HPUI: Hand Proximate User Interfaces for One-Handed Interactions on Head Mounted Displays

Shariff AM Faleel, Michael Gammon, Kevin Fan, Da-Yuan Huang, Wei Li and Pourang Irani

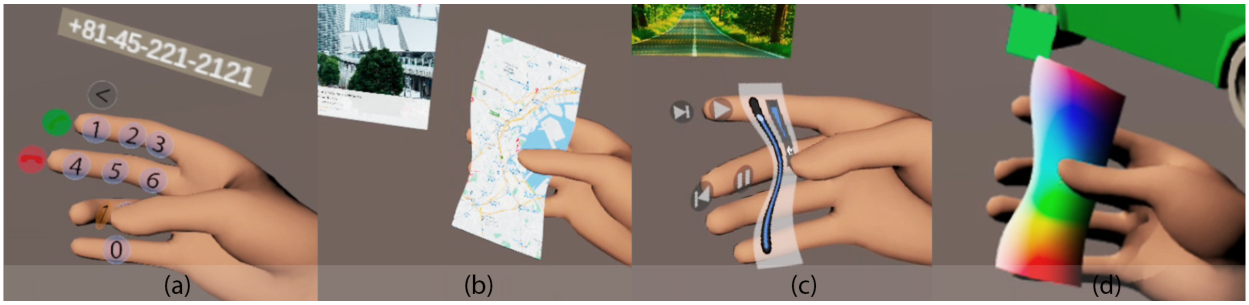


Fig. 1: Sample applications using Hand Proximate User Interfaces. (a) Phone dial user interface. (b) Map application, with additional information showing up off the hand. (b) Media player application, with display off the hand. (d) Colour selection app for assigning a colour to an object in the environment.

Abstract—We explore the design of Hand Proximate User Interfaces (HPUIs) for head-mounted displays (HMDs) to facilitate near-body interactions with the display directly projected on, or around the user's hand. We focus on single-handed input, while taking into consideration the hand anatomy which distorts naturally when the user interacts with the display. Through two user studies, we explore the potential for discrete as well as continuous input. For discrete input, HPUIs favor targets that are directly on the fingers (as opposed to off-finger) as they offer tactile feedback. We demonstrate that continuous interaction is also possible, and is as effective on the fingers as in the off-finger space between the index finger and thumb. We also find that with continuous input, content is more easily controlled when the interaction occurs in the vertical or horizontal axes, and less with diagonal movements. We conclude with applications and recommendations for the design of future HPUIs.

Index Terms—On-hand projected interfaces, Deformable UIs, Virtual Reality

1 Introduction

In recent years, head mounted displays (HMDs) have gained consumer level traction for productivity and entertainment applications [40,45,52]. They have broadened the possibilities of personal computing through immersion in virtual worlds as seen with Virtual Reality (VR) systems, or by integrating our digital world into the user's physical environment using Mixed Reality (MR) systems. On HMDs, mid-air displays and free-hand interactions are a common input modality as seen with the consumer-ready Microsoft Hololens and Oculus Quest. However, such forms of interaction can be uncomfortable, causing fatigue over time [4, 15, 28] and unnatural to interact with, both socially [1, 17, 27, 58] and physically [4, 14, 17, 27].

To address this challenge, we explore Hand Proximate User Interfaces (HPUIs), virtual user interfaces positioned on and around a user's hand (Fig. 1). These can be imagined as similar to the smartphone's interface, but without the physical device, and with the possibility of further integration with the massive interaction and display space offered by HMDs in general. Since our hands are already the primary means of physical interaction with devices and the world, the familiarity and dexterity already exists for most users, along with high levels of social acceptance [23, 58, 62], making it a natural place to anchor a display. On-hand interfaces are presumed to be useful as a more comfortable alternative to navigating virtual menus (both physically and socially), while also offering unique interaction and display opportunities due to the novelty of the hand's anatomy as an input/output space. HPUIs may also be useful for bringing in regular mobile interactions while immersed in VR tasks, such as checking notifications and replying to messages (Fig. 1a)Furthermore, menus and UI controls in VR can be placed on the fingers, such as for controlling a video timeline (Fig. 1c) or manipulating a color palette (Fig. 1d).

While prior work has explored on-hand input with the visualization displayed elsewhere (indirect input) [9, 12, 34, 60, 61, 63, 66], having the actual display anchored to the hand (direct input) and its design space, remains largely unexplored - particularly the use of a single hand as a joint input/output space. While displaying content on a non-folding palm has been demonstrated [22–25, 33, 47, 54, 64], a one-handed HPUI with direct input using the thumb necessitates an output space on the fingers themselves. This UI offers a dynamic, deformable surface with discontinuities between fingers. The best way to display content and allow uninstrumented one-handed interaction is not immediately obvious.

We propose a framework for displaying and interacting with content on the hand and investigate the key factors influencing the design of an HPUI. We first explore the merit of directly displaying and interacting with discrete items (representative of app icons, for example) on and between the deformable fingers. From our first study we observe that interacting with targets displayed directly on the hand (direct input) is more efficient than when the users have to primarily target at an item without a display (indirect input). Furthermore, our results show that

- *Shariff AM Faleel, University of Manitoba. mohommas@myumanitoba.ca*
- *Michael Gammon, University of Manitoba. michael.gammon@umanitoba.ca*
- *Kevin Fan, Human-Machine Interaction Lab, Huawei Canada. szu.wen.fan@huawei.com*
- *Da-Yuan Huang, Human-Machine Interaction Lab, Huawei Canada. dayuan.huang@huawei.com*
- *Wei Li, Human-Machine Interaction Lab, Huawei Canada. wei.li.crc@huawei.com*
- *Pourang Irani, University of Manitoba. pourang.irani@cs.umanitoba.ca*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

icons being displayed off-fingers are also more efficient to interact with in direct instead of indirect input mode. Informed by our results, we then propose methods for displaying and interacting with a continuous workspace on the hand. We explore the case of a display surface that deforms with movements of each finger, offering tactile feeling. We demonstrate how such a display surface affords interactivity for both discrete (pointing) and continuous (dragging) tasks. We finally present a number of applications for hand-proximate UIs on HMDs.

In this paper we offer the following contributions: (i) a design framework for interactive on-hand displays, which we refer to as a hand-proximate user interface (HPUI); (ii) a demonstration of the potential for direct one-handed HPUIs for discrete item selection, as well as (iii) continuous input/output, and (iv) a validation of our interfaces and results that can inform the design of future HPUIs.

2 Related Work

We briefly review work relevant to HPUIs, namely research including indirect on-hand interaction as well as work on interactive on-hand displays. Given the resulting deformable display on the fingers, we also briefly review work on deformable interfaces in general.

2.1 On-Hand Input

Prior research has presented on-hand interaction as a controller for a variety of tasks, from coarse input such as simple buttons and switches [34] and TV remote controls [12], to high precision tasks including full QWERTY keyboards [60, 63, 66] and handwriting detection [9, 61]. These works cite the hand as a beneficial input space for HMDs for numerous reasons. For example, the tactile/haptic feedback offered by the hand makes the fine manipulation of UI elements such as buttons, sliders, and dials much more precise than it otherwise would be using midair input [34]. This of course can be achieved instead by an external physical object (commonly tablets or paddles such as the HARP [37]). Kohli & Whitton [34] argue that such props are impractical to carry around, and unnecessary when the hand can provide the aforementioned benefits as an input device along with the advantage afforded by human proprioception. Similarly, Wang et al. [61] aptly noted that the human palm offers unique affordances that benefit eyes-free interaction. Our developed sense of proprioception allows the palm to be used as a remote control input surface. In such cases, the input on the hand was indirect, without a display affixed to it, and instead external to the palm.

In addition to precision and performance benefits, on-hand interaction has also rated highly in comfort [23, 27, 59] and social acceptability [27, 58, 59] when compared to mid-air and on-body interaction. Tung et al. [58] investigated game input for smart glasses in public spaces and found that users preferred interacting with their palms over using wearable devices, largely because the interactions were less noticeable and thus more socially acceptable. Furthermore, several works have investigated the natural segments of the hand's fingers as "comfort regions" rated by participants [11, 29]. These have typically found thumb-to-finger interactions on the index and middle fingers to be most appropriate for a majority of users. We expand on the functional interaction space described by Huang et al. [29] and Dewitz et al. [11].

2.2 On-Hand Display

Although many of the above works argue that a primary benefit of the hand is for eyes-free interaction, Gustafson et al. [20] demonstrated that for imaginary palm-based interfaces, tactile clues are the second-most important mechanism by which users interact with the palm. They found that hand-based visual landmarks were actually the most important features for fluid input [20, 21]. With tactile cues alone users were able to discriminate at least 16 discrete on-finger buttons, meaning that the hand is indeed a promising interface for eyes-free interaction. However, exploring it as a joint input-output space could possibly yield even more impressive results. Similarly, Steimle et al. [50] found that body landmarks (anatomical as well as things like jewellery and tattoos) facilitate the localization of interactive elements by leveraging human sensory and motor capabilities. This suggests that the anchoring of virtual UI elements on the hand could provide a more intuitive interaction experience in comparison to mid-air input.

Many works [22–25, 33, 47, 54, 64] have used RGB projectors to display ambient information on the user's palm or body, sometimes with a basic touch-based UI hosting a simple menu. Muller et al. [41] conceive of an AR palm-based interface, and design a possible interface [42] which allows for accessing layers of data by shifting the hand back-and-forth in mid-air. This latter work focuses on the degree of interactive freedom offered by the elbow joint rather than considering the hand as a touch interface. Azai et al. [3] also developed an intriguing two-handed off-skin HPUI and argue that the mid-air interface allows for a larger display, but at the cost of tactile feedback or proprioception. None of the above address the latent challenges for one-handed interactive displays on the palm, which have to adapt to the natural deformation of the fingers for both the interaction and presentation.

Prior work concerning on-hand displays (as opposed to on-hand input-only) primarily focus on two-handed interfaces where the UI is anchored to the non-dominant hand, and interacted with by the free hand. Three notable exceptions being Xu et al's. [65] one-handed VR interface where the output exists as a semi-circle around the hand, an elicitation study [16] on HPUIs for mobile interactions, and TULIP menu [7] where menu items are assigned to each finger with selection done using pinch gestures. Our work can be seen as an extension of these works. Allowing for one-handed input offers numerous benefits [31, 32] and frees the other hand for either using a controller (in VR, for example) or holding other critical items (physical objects, in MR). Furthermore, Karlson et al. [31] found that one-handed device usage was overwhelmingly preferred for all tasks surveyed, and two-handed usage was often a consequence of interfaces that necessitate it, rather than by preference. It is possible that these findings could extend from physical devices to HPUIs. In an elicitation study on HPUIs [16], participants derived single-handed gestures for applications which require two-handed input on current mobile devices. Accordingly, in this paper we exclusively consider one-handed HPUIs. One of the main problems with single-handed interaction on current mobile devices is the trade-off between display real-estate and reachability for continuous displays due to the limited reachable screen space during such interactions [26, 44]. In section 6.2 we propose a single-handed continuous deformable UI that is anchored to the hand which could help leverage the benefits of a one-handed UI with a sufficiently large interaction space.

In addition to the lack of single-handed on-hand input, we also found little insight into the problem of projecting information onto a discontinuous physical surface which is constantly deforming during interaction. The works which do use the hand as an output space either consider planar displays in the around-hand region (see Table 1), or projection of simple discrete interface components onto the mostly static, continuous space that is the palm. However, since we are interested in one-handed interaction, it becomes necessary to consider the fingers themselves as the main interactable output space [11, 29]. Based on the current literature, it is unclear what projection on such a discontinuous, dynamic surface should look like - a non-requisite exploration when projecting only on a non-deformable palm. Regardless of whether designers choose a discrete or continuous workspace on the fingers (see Sect. 3.5), we have a dynamic display which in some cases may necessitate the exploration of new design space dimensions.

2.3 Deformable Displays and Interfaces

Of interest to our work is that of deformable displays and interfaces which have seen relatively strong interest recently [2, 18, 19, 35, 38, 46, 51, 53, 55]. For example, Lindlbauer et al. [38] explored a physical shape changing interface which also includes "optical deformations" of the displayed content that match the physical deformations for more accurate and realistic representations. This compliments 3D graphics with tactile sensation and natural depth cues. Rather than looking at a typical user interface, the authors explored a few individual applications (ie. weather app, games) for which physical deformation of the display could enhance user experience. The physical display is made to automatically deform to match the virtual content - for example, a weather app with on-screen waves which physically deform the screen as they pass by. Gomes & Vertegaal [19] explored shape changes in a multi-segmented mobile device for triggering viewport transformations in its GUI. These allow for investigating the utility of display shape actuation as a method of providing notifications to users based on urgency [18]. Another example of self-actuating mobile displays is presented in [46]. Researchers have also proposed deformation properties as an input method [35] [55].

In theory one could also consider on-body RGB projected interfaces [22–25, 47, 54, 64] as "deformable" in a literal sense, although these are unintentional deformations which are not meant to serve some greater interaction or visualization purpose. The deformations are a consequence of projecting light onto a dynamic surface - the deformation isn't built into the system, and although the resulting output *appears* deformed, the content itself is not being deformed in some controlled way. For these reasons, such displays are not what is meant when we discuss deformable interfaces.

The literature on deformable interactive displays primarily concerns devices with screens either built-in or projected onto them. There seems to be a gap in the research regarding interacting with virtual deformable surfaces, and it is unclear how factors such as tactile discontinuities affect key tasks such as discrete pointing or continuous dragging.

3 Factors influencing the design of HPUIs

We describe several key factors that can describe and influence the design of HPUIs, namely: output spaces, input spaces, workspace styles, display location, and display frame of reference.

3.1 Output Spaces

In this work, we focus on a subset of possible HPUIs, namely those where the input comes from the same hand which serves as the anchor for the visualization. This distinction naturally divides the hand-proximate space into two broader categories:

Interactable Output Spaces: Spaces where visual content can be anchored for displaying information which can be reached by the thumb (ie. joint input-output spaces).

Non-Interactable Output Spaces: Spaces where visual interfaces can be anchored for displaying information that cannot be reached by the thumb (ie. output-only spaces).

Referring to Table 1, most of the "around hand" locations which use the broader hand as an anchor are treated as non-interactable output spaces, as they are not easily reachable by the thumb and fingers. While we depict such space as a plane in a 2D image, the "around hand" region also includes the unreachable locations in and out of the page - ie. a display floating above the palm, or a sphere surrounding the entire hand. Similarly, the "on-palm" location is not easily interacted with using the thumb, but can be reached using the fingers. In such cases we consider this space as an output-only space as we typically do not use the fingers for interaction in this way [11]. The "above finger"/"between finger" locations can mostly be reached with the thumb by bending the fingers inward and thus can be thought of as joint input-output space. However, it is not immediately obvious whether the lack of tactile feedback in such regions will encumber performance on selection tasks. Finally, the "on-finger" location is prime real estate for interactable displays, and is the main motivator for using the hand as a joint input-output space.

3.2 Input Spaces

We further divide the input spaces for HPUIs into two broader categories:

On-Fingers: Icons or general input spaces that are located on the hand itself. A touch event on such input spaces requires contact between the thumb and another part of the hand, providing tactile feedback to the user.

Off-Fingers: Icons or general input spaces that float above/around the hand, or in between the fingers. A touch event occurs by passing the thumb through the input space, and no tactile feedback is provided.

Making this distinction prompts the question of performance differences between the two spaces - presumably on-finger input spaces will be better suited for eyes-free interaction due to our unconscious awareness (proprioception) of the hand, whereas off-finger interactions will require visual cues to be efficient and accurate, an idea we explore in Sect. 5.

3.3 Display Location

The on and around-hand regions can be broken down into five main display locations, each of which is associated with a Frame of Reference (see Sect. 3.4). We refer the reader to Table 1 for the visual representations of this factor.

Around Hand: The mid-air region surrounding the perimeter of the hand seen in Table 1, but also the greater three-dimensional sphere around the hand which extends in and out of the page (ie. above-palm).

Between Fingers: The mid-air region between any two adjacent fingers.

On Palm: The relatively flat on-skin region in the palm of the hand, as well as the back of the hand.

On Finger: The on-finger regions along the front, back, and sides of the index, middle, ring, and pinky fingers.

Above Finger: The mid-air regions directly above the tips of the index, middle, ring, and pinky fingers.

We typically think of the around-hand regions shown in Table 1 as non-interactable spaces or, if they are interactable, requiring two hands, such as in Azai et al. [3] or the MRTK [1]. The between-finger locations are reachable by the thumb and are thus candidates for off-finger direct input. However, it is unclear how lack of tactile feedback will affect input performance in these areas. The on-palm region is treated similarly to around-hand, where for one-handed thumb input most of the surface is unreachable. The surface could potentially be used for touch input via the fingers, although the literature suggests that interactive elements on the palm can not be reached reliably by the fingers [11]. We consider the on-finger regions the main input space for one-handed HPUIs. Finally, although both above-finger and between-finger regions are off-fingers, we treat them separately as they each have their own frame-of-reference.

3.4 Display Frame of Reference

Each of the display locations outlined in Sect. 3.3 correspond with one of the following frames of reference, which dictate the parts of the hand responsible for the relative position and orientation of those displays. We consider four main frames of reference on the hand.

Palm: Elements whose orientation and position are set relative to the palm of the hand. Such elements rotate with the wrist and have a constant distance from the palm.

Individual Phalanx: Elements whose orientation and position are set relative to an individual segment (phalanx) of a finger, such as the finger tip (distal phalanx), or the segment between the first and second knuckles of a finger (proximal phalanx).

Whole Finger: Elements whose orientation and position are set relative to one whole finger (ie. a planar display whose angle changes such that the middle finger always stays completely behind it).

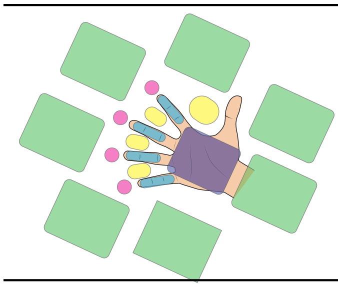
Multiple Fingers: Elements whose orientation and position are set relative to multiple whole fingers (ie. a planar display whose angle changes such that all fingers stay behind it at all times).

3.5 Workspace Styles

To investigate how to organize HPUIs which take advantage of the hand's inherent anatomy and dexterity, we identify two broader workspace styles:

Discrete Workspace: Interfaces made up of individual elements which are unrelated to each other. The individual elements' positions

[1] Mixed Reality Toolkit (https://github.com/microsoft/MixedRealityToolkit)



	Output Space	Input Space	Display location	Frame of Reference
	Non-Interactable*	Off-fingers	Around the hand	Palm
	Interactable	Off-fingers	Between fingers	Multiple Fingers
	Interactable	On-fingers	On finger	Individual Phalanx
	Non-Interactable*	On-fingers	On palm	Palm
	Interactable	Off-fingers	Above fingers	Whole Finger

Table 1: Visual representation of the HPUI design dimensions “Location” and “Frame of Reference”, with the table providing an overview of the taxonomy. (* = Though this space can be reached using other fingers, since this work is concerned with single handed thumb to finger interaction, this is considered as a non-interactable space)

and distances relative to one another can be laid out freely without loss of information (ie. the home screen of a smartphone).

Continuous Workspace: Interfaces whose elements are intrinsically linked to one another and cannot be arbitrarily rearranged or separated. The spatial positions and distances between such elements purvey important information (ie. a map application or video player).

For a HPUI, a discrete UI is most simply thought of as an icon-based layout where icons are either fastened to the skin (ie. along the fingers, palm, etc.) or which hover in midair with positions that are fixed relative to a landmark on the hand. We explore performance in such workspaces in Sect. 5. What a continuous display should look like for such a UI is less clear. The most obvious option is a simple, planar continuous display (similar to a flat smartphone screen) which stays anchored to the hand and can be pierced by the thumb and fingers for interaction. Alternatively, since the output space of the hand is dynamic, we also consider extending this dimension of the design space to include a deformable surface which moves along with the hand itself. Since we know that users typically prefer one-handed mobile interaction [32], a deformable display could offer the benefits of a larger display space while keeping all points of the continuous surface reachable with one hand (and affording more tactile feedback during interactions). We further discuss continuous workspaces in Sect. 6.

4 Apparatus

Given the current state of hand-tracking algorithms on HMDs being error-prone under self-occlusion, we instead resorted to studying our HPUI designs using a Vicon system and marker glove. This was used only for a proof-of-concept as it provides the necessary fidelity for our investigation. We expect our results to apply when HMDs attain the degree of fine hand-tracking needed for a self-contained solution. All our demos (see video) where done using the Oculus hand tracker.

We used 13 Vicon motion-capture cameras streaming to the Nexus 2 software to track finger movements using a felt glove and IR markers (see Fig. 2). The joint positions from the tracked skeleton on Nexus 2 were then streamed to Unity in real time, and were used to update the joint angles of a 3D hand model that we constructed in Blender. The hand model scales and stretches itself to match the hand of the participant wearing the glove. The result of this setup is a virtual 3D hand that mimics the user’s hand motion in real time. The position of the Oculus Quest headset was also tracked using Vicon in a similar manner. This allows users to see the hand in virtual reality at the same distance, angle and viewpoint as they would their real hand. In Study 2 we used an Oculus Link that allowed running our experimental application on the PC instead of directly on the headset.

Interaction with interface elements was detected using collisions between GameObjects in Unity. Colliders allow detection of interactions with elements that are off-skin (near the hand but not directly on-finger). A sphere was used as the collider on the thumb’s tip, and a rectangular cuboid was used as the collider on different elements/targets.

To explore the design space of hand-proximate user interfaces where the hand is used as a joint input/output space, we ran two exploratory studies. The first study involves repeated target-selection on discrete

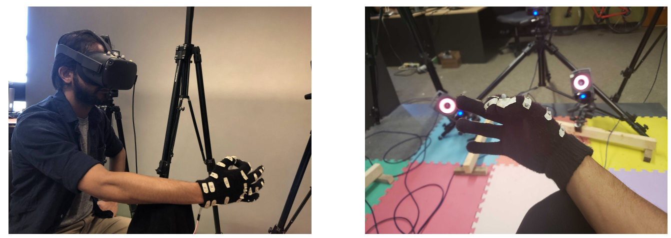


Fig. 2: The tracking setup consisting of 13 Vicon cameras and a felt glove with IR tracking markers attached.

icons fixed on and around the fingers. In our second study we explore the selection of targets on a continuous surface. We additionally investigate the performance of continuous input on such a surface.

5 Study 1: Discrete Target Selection

In prior work [6,8,9,13,20,21,29,43,56,60,63,66], on-hand interaction was used as an indirect input device (in which the visual space is disjoint from the input). Proprioception could allow for more accurate eyes-free interaction than with an external device [34, 56, 61]. This is also a natural result of the input device having nerve endings, where we can feel the location of touch events on the input surface. With our proposal of HPUIs being used as direct input spaces, the aim of this study is to explore the performance differences between direct and indirect input, ie. when the display is directly on the hand vs. away from it. Considering deformations the hand undergoes and proprioception, it is not obvious if displaying directly on the hand would have an advantage. The indirect condition is used as a baseline to inform if there is performance difference when the location of the target is visible on the hand, instead of relying on proprioception or muscle memory. In particular, we look at these differences when comparing on-finger and off-finger input spaces, as we anticipate the lack of tactile feedback for off-finger input lending itself better to direct visual input. We focus on discrete target selection in this study.

5.1 Participants

Eight participants volunteered for the study (2 Female, age between 21 and 38 ($M = 28.6, SD = 6.47$)). Four participants did not have any prior experience wearing a VR headset. Two participants were left handed, one of them uses the right hand when interacting with smartphones and the other uses both hands. None were color blind.

5.2 Task and study design

To facilitate a target selection task, we lay out the interface using an icon-based format as shown in Fig. 5c. A total of 22 targets are placed on and around the hand as interface elements that can be interacted with: one target on each of the 11 finger segments (phalanx) (targets 2-12 in Fig. 5c), and 11 off-finger targets around the finger (anchored into the corresponding phalanx’s frame of reference) (targets 13-23 in Fig. 5c). We avoided placing any targets between the proximal phalanges as it could require extreme flexion of the fingers for the thumb to reach that space. Pilot studies showed that not all participants were able to reach

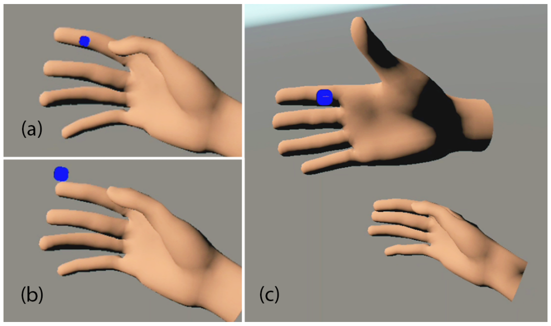


Fig. 3: Screen captures from Study 1. (a) The direct input mode, where an on-finger target appears directly on the hand model driven by the Vicon system. (b) An off-finger target being displayed in the direct input mode. (c) The indirect input mode, where the targets appear on a static hand model that is fixed in the participant's field of view. The hand model controlled by the user is visible at the bottom but does not display the target on it.

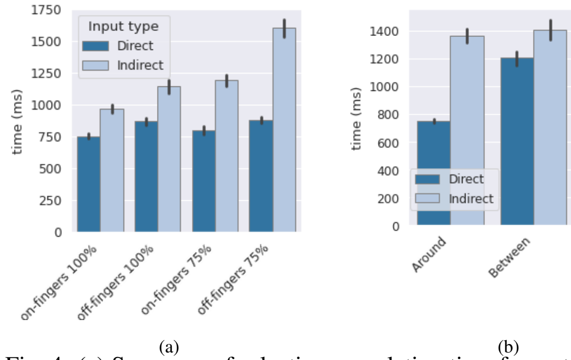


Fig. 4: (a) Summary of selection completion time from study 1. For each condition (b) Summary of breakdown of the off-finger target performance; where "Around" targets are targets 13-17 from Fig. 5c and "Between" are targets 18-23 in Fig. 5c. Standard error bars are displayed for both.

the off-finger region beyond the pinky finger, and thus no targets were placed there.

Users complete the task as follows. We position the "Start" button at the base of the user's index finger (targets 1 in Fig. 5c), which disappears when selected, causing a target to appear at random. Once this target is successfully selected, it disappears and the "Start" button reappears. If the user misses a target, they continue to attempt selection until they are successful. This process repeats until each of the targets have been selected across all trials. The "Start" button is at the base of the index finger to ensure that targets on and around the fingers are not occluded by the thumb when the trial begins.

The experiment is designed with the following conditions: (1) *Input Mode*: direct or indirect, (2) *Button Size*: 100% or 75% of finger width, (3) *Location*: on-finger or off-finger. In the direct condition the user directly controls the virtual hand using their own hand, and the targets are placed onto the corresponding fingers (Fig. 3a). In the indirect condition the user still controls a virtual hand using their own hand (and are able to see their movements in real time), but the targets are placed on a second, static virtual hand model (Fig. 3c). The user is to note where the target lies on the static hand, and uses the thumb to select that spot on their own dynamic hand. The purpose of these two visual feedback conditions is to investigate whether using the hand as a joint input/output space really provides any benefit. The described indirect condition is treated as an ideal indirect condition: The participant does not have to imagine where on the hand the target is and their own hand is in view; ie, the only difference between the direct and indirect conditions is if the targets are being displayed on the virtual hand that functions as a proxy to the users physical hand. The on and off-finger targets (Fig. 3b) were treated as separate conditions in order to investigate how the different input spaces (Sect. 3.2) affect performance.

We record the time to select a target (*Completion Time*), as well as how far off the target center selection occurs *(Selection Spread)*. To get a better understanding of the accuracy of interacting with each of the aforementioned locations, for each target, the positions within a finger segment (ie. between two finger joints) are randomized along the finger (y-axis in Fig. 5). For the off-finger targets, the average range of the randomization of the adjacent on-finger targets are used. The "Start" button is always in a fixed position. Half the participants started with one of the two Input Modes. All other conditions were randomly presented. This resulted in a $2 \times 2 \times 2$ factorial design. Each condition was repeated 3 times, resulting in a total of 8 conditions $\times$ 11 targets $\times$ 3 trials $= 264$ trials per participant. This resulted in a total of 2112 trails collected from 8 participants.

5.3 Procedure

Each participant is seated and wears the tracked glove (Fig. 2). A calibration routine is performed with the Vicon for accurate tracking of the hand skeleton. We explain the study procedure and the participant then wears the sanitized Oculus headset that is running our Unity application. Throughout the study, they are asked to keep their hand fixed on the stand in front of them to ensure it stays within the Vicon's capture volume, the hand posture was not controlled in any other way. For each visual feedback condition, the participant is presented with a set of practice trials to gain familiarity with the system. They start the actual trials when comfortable. At the end of each condition, they rank the Input Modes on a Likert Scale, with 1 being very uncomfortable and 5 being very comfortable, as well as their preference ratings. Participants are allowed breaks at any time during the experiment, as long as they occur prior to hitting the "Start" button.

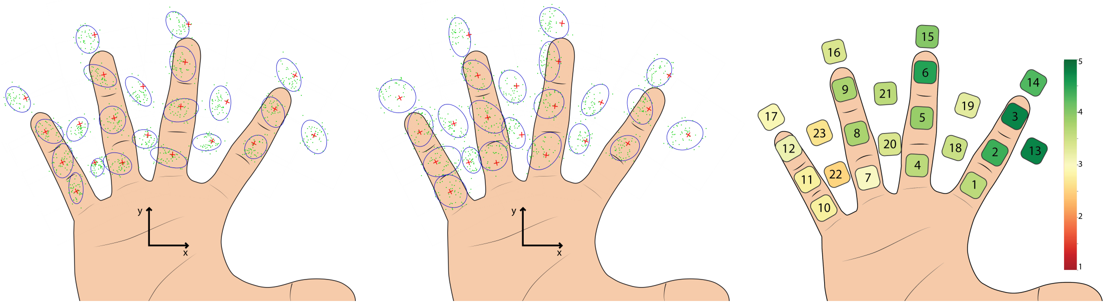
5.4 Results

Of the 2112 total trials, 33 trials were were outliers (1.5%), as defined by 3 std. deviations from the mean completion time. We removed these outliers from our analyses.

5.4.1 Completion time

Fig. 4a shows average Completion Time across all conditions. As the data met the normality and homogeneity of variance assumptions, repeated measures ANOVA analysis on mean completion time was performed. We observed a main effect of Input Mode ($F(1,7) = 21.945, p < 0.01$), Location ($F(1,7) = 13.069, p < 0.01$) and Button Size ($F(1,7) = 30.059, p < 0.001$). The Direct condition (mean=823 ms, se=12.6 ms) showed a significantly shorter average completion time than the Indirect condition (mean=1224 ms, se=25.1 ms). Average completion time On-Finger (mean=926 ms, se=16.6 ms) was significant lower than when targets were Off-Finger (mean=1119 ms, se=23.21 ms). Performance was more efficient with the 100% width (mean=931 ms, se=17 ms) targets in contrast to the 75% width targets (mean=1113 ms, se=22.9ms). With pairwise comparisons using the Bonferroni adjustment, we also observed a two-way interaction effect for Input Mode $\times$ Button Size ($p < 0.01$). There was no significant difference between the target widths in the direct condition (75% target width (mean=837 ms, se=18.8 ms) & 100% target width (mean=810 ms, se=16.8 ms)), whereas in the indirect condition (75% target width (mean=1394 ms, se=40.34) & 100% target width (mean=1054 ms, se=29.2)) we observe a significant difference. We did not observe any other 2-way or 3-way interaction effects.

Our initial hypothesis was that the performance of off-finger targets would not be significantly different from the on-finger targets in the direct condition. To further understand this discrepancy we compare the performance of the off-finger targets by grouping them based on if they are between the fingers (excluding the thumb) (targets 18-23 in Fig. 5c) or around the 4 fingers (targets 13-17 in Fig. 5c). The summary of the results can be seen in Fig. 4b. What we observe is the around finger targets performed comparable to the on-finger targets. This partially supports our initial assumption. Since we did not control the posture of



(a) Direct Input Mode (b) Indirect Input Mode (c) Average comfort rating for targets at each location.

Fig. 5: (a) and (b) shows the distribution of contact points for each target location. The green dots are the center of the contact. The blue ellipses are fitted to the points belonging to each location. The red cross is the center of the target location, oriented along the x-axis and y-axis of the target (y-axis being along the finger). (c) shows the average comfort rating for the targets at each location. Note that the open fingers in the figure is for illustration only.

the hands, the exact cause of the poorer performance of the between finger targets still remains an open question.

5.4.2 Selection Spread

We analyzed the selection spread for each target position, which was motivated by our expectation that there might be a bias towards the right side of the targets [30]. We did not observe this bias in the collected data. The selection spread is the distance measured in millimeters, from the center of the target to the collision center, on the targets plane. Fig. 5 shows the summary of the selection spread for each target location. There was no significant difference among the mean spread for both Input Modes. There was only a trend showing, the mean distance and the standard deviation were smaller in the Direct Input Mode (mean selection distance 2.9mm and mean of standard deviation across target locations = 6.1mm) compared to the Indirect Input Mode (mean selection distance 3.3mm and mean of standard deviation across target locations = 8mm). Also, on average, the point of interaction is skewed towards the left (negative direction along x-axis in Fig. 5) (mean x of x-axis $= -1.6mm$). Note that this cannot be interpreted as a measure of accuracy as the target location is recorded upon successful selection and not the first attempt at selection.

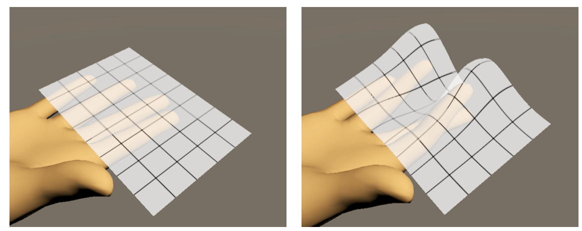
5.4.3 Comfort and convenience

The average user rating of each target location is shown in Fig. 5c. The average rating of on-finger and off-finger targets were 3.65 and 3.4, respectively (no sig. difference). The findings for on-finger closely resemble that of Huang et al. [29] and Dewitz et al. [11]. We notice that the more critical factor is not whether the targets are on- or off-finger, but rather how close they appear to the thumb. These results relate to the thumb reachable areas as regions closer to the pinky finger are more difficult to reach than the index or middle fingers, for example.

The overall comfort and convenience ratings between the direct and indirect visual feedback conditions were compared using a Wilcoxon signed rank test. There was a significant difference between the direct (comfort mean=6.5, convenience mean=6.3) and indirect (comfort mean=4.6, convenience mean=4.6) conditions for both overall comfort ($Z = 0.0, p = 0.01$) and overall convenience ($Z = 2.5, p = 0.03$). Additionally, all except one participant ranked the direct condition as the most preferable input mode.

6 Design of a Continuous HPUI

Informed by Study 1 results, we now turn our attention to designing a one-handed continuous HPUI. The constant deformations of the hand requires the continuous interfaces to be considered separately (Sect. 3.5) and is necessary for a more complete UI, which would otherwise be limited to having only discrete input elements. We specifically aim to optimize the degree of tactile feedback, as well as the space available for displaying content.



(a) (b)

Fig. 6: The deformed display. (a) Initially the hand is held flat for calibration. The points of the fingers and the points of the display mesh, relative to the hand, are used for calculating the deformations. (b) Once calibrated, the position of the display mesh is updated based on the movements of the hand segments.

6.1 Continuous HPUI Considerations

When considering a continuous HPUI that can be directly interacted with only a single hand, the display must be thumb-reachable. Prior work [11,29] suggests that the dynamic region on and around the fingers (especially, index and middle fingers) should be the most comfortable locations for one-handed direct input. For a continuous interface, this further prompts the question of what should happen to the interface as the fingers move. There are different ways in which continuous interfaces can be implemented with HPUI. Similar to smartphones [26] and tablets [44], planar surfaces anchored to the fingers or the palm (see Sect. 3.4) would suffer from the same issue of the limited thumb reachability. Not being able to easily slide such a planar virtual surface around the hand, like one could do with a smartphone, further compounds this problem. We observe that an interface where the display deforms along with the overall shape of all four fingers while they move (Fig. 6) could be a better solution. This interface not only offers the most opportunity for tactile feedback, but also allows for intuitive interaction with the display (ie. a target on the tip of the index will remain there as the index bends toward the thumb for a touch event). In any other case, interface elements are at best loosely anchored to certain parts of the fingers which could be confusing if we want to take advantage of proprioception and the hand's natural anatomy. Hence, we focus on the deformable surface for study 2. Another important consideration is the fact that we only obtain a small space for interaction if only the space on the fingers are considered. Since results show interacting with elements closer to the pinky finger are less desirable, it is worthwhile exploring the use of the thumb-reachable space between the index finger and thumb to expand the available UI for interactions.

6.2 Implementation of the Deformable Continuous Display

Our deformable HPUI (see Fig. 6a) was designed on the premise of using a dynamic deformable surface (the hand) as a joint input and output space. Our deformable HPUI also offers the advantage of being able to anchor key points of interest such as icons, sliders, etc. to fixed points on the fingers without using a strictly discrete interface. Important elements are able to stay anchored while the rest of the continuous display deforms around them.

The deformable UI is built using thin-plate spline interpolations [5] in three dimensions[2]. To begin using the interface, the user flattens their palm with the fingers together and begins the calibration routine. This stores the calibration x/y/z positions of each of the finger joints in the coordinate system of the palm and corresponds with a planar display which then appears covering the fingers. In each subsequent frame, the x/y/z deviations ($\Delta x, \Delta y, \Delta z$) of these finger joints from their calibration coordinates in palm-space are used to create three separate thin-plate splines. Each of these splines can be thought of as the least bent surface which passes through the corresponding displacement value (ie. $\Delta x, \Delta y$, or Δz) for each hand keypoint $P(x,y)$, and is governed by the equation:

$$f(x,y) = a_1 + a_2 x + a_3 y + \sum_{i=1}^{n} w_i r^2 log(r) \quad (1)$$

where a_1, a_2, a_3 are coefficients which define the closest flat plane to our desired interpolation, r is the distance between a calibration point and the corresponding input point along the relevant dimension, n is the number of keypoints, and w_i is an unknown coefficient which must be solved for. Note that since we calculate three splines independently (one for each dimension), the distance r in this case actually just refers to the corresponding 1-dimensional distance between a given keypoint and its calibration position in palm space (Δx, Δy, or Δz).

In practice, the splines are actually calculated using the alglib.net[3] library in C#, which is passed a list of constant calibration coordinates (x_{cal}/y_{cal}) along with the relevant list of displacements from calibration for a particular spline ($\Delta x, \Delta y$, or Δz). Once calculated, we are then able to map each of the vertices making up the planar display to a new coordinate representing smoothly deformed display which follows the movement of the hand and fingers as follows:

$$\begin{aligned} x &= x_{cal} + \Delta x_{int} \\ y &= y_{cal} + \Delta y_{int} \\ z &= z_{cal} + \Delta z_{int} \end{aligned} \quad (2)$$

where $\Delta x_{int}/\Delta y_{int}/\Delta z_{int}$ are the spline-interpolated values for the displacements from calibration at a given x_{cal}, y_{cal}. If strict enough interpolation parameters are chosen, then for the keypoints used to build the spline $\Delta \sim \Delta_{int}$, and the display should stay fastened to those points as they move through space.

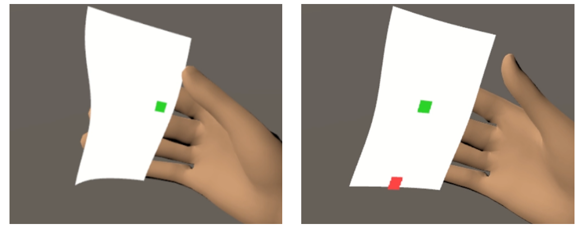
The implementation of Study 1 is similar to the system described in Sect. 4 and Sect. 5.2. Each individual pixel on the surface was made to be a object with colliders, which the thumb can then interact with. The number of pixels was made to be equal to the number of vertices on the deformable surface. Each of these pixels' positions and rotations follow the position and normal of a corresponding vertex on the surface. The targets used throughout this study were squares made from a grid of 3×3 pixels. To avoid the bottleneck created by the computational cost of the splines described above we used the Oculus link where the computation is done on the connected workstation.

7 Study 2: Evaluating HPUI on a continuous workspace

We evaluate our deformable HPUI on a continuous workspace. We specifically explore how different regions of the surface perform in selection tasks (ie. discrete elements on a continuous deformable surface), and how easily the thumb can drag content (ie. continuous input on a continuous deformable surface) on the HPUI. We also compare performance when interacting with content over hand region (including the index, middle, and ring fingers) to the above-hand region of equal size. We separate this exploration across two tasks.

[2]Our method was partially inspired by Dr. Herve Lombaert's webpage, which can be found at: https://profs.etsmtl.ca/hlombaert/thinplates/

[3]https://www.alglib.net/



(a) (b)

Fig. 7: Implementation of study 2. (a) The selection task (task 1): The participants have to select the green target. (b) The drag task (task 2): The participants have to select the green target and drag it to the location of the red target.

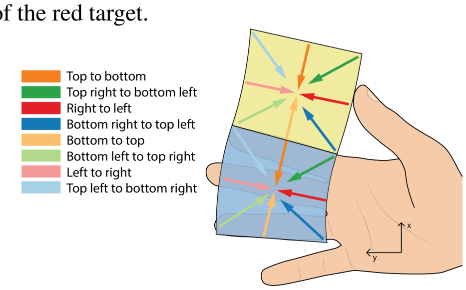


Fig. 8: The layout used for Study 2. The yellow region is the Above Index region, the blue region is the Below Index region. The arrows represent the directions used in task 2 of Study 2 (see Sect. 7.3).

7.1 Participants

Eight (8) participants volunteered (2 Female, age between 21 and 38 ($M = 28.8, SD = 6.0$)). Three participants did not have prior experience wearing an Oculus VR headset. Also three participants were left handed, one of them uses both hands to interact with their smartphone and the others use both hands. None were color blind.

7.2 Task 1: One-handed target selection on a continuous deformable surface

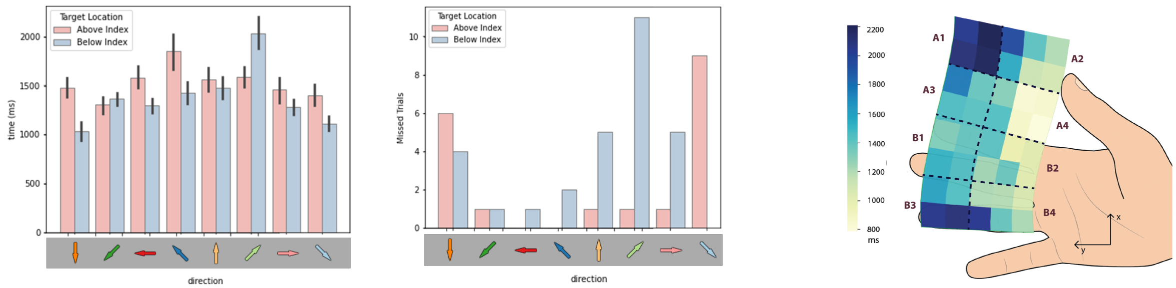
While we examined target selection of discrete items in Study 1, here we investigate selection of items on a continuous workspace. For each trial the participant initiates the task by moving their thumb to the "Start" button, that appears floating above the base of the index finger. We place the "Start" button at this location to avoid the thumb occluding targets on the workspace. Participants are to select the green target with their thumb, the target turns blue for feedback, and then the "Start" button reappears.

To ensure that the targets appear uniformly across the entire display surface, the display is divided into a 5×8 grid. For each grid cell, participants will be presented with at least 3 trials, where the targets are placed randomly within the cell. Targets are of size 3×3 pixels. Additionally, the surface is placed such that half of the cells are covering the fingers (where tactile feedback can be utilized), and the other half of the cells are in the space between the index finger and thumb, where the interactions take place off-fingers without tactile feedback. We refer to these regions as Below Index (finger) and Above Index respectively (see Fig. 8). To avoid the task being hindered by the participant's thumb range-of-motion, if the participant fails to reach a target within 5 seconds the trial registers a time-out and another trial is placed in queue, in a cell for which the participant has not failed a trial. The size of the surface is relative to each user's hand (width of their index finger), which we obtain during a calibration stage.

Task 1 included the following conditions: (1) *Target Location*: Below or Above Index Finger, (2) *Target Position*: Position in the 5 x 8 grid. Each participant performed 3 trials per grid, resulting in 120 trials for task 1 (3 trials $\times$ 40 grid cells).

7.3 Task 2: One-handed continuous input on a deformable surface

In the second (dragging) task, we explore whether continuous input is impacted by the deformable surface. Given the anatomy of the



(a) (b) (c)

Fig. 9: Results for study 2 task 2. (a) Average time for each direction (b) Number of missed trials (c) The results of the selection task showing the average selection time for each grid cell in milliseconds and the subdivisions of the grids used in our analysis, see Sect. 7.5.1

thumb we expect to see directional preferences. In line with task 1, we compare the above and below-index regions to test whether access to tactile feedback has an impact on performance. Each trial starts with the deformable display showing both a green and a red target. The participant must select and drag the green target across the surface to the red target's location using the thumb (Fig. 7b). The participants perform trials for 8 different directions: the two 45 degree diagonal directions, up to down, left to right and their inverse directions in both regions (Above Index and Below Index) as seen in Fig. 8. The distance to drag is fixed at 15 pixels in all conditions. If the participant fails to reach a displayed green target within 5 seconds, the trial is removed under the assumption that the user cannot reach that region. Each dismissed target is added back to the cue for a maximum of 2 other appearances. We record dragging time, number of attempts to reach a missed target, and unselected targets.

Task 2 included the following conditions: (1) *Target Location*: Below or Above Index Finger, (2) *Dragging Direction*: One of 8 possible directions. This results in 16 conditions (2 target locations $\times$ 8 directions). For each condition, a participant performed 5 trials, resulting in $16 \times 5 = 80$ trials per participant.

7.4 Procedure

All participants completed the tasks in order. They first performed the Vicon calibration process to ensure accurate hand skeleton tracking. Then they were asked to wear the sanitized Oculus headset connected to a PC with our Unity application running. Similar to Study 1 (Sect. 5.2), they were presented with a set of practice trials. Before each set of trials, the participant was asked to keep their fingers straight for the calibration of the deformable surface (as in Fig. 6a), described in Sect. 6.2. The participants were allowed breaks between each set of trials. Once all trials were completed, participants completed a short survey.

7.5 Results

7.5.1 Task 1

A total 1113 number of trials were recorded, out of which 39 were missed and thus not included in the analysis. 5% of the valid trials were outliers and were excluded from the analysis. Fig. 9c shows the mean selection time for each cell. An ANOVA test to assess the performance of the Above Index region with the Below Index region showed no significant effect ($F(1,16) = 1.43, p = 0.248$). To further compare the difference in performance time seen in Fig. 9c, the grids were divided into subregions (Fig. 9c). A repeated measures ANOVA with these subregions as a factor shows a significant overall effect $F(7,64) = 21.59, p < 0.001$). A pairwise comparison with Bonferroni adjustment was conducted on the sub-regions. The largest differences were between A1 (mean=2131.9 ms, se=93.7 ms) and A4 (mean=1027.2 ms, se=38.9) (Fig. 9c). A1 has a significant difference with all other sub-regions as it had the highest average compared to other regions (A2 (mean=1415.2 ms, se=45.6 ms), A3 (mean=1501.2 ms, se=61.0ms), B1 (mean=1423.86 ms, se=44.4 ms), B2 (mean=1175.9 ms, se=35.6 ms), B3 (mean=1703.4 ms, 77.2 ms) and B4 (mean=1343.9 ms, se=50.4 ms)). The A4 subregion, which had the lowest average mean time, had a significant effect with all sub-regions except B2. Within the Below Index region, only the B3 sub-region showed a significant difference with other sub-regions in the Below Index region, with B1 and B2.

7.5.2 Task 2

For task 2, 774 trials were recorded, of which 57 were missed or not completed (participant could not reach the target). Analysis was conducted on the 717 valid trials. Repeated measures ANOVA was conducted on the mean completion time of the drag task. An overall effect of direction was observed ($F(1,56) = 9.282, p < 0.001$), but no overall effect was seen for the target location ($F(1,8) = 2.925, p = 0.12$). An interaction effect was seen between the direction and target location ($F(7,56) = 2.987, p = 0.01$). A summary of the result can be seen in Fig. 9a. Fig. 9b shows the number of trials that were not completed. The primary contributor to incomplete trials as well as high completion times are when a drag task is being initiated from one of the outer extreme target locations: bottom left (B3) to top right (B2) (mean=2036.3ms).

7.5.3 Preference

For each task participants were asked to rank their preference of interacting with the different Target Locations (Above Index and Below Index). For the selection task, 4 out of 8 participants preferred the Below Index region. For the dragging task, 5 out of 8 participants preferred the Below Index region.

8 Discussion

8.1 Interpretation of Results

8.1.1 Study 1

Overall, the results from Study 1 show that interacting with icons being displayed directly on the hand is more efficient than indirect input and participants found direct input mode to be "natural". This reflects the results from previous studies that compare direct vs indirect interactions [7, 39]. Our results further show direct visual feedback has higher performance even when tactile feedback is not present despite human prorpioception. In the indirect input mode, it is possible to rely on our proprioception to select on-finger targets, but off-finger targets lack this advantage. This is particularly evident with the targets around the fingers. One consideration of direct input is occlusion of targets by the thumb over the UI. While this is a common problem with devices such as smartphones, in VR or even MR, it is easy to solve such issues by having additional visual feedback showing the occluded targets.

8.1.2 Study 2

The results from task 1 show that when the targets are closer to the thumb, the selection times are much lower even when the targets are above the index finger. While this suggests that the thumb reach is a factor, the position of the "Start" button and the milder deformations near the base of the fingers could also be contributing factors. The results of task 2 show a stronger relationship with the thumb's reach. We see optimal performance when the drag task is initiated in one of the regions closer to the thumb (E.g. B2 or A4 in Fig. 9c) and the end target is not in one of the extreme corners of the display. Participants also reflected this in their survey responses. While some participants noted the tactile feedback providing a more natural experience, other participants cited the thumbs reach as a factor that influenced their preference for the Above Index region. Still, the effect of the deformation on the performance is unclear. Some participants noted that the target in region B3 was less reachable, which was one of the reasons some participants said they preferred the dragging on the Above Index region. One explanation could be the spline implementation itself - under extreme finger flexion the interpolation algorithm can cause the display's far edges to behave in unexpected ways and thus hinder performance in these areas. This is partially due to keeping some of the interpolation parameters loose to save computation time, and partially due to a tendency for the display edges to stay in their initial positions. The latter factor is a consequence of the display edges being located outside the tracked finger keypoints. Further iterations on a fluid and natural interpolation could likely circumvent this problem, and improve performance where reachability is an issue. Even though we chose to focus on the deformable display to enable better thumb reachability, we observe a lower performance (Fig. 9a & Fig. 9c), and at times not being able to successfully complete tasks when it involves extreme regions (Fig. 9b). It would be beneficial to explore the difference between the approaches for continuous interfaces, which would allow for a more direct comparison with results from studies exploring thumb reach on smartphones [26] and tablets [44]. A potential solution could lie in the combination of these techniques. Following our observations from both the studies, it is not obvious how previous work on optimizing single-handed input on rigid planar surfaces (eg: ForceRay [10], BazelCursor [36]) can be applied here and would require further investigation.

8.2 Applications

To justify our work and provide a basis for the development of our prototypes, we briefly outline some potential HPUI use cases along with some features and interaction methods that could be made possible in the future. We implemented all our applications using the methods above on the Oculus quest using its built-in hand tracker[4] (see Fig. 1). The hand tracking on current commercial platforms (eg: Oculus Quest, Hololens) are not robust, we use the Oculus Quest only to showcase the potential applications of HPUI. We expect that this technology will see improvements in the near future making HPUI more applicable.

One can envision novel features and interaction techniques which result from blending the massive interaction space of HMDs with a hand-proximate user interface: the user is surrounded by available applications to be "grabbed" from mid-air and made active on the HPUI for interaction.Similarly, this could be integrated into seamless interaction experiences (ie. Gluey [48], Ubi-Finger [57]) - users could grab or point at things in the room such as the thermometer or lights to 'attract' the temperature and light control menus on or around their hand. This allows for 2D interface interactions (eg: menus) on one hand and interacting with the world using the other hand such as in BiShare [67]. Such features would offer benefits boasted by HMDs while allowing for much of the physical and social comfort one experiences when using a smartphone. The advantage of having HPUI over a physical device is that the user is not burdened by having to hold a physical device, and the hand(s) with the HPUI on it can seamlessly transition between HPUI, interacting with the world and using gestures. HPUI can also be used to improve existing interactions with HMD. In the following we discuss some of these scenarios while considering the discrete and continuous interfaces and interactions.

Discrete Input/Discrete Output: With a HPUI interface available, it would be easy to view and interact with notifications (or make calls) from inside the virtual environment without breaking immersion (Fig. 1a). HPUIs could also simplify the way we interact with virtual menus and discrete icons built into existing systems.

Discrete Input/Continuous Output: The essential mobile applications should be able to carry over for use on HPUIs. One example is interacting with a continuous on-hand map in MR (Fig. 1b), where additional output is shown in the output-only around-hand region.

[4]https://github.com/ahmed-shariff/SampleHPUI

Continuous Input/Continuous Output: Continuous sliders and dials in applications like media players can be placed in the interactable space on the fingers, while the media itself is displayed in the around-hand output-only region. We can also imagine design applications which require a 2D colour palette which could be placed in the deformable region on the fingers which could also be used for choosing colours to be augmented on items in the room. These examples were implemented and can be seen in Fig. 1c and Fig. 1d respectively.

Finally, there are use cases where it is simply desirable to not use a physical device. Participants in [6] expressed that a hand-input system could be useful in the kitchen when looking up recipes, when it is not hygienic to touch devices.

8.3 Limitations & Future Work

8.3.1 Study limitations

Our studies, while approved by the ethics research board and conducted with care during the pandemic, could benefit from a larger sample size. Additionally, the studies are conducted in a controlled environment with the hand being rested in a fixed position at all times. This was necessary to not introduce unwanted tracking effects in our studies. The results may vary when such restrictions are removed.

8.3.2 HPUI factors

We have explored only a small subset of the factors described in Sect. 3. Future studies will be needed to further understand the optimal configurations and their combinations for different tasks and contexts. Also, this exploration focuses only in VR, but HPUIs can be used with MR applications as well. Whether the results and observations made in these studies are directly applicable to MR applications is still an open question. How HPUIs can be used in conjunction with other input modalities (such as 3D interactions) is another factor that warrants further exploration. Another factor is the fatigue of using HPUI compared to other modalities. While it is known that interacting with the hand closer to the body is less fatiguing [28], and we are drawing inspiration from common smartphone usage, further exploration is needed to quantify the fatigue of using different modalities in MR and VR.

8.3.3 Content Layout on HPUIS

When considering the hand as a joint input/output space, one realizes that the potential input space for the hand is actually much smaller than the potential output space. The input space is mostly restricted to the fingers themselves (back and front) and the area around them for direct thumb selection. This space can be enlarged if we also consider selections made by the other fingers [49], although this for the most part is much less natural [11]. In future work we intend on exploring how best to lay out application content along the various regions of a HPUI, including output-only spaces.

9 Conclusion

The current-day landscape of HMD applications mainly use the hand as an input-only device, but the usual interactions do not scale well when considering mobile applications. Previous works using the hand as an output space typically treat it only as an anchor or tactile surface to interact with. We propose HPUIs as a solution that leverages the hands dexterity and unique anatomy as a joint input/output space, allowing users to perform single-handed input similar to their current mobile devices. We define a broader design space of HPUIs, and conduct two studies exploring this space in VR. In the first study, where we look at discrete input/output, we observe that direct input on HPUIs performs better than using the hand only as an input. In the second study we explore continuous interfaces, namely our novel deformable interface. The results show that the thumb reach is one of the main factors influencing the performance. With this knowledge we inform future research in this area, and provide an empirical framework on which further iterations of HPUIs can be based upon.

REFERENCES

[1] F. Alallah, A. Neshati, Y. Sakamoto, K. Hasan, E. Lank, A. Bunt, and P. Irani. Performer vs. observer. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology - VRST '18*, 2018. doi: 10.1145/3281505.3281541

[2] J. Alexander, A. Roudaut, J. Steimle, K. Hornbæk, M. B. Alonso, S. Follmer, and T. Merritt. Grand challenges in shape-changing interface research. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 2018. doi: 10.1145/3173574.3173873

[3] T. Azai, M. Otsuki, F. Shibata, and A. Kimura. Open palm menu. In *Proceedings of the 9th Augmented Human International Conference on - AH '18*, 2018. doi: 10.1145/3174910.3174929

[4] I. Belkacem, I. Pecci, and B. Martin. Pointing task on smart glasses: Comparison of four interaction techniques. *CoRR*, abs/1905.05810, 2019.

[5] F. L. Bookstein. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[6] I. Bostan, O. T. Buruk, M. Canat, M. O. Tezcan, C. Yurdakul, T. Göksun, and O. Özcan. Hands as a controller: User preferences for hand specific on-skin gestures. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, DIS '17, p. 1123–1134. ACM, New York, NY, USA, 2017. doi: 10.1145/3064663.3064766

[7] D. Bowman and C. Wingrave. Design and evaluation of menu systems for immersive virtual environments. In *Proceedings IEEE Virtual Reality 2001*, - 2001. doi: 10.1109/vr.2001.913781

[8] E. Chan, T. Seyed, W. Stuerzlinger, X.-D. Yang, and F. Maurer. User elicitation on single-hand microgestures. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, p. 3403–3414. ACM, New York, NY, USA, 2016. doi: 10.1145/2858036.2858589

[9] L. Chan, R.-H. Liang, M.-C. Tsai, K.-Y. Cheng, C.-H. Su, M. Y. Chen, W.-H. Cheng, and B.-Y. Chen. Fingerpad: Private and subtle interaction using fingertips. In *Proceedings of the 26th annual ACM symposium on User interface software and technology - UIST '13*, 2013. doi: 10.1145/2501988.2502016

[10] C. Corsten, M. Lahaye, J. Borchers, and S. Voelker. Forceray: Extending thumb reach via force input stabilizes device grip for mobile touch input. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, p. 1–12. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290605.3300442

[11] B. Dewitz, F. Steinicke, and C. Geiger. Functional workspace for one-handed tap and swipe microgestures. In *Mensch und Computer 2019 - Workshopband*. Gesellschaft für Informatik e.V., Bonn, 2019. doi: 10.18420/muc2019-ws-440

[12] N. Dezfuli, M. Khalilbeigi, J. Huber, F. Müller, and M. Mühlhäuser. Palmrc: Imaginary palm-based remote control for eyes-free television interaction. In *Proceedings of the 10th European Conference on Interactive TV and Video*, EuroITV '12, p. 27–34. ACM, New York, NY, USA, 2012. doi: 10.1145/2325616.2325623

[13] N. Dezfuli, M. Khalilbeigi, J. Huber, M. Özkorkmaz, and M. Mühlhäuser. Palmrc: Leveraging the palm surface as an imaginary eyes-free television remote control. *Behaviour & Information Technology*, 33(8):829–843, 2013. doi: 10.1080/0144929x.2013.810781

[14] J. Dudley, H. Benko, D. Wigdor, and P. O. Kristensson. Performance envelopes of virtual keyboard text input strategies in virtual reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 10 2019. doi: 10.1109/ismar.2019.00027

[15] B. Ens, A. Byagowi, T. Han, J. D. Hincapié-Ramos, and P. Irani. Combining ring input with hand tracking for precise, natural interaction with spatial analytic interfaces. In *Proceedings of the 2016 Symposium on Spatial User Interaction*, pp. 99–102, 2016.

[16] S. A. M. Faleel, M. Gammon, Y. Sakamoto, C. Menon, and P. Irani. User gesture elicitation of common smartphone tasks for hand proximate user interfaces. In *Proceedings of the 11th Augmented Human International Conference*, 5 2020. doi: 10.1145/3396339.3396363

[17] J. Franco and D. Cabral. Augmented object selection through smart glasses. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*, MUM '19. ACM, New York, NY, USA, 2019. doi: 10.1145/3365610.3368416

[18] A. Gomes, A. Nesbitt, and R. Vertegaal. Morephone. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, pp. 583–592, 2013. doi: 10.1145/2470654.2470737

[19] A. Gomes and R. Vertegaal. Paperfold. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction - TEI '14*, pp. 153–160, 2015. doi: 10.1145/2677199.2680572

[20] S. Gustafson, C. Holz, and P. Baudisch. Imaginary phone. In *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11*, 2011. doi: 10.1145/2047196.2047233

[21] S. G. Gustafson, B. Rabe, and P. M. Baudisch. Understanding palm-based imaginary interfaces: The role of visual and tactile cues when browsing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, p. 889–898. ACM, New York, NY, USA, 2013. doi: 10.1145/2470654.2466114

[22] C. Harrison, H. Benko, and A. D. Wilson. Omnitouch: Wearable multitouch interaction everywhere. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, p. 441–450. ACM, New York, NY, USA, 2011. doi: 10.1145/2047196.2047255

[23] C. Harrison and H. Faste. Implications of location and touch for on-body projected interfaces. In *Proceedings of the 2014 conference on Designing interactive systems - DIS '14*, 2014. doi: 10.1145/2598510.2598587

[24] C. Harrison, S. Ramamurthy, and S. E. Hudson. On-body interaction. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction - TEI '12*, 2012. doi: 10.1145/2148131.2148148

[25] C. Harrison, D. Tan, and D. Morris. Skinput: Appropriating the body as an input surface. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*, 2010. doi: 10.1145/1753326.1753394

[26] K. Hasan, J. Kim, D. Ahlström, and P. Irani. Thumbs-up. In *Proceedings of the 2016 Symposium on Spatial User Interaction - SUI '16*, 2016. doi: 10.1145/2983310.2985755

[27] H. Havlucu, M. Y. Ergin, İ. Bostan, O. T. Buruk, T. Göksun, and O. Özcan. It made more sense: Comparison of user-elicited on-skin touch and freehand gesture sets. In N. Streitz and P. Markopoulos, eds., *Distributed, Ambient and Pervasive Interactions*, pp. 159–171. Springer International Publishing, Cham, 2017.

[28] J. D. Hincapié-Ramos, X. Guo, P. Moghadasian, and P. Irani. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1063–1072, 2014.

[29] D.-Y. Huang, L. Chan, S. Yang, F. Wang, R.-H. Liang, D.-N. Yang, Y.-P. Hung, and B.-Y. Chen. Digitspace: Designing thumb-to-fingers touch interfaces for one-handed and eyes-free interactions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, p. 1526–1537. ACM, New York, NY, USA, 2016. doi: 10.1145/2858036.2858483

[30] H. Jiang, D. Weng, Z. Zhang, and F. Chen. Hifinger: One-handed text entry technique for virtual environments based on touches between fingers. *Sensors*, 19(14):3063, 2019. doi: 10.3390/s19143063

[31] A. Karlson, B. B. Bederson, and J. Contreras-Vidal. Studies in one-handed mobile design : Habit , desire and agility. 2006.

[32] A. K. Karlson, B. B. Bederson, and J. L. Contreras-Vidal. Understanding single-handed mobile device interaction. Technical report, University of Maryland, 2006.

[33] S. Kim, S. Takahashi, and J. Tanaka. A location-sensitive visual interface on the palm: Interacting with common objects in an augmented space. *Personal and Ubiquitous Computing*, 19(1):175–187, 2014. doi: 10.1007/s00779-014-0769-0

[34] L. Kohli and M. Whitton. The haptic hand: providing user interface feedback with the non-dominant hand in virtual environments. In *Proceedings of Graphics Interface 2005*, GI 2005, pp. 1–8. Canadian Human-Computer Communications Society, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 2005.

[35] S.-S. Lee, S. Kim, B. Jin, E. Choi, B. Kim, X. Jia, D. Kim, and K. pyo Lee. How users manipulate deformable displays as input devices. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*, pp. 1647–1656. ACM, 2010. doi: 10.1145/1753326.1753572

[36] W. H. A. Li, H. Fu, and K. Zhu. Bezelcursor: Bezel-initiated cursor for one-handed target acquisition on mobile touch screens. 8(1):1–22, Jan. 2016. doi: 10.4018/IJMHCI.2016010101

[37] R. Lindeman, J. Sibert, and J. Hahn. Hand-held windows: towards effective 2d interaction in immersive virtual environments. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*, 1999. doi: 10.1109/vr.1999.

756952

[38] D. Lindlbauer, J. E. Grønbæk, M. Birk, K. Halskov, M. Alexa, and J. Müller. Combining shape-changing interfaces and spatial augmented reality enables extended object appearance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, pp. 791–802, 2016. doi: 10.1145/2858036.2858457

[39] I. S. MacKenzie. *Fitts' Throughput and the Remarkable Case of Touch-Based Target Selection*, pp. 238–249. Human-Computer Interaction: Interaction Technologies. Springer International Publishing, 2015. doi: 10.1007/978-3-319-20916-6_23

[40] T. Merel. The reality of vr/ar growth. *Tech Crunch*.

[41] F. Müller, N. Dezfuli, M. Mühlhäuser, M. Schmitz, and M. Khalilbeigi. Palm-based interaction with head-mounted displays. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct - MobileHCI '15*, 2015. doi: 10.1145/2786567.2794314

[42] F. Müller, M. Khalilbeigi, N. Dezfuli, A. S. Shirazi, S. Günther, and M. Mühlhäuser. A study on proximity-based hand input for one-handed mobile interaction. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction - SUI '15*, 2015. doi: 10.1145/2788940.2788955

[43] R. Ono, S. Yoshimoto, and K. Sato. Palm+act. In *SIGGRAPH Asia 2013 Emerging Technologies on - SA '13*, 2013. doi: 10.1145/2542284.2542298

[44] A. Oulasvirta, A. Reichel, W. Li, Y. Zhang, M. Bachynskyi, K. Vertanen, and P. O. Kristensson. Improving two-thumb text entry on touchscreen devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, p. nil, 4 2013. doi: 10.1145/2470654.2481383

[45] P. Rosedale. How the new vr screen could end the smartphone. *Tech Crunch*.

[46] A. Roudaut, A. Karnik, M. Löchtefeld, and S. Subramanian. Morphees. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, pp. 593–602, 2013. doi: 10.1145/2470654.2470738

[47] S. Sekar, S. K. Vasudevan, K. Velusamy, M. Purohit, and N. Venkatapathy. Ubiquitous palm display and fingertip tracker system using opencv. *Journal of Computer Science*, 10(3):382–392, 2014. doi: 10.3844/jcssp.2014.382.392

[48] M. Serrano, B. Ens, X.-D. Yang, and P. Irani. Gluey. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '15*, - 2015. doi: 10.1145/2785830.2785838

[49] M. Soliman, F. Mueller, L. Hegemann, J. S. Roo, C. Theobalt, and J. Steimle. Fingerinput: Capturing expressive single-hand thumb-to-finger microgestures. In *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*, ISS '18, p. 177–187. ACM, New York, NY, USA, 2018. doi: 10.1145/3279778.3279799

[50] J. Steimle, J. Bergstrom-Lehtovirta, M. Weigel, A. S. Nittala, S. Boring, A. Olwal, and K. Hornbak. On-skin interaction using body landmarks. *Computer*, 50(10):19–27, 2017. doi: 10.1109/mc.2017.3641636

[51] J. Steimle, A. Jordt, and P. Maes. Flexpad. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, pp. 237–246, 2013. doi: 10.1145/2470654.2470688

[52] S. Steudel. Ar glasses have the potential to replace the smartphone within 10-15 years from now. *AiThority*.

[53] M. Sturdee and J. Alexander. Analysis and classification of shape-changing interfaces for design and application-based research. *ACM Computing Surveys*, 51(1):1–32, 2018. doi: 10.1145/3143559

[54] E. Tamaki, T. Miyaki, and J. Rekimoto. Brainy hand. In *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems - CHI EA '09*, 2009. doi: 10.1145/1520340.1520649

[55] G. M. Troiano, E. W. Pedersen, and K. Hornbæk. User-defined gestures for elastic, deformable displays. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces - AVI '14*, pp. 1–8, 2014. doi: 10.1145/2598153.2598184

[56] H.-R. Tsai, T.-Y. Wu, D.-Y. Huang, M.-C. Hsiu, J.-C. Hsiao, Y.-P. Hung, M. Y. Chen, and B.-Y. Chen. Segtouch: Enhancing touch input while providing touch gestures on screens using thumb-to-index-finger gestures. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '17, p. 2164–2171. ACM, New York, NY, USA, 2017. doi: 10.1145/3027063.3053109

[57] K. Tsukada and M. Yasumura. Ubi-finger: A simple gesture input device for mobile and ubiquitous environment. *Journal of Asian Information, Science and Life (AISL)*, 2(2):111–120, 2004.

[58] Y.-C. Tung, C.-Y. Hsu, H.-Y. Wang, S. Chyou, J.-W. Lin, P.-J. Wu, A. Valstar, and M. Y. Chen. User-defined game input for smart glasses in public space. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, p. 3327–3336. ACM, New York, NY, USA, 2015. doi: 10.1145/2702123.2702214

[59] J. Wagner, M. Nancel, S. G. Gustafson, S. Huot, and W. E. Mackay. Body-centric design space for multi-surface interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, 2013. doi: 10.1145/2470654.2466170

[60] C.-Y. Wang, W.-C. Chu, P.-T. Chiu, M.-C. Hsiu, Y.-H. Chiang, and M. Y. Chen. Palmtype: Using palms as keyboards for smart glasses. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '15, p. 153–160. ACM, New York, NY, USA, 2015. doi: 10.1145/2785830.2785886

[61] C.-Y. Wang, M.-C. Hsiu, P.-T. Chiu, C.-H. Chang, L. Chan, B.-Y. Chen, and M. Y. Chen. Palmgesture: Using palms as gesture interfaces for eyes-free input. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '15, p. 217–226. ACM, New York, NY, USA, 2015. doi: 10.1145/2785830.2785885

[62] M. Weigel, V. Mehta, and J. Steimle. More than touch: Understanding how people use skin as an input surface for mobile computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, p. 179–188. ACM, New York, NY, USA, 2014. doi: 10.1145/2556288.2557239

[63] E. Whitmire, M. Jain, D. Jain, G. Nelson, R. Karkar, S. Patel, and M. Goel. Digitouch. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):1–21, 2017. doi: 10.1145/3130978

[64] C. Winkler, J. Seifert, D. Dobbelstein, and E. Rukzio. Pervasive information through constant personal projection. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, 2014. doi: 10.1145/2556288.2557365

[65] X. Xu, A. Dancu, P. Maes, and S. Nanayakkara. Hand range interface. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '18*, 2018. doi: 10.1145/3229434.3229449

[66] Z. Xu, P. C. Wong, J. Gong, T.-Y. Wu, A. S. Nittala, X. Bi, J. Steimle, H. Fu, K. Zhu, and X.-D. Yang. Tiptext: Eyes-free text entry on a fingertip keyboard. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, UIST '19, p. 883–899. ACM, New York, NY, USA, 2019. doi: 10.1145/3332165.3347865

[67] F. Zhu and T. Grossman. Bishare: Exploring bidirectional interactions between smartphones and head-mounted augmented reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, p. nil, 4 2020. doi: 10.1145/3313831.3376233