarding bi-manual interactions in the context of proprioception in peripersonal spaces. Hence, we compare bi-manual spatial interactions using our study interface with actual physical interactions. We test this through a MR-based SUI [1,2,6] spatial interaction interface wherein the visual feedback is setup between user and their hands (Fig. 2).

3 Interaction Design Consideration

We draw inspiration from prior works [31,27] to explore a co-located visuo-motor space for performing controlled bi-manual experiments in the user's peripersonal space. For this, we design and prototype one SUI per experiment (Fig. 1).

3.1 Factors Affecting Interaction Design. Below, we discuss some fundamental factors that influenced our design decisions for the exploration of the vision–proprioception spectrum:

3.1.1 Anatomy. Peripersonal space varies with the user's body structure, further classified based on the body part in action—*hands, face,* and *trunk* [18,75,76]. In fact, Galigani et al. [77] make a note that active tool usage in the user's peripersonal space enhances their proxemic perception, thereby improving spatial manipulation abilities. In our work, we focus on the peripersonal space defined by the upper limbs (Fig. 2(*b*)), mainly to observe and identify a comfortable interaction distance to perform precise motor tasks in MR systems and also to analyze its relation to the anatomical peripersonal space, specifically the hand, i.e., the *perihand space*.

3.1.2 Interaction Space. There are two key factors for the interaction work space that demand attention. First is the *location of the interaction space* with respect to the body (specifically the torso). This volume should represent what is currently known as the peripersonal space in the literature. The second factor is the *volume of the interaction space,* which is defined by the physical limits in sagittal, transversal, and coronal directions [78] from the

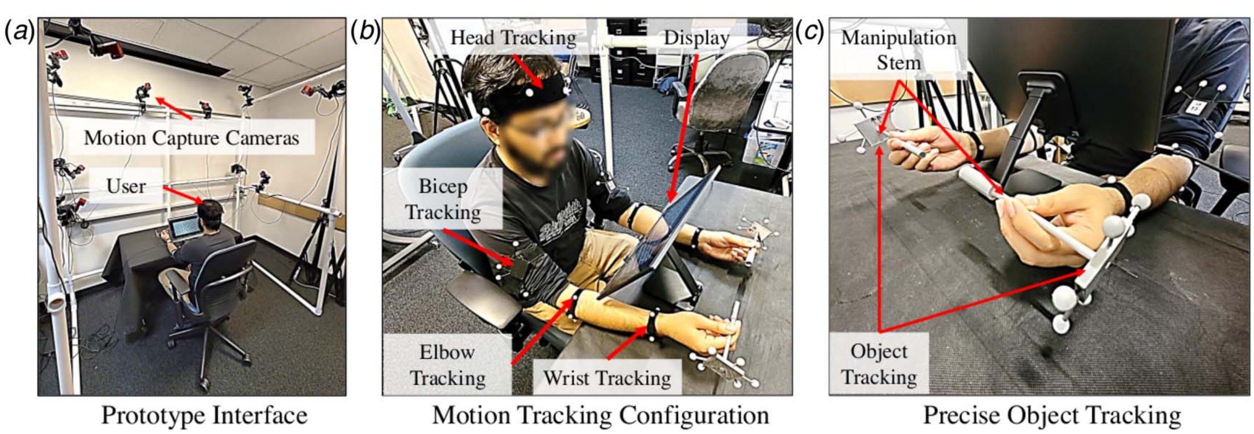

Fig. 2 Our prototype mixed-reality setup comprises (*a*) a screen placed between the user's torso and hands and also, ten high-precision OPTITRACK motion capture cameras (*b*) tracking the object, wrist, elbow, biceps on both arms, and the head (middle). (c) The object tracking configuration supports comfortable precision grip for user comfort and tracking efficiency.

human body. The manipulation is meant to occur within this volume. We utilize the *action field theory* [79] to determine the extent of motor control based on the proximity of the object to the user and place where the manipulation action takes place. This theory is of essence from the perspective of interaction amplitude, i.e., *coarse* and fine motor strategies adopted by the user to manipulate an object in 3D space based on their distance from the object [80].

3.1.3 Visuo-Motor Configuration. The average proximity of a user to the manipulated object is somewhere between 1.5 ft and 4 ft, which falls between the proxemic ranges of *intimate to peripersonal space*. It can be generalized as a space bounded by a user's *arm length* [81,18]. However, current MR interaction design methodology integrates actions that are either performed at the boundary of one's peripersonal space (arm's length) or the visuo-motor space is disconnected (Fig. 1), which contradicts the need for a co-located action–perception space for precise spatial actions. Prior research [82] also emphasizes on matching the physical and mental representations of the user's proprioceptive interaction space to performing spatial manipulation tasks.

3.2 System Design and Development.

We build on the aforementioned factors to design our experimental setup as follows:

3.2.1 Hardware Setup. The experimental hardware setup consists of 10 Optitrack Flex 13 motion capture cameras (field of view: 56 deg; refresh rate: 120 Hz) mounted on a custom gantry built using PVC pipes, measuring 7 ft × 4 ft × 8 ft in volume (Fig. 2(*a*)). The cameras were synchronized and operated through an Alienware 15R3 laptop computer with an Intel Core i7-7700HQ CPU (2.8GHz), 16GB of GDDR5 RAM, and an NVIDIA GeForce GTX 1060 graphics card having 6 GB video memory, running 64-bit Windows 10 Professional Operating System, which also ran our user evaluation interface. The application was mirrored on a portable monitor.

3.2.2 Setup Design. Our experimental setup (Fig. 2(*a*)) was designed to facilitate a visually and perceptually coherent interaction space in the user's peripersonal space to help enable precise spatial interactions in a MR environment. To this effect, we designed the setup to integrate a visuo-motor configuration having the following sequential arrangement: Eyes followed by screen followed by hands, thus maintaining an interaction proximity within the user's peripersonal space. Our MR-based configuration is specifically an augmented virtuality (AV) interface satisfying four out of six notions for MR, namely, *continuum, collaboration, combination*, and *alignment* [6]. The AV interface includes a virtual environment with 3*D* objects manipulated through physical user input behind the display. The interface was designed with the intention of (a) co-locating the virtual and physical (motor) peripersonal space for high-precision tasks and (b) reducing occlusion caused by placing the user's hands in front of the screen [23,83,84] as observed for traditional mixed-reality systems.

3.2.3 Interaction Space. Inspired by the *action-field theory* [79], our interaction space is located within the range of 45–60 cm (1.5–2 ft) from the torso (the range signifies anthropometric variations across different users). As for the volume (i.e., the physical limits), we followed an iterative approach starting from a standard table-top dimensions that led to us constructing the motion capture camera mounting frame. The idea was to determine a reasonably small to medium working volume (focused on the upper limbs and the peri-hand space), suitable for precise manipulations while maintaining robust tracking.

3.2.4 Tracking Methodology. Our MR-based user evaluation interface is based on Unity3D in tandem with the OPTITRACK MOTIVE API for streaming motion capture data to the former. We tracked user input through reflective markers mounted on custom designed and 3D-printed marker configurations attached to a proxy stem (Fig. 2(*c*)). The re-designed motion capture setup ensured good tracking coverage and minimized any blind spots due to the compact nature of the interaction space. We were able to track very small objects (0.5–2.5 cm in diameter and up to 4.5 cm in length) with a mean tracking error of 0.035 mm as per the API.

4 Experiment Design

4.1 Overview.

To explore the effect of action–perception blending for performing precise spatial tasks, we conducted the following experiments.

4.1.1 Experiment 1 (Real). Our aim for the first experiment was to observe, understand, and analyze the motor strategies followed by users to perform precise spatial manipulation tasks in their peripersonal space. The users were asked to manipulate a set of with 3D-printed shapes (Fig. 2(*c*)) akin to any other object manipulation task in the *physical world* with a natural and directly perceptible co-located visuo-motor space. This experiment serves as our *real-world* reference for evaluating the effect of action–perception blending for precise virtual task performed close to the user's body.

4.1.2 Experiment 2 (Representative). In this experiment, the precise spatial tasks are identical to the ones in experiment 1 with the exception of using the 3D-printed shapes as proxies to manipulate their virtual counterparts using an MR-based spatial interface. The motion capture system (Sec. 3.2) tracks user head movement and the 3D-printed objects, which are reflected in the manipulation of the virtual scene. The visuo-motor configuration (Sec. 3) for the virtual interface is designed with the intent of maintaining an interaction experience similar to the physical world (if the screen was absent). Our interface for experiment 2 serves as a prototypical "representation" of an MR interface facilitating a co-located action–perception space with appropriate visual and tactile feedback.

4.1.3 Experiment 3 (Generic). This experiment is identical to experiment 2 from the point of performing precise virtual tasks using a MR interface. However, the feedback in this interface is purely visual i.e. the users user *generic* proxies to manipulate the virtual shapes without any tactile feedback.

4.2 Task Selection.

In this work, we chose the classic *peg-in-the-hole* assembly task, which is a typical task to study spatial manipulation [85] and psycho-physical evaluation in virtual environments [86,87]. Further, this task involves enough visuo-motor complexity for our specific interest in the user's ability to match object location and orientation with high precision. In addition, HCI literature has shown studies using peg-and-hole assembly tasks for evaluating spatial user performance in terms of precision and accuracy [59,88,89]. Based on peg-in-the-hole task, the haptic feedback for Real and Representative (Repr.) experiments is facilitated through the interaction between the 3D-printed peg and hole objects during the insertion phase, completing the object assembly.

4.3 Task Variables.

While peg-and-hole assembly tasks are moderately complex, we added further variables to our experiment design specifically in terms of shape geometry to test the effect of action–perception blending on precise virtual tasks across diverse conditions.

4.3.1 Shape Complexity. We took cues from day-to-day objects that integrate the broader peg-and-hole insertion action and designed new shapes (Fig. 3) based on (a) number of peg–hole pairs and (b) rotational asymmetry. Keeping the *cylinder* as

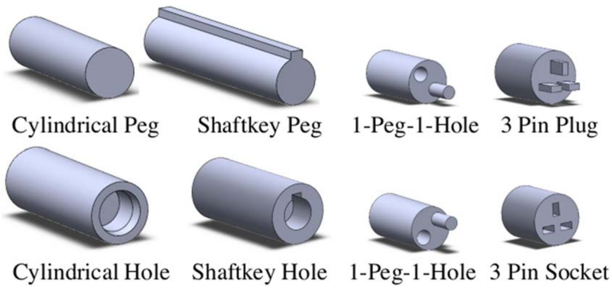

Fig. 3 (Left to Right) Peg-in-the-hole user evaluation shapes in the increasing order of insertion difficulty, i.e., single to multiple peg–hole configuration

our reference (single peg–hole and rotational symmetry), we first designed the *shaftkey* that had a rotational constraint due to the key-way. Shaftkeys are typically found in electric motor assemblies and mechanical contraptions. Further, the alternate peg–hole pair in *1-peg-1-hole* (found in automotive trickle chargers) added positional (two pegs) as well as rotational complexity to the insertion task. Finally, the *3-pin socket* was designed for the insertion task using three peg–hole pairs having different orientations. The new peg–hole pairs were also designed to have varying clearance fits: (a) *cylinder* (peg ∅: 10 mm; hole ∅: 10.6 mm, chamfer: 3 × 0.5 mm (tight), 3 × 1 mm (medium), 3 × 1.5 mm (loose)), (b) *shaftkey* (peg ∅: 10 mm; hole ∅: 10.6 mm (tight), 11 mm (medium), 11.4 mm (loose)), (c) *1-peg-1-hole* (peg ∅: 5 mm; hole ∅: 5.6 mm (tight), 6 mm (medium), 6.4 mm (loose)), and finally (d) *3-Pin Socket* (rectangular peg: 3 × 6 mm, 6 × 3 mm, 6 × 3 mm; rectangular hole: 3.4 × 6.4 mm, 6.4 × 3.4 mm, 6.4 × 3.4 mm (tight); 3.5 × 6.5 mm, 6.5 × 3.5 mm, 6.5 × 3.5 mm (medium); 3.6 × 6.6 mm, 6.6 × 3.6 mm, 6.6 × 3.6 mm (loose)).

4.3.2 Visual Perception. We implemented head tracking (Fig. 2(*b*)) for our MR interface that was mapped to the virtual scene camera to help users explore the virtual scene and virtual shapes as they would do in the physical world (in absence of the screen). We iteratively configured the virtual cameras (FoV: 33.4 deg) to appropriately scale the visual rendering of the physical object being tracked, based on its distance from the user. We verified the same experimentally by placing the display mid-way between the user and the physical object. We physically measured the virtual object rendering and found it to be half the size of its physical counter part. We repeated this for multiple distances and the size perception remained consistent based on the law of similar triangles for the view frustum emanating from the camera.

5 Experiment

We conduct a set of controlled lab experiments (Sec. 4.1) to test the effect of our proposed experimental setup on user performance for bi-manual peg-and-hole object assembly tasks.

5.1 Participants.

We recruited 39 participants (18–30 years old) who were graduate and undergraduate students from engineering, architecture, and visualization majors. This was a between-subjects experiment where participants were equally distributed across the three experiments (13 participants per group)—Real, Repr., and Generic. To minimize any bias in the experimental data, we also verified and obtained written confirmation through a pretest questionnaire from each participant regarding any injuries on the upper half of their body that may affect user performance.

5.2 Procedure

5.2.1 Calibration. The motion capture camera system was calibrated every two to three participants (3–4 h) to ensure robust tracking.

5.2.2 Tracker Assignment. We begin by assigning individual tracking to each of the peg-and-hole object pairs. We also wanted to gain a deeper perspective into how user hand kinematics play a role in precise spatial tasks. Therefore, each participant wore tracking markers at their wrist, elbow, and biceps for both arms (Fig. 2(*b*)). To integrate both object and body tracking, each of the joints were tracked as rigid bodies and care was taken to avoid any tracking confusion by ensuring that no two tracked entities have the same marker configuration. The users also wore a head band, which was tracked to manipulate the application scene. Due to the highly sensitive motion tracking, we applied exponential smoothing to the object and head tracking for a *jitter-free* interaction experience. After assigning the trackers, participants were given a brief walk-through of the experimental setup. They were allowed to adjust their seating as well as the placement of the display screen as per their comfort. We found that every user placed the screen close to their torso (within their respective peripersonal space).

5.2.3 Practice. From the point of standardization, the participants were asked to manipulate the shapes (virtual and physical) through a stem (Fig. 2(*c*)) on which the 3D-printed peg–hole pairs were attached on one end. In addition for Repr. and Generic experimental groups, participants were also informed of the head tracking and encouraged to use it for exploring the virtual scene and objects during the peg-and-hole assembly task. Following this, the participants practiced with the physical or virtual shapes based on their experimental group to get acquainted with the setup before starting with the user study trials.

5.2.4 Study Task. Each participant was asked to insert the peg into the hole (Fig. 4) across different geometries and fits (Fig. 3). We varied the study shapes during each study session and trial sequence in increasing order of shape complexity—cylinder, shaftkey, 1-peg-1-hole, and 3-pin socket. The fits were randomized for each shape. To control the trial data for the insertion task only, participants started from a default position (shapes in vertical orientation) after an explicit indication from the study coordinator and the task completion was indicated by the participants to stop recording the trial data.

While there was no time limit for assembling the peg into the hole, participants were able to complete each trial in 5–10 s except for Repr. control group where it took relatively more time to complete the task. We analyze and discuss this user behavior in Sec. 6.

5.3 Data Collection and Metrics.

Each user session per experimental group took 45–60 min for completion. For a given experiment, we recorded user data for six trials per shape per fit for every participant, and each participant performed 72 trials across all 12 shapes (four shapes and three fits per shape). For each user trial, we recorded the (a) motion trajectory (position and orientation of object coordinate frame) for peg and hole objects followed by each joint on both arms, and the head (b) task completion time. We also collected self-reported user feedback using the NASA-TLX metrics [90]. To conduct a deeper analysis of user behavior for precise spatial task, we further

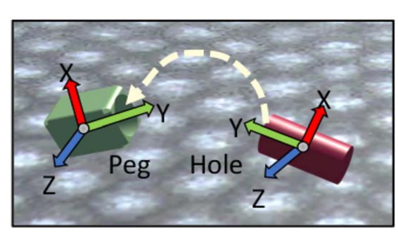

Fig. 4 Virtual scene for the peg-in-the-hole assembly task

processed the trajectory data to compute the following user evaluation metrics.

5.3.1 Cumulative Kinetic Energy. Existing metrics for evaluating precise actions are found in medical surgery [91–93] with the assumption of human motion being linear and in the Euclidean space. However, hand motion is inherently nonlinear, and there is a need for metrics that connect motor strategies taken by the users to the quantitative measure. Therefore, we chose to compute *kinetic energy* between successive motion frames (position and orientation) to spatially move the objects from its initial to final insertion phase for peg-and-hole assembly, computed as follows:

$$K.E. = \frac{1}{2}\Phi^T M \Phi \tag{1}$$

where Φ is the twist vector $\in \mathbb{R}^6$ of the object coordinate frame and M is the inertia matrix of the object being manipulated $\in \mathbb{R}^{6X6}$. To get a holistic view of the total energy expended for a peg–hole task, we compute the *cumulative kinetic energy*.

5.3.2 Change Detection. On analyzing the cumulative K.E. versus time plots (Fig. 5), we observed transition phases from higher cumulative K.E. to a lower energetic phase and vice versa for the time taken during the insertion process. We found this to be consistent across all shapes and their fits across the three control groups. We quantify these energetic transitions using the *change detection* approach [94,95]. We compute this using MATLAB's *findchangepts* function. However, the function alone was not sufficient as it often detected additional points close to each other where energetic transitions may or may not have occurred. Therefore, we further processed these detected points and *pruned* them by putting a maximum threshold (N) to the number of points detected as an input to the MATLAB function followed by a vector analysis between consecutive points with a window size of 3 ($\mathbf{p_{i-1}}$, $\mathbf{p_i}$, $\mathbf{p_{i+1}}$) to further isolate the energetic transition points as follows:

$$|(\hat{\mathbf{v}}_1 \cdot \hat{\mathbf{v}}_2) - 1| \geq \delta \tag{2}$$

Here, $\hat{\mathbf{v}}_1$ and $\hat{\mathbf{v}}_2$ are two consecutive vectors connecting points $\mathbf{p_{i-1}}$ and $\mathbf{p_i}$, and $\mathbf{p_i}$ and $\mathbf{p_{i+1}}$; δ is the threshold parameter. We

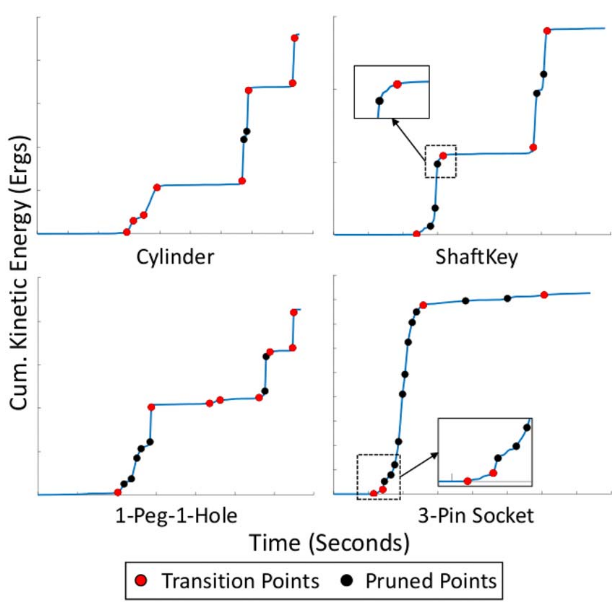

Fig. 5 Illustration of the change detection algorithm to identify high- to low-energy phases and vice versa (based on large shifts in the tangent), for the cumulative kinetic energy versus time plots for real experimental group across the tightest fit of all shape geometries

iteratively identified the following values of N and delta for each shape geometry across experiment groups: (a) *cylinder* and *shaftkey* (N: 10 (Real, Repr., Generic); δ: 0.01 (Real), 0.05 (Repr.), 0.005 (Generic)), (b) *1-peg-1-hole* and *3-Pin Socket* (N: 15 (Real, Repr., Generic); δ: 0.05 (Real), 0.0005 (Repr., Generic)). We observe that the parameters are same for all fits of a given shape and consistent for shapes having *single* peg-and-hole and *multiple* peg–hole pairs across all experimental groups.

6 Results

We recorded 936 user trials across the four shapes for a given experimental group and overall 2808 user trials were recorded for all three experiments (Sec. 4.1). In the subsections below, we perform a qualitative evaluation of user performance and behavior for the peg-in-the-hole task and further support our observations with quantitative analysis.

6.1 Observational Analysis of Motor Strategies. We begin with an in-depth video analysis of user study sessions across the three experiments with the aim of (a) observing and identifying key motor strategy categories for performing the experimental tasks and (b) generating specific hypothesis for quantitative analyses.

6.1.1 Methodology

- *Video Recording and Verification:* Before conducting actual user trials, the video camera was arranged (position, orientation, zoom, and focus) in a manner such that it recorded the upper body movements of the participants both in front as well as behind the display screen (Fig. 2(*b*)). We further reviewed the recordings collected from pilot experiments for video clarity and our ability to identify distinctive user actions at slower (0.25x, 0.5x) playback speeds.
- *Coder Recruitment:* Each of the three experiments was assigned to one study coordinator involved with the user studies such that the individual analyzing videos for a given experimental group had minimal to no involvement in conducting the experiment. This is to ensure minimize any biases while analyzing user behavior.
- *Coding Scheme:* The three coders performed a preliminary analysis on 6 of 13 user sessions videos per group and came up with individual coding schemes based on their observations of user behavior for a given experimental group. Following this, the three study coordinators convened to discuss their schemes, and to our surprise, the codes were very similar that made it relatively to finalize on a common scheme to be used for further video analysis across all experimental groups. The codes included but not limited to *palmDown, rotatingPeg, seeHole, etc.,* highlighting the palm facing toward the table while holding the peg–hole object pairs followed by the peg being rotated to align with hole and glancing the hole to ensure proper alignment before insertion, respectively.
- *Intercoder Reliability:* Once the coders completed analyzing user videos for their respective experiment group, they re-convened to cross-verify each other's analysis as well as gain any new perspective that might have been overlooked by any of the three coders. Here, each coder took a pass at four videos (about 25% of sample size per group) from another group with the purpose of minimizing any oversight, bias, or unintentional errors in prior analysis. Following this, the coders took a final pass at the videos and further refined their prior analysis of identifying the key motor strategies discussed further.

6.1.2 Analysis of User Behavior. Drawing from the video analysis, we organized it into three broader categories.

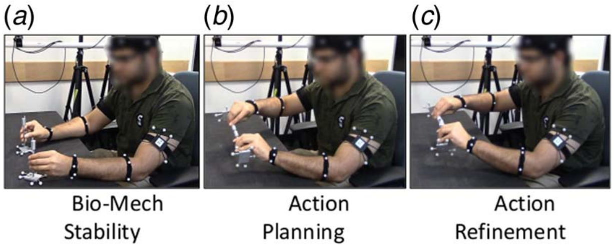

Fig. 6 Screen capture of user study session from the real experimental group showcasing the key phases of user motor strategies for peg-and-hole assembly tasks in the user's peripersonal space: (*a*) bio-mech stability, (*b*) action planning, and (c) action refinement

- *Bio-Mechanical Stability:* One of the key motor strategies followed commonly by users across all experimental groups was to stabilize their upper body, especially the limbs (Fig. 6(*a*)). A key observation was made by the coders about all users resting their elbows on the table and close to their body before proceeding with performing the peg and hole assembly task. In few cases, the participants stabilized their elbow on the non-dominant limb where they held the hole object,*** whereas the dominant limb was suspended in mid-air with a prehensile pose showing the user's *intent to perform an action* after ensuring bio-mechanical stability of the other limb. Bio-mechanical stability was observed to be a crucial and one of the first movements executed by all participants to ensure stable grounding of their shoulders and elbows before performing fine wrist and finger level movements to properly align and assemble the peg-and-hole objects.
- *Action-Planning:* The planning phase often overlapped with the stabilization phase and we identified some key strategies (Fig. 6(*b*)) observed for users across the three experiments.
  — *Grasp:* Users typically grasped the 3D-printed peg and hole objects using three to four finger grip which is indicative of precise movements [96,97]. Once stabilized, the grasp allowed the users to make fine motor movements such as rotating (finger level precision) and adjusting the peg to align (wrist level precision) with the hole in most scenarios before insertion. This was observed for all shapes except the cylinder due to rotational asymmetry in their geometry.
  — *Peek:* While grasp helped align the shapes, users were often observed to glance into the peg and hole after alignment for a follow-up confirmation of their prior alignment motor strategy. In this scenario, the users turned the peg and hole face towards them as well as aligned their head for an appropriate and non-skewed visual perception. Following this, the users initiated the insertion action phase which came with its own challenges as the geometric complexity increased with the study shapes.
- *Action Refinement:* Unlike cylinder that had rotational symmetry, the remaining three shapes were found to be challenging by the users to successfully insert the peg into the hole in their first attempt. Rotational asymmetry often resulted in the peg–hole pairs intersection due to improper assembly, which forced the users to retract the shapes. Further, the participants *re-strategized* their actions by adjusting the *grasp* (rotational alignment) and re-confirming it with by a brief *peeking* until the two objects were fully inserted. This was a crucial phase as the participants refined their fine motor actions only and no adjustments were made to coarse actions that were performed for bio-mechanical stability. For the Repr. experimental group, participants were found to struggle relatively more than the other two groups, and this was due to a visuo-tactile disparity where the participants often found their proprioceptive perception not matching the visual feedback, thereby leading to relatively larger task completion times and user frustration.

We further confirmed this our observation analyses by reviewing the velocity profiles (Fig. 7) and found maximum activity occurring at the object level. Based on this, we analyze user performance for the manipulations performed at peg level in the subsections below.

6.2 Quantitative Analysis of Motor Strategies. We found the cumulative K.E., change detection, and completion time to be nonnormally distributed using Kolmogorov–Smirnov test. Further, we analyzed the performance metrics for a given shape geometry across the three control groups using a nonparametric Kruskal–Wallis test. On the basis of our observational analyses, we make the following hypotheses:

H1 Representative (Fig. 1) should enable similar motor strategies (cumulative K.E and change detection) with respect to Real as they both afford kinesthetic feedback. Hence, there should be no statistical significance.
H2 Generic (Fig. 1) should exhibit higher consumption of energy and motor strategy (cumulative K.E. and change detection) with respect to Real due to the absence of kinesthetic feedback in the former. Here, we should observe statistical significance.
H3 Representative should enable same performance (completion time and docking accuracy) with respect to Real. We should not observe statistical significance.
H4 Generic should have less accuracy and take relatively more time to complete the docking with respect to Real. This should lead to statistical significance for both completion time and accuracy.
H5 Generic should exhibit higher energy consumption for both view manipulation (cumulative K.E. for the head) and peg with respect to Real. We should observe statistical significance for number of change points detected and cumulative kinetic energy for both the head and the peg.

6.3 Energy Consumption. We observe an overall statistical significance (Fig. 8) for each shape across all control groups (cylinder: $p = 6.4 \times 10^{-9}$; shaftkey: $p = 4.91 \times 10^{-4}$; 1-peg-1-hole: $p =$

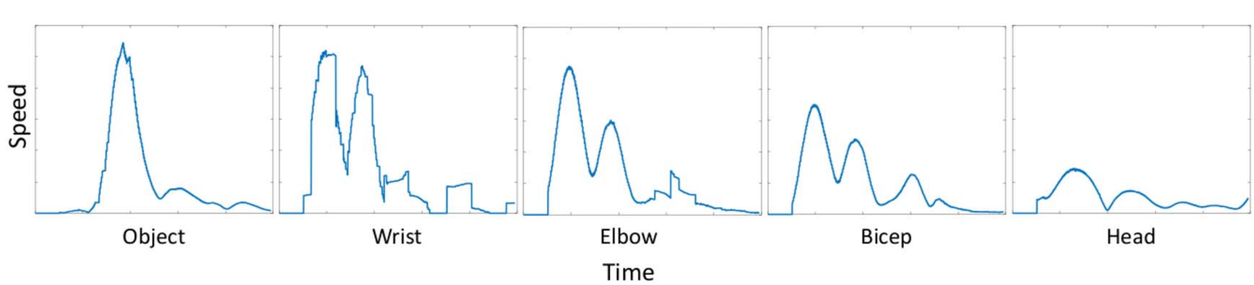

Fig. 7 Speed profiles at different joints of the upper limbs and head denoting the increased activity below the elbow post bio-mechanical stability to perform fine motor actions at the wrist and finger (object) level

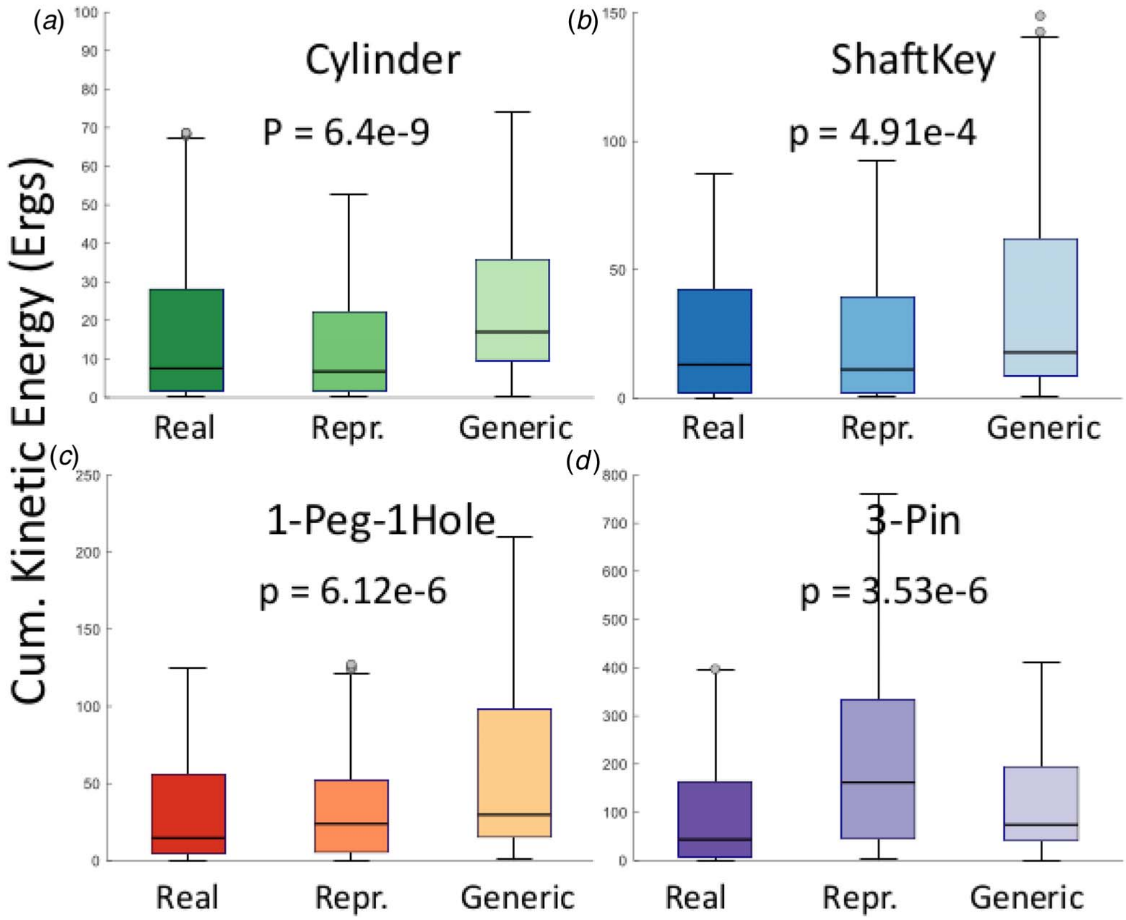

Fig. 8 Statistical analysis for the cumulative kinetic energy computed for the peg for all shape geometry across all control groups. We observe statistical significance for all shapes with higher mean energy consumption for the generic except for the 3-pin socket assembly.

6.12×10^{-6}; 3-pin: $p=3.53\times10^{-6}$). On further post hoc analysis using Tukey's multiple comparison test, we found the mean kinetic energies to be similar for Real and Generic across shaftkey and 1-peg-1-hole having p values as 0.8 and 0.2, respectively. This was also observed during the video analysis as users in Repr. group struggled to assemble these shapes due to vision–perception disparity. The mean energy consumption was found to be consistently higher for Generic across all shapes except the 3-pin socket. This is attributed to the user familiarity with the three pin sockets as a daily object and their ability to assemble it solely based on proprioception (overcoming the disparity in Repr.). The aforementioned analyses confirm our hypothesis *H1* only for the shaftkey and 1-peg-1-hole shapes and is rejected for the remaining shapes. However, our hypothesis *H2* holds through for all shapes with higher mean cumulative K.E. for Generic than Real. The anomalous user behavior for the 3-pin socket is further corroborated in NASA-TLX [98] with consistently higher mental (13.9), physical (13.5), temporal effort (14.2), and frustration (13.1) as well as, lower overall performance (7.8) from user-reported scores.

6.4 Energetic Transitions. We observed a change detection trend similar to the cumulative kinetic energy (Sec. 6.3) across the *single* (cylinder and shaftkey) and *multiple* (1-peg-1-hole and 3-pin socket) peg–hole pairs (Fig. 9). For cylinder and shaftkey, the mean energetic transitions were found to be higher for Real and Repr. than Generic, which shows that tactile perception led to additional effort from the users as we observed in *action–refinement* motor strategy followed to ensure proper assembly of the peg–hole object pairs. This confirms the hypothesis *H2* but with higher mean transitions for Real (cylinder: $p=9.56\times10^{-10}$; shaftkey: $p=9.75\times10^{-10}$), also, rejects *H1* with higher mean transition points for Repr. with respect to real due to the vision–proprioception disparity (cylinder: $p=9.56\times10^{-10}$; shaftkey: $p=9.56\times10^{-10}$; 1-peg-1-hole: $p=9.56\times10^{-10}$; 3-pin: $p=9.56\times10^{-10}$). In case of 1-peg-1-hole and 3-pin socket, Real had the lowest transitions, which shows that participants re-strategized their docking at a higher frequency for virtual manipulation. Thus, rejecting the hypothesis *H1* (1-peg-1-hole: $p=9.56\times10^{-10}$; 3-pin: $p=9.56\times10^{-10}$), but confirming hypothesis *H2*.

6.5 Completion Time and Accuracy. Similar to change detection, we observe similar completion time trends for *single* and *multiple* peg–hole pairs (Fig. 10). For cylinder and shaftkey, the mean task completion time increases from Real to Generic, thus confirming the hypothesis *H4* (cylinder: $p=9.56\times10^{-10}$; shaftkey: $p=9.75\times10^{-10}$). We also observe a higher mean task completion time for Repr. with respect to Real, rejecting *H3*. In case of 1-peg-1-hole and 3-pin socket, the mean task completion time for Repr. (σ : 11.71 and 28.2 s, respectively) is higher than Real (5.15 and 4.26 s) and Generic (σ : 11.51 and 11.06 s), which is surprising and contradictory to *H3*. This is also observed for NASA-TLX scores on a 21-point scale across 1-peg-1-hole and3-pin socket for Repr. having higher mental (10.3 and 13.9, respectively), physical (10.2 and 13.5, respectively), temporal (11 and 14.3, respectively), and overall effort (10.7 and 7.8, respectively) along with lower performance (12.2 and 14.8, respectively) and higher frustration (10.8 and 13.1, respectively) scores. We also observed the same during the studies as the participant struggle

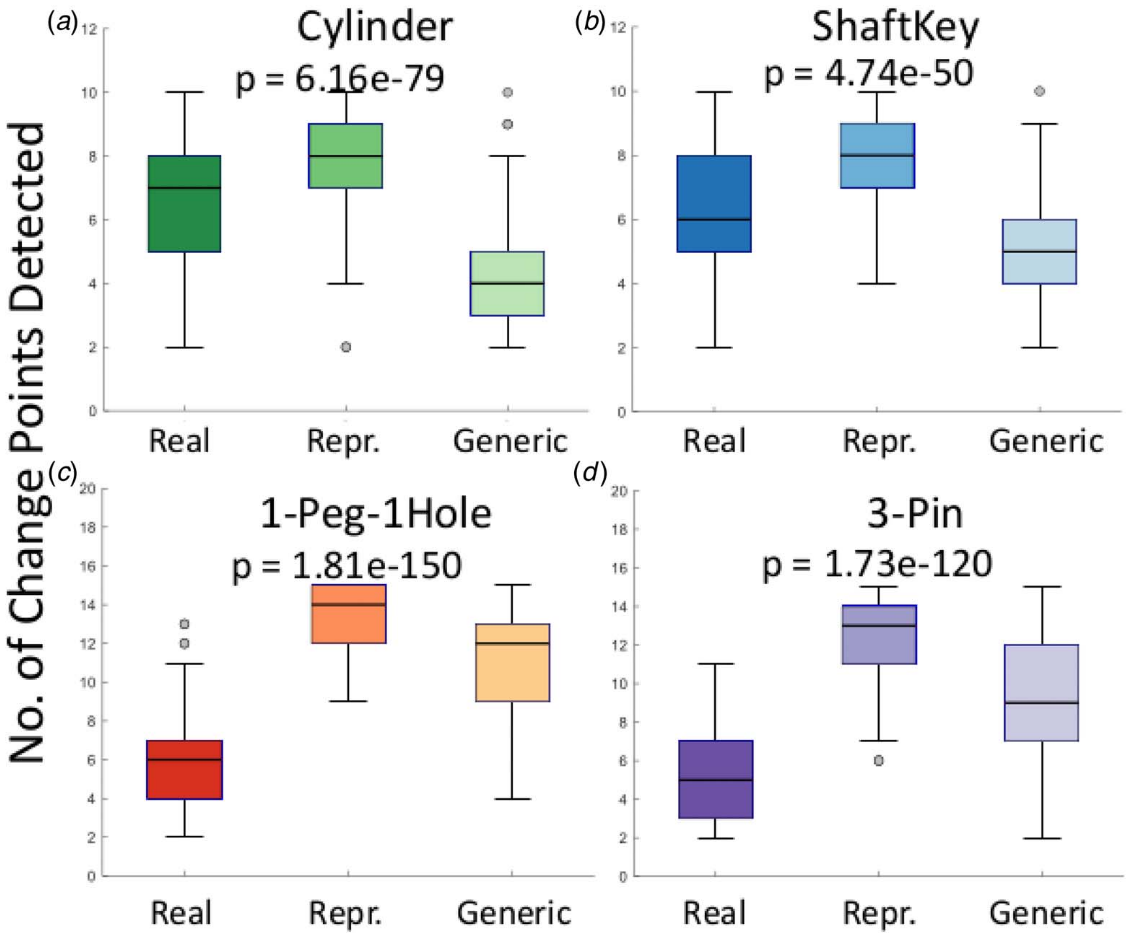

Fig. 9 Statistical analysis for the change detection computed for the peg for all shape geometries across all control groups. We observe statistical significance for all shapes with higher mean transitions for Repr. group.

increased with shape difficulty. The mean completion time for Generic was observed to be higher than Real (cylinder: $p = 9.56 \times 10^{-10}$; shaftkey: $p = 9.75 \times 10^{-10}$; 1-peg-1-hole: $p = 9.56 \times 10^{-10}$; 3-pin: $p = 2.23 \times 10^{-8}$). Thus, confirming the hypothesis *H4*.

While each participant was allowed to assemble the peg and hole, the accuracy should have been same for all shapes across control groups with kinesthetic feedback and different for Generic. As expected, we observed similar mean accuracy values for Real and Repr. confirming hypothesis *H3* ($p = 0.08$) and higher mean accuracy for Generic confirming hypothesis *H4* ($p < 0.05$). This is an artifact of the Real and Repr. setup itself, that the error is simply the inaccuracy error of tracking. While for Generic, the error is purely caused by visual cues.

6.6 Vision–Proprioception Blending. What really surprised us (as evident from our initial hypotheses) is the significant time disparity between Repr. and Generic for 1-peg-1-hole and 3-pin socket cases (Sec. 6.5). This is even after considering the fact that the mean docking accuracy was found lower for Generic. There are three potential factors for this that we considered, namely: (1) size disparity (difference between perceived dimensions of parts visually and through touch), (2) tracking disparity (angular and positional errors caused during multi-object and arm tracking), and (3) visual-tactile preference. Here, the third factor is particularly interesting.

6.6.1 Size Disparity. Since our visual feedback was in a completely virtual space (rather than a typical mixed-reality setup), there was always a possibility that the sizes of the objects rendered could be different from those being held and manipulated. This could lead to a perceptual mismatch between what the users were seeing and what they were feeling in their hands.

6.6.2 Tracking Disparity. In general, all vision-based systems have tracking errors especially when multiple objects are being tracked (in our case: two hand-held objects, six joints on the arms, and the head). In addition, in our case, noise was inevitably introduced due to reflection from the participants' apparel and hair.

6.6.3 Visual-Tactile Preference. Given the choice between visual and haptic perception, humans tend to use the former for coarse manipulations, but rely on kinesthetic perception for precise docking tasks [10]. This can be a potential cause for the additional time that we observe in Repr. users.

Each of these factors contribute to the quality of the blend between the action space (shoulder–elbow–wrist–hand–object) and the perception space (visual and proprioceptive). We made significant efforts to resolve the size disparity out the very outset of our experimental setup design (Sec. 4). Therefore, it is likely that the increased effort in Repr. users is a combination of the tracking disparity and the visual-tactile preference as also shown by a consistently wider spread and higher mean scores in the NASA-TLX. The tracking error for the Real and Repr. amounts to 3–5 deg and are identical because in both cases, the user is physically inserting a peg in a hole. In contrast, the main cue for the user to asses success of insertion was based on collision in the Generic interface, thereby accounting for higher errors (5 − −20 deg). Therefore, it is reasonable to assume that tracking disparities may have minor influences over user performance in the Representative control group.

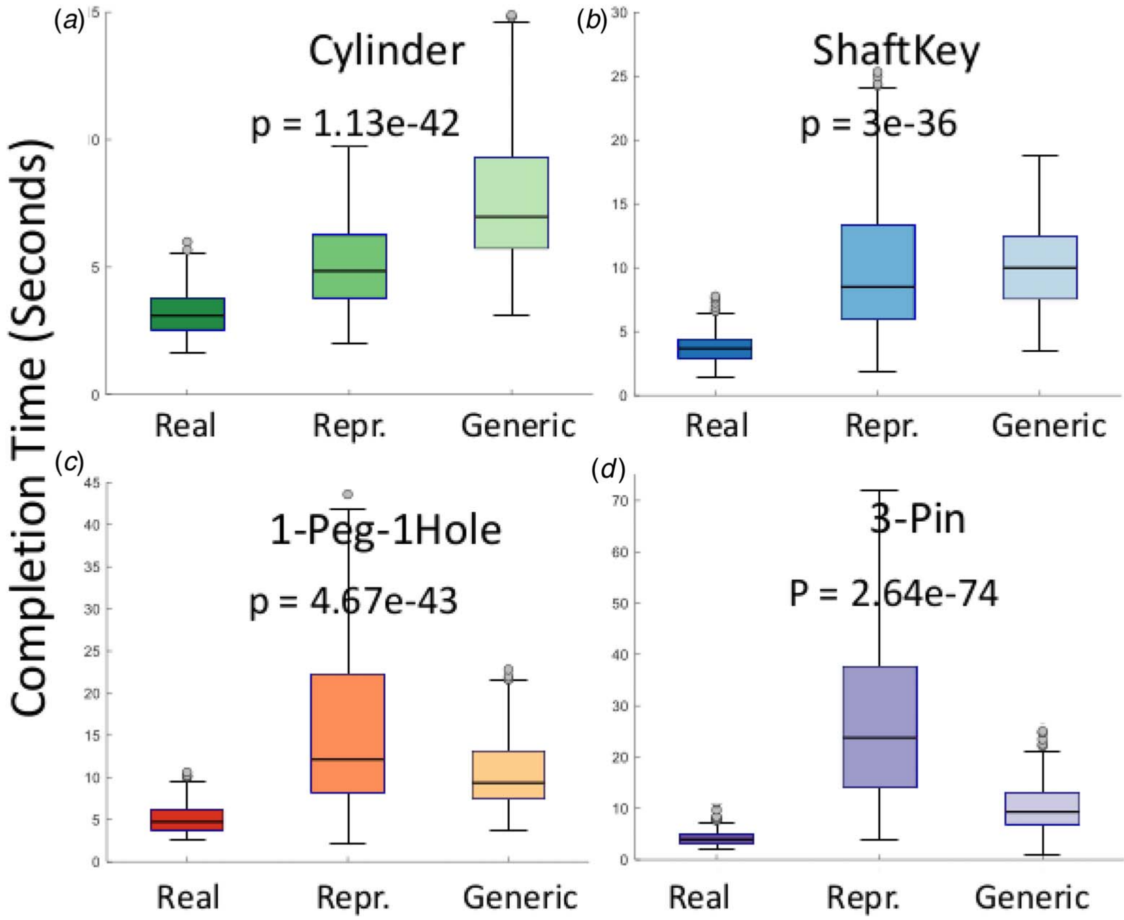

Fig. 10 Statistical analysis for the completion time computed for the peg for all shape geometry across all control groups. We observe statistical significance for all shapes with higher mean completion time for Generic control group in case of cylinder and shaftkey. However, for 1-peg-1-hole and 3-pin socket, the higher mean completion time is for the Repr. control group.

We believe that the visual-tactile preference may have played a bigger role. To verify this quantitatively, we conducted an analysis of user's head movement by computing the cumulative K.E. consumed with each control group (Fig. 11). We assumed a spherical representation of the head to compute the cumulative K.E. [99]). We observe that Real, Repr., and Generic are all statistically different from each other with mean cumulative K.E. ordered as Real > Generic > Repr. for shapes except the 3-pin socket where it is Real > Repr. > Generic. In effect, users moved very little while using the Repr. interface compared to Real and Generic. This suggests that proprioception kicked in as a preferred means for assessing the success of insertion closer to the end of the task. This was also observed in our video analysis wherein users relied on visual feedback until the initial shape alignment and relied more on their tactile perception of the peg–hole contact to insert. This is a critical factor that needs a detailed study in its own right.

Incidentally, this also may connect with the third potential cause, that is, the preference for perceptual modality as observed for higher energy consumed for the head for the Real group, which is purely physical interaction. Despite lower head movement, the completion times were higher for the virtual interfaces. From the point of motor strategies, we observed that most users performed the docking task by first *stabilizing* the nondominant arm (which controlled the hole), followed by grasping the peg and hole stems through a two to three finger precision grip with their palms facing down towards the table. Further, the users adjusted the hole into their field of view and then manipulated the peg to complete the docking.

7 Discussion

Our study revealed some important design and technological considerations for future interfaces that seek to enable high-precision tasks. Therefore, we draw a deeper understanding of our research questions (Sec. 1.2) by highlighting following three important considerations.

7.1 Motion Tracking in Peripersonal Space. While handheld controllers are fairly common and robust, other technological aspects need careful consideration. This is especially true since current AR/VR/XR headsets are primarily designed for interaction at a distance. Despite prior implementations of MR interfaces in the peripersonal space (Sec. 1.1), a major challenge that we faced was the lack of guidelines for implementing a close-range motion capture system to facilitate a co-located vision-proprioception space for precise spatial interactions. By using our setup, we encountered a tracking error of 3–5 deg with our interfaces for Real and Repr. experimental groups (Sec. 6.6), which affected user performance for the latter in terms of higher energetic transitions (Fig. 9) and resulting in higher task completion time (Fig. 10). Despite this, our setup performed relatively better than existing commercially available tracking devices. This draws a direct understanding of the quantitative extent to which technological factors affect precise control in spatial interactions (Q2), which until now has largely remained unexplored. Therefore, despite the available resources that allow tracking of user movement and spatial actions, there is not any hardware setup that solely focuses

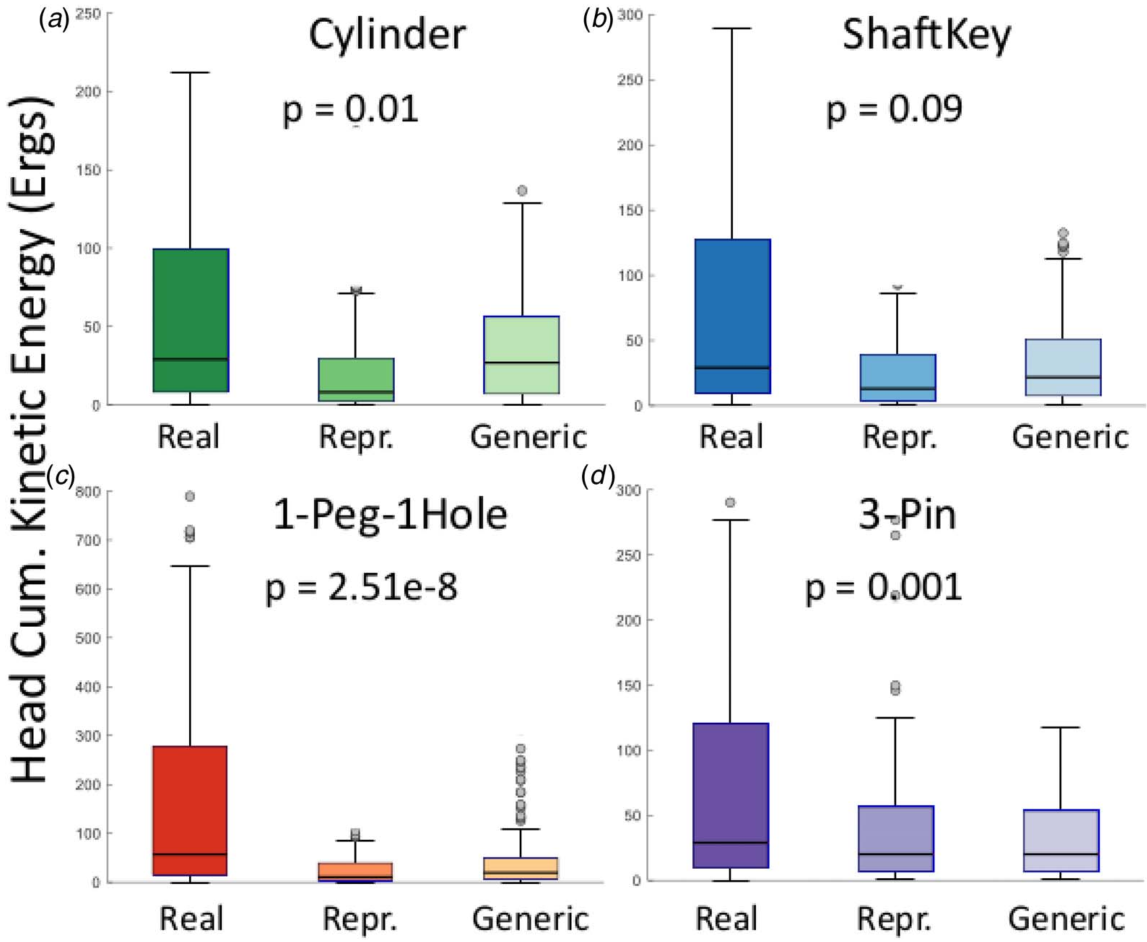

Fig. 11 Statistical analysis for the cumulative kinetic energy computed for the Head across all control groups. We observe statistical significance for all groups with higher mean cumulative K.E. for Real control group.

on tracking fine motor movements. Hence, our quantitative and qualitative assessments highlight the critical importance of precise tracking for an uninhibited and intuitive spatial manipulation experience for interaction close to the body. We see tremendous opportunity for improvement in commodity sensors to enable robust tracking at sub-millimeter scale.

7.2 Vision versus Proprioception in Precise Tasks. While our behind-the-display interaction approach did not support a see-through display, it facilitated a co-located visuo-motor space for most users combining the user's proprioceptive bi-manually coordinated actions. However, the most crucial challenge that we faced was for the Repr. interface, wherein the participants repeatedly encountered vision–proprioception disparity despite visual and tactile perception being provided for the peg-and-hole assembly task. The unexpected increase in completion time for the Repr. interface (Fig. 10) and higher energetic transitions (Fig. 9) is a critical example that points toward greater care in blending the visual and proprioceptive spaces. The action-specific visual perception [100,101] of manipulating such small objects will need co-locating the human motor zone with small-sized ultra-high resolution see-through displays. There is a promising research potential in leveraging 4K visual displays that mobile devices already possess for configuring AR and VR systems. Second, our study (Sec. 6.6) also suggests that instead of a physically realistic force and tactile feedback, what is really needed is a consistently mapped combination of visual and tactile cues that support each other seamlessly and get out of the user's way when needed. In summary, it is important that design of future interaction spaces reflect and facilitate the innate blending of action-proprioception space (Q3) for performing high-precision interactions in intimate spaces where the sense of touch overshadows visual feedback in many cases [102].

7.3 Energetics of Spatial Interactions. High-precision manipulation tasks often require a higher-degree of hand–eye coordination whose effect can be seen in the motion trajectory of the action performed. From a kinematics standpoint, the movement and therefore manipulative ability enabled by the shoulder-arm-wrist complex is primarily responsible for *ballistic* actions [103] in the outer regions of the peripersonal space leading up to the extra-personal space (beyond arm's length). This is the space in which current spatial interactions are typically designed and studied. In contrast, finger-level manipulations can afford to perform highly precise activities (such as hand writing and use sharp and fine tools) because of the highly redundant degrees-of-freedom [104]. A direct consequence of these kinematic differences is that the types of metrics (completion time and accuracy) typically used in spatial manipulation literature are not nearly enough to capture fine-grained spatial manipulations. We propose that body kinematics plays a major role in characterizing task precision as opposed to Euclidean metrics such as path efficiency and deviation proposed in related areas such as minimally invasive surgery [91–93]. In our prior attempts, none of these metrics were able to clearly elucidate the precise tasks studied in this article, i.e., they evaluated poorer user performance for Real experimental group followed by Repr. and finally, Generic that is counterintuitive to our assumption of Real being the gold standard for the remaining two experiment groups. To that effect, our proposed kinetic energy metric was able to clearly show differences across the three interfaces studied in this article as well facilitate with logical insights that could be corroborated with the motor strategies extracted from the video analysis (Sec. 6.1). In this article, we build upon existing knowledge on measuring precise actions and take initiative in proposing new metrics that focus on "how the action is performed" than focusing on the end outcomes of a given precise action. We believe that this will encourage future research on exploring metrics that appropriately precise spatial actions in the user's peripersonal space (Q1). In fact, there is a rich set of research questions yet to be asked regarding the synchronization of head, arm, and finger level movements in precise spatial tasks.

8 Conclusion

The goal of this article was to study precision in spatial manipulation in the peripersonal space and understand motor strategies in this space. To this end, we developed three variants (real, representative, and generic) of a tablet-based virtual manipulation system with camera-based motion tracking of hand-held objects. The data collected through the peg–hole task using these variants shed light on the common motion strategies used while performing precise manipulation—*bio-mechanical stability, object grasp, exploring the scene and objects through brief glances,* and *refining actions to ensure proper and intersection free alignment* of peg–hole pairs. Moreover, we find that while the gross strategies remained the same across the three variants, the main change was observed in terms of finer finger-level manipulation, the cognitive load, the time taken, and the accuracy. The main underlying principle at play here is the quality of blending between the visual and tactile feedback in precise tasks and need a deeper investigation. Another critical issue that was revealed in our study was the need for defining metrics that align with the unique way in which our hands and fingers work. Here, our energy-based metrics seem to capture important information regarding the task. However, further research should be done to investigate other metrics, perhaps with the help of extremely precise tracking of hands, fingers, objects, and head. Overall, not only do our findings align with the current psychology and motor behavior of spatial manipulation, but they also pave way for research in the design, development, and evaluation of future spatial interfaces for fine-grained tasks, especially for 3D modeling and design applications. We believe that the learning outcomes of our study will motivate the development of better technology, principled approaches for interaction design, as well as new avenues for exploring high-precision spatial manipulation.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant #2008800 and partially supported by AggieE_Challenge program, Texas A&M University Triads for Transformation (T3) program, and Orthopaedic Research and Education Foundation (Award #0014566A). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

References

[1] Milgram, P., and Kishino, F., 1994, "A Taxonomy of Mixed Reality Visual Displays," IEICE Trans. Inf. Syst., **77**(12), pp. 1321–1329.

[2] Milgram, P., and Colquhoun, H., 1999, "A Taxonomy of Real and Virtual World Display Integration," Mixed Reality: Merging Real Virtual Worlds, **1**(1999), pp. 1–26.

[3] Billinghurst, M., 2020, "Rapid Prototyping for AR/VR Experiences," Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, CHI EA '20, Honolulu, HI, Apr. 25–30, Association for Computing Machinery, pp. 1–4.

[4] Rokhsaritalemi, S., Sadeghi-Niaraki, A., and Choi, S.-M., 2020, "A Review on Mixed Reality: Current Trends, Challenges and Prospects," Appl. Sci., **10**(2), p. 636.

[5] Çltekin, A., Lochhead, I., Madden, M., Christophe, S., Devaux, A., Pettit, C., Lock, O., Shukla, S., Herman, L., Stachoň, Z., Kubíček, P., Snopková, D., Bernardes, S., and Hedley, N., 2020, "Extended Reality in Spatial Sciences: A Review of Research Challenges and Future Directions," ISPRS Int. J. Geo-Inf., **9**(7), p. 439.

[6] Speicher, M., Hall, B. D., and Nebeling, M., 2019, "What is Mixed Reality?" Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, May 4–9, Association for Computing Machinery, pp. 1–15.

[7] Abidi, M. H., Ahmad, A., Darmoul, S., and Al-Ahmari, A. M., 2015, "Haptics Assisted Virtual Assembly," IFAC-PapersOnLine, **48**(3), pp. 100–105.

[8] Abidi, M. H., Al-Ahmari, A., Ahmad, A., Ameen, W., and Alkhalefah, H., 2019, "Assessment of Virtual Reality-Based Manufacturing Assembly Training System," Int. J. Adv. Manuf. Technol., **105**(9), pp. 3743–3759.

[9] Das, A. N., and Popa, D. O., 2011, "Precision Evaluation of Mmodular Multiscale Robots for Peg-In-Hole Microassembly Tasks," 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, CA, Sept. 25–30, pp. 1699–1704.

[10] Park, H., Bae, J.-H., Park, J.-H., Baeg, M.-H., and Park, J., 2013, "Intuitive Peg-In-Hole Assembly Strategy With a Compliant Manipulator," 44th International Symposium on Robotics (ISR), Kintex, Seoul, South Korea, Oct. 24–26, pp. 1–5.

[11] Marvel, J. A., Bostelman, R., and Falco, J., 2018, "Multi-Robot Assembly Strategies and Metrics," International Conference on Intelligent Robots and Systems, Vancouver, British Columbia, Canada, Sept. 24–28.

[12] Li, X., Li, R., Qiao, H., Ma, C., and Li, L., 2017, "Human-Inspired Compliant Strategy for Peg-in-Hole Assembly Using Environmental Constraint and Coarse Force Information," 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4743–4748.

[13] Bowman, D. A., Kruijff, E., LaViola, J. J., and Poupyrev, I., 2004, *3D User Interfaces: Theory and Practice*, Addison Wesley Longman Publishing Co., Inc., Redwood City, CA.

[14] Argelaguet, F., and Andujar, C., 2013, "A Survey of 3D Object Selection Techniques for Virtual Environments," Comput. Graph., **37**(3), pp. 121–136.

[15] Fekri, A. M. Y., and Wanis, I. A., 2019, "A Review on Multimodal Interaction in Mixed Reality Environment," IO Conference Series: Materials Science and Engineering, Bangkok, Thailand, Feb. 4–9, Vol. 551, p. 012049.

[16] Klemmer, S. R., Hartmann, B., and Takayama, L., 2006, "How Bodies Matter: Five Themes for Interaction Design," ACM Conference on Designing Interactive Systems (DIS '06), Brisbane, Australia, June 4–8, pp. 140–149.

[17] Dourish, P., 2001, *Where the Action is: The Foundations of Embodied Interaction*, The MIT Press, Camdridge, MA.

[18] Hall, E. T., 1966, *The Hidden Dimension*, Vol. 609, NY: Doubleday, Garden City.

[19] Wang, R., Paris, S., and Popović, J., 2011, "6D Hands: Markerless Hand-Tracking for Computer Aided Design," 24th ACM Symposium on User Interface Software and Technology (UIST), Santa Barbara, CA, Oct. 16–19, pp. 549–558.

[20] Song, P., Goh, W. B., Hutama, W., Fu, C.-W., and Liu, X., 2012, "A Handle Bar Metaphor for Virtual Object Manipulation With Mid-Air Interaction," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12, Austin, TX, May 5–10, ACM, pp. 1297–1306.

[21] Vinayak, Murugappan, and Ramani, K., 2013, "Handy-Potter: Rapid Exploration of Rotationally Symmetric Shapes Through Natural Hand Motions," ASME J. Comput. Inf. Sci. Eng., **13**(2), p. 021008.

[22] Guenter, B., Finch, M., Drucker, S., Tan, D., and Snyder, J., 2012, "Foveated 3D Graphics," SIGGRAPH Asia 2012 Computer Animation Festival, SA '12, Singapore, Nov. 29–Dec. 1.

[23] Azuma, R. T., 2017, "Making Augmented Reality a Reality," Image and Metrology 2017, San Francisco, CA, June 26–29.

[24] Laitin, E. L., Tymoski, M. J., Tenhundfeld, N. L., and Witt, J. K., 2019, "The Uphill Battle for Action-Specific Perception," Atten., Percept., Psychophys., **81**(3), pp. 778–793.

[25] Witt, J. K., 2020, "Action's Influence on Spatial Perception: Resolution and a Mystery," Curr. Opin. Psychol., **32**, pp. 153–157.

[26] Armbruster, C., and Spijkers, W., 2006, "Movement Planning in Prehension: Do Intended Actions Influence the Initial Reach and Grasp Movement?," Motor Control, **10**(4), p. 311.

[27] Hilliges, O., Kim, D., Izadi, S., Weiss, M., and Wilson, A., 2012, "Holodesk: Direct 3D Interactions With a Situated See-Through Display," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12, Austin, TX, May 5–10, ACM, pp. 2421–2430.

[28] Lee, J., Olwal, A., Ishii, H., and Boulanger, C., 2013, "Spacetop: Integrating 2D and Spatial 3D Interactions in a See-Through Desktop Environment," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '13, Paris France, Apr. 27–May 2, ACM, pp. 189–192.

[29] Weichel, C., Lau, M., Kim, D., Villar, N., and Gellersen, H. W., 2014, "Mixfab: A Mixed-Reality Environment for Personal Fabrication," Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems, CHI '14, Toronto, Ontario, Canada, Apr. 26–May 1, ACM, pp. 3855–3864.

[30] Gibson, J. J., 2014, *The Ecological Approach to Visual Perception: Classic Edition*, Psychology Press, UK.

[31] Malmberg, F., Vidholm, E., and Nyström, I., 2006, "A 3D Live-Wire Segmentation Method for Volume Images Using Haptic Interaction," *Discrete Geometry for Computer Imagery*, A. Kuba, L. G. Nyúl, and K. Palágyi, eds., Springer, Berlin/Heidelberg, pp. 663–673.

[32] Petersen, N., and Stricker, D., 2009, "Continuous Natural User Interface: Reducing the Gap Between Real and Digital World," 2009 8th IEEE International Symposium on Mixed and Augmented Reality, Orlando, FL, Oct. 19–20, pp. 23–26.

[33] Regazzoni, D., Rizzi, C., and Vitali, A., 2018, "Virtual Reality Applications: Guidelines to Design Natural User Interface," Volume 1B: 38th Computers and Information in Engineering Conference of International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Quebec City, QC, Canada, Aug. 26–29, Paper No. V01BT02A029.

[34] Masliah, M. R., and Milgram, P., 2000, "Measuring the Allocation of Control in a 6 deg-of-Freedom Docking Experiment," Proceedings of the Annual ACM Conference on Human Factors in Computing Systems, CHI '00, The Hague, The Netherlands, Apr. 1–6, ACM, pp. 25–32.

[35] Freeman, E., Brewster, S., and Lantz, V., 2016, "Do That, There," Proceedings of the Annual ACM Conference on Human Factors in Computing Systems, CHI '16, San Jose, CA, May 7–12, ACM Press, pp. 2319–2331.

[36] Mumm, J., and Mutlu, B., 2011, "Human-Robot Proxemics," 6th Annual Conference for Basic and Applied Human–Robot Interaction Research, HRI '11, Lausanne, Switzerland, Mar. 6–9, ACM Press, p. 331.

[37] Kokubo, Y., Sato-Shimokawara, E., and Yamaguchi, T., 2016, "Relationship Between Personal Space and the Big Five With Telepresence Robot," 9th International Conference on Intelligent Robotics and Applications, ICIRA '16, Tokyo, Japan, Aug. 22–24.

[38] Dickinson, P., Gerling, K., Hicks, K., Murray, J., Shearer, J., and Greenwood, J., 2019, "Virtual Reality Crowd Simulation: Effects of Agent Density on User Experience and Behaviour," Virtual Reality, **23**(1), pp. 19–32.

[39] Zhang, J., Janeh, O., Katzakis, N., Krupke, D., and Steinicke, F., 2019, "Evaluation of Proxemics in Dynamic Interaction With a Mixed Reality Avatar Robot," Proceedings of the International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments (ICAT-EGVE) 2019, Tokyo, Japan, Oct. 20–23, pp. 1–8.

[40] Grønbæk, J. E., Knudsen, M. S., O'Hara, K., Krogh, P. G., Vermeulen, J., and Petersen, M. G., 2020, "Proxemics Beyond Proximity: Designing for Flexible Social Interaction Through Cross-Device Interaction," Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20, Honolulu, HI, Apr. 25–30, Association for Computing Machinery, pp. 1–14.

[41] Pederson, T., Janlert, L.-E., Jalaliniya, S., and Surie, D., "Proximity as Key Property in the Egocentric Interaction Paradigm."

[42] Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., and Wang, M., 2011, "Proxemic Interactions: The New Ubicomp?," Interactions, **18**(1), pp. 42–50.

[43] Ballendat, T., Marquardt, N., and Greenberg, S., 2010, "Proxemic Interaction," ACM International Conference on Interactive Tabletops and Surfaces – ITS '10, Saarbrucken, Germany, Nov. 7–10, ACM Press, p. 121.

[44] Marquardt, N., Diaz-Marino, R., Boring, S., and Greenberg, S., 2011, "The Proximity Toolkit," Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology - UIST '11, Santa Barbara, CA, Oct. 16–19, ACM Press, p. 315.

[45] Marquardt, N, and Nicolai, 2013, "Proxemic Interactions With and Around Digital Surfaces," ACM Interactive Tabletops and Surfaces 2013 Conference (ITS), Dresden, Germany, Oct. 6–9, ACM Press, pp. 493–494.

[46] Paul, C. L., and Bradel, L., 2018, "Size Matters," Proceedings of the 2018 International Conference on Advanced Visual Interfaces, AVI '18, New York, NY, May 29–June 1, ACM Press, pp. 1–5.

[47] Li, A. X., Lou, X., Hansen, P., and Peng, R., 2016, "Improving the User Engagement in Large Display Using Distance-Driven Adaptive Interface," Interact. Comput., **28**(4), pp. 462–478.

[48] Jakobsen, M. R., and HornbÆk, K., 2014, "Up Close and Personal," ACM Trans. Comput.-Hum. Interact., **21**(2), pp. 1–34.

[49] Ledo, D., Greenberg, S., Marquardt, N., and Boring, S., 2015, "Proxemic-Aware Controls," Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services, MHCI '15, New York, NY, Aug. 24–27, ACM Press, pp. 187–198.

[50] Babic, T., Reiterer, H., and Haller, M., 2018, "Pocket6: A 6DOF Controller Based on a Simple Smartphone Application," 6th ACM Symposium on Spatial User Interaction (SUI 2018), Berlin, Germany, Oct. 13–14, Association for Computing Machinery, pp. 2–10.

[51] Huo, K., Cao, Y., Yoon, S. H., Xu, Z., Chen, G., and Ramani, K., 2018, "Scenariot: Spatially Mapping Smart Things Within Augmented Reality Scenes," Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI '18, Montreal, QC, Canada, Apr. 21–26, Association for Computing Machinery, pp. 1–13.

[52] Fogtmann, M. H., Fritsch, J., and Kortbek, K. J., 2008, "Kinesthetic Interaction: Revealing the Bodily Potential in Interaction Design," Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat, OZCHI '08, Cairns, Australia, Dec. 8–12, ACM, pp. 89–96.

[53] Makin, L., Barnaby, G., and Roudaut, A., 2019, "Tactile and Kinesthetic Feedbacks Improve Distance Perception in Virtual Reality," Proceedings of the 31st Conference on l'Interaction Homme-Machine: Adjunct, IHM '19, Grenoble, France, Dec. 10–13.

[54] Cockburn, A., Quinn, P., Gutwin, C., Ramos, G., and Looser, J., 2011, "Air Pointing: Design and Evaluation of Spatial Target Acquisition With and Without Visual Feedback," Int. J. Hum.-Comput. Stud., **69**(6), pp. 401–414.

[55] Yan, Y., Yu, C., Ma, X., Huang, S., Iqbal, H., and Shi, Y., 2018, "Eyes-Free Target Acquisition in Interaction Space Around the Body for Virtual Reality," Proceedings of the 2018 CHI Conference on Human Factors in Computing System, CHI '18, Montreal, QC, Canada, Apr. 21–26, ACM, pp. 42:1–42:13.

[56] Plaumann, K., Weing, M., Winkler, C., Müller, M., and Rukzio, E., 2018, "Towards Accurate Cursorless Pointing: The Effects of Ocular Dominance and Handedness," Pers. Ubiquitous Comput., **22**(4), pp. 633–646.

[57] De Boeck, J., Raymaekers, C., and Coninx, K., 2006, "Exploiting Proprioception to Improve Haptic Interaction in a Virtual Environment," Presence: Teleoper. Virtual Environ., **15**(6), pp. 627–636.

[58] Gao, B., Lu, Y., Kim, H., Kim, B., and Long, J., 2019, *Virtual, Augmented and Mixed Reality. Multimodal Interaction*, J. Y. Chen, and G. Fragomeni, eds., Springer International Publishing, Switzerland, pp. 44–58.

[59] Hinckley, K., Pausch, R., Proffitt, D., Patten, J., and Kassell, N., 1997, "Cooperative Bimanual Action," Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, CHI '97, Atlanta, GA, Mar. 22–27, ACM, pp. 27–34.

[60] Leganchuk, A., Zhai, S., and Buxton, W., 1998, "Manual and Cognitive Benefits of Two-Handed Input: An Experimental Study," ACM Trans. Comput.-Hum. Interact., **5**(4), pp. 326–359.

[61] Schmidt, B., 2002, *When are Two Hands Better Than one? : A Study of Bimanual Interaction*, Computer Science and Software Engineering.

[62] Sakurada, T., Ito, K., and Gomi, H., 2016, "Bimanual Motor Coordination Controlled by Cooperative Interactions in Intrinsic and Extrinsic Coordinates," Eur. J. Neurosci., **43**(1), pp. 120–130.

[63] Gribnau, M. W., and Hennessey, J. M., 1998, "Comparing Single- and Two-Handed 3D Input for a 3D Object Assembly Task," CHI 98 Conference Summary on Human Factors in Computing Systems, Los Angeles, CA, Apr. 18–23, ACM,pp. 233–234.

[64] Buxton, W., and Myers, B., 1986, "A Study in Two-Handed Input," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '86, Boston, MA, Apr. 13 –17ACM, pp. 321–326.

[65] Swinnen, S., Young, D., Walter, C., and Serrien, D., 1991, "Control of Asymmetrical Bimanual Movements," Exp. Brain Res., **85**(1), pp. 163–173.

[66] Hinckley, K., Pausch, R., Proffitt, D., and Kassell, N. F., 1998, "Two-Handed Virtual Manipulation," ACM Trans. Comput.-Hum. Interact., **5**(3), pp. 260–302.

[67] Balakrishnan, R., and Kurtenbach, G., 1999, "Exploring Bimanual Camera Control and Object Manipulation in 3D Graphics Interfaces," Proceedings of the ACM CHI 99 Human Factors in Computing Systems Conference, CHI '99, Pittsburg, PA, May 15–20, ACM, pp. 56–62.

[68] Balakrishnan, R., and Hinckley, K., 2000, "Symmetric Bimanual Interaction," Proceedings of the SIGCHI conference on Human Factors in Computing Systems, CHI '00, The Hague, The Netherlands, Apr. 1–6, ACM, pp. 33–40.

[69] Xia, X., Irani, P., and Wang, J., 2007, "Evaluation of Guiard's Theory of Bimanual Control for Navigation and Selection," Ergonomics and Health Aspects of Work with Computers, EHAWC '07, Beijing, China, July 22–27, Springer-Verlag, pp. 368–377.

[70] Song, J., Cho, S., Baek, S.-Y., Lee, K., and Bang, H., 2014, "Gafinc: Gaze and Finger Control Interface for 3D Model Manipulation in {CAD} Application," Computer-Aided Design, **46**, pp. 239–245.

[71] Ulinski, A., Wartell, Z., and Hodges, L. F., 2007, "Bimanual Task Division Preferences for Volume Selection," Proceedings of the 2007 ACM Symposium on Virtual reality Software and Technology, VRST '07, Newport Beach, CA, Nov. 5–7, ACM, pp. 217–218.

[72] Surale, H. B., Gupta, A., Hancock, M., and Vogel, D., 2019, "Tabletinvr: Exploring the Design Space for Using a Multi-Touch Tablet in Virtual Reality," Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19, Glasgow, Scotland, UK, May 4–9, pp. 1–13.

[73] Montano-Murillo, R. A., Nguyen, C., Kazi, R. H., Subramanian, S., DiVerdi, S., and Martinez-Plasencia, D., 2020, "Slicing-Volume: Hybrid 3D/2D Multi-Target Selection Technique for Dense Virtual Environments," IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR), Atlanta, GA, Mar. 22–26, pp. 53–62.

[74] Brandl, P., Forlines, C., Wigdor, D., Haller, M., and Shen, C., 2008, "Combining and Measuring the Benefits of Bimanual Pen and Direct-Touch Interaction on Horizontal Interfaces," Proceedings of the Working Conference on Advanced Visual Interfaces, AVI '08, Napoli, Italy, May 28–30, ACM, pp. 154–161.

[75] Serino, A., Noel, J.-P., Galli, G., Canzoneri, E., Marmaroli, P., Lissek, H., and Blanke, O., 2015, "Body Part-Centered and Full Body-Centered Peripersonal Space Representations," Sci. Rep., **5**, p. 18603.

[76] di Pellegrino, G., and Làdavas, E., 2015, "Peripersonal Space in the Brain," Neuropsychologia, **66**, pp. 126–133.

[77] Galigani, M., Castellani, N., Donno, B., Franza, M., Zuber, C., Allet, L., Garbarini, F., and Bassolino, M., 2020, "Effect of Tool-Use Observation on

Metric Body Representation and Peripersonal Space," Neuropsychologia, **148**, p. 107622.

[78] Marieb, E. N., and Hoehn, K., 2007, *Human Anatomy & Physiology*, Pearson Education, London, UK..

[79] Bufacchi, R. J., and Iannetti, G. D., 2018, "An Action Field Theory of Peripersonal Space," Trends Cogn. Sci., **22**(12), pp. 1076–1090.

[80] Popovici, I., Schipor, O.-A., and Vatavu, R.-D., 2019, "Hover: Exploring Cognitive Maps and Mid-Air Pointing for Television Control," Int. J. Hum.-Comput. Stud., **129**, pp. 95–107.

[81] Hans, A., and Hans, E., 2015, "Kinesics, Haptics and Proxemics: Aspects of Non-Verbal Communication," IOSR J. Humanit. Soc. Sci. (IOSR-JHSS), **20**(2), pp. 47–52.

[82] Klatzky, R. L., 1998, *Allocentric and Egocentric Spatial Representations: Definitions, Distinctions, and Interconnections*, Springer, Berlin, Heidelberg, pp. 1–17.

[83] Bach, B., Sicat, R., Beyer, J., Cordeil, M., and Pfister, H., 2018, "The Hologram in My Hand: How Effective Is Interactive Exploration of 3d Visualizations in Immersive Tangible Augmented Reality?" IEEE Trans. Vis. Comput. Graph., **24**(1), pp. 457–467.

[84] Feng, Q., Shum, H. P. H., and Morishima, S., 2018, "Resolving Occlusion for 3D Object Manipulation With Hands in Mixed Reality," roceedings of the 24th ACM Symposium on Virtual Reality Software and Technology, VRST '18, Tokyo, Japan, Nov. 28 –Dec. 1, ACM, pp. 119:1–119:2.

[85] Hannaford, B., Wood, L., McAffee, D., and Zak, H., 1991, "Performance Evaluation of a Six-Axis Generalized Force-Reflecting Teleoperator," IEEE Trans. Syst., Man, and Cybern., **21**(3), pp. 620–633.

[86] Unger, B., Nicolaidis, A., Berkelman, P., Thompson, A., Klatzky, R., and Hollis, R., 2001, "Comparison of 3-D Haptic Peg-in-Hole Tasks in Real and Virtual Environments," Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems, Maui, HI, Oct. 29–Nov. 3, vol. 3, pp. 1751–1756.

[87] Unger, B., Nicolaidis, A., Berkelman, P., Thompson, A., Lederman, S., Klatzky, R., and Hollis, R., 2002, "Virtual Peg-in-Hole Performance Using a 6-DOF Magnetic Levitation Haptic Device: Comparison With Real Forces and With Visual Guidance Alone," 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, HAPTICS 2002, Orlando, FL, Mar. 24–25, pp. 263–270.

[88] Speicher, M., Daiber, F., Kiefer, G.-L., and Krüger, A., 2017, "Exploring Task Performance and User's Preference of Mid-Air Hand Interaction in a 3D Docking Task Experiment," 5th ACM Symposium on Spatial User Interaction (SUI 2017), Brighton, UK, Oct. 16–17, p. 160.

[89] Besançon, L., Issartel, P., Ammi, M., and Isenberg, T., 2017, "Mouse, Tactile, and Tangible Input for 3D Manipulation," Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17, Denver, CO, May 6–11, pp. 4727–4740.

[90] Hart, S. G., and Staveland, L. E., 1988, "Development of NASA-tlx (Task Load Index): Results of Empirical and Theoretical Research," Adv. Psychol., **52**, pp. 139–183.

[91] Chmarra, M. K., Jansen, F. W., Grimbergen, C. A., and Dankelman, J., 2008, "Retracting and Seeking Movements During Laparoscopic Goal-oriented Movements. Is the Shortest Path Length Optimal?," Surg. Endosc., **22**(4), pp. 943–949.

[92] Williams, M. R., and Kirsch, R. F., 2008, "Evaluation of Head Orientation and Neck Muscle Emg Signals as Command Inputs to a Human–Computer Interface for Individuals With High Tetraplegia," IEEE Trans. Neural Syst. Rehabil. Eng., **16**(5), pp. 485–496.

[93] Wang, Z., and Fey, A. M., 2018, "Human-Centric Predictive Model of Task Difficulty for Human-in-the-Loop Control Tasks," PLoS. One., **13**(4), p. e0195053.

[94] Lavielle, M., 2005, "Using Penalized Contrasts for the Change-Point Problem," Signal Process., **85**(8), pp. 1501–1510.

[95] Killick, R., Fearnhead, P., and Eckley, I. A., 2012, "Optimal Detection of Changepoints With a Linear Computational Cost," J. Am. Stat. Assoc., **107**(500), pp. 1590–1598.

[96] Feix, T., Romero, J., Schmiedmayer, H., Dollar, A. M., and Kragic, D., 2016, "The Grasp Taxonomy of Human Grasp Types," IEEE Trans. Human-Mach. Syst., **46**(1), pp. 66–77.

[97] Kroemer, K. H. E., 1986, "Coupling the Hand With the Handle: An Improved Notation of Touch, Grip, and Grasp," Human Fact.: J. Hum. Fact. Ergon. Soc., **28**(3), pp. 337–339.

[98] Hart, S. G., 2006, "Nasa-Task Load Index (nasa-tlx); 20 Years Later," Human Factors and Ergonomics Society 46th Annual Meeting, Baltimore, MD, Sept. 29–Oct. 4.

[99] Algazi, V., Avendaño, C., and Duda, R., 2002, "Estimation of a Spherical-Head Model From Anthropometry," AES: J. Audio Eng. Soc., **49**(6), p. 0.

[100] Witt, J. K., and Proffitt, D. R., 2008, "Action-Specific Influences on Distance Perception: a Role for Motor Simulation.," J. Exp. Psychol.: Hum. Percept. Perform., **34**(6), p. 1479.

[101] Witt, J. K., Proffitt, D. R., and Epstein, W., 2005, "Tool Use Affects Perceived Distance, But Only When You Intend to Use It," J. Exp. Psychol.: Hum. Percept. Perform., **31**(5), p. 880.

[102] Di Luca, M., and Mahnan, A., 2019, "Perceptual Limits of Visual-Haptic Simultaneity in Virtual Reality Interactions," 2019 IEEE World Haptics Conference (WHC), Tokyo, Japan, July 9–12, pp. 67–72.

[103] Schwarz, R. J., and Taylor, C., 1955, "The Anatomy and Mechanics of the Human Hand," Artif. Limbs, **2**(2), pp. 22–35.

[104] Ma'touq, J., Hu, T., and Haddadin, S., 2018, "Sub-millimetre Accurate Human Hand Kinematics: From Surface to Skeleton," Comput. Methods Biomech. Biomed. Eng., **21**(2), pp. 113–128.