Int. J. Human–Computer Studies 184 (2024) 103219

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

PalmSpace: Leveraging the palm for touchless interaction on public touch screen devices

Pinku Deb Nath, William Delamare, Khalad Hasan *

University of British Columbia, Kelowna, BC, Canada
Univ. Bordeaux, ESTIA Institute of Technology, Bidart, F-64210, France

ARTICLE INFO

Keywords:
Hand palm
Mid-air
On-body input
Public touch display
Touchless interaction

ABSTRACT

Touchscreen is the primary solution to interact with public devices such as Automated Teller Machines (ATMs). However, the touch modality raises health concerns since users have to touch the screens, and therefore risking the spread of contagious diseases. We design PalmSpace, an alternate input technique leveraging users' hand palms to interact with public devices. With Palmspace, UI elements are mapped onto the users' palms and can be accessed by touching various locations directly on the palm. We conduct a series of user studies to evaluate several design options, such as interface layout, item size, preferred item location, and suitable feedback for items. Based on the results, we design PalmSpace and compare its performance with mid-air input. We show that PalmSpace is a potential solution to interact with public devices without using their touchscreen. We conclude with design guidelines for using the palm as an alternative input space for touchscreen devices.

1. Introduction

Public touch interactive devices are becoming predominant in public spaces for a wide range of activities. For instance, users can purchase transport tickets in transit stations, manage financial transactions in banks, or order food in restaurant kiosks to name a few. Due to its convenience and intuitive ease of use, the touch modality is the primary input paradigm to interact with such public devices. However, there is a growing concern that public touchscreens can become hubs for various health hazards – including infectious diseases – due to the large number of users interacting with them on a daily basis (NIH, 2021). In addition, alternative forms of interaction, such as touchless methods, can lead to more durability as there are no mechanical components that users need to interact with frequently and that would degrade or be prone to damage over time. This warrants an exploration of alternative inputs to interact with public touchscreen devices.

Researchers have explored ways to shift the input space by leveraging around-device interaction. For instance, prior research demonstrated ways to use on-body (e.g., arm) as an alternative input space to interact with mobile and wearable devices (Harrison et al., 2012, 2011; Laput et al., 2014; Harrison et al., 2010). They also investigated ways to map smartphone UI elements and layouts onto users' palm (Gustafson and Irani, 2007), even using the space on and above the back of the hand as an extended input space for smartwatches (Sridhar et al., 2017). All these solutions focus on using on-body surfaces for interacting with personal devices such as smartphones or smartwatches and are not designed for public touch interactive devices. In addition, they were investigated in a context where users had to look directly at their palms. Such external instrumentation to switch the display space may not be practical for interacting with public touch interactive devices already equipped with a screen. Consequently, we envision a system where palms can be used as an alternate mirrored input space when facing toward the devices' camera, keeping users focused on the existing device's screen without the need for additional hardware, and hence limiting the changes required to switch from touch interaction to palm interaction. Here, UI elements can be accessed directly by touching various locations on the palm, making it an intuitive and convenient solution to access content on display from a distance. To the best of our knowledge, leveraging the palm to mirror a touchscreen input space and access public device functionalities has never been explored.

In this paper, we propose PalmSpace, a technique that leverages the palm of a user's hand as an alternative input space for public touchscreens interaction (Fig. 1a). With PalmSpace, UI elements are anchored on users' non-dominant hands, mimicking the actual physical screen content. Users can use their dominant hand to select UI elements. Public service terminals are now mostly equipped with cameras that face the user directly at shoulder-to-face height. Thus, the user holds their palm at approximately shoulder height to use PalmSpace, with the palm facing away from them toward the camera, and use the index finger of their other hand to select options. Consequently,

---

* Corresponding author at: University of British Columbia, Kelowna, BC, Canada.
E-mail address: khalad.hasan@ubc.ca (K. Hasan).

https://doi.org/10.1016/j.ijhcs.2024.103219
Received 11 May 2023; Received in revised form 11 December 2023; Accepted 2 January 2024

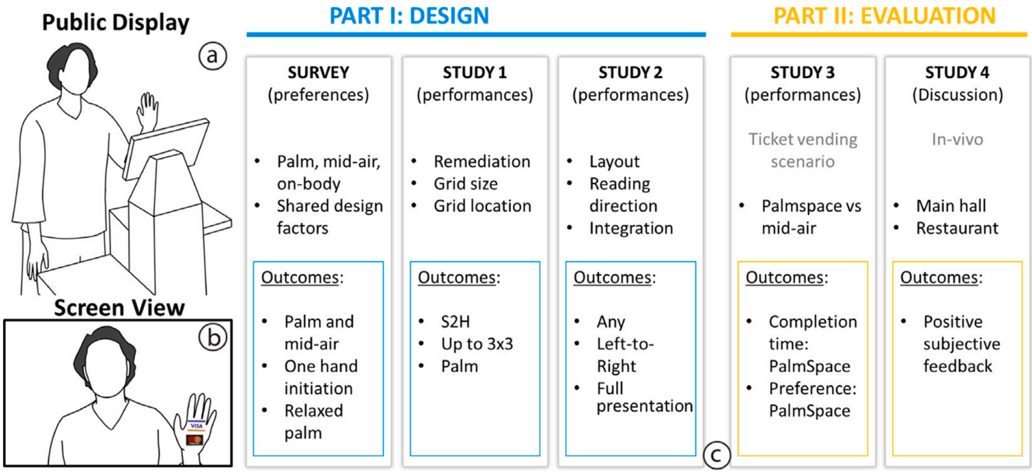

Fig. 1. (a–b) PalmSpace: an input technique leveraging users' palms to interact with public devices. (c) Design and evaluation studies.

PalmSpace can be used with current public displays without needing customized hardware prototypes. To design PalmSpace, we first conducted a survey to refine factors such as users' preferences for using different input spaces (e.g., mid-air, palm, and on-body), solutions to initialize UIs on the palm, or solutions to trigger items' selection. Results reveal that users prefer palm and mid-air over on-body locations, one-handed over two-handed initialization, and dwell and depth-based tap selection triggers over other hand gestures. We conducted two user studies to design PalmSpace, followed by two evaluative studies, in-lab and in-the-wild. The design studies examine (i) appropriate item size and location on the palm, (ii) suitable item mapping from the touchscreen onto the palm, and (iii) suitable layout to place the items. During the in-lab evaluation study, we asked participants to pretend to purchase tickets via a custom transit system vending application. Results reveal that participants were more efficient with PalmSpace than with mid-air interaction. Finally, we installed PalmSpace in two public spaces – a restaurant and a university main hall – to gather valuable feedback regarding the concept of using Palm for interacting with public touchscreen devices.

Our contributions include: (1) PalmSpace, a novel interaction technique that leverages the palm as an input space for public touchscreens; (2) an evaluation of design factors related to palm interaction; and (3) a demonstration of PalmSpace's benefits via in-lab and in-the-wild studies.

2. Related work

We review previous work that influenced PalmSpace, including mid-air, body-based and palm-based interactions.

2.1. Mid-air interaction

Prior research has shown that mid-air interactions can be a potential alternative to touchscreens, especially for interacting with mobile devices that have limited screen real estate (Hasan et al., 2017; Chen et al., 2014a; Song et al., 2014; Niikura et al., 2010b; Hasan et al., 2016). For instance, studies have demonstrated that mid-air space can be used for elementary interactions, such as item selection (Harrison and Hudson, 2009) and navigation (Butler et al., 2008). In-air space has also been utilized for multi-touch operations like typing text on devices without touchscreens (Butler et al., 2008), as well as for more complex tasks such as browsing e-commerce applications (Hasan et al., 2017). Researchers have also explored the use of in-air interactions in conjunction with other standard input methods. For example, Takashima et al. (2015) used mid-air space in combination with touch to enable map and document scrolling. This in-air space has been shown to offer benefits over traditional touch input, including minimizing screen occlusion (Ketabdar et al., 2010; Chen et al., 2014b; Niikura et al., 2010b), providing access to larger information content (Hasan et al., 2013), and allowing users faster access to items (Hasan et al., 2017). Cockburn et al. (2011) explored three mid-air pointing interfaces: Raycasting, which involved directing the pointer at a virtual screen; the 2D plane, where users were required to move the pointer across a 2D plane; and the 3D volume, where the pointer was translated within a 3D space. Results from their study indicated that the 2D plane technique outperformed the others: while raycasting was fast but less accurate, the 3D volume approach was slower and less accurate. In addition, Uddin et al. (2016) demonstrated that leveraging hands for designing new interactions is a relatively unexplored area. They also showed that hand-centric interfaces are viable, potentially quicker than conventional methods, and preferred by users. In this work, we capitalize on these benefits to propose an off-screen hand-based input technique.

Several earlier projects have leveraged different commercial solutions for tracking users' activities around the device. For instance, Vicon (Hasan et al., 2013, 2015c), OptiTrack (Gustafson et al., 2010; Jones et al., 2012; Spindler et al., 2014), Microsoft Kinect (Hausen et al., 2013), or LeapMotion (Bachmann et al., 2018; Breslauer et al., 2019; LeapMotion, 2021) have been used to demonstrate mid-air applications and use cases as they offer accurate hand tracking. In addition, researchers have used external sensors attached to mobile and wearable devices for tracking users' mid-air activities. For instance, IR sensors (Butler et al., 2008; Kratz and Rohs, 2009), omnidirectional mirrors (Yang et al., 2013), depth-sensing cameras (Chen et al., 2014a), acoustic sensors (Han et al., 2017), and radar sensors (Lien et al., 2016) have been shown to be promising solutions to track in-air hand and finger movements. We decided to prototype PalmSpace using the device's camera, as it eliminates the need for additional instrumentation since public touch displays are typically equipped with cameras.

Researchers have investigated various visual and non-visual feedback mechanisms to convey users' mid-air activities. For example, prior work explored different visual presentations of users' hand and finger movements around a smart device (e.g., smartphones) (Hasan et al., 2013, 2015c,a; Baudisch et al., 2004; Irani et al., 2006). One common approach is to dedicate a small space on the screen to show a miniature view or overview of the entire interaction space, with abstract shapes (such as a circle) representing users' hands/fingers on the overview. Contextual cues, such as arrows or wedges, have also been used to indicate off-screen items (Gustafson et al., 2008; Hossain et al., 2012; Burigat et al., 2012). Additionally, researchers have explored different parameters for output using LEDs, including intensity, direction, and distance of the light from the LEDs (Freeman et al., 2016; Müller et al., 2014; Qin et al., 2011). Non-conventional output methods, such as fog emitting from a fixed source, have also been used in unique ways to provide mid-air feedback (Martinez Plasencia et al., 2014; Rakkolainen, 2007).

Researchers have explored various methods for providing non-visual feedback for mid-air item selection (Freeman et al., 2014; Gustafson et al., 2013). For example, tactile feedback has been investigated using pressure created by synchronizing a large number of ultrasound transducers (Hoshi et al., 2010; Hoshi, 2011; Hoshi et al., 2009). Vibrations have also been explored as a channel for providing feedback, such as for typing in mid-air space (Niikura et al., 2010b). Another approach is using bursts of air vortex to provide tactile feedback on the users' hands (Sodhi et al., 2013). Researchers further investigated relevant factors such as vortex velocity, feedback delay, and dimensions of the vortex generator for effective usage of air vortex for mid-air interaction (Gupta et al., 2013). In PalmSpace, on-screen visual feedback was used to avoid the need for additional instrumentation.

While in-air interactions have shown promise, further investigation is necessary to explore users' acceptance of their usage in public and private spaces, such as transportation and homes. Consequently, researchers have conducted studies examining the acceptance of mid-air interactions in different settings, including public spaces (Jones et al., 2012; Kratz et al., 2012; Montero et al., 2010; Rico and Brewster, 2010b) and lab environments (Ahlström et al., 2014). These studies have shown that acceptance of mid-air interactions depends on gesture properties, such as gesture size (Montero et al., 2010), gesture duration, and gesture distance from the device (Ahlström et al., 2014), where less noticeable gestures tend to be more acceptable. Moreover, they found that users' decision to use gestures is influenced by how these gestures are perceived by surrounding people (Ahlström et al., 2014).

While most research on the social acceptability of mid-air input has been investigated from the users' (performers') perspective, a few prior studies have explored social acceptability from the perspective of observers who witness performers using gestures in different private and public contexts. They found that performers' perspectives of social acceptability do not always match those of observers and can even change based on input modalities (Alallah et al., 2018). Therefore, it is important to carefully consider both performer and observer perspectives when designing mid-air input techniques.

2.2. Body- and palm-based interaction

Previous work has explored body-based interaction, including touch or gestures on different body parts. For instance, Harrison et al. (2010) showed the skin to be a promising input surface for interacting with digital content projected on users. In a similar work, Weigel et al. (2014) investigated important factors for designing on-skin interactions where they revealed that the forearm and the hand are the most preferred locations. Other body areas, such as the skin surface on the arm and the back of the hand, can also be used as both input and output medium using specialized hardware. For instance, smartwatches with pico-projector have been used to project an interactive surface on the arm (Xiao et al., 2018; Laput et al., 2014), and electric field sensing used to detect gestures on the arm around the smartwatch (Zhou et al., 2016). Wearable rings emitting AC signals also used to detect touches made by the finger with the ring (Zhang et al., 2016).

Finger and palm spaces have been revealed to be promising input spaces, as a recent study showed that people were more comfortable with interaction using their fingers than using in-air gestures (Oh et al., 2020). For instance, researchers have explored the use of the palm as an additional input modality to trigger pre-defined functions (Le et al., 2018), perform 3D rotation (Kratz et al., 2012), or use it as an input space for augmenting keyboards to type on smart wearables such as Google Glass (Wang et al., 2015). Prior research has also investigated human factors related to palm interaction, such as visual dependency or memorability while accessing palm-based content. For example, Gustafson et al. (2011) revealed that the palm could be used as an imaginary interface to interact with a phone, where users could recall the phone UI leveraging their spatial memory. They also showed that users are more accurate with their palms than other surfaces such as touchscreens. Similarly, Dezfuli et al. (2014) showed that participants could touch specific points on their hand surface without looking at them.

Researchers have also investigated the social acceptance of on-body gestures (Rico and Brewster, 2010b; Ahlström et al., 2014). For instance, Rico and Brewster (2010b) conducted a survey on the social acceptability of various on-body gestures, such as touching the nose, tapping the cheek, or squeezing the forearm and found that user preference for gestures was contextual to the location and the audience around the users.

Inspired by these promising results, we focus our investigation on leveraging the palm as a surface for interacting with public touchscreen devices. More specifically, we explore a novel palm-based interaction where a user puts their non-dominant hand's palm facing toward a device camera while interacting with the palm using their dominant hand's finger. To the best of our knowledge, the use of the palm to mirror a distant screen has not been previously explored.

3. Survey

To understand how users prefer to interact with public devices using their palm as an alternative to touchscreens, we conducted interviews investigating (i) how users want to interact with public devices while using the palm as an alternative to touchscreens (e.g., number of hands for interaction, UI initialization); (ii) how to display UI elements (e.g., remediation, anchor style, preferred location); and (iii) what are their preferred selection methods.

3.1. Participants

We recruited 33 participants (11 females, mean age 30.09 years, s.d. 12.24, all right-handed) via emails, social networking websites, and word of mouth and conducted interviews online via Zoom. All participants had prior experience using public touchscreen devices such as Bank ATMs (mean 5.61 years, s.d. 2.44).

3.2. Procedure

The interviewer met a group of 1–6 participants at a time using Zoom. We used a PowerPoint presentation and a narrative script to walk participants through our design space. The scenario started with a user walking to a public touchscreen and initializing the device for interactions. Next, the scenario showcased different options for interacting with the device without using the touchscreen (e.g., selection trigger via mid-air tap). Participants could then access a Qualtrics (2021) survey to provide information about their demographic and select their preferences regarding each option. The presentation and question order remained the same for each participant and we used the following attributes to define the interaction space:

- **Initialization** refers to how the prototype should detect a participant's intent for interaction. The factors for initialization include (i) the *number of hands* (one-handed or two-handed); and (ii) the *dynamicity* (posture or gesture) to initialize a window. Based on prior work (Aigner et al., 2012; Delamare et al., 2019; Surale et al., 2019), we included *Open Palm*, *Closed Fist*, *Peace Sign*, *Relaxed Palm*, *Fist-to-Open-Palm* and *Finger Combination* (Fig. 2).
- **Remediation** Depending on the chosen visual feedback representation (Argelaguet et al., 2016), the system can impact the sense of body ownership and agency, and hence the sense of embodiment (Kilteni et al., 2012). For instance, the user interface can be shown with a superimposition of a palm illustration.

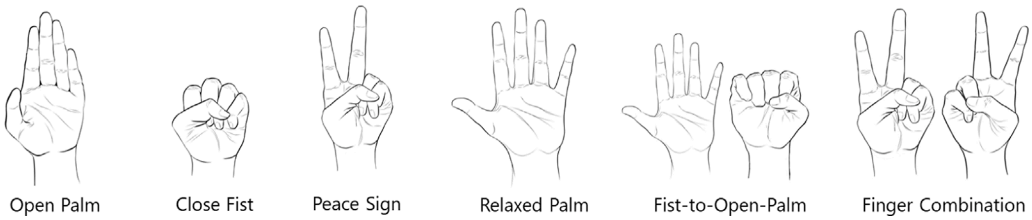

Fig. 2. Dynamicity to initialize a window.

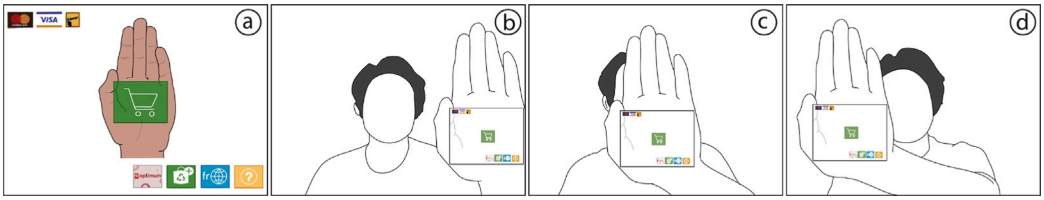

Fig. 3. Remediation. (a) Hand-to-Screen: The user interface is presented with a superimposed image on the palm. and (b–d) Screen-to-Hand with relative anchoring. With Screen-to-Hand, the user interface is virtually augmented onto the user's palm while enabling the current camera feed to be displayed and shown to the user.

We refer to these situations as *Hand-to-Screen (H2S)* remediation (see Fig. 3a). Alternatively, we can virtually augment the user interface onto the user's palm while showing the current camera feed to the user to provide a stronger sense of body ownership. We refer to this as *Screen-to-Hand (S2H)* remediation (see Fig. 3b). We can also have a *Vignette* remediation, i.e., a combination of S2H and H2S. With a *Vignette*, users see a superimposition of the palm on the screen (H2S) and a small live camera feed showing the S2H. For each remediation, we also distinguish between two control-display ratios: the palm and the screen with the same height and the screen larger than the palm.

- **Anchoring** refers to whether the UI position is absolute, i.e., fixed in a given location or relative, i.e., moves with the hand — see Fig. 3(b–d). An *Absolute* anchoring can be uncomfortable as it imposes a fixed location for interaction. A *Relative* anchoring allows users to adjust the input space to a comfortable position, providing an increased sense of agency (Kilteni et al., 2012) compared to absolute anchoring.
- **Trigger** refers to actions to validate a selection. Examples of mid-air triggers include pinch gesture with the Microsoft Hololens (HoloLens, 2021), dwell (Hasan et al., 2013), pinch (Hasan et al., 2015a), or quick finger lift (Hasan et al., 2013) to name a few. We include *Dwell* where users required to keep their finger static in a location for an amount of time, *Depth tap* which is triggered by a rapid finger movement in depth (i.e., to the camera), *Double tap* that requires the user to tap twice in mid-air, *Shooting gesture* where the user mimics a shooting pose with finger, and *Pinch* that requires the user to pinch with index and thumb.
- **Interaction Space** refers to the input area in motor space. We explore the palm space, the mid-air space (i.e., in front of the user), and the on-body space (i.e., any part of the torso).

3.3. Results

We report the results of our analysis using Friedman and Wilcoxon tests and Bonferroni-corrected p-values for each factor: *Initialization*, *Remediation* (and the CD ratio), *Anchoring*, *Trigger*, and *Interaction Space*. Note that for mean ranks, the lower the number, the higher the user preference (see Fig. 4).

Initialization: Participants prefer using one hand to initialize the interaction with the UI over two hands ($z = -5.40$, $p < 0.001$, $r = 0.88$). We also found that the dynamicity (posture/gesture) has a significant effect on users' preferences ($\chi^2(5, N = 33) = 122.28, p < 0.001$). Relaxed and open palm are the most preferred solutions, with a mean ranking of 1.70 and 1.88, respectively. All pairwise comparisons are significant except between Relaxed and Open Palm, and Peace Sign and Close Fist.

Remediation: Remediation has a significant effect on users ranking ($\chi^2(2, N = 33) = 18.2, p < 0.001$): H2S (mean rank 1.49) and S2H (mean rank 1.97) are significantly preferred over Vignette (mean rank 2.57). No other comparisons are significant. We also did not find any significant differences regarding the CD ratio (i.e., screen larger than palm vs. approximately the same height) for both S2H and H2S conditions ($z = -0.52, p = .60$).

Anchoring: For both remediation options, participants have a preference for Relative anchoring in which the UIs would move with the palm as opposed to Absolute anchoring (55% for Relative Anchoring and 45% for Absolute Anchoring). We did not observe any significant differences.

Trigger: Trigger has an significant effect on users ranking ($\chi^2(4, N = 33) = 41.63, p < 0.001$). Depth Tap (mean rank 1.81) is significantly preferred over the other trigger mechanisms. Double Tap and Dwell have a mean rank of 2.56 and 3.00, respectively, with no significant differences between them. No other pairwise comparisons are significant except between Pinch and Double Tap.

Interaction Space: We found differences between the three interaction spaces ($\chi^2(2, N = 33) = 53.70, p < 0.001$): Palm (mean rank 1.52) and Mid-Air (mean rank 1.55) are significantly preferred over Body (mean rank 2.94). No other pairwise comparisons are significant.

3.4. Discussion

One hand with a relaxed or open palm is the preferred solution to initialize interaction. Without a strong preference for using either H2S or S2H, we plan to rely on a follow-up investigation regarding performances to determine the best option. For our next steps and studies, we use relative anchoring (i.e., UIs move with the palm), palm, and mid-air input spaces as they were the preferred options for public touchscreen devices. Though participants ranked depth tap and double tap higher than other triggers, they are difficult to detect accurately without depth-sensing cameras. Since we aim for a solution requiring no additional instrumentation, we continue our studies with dwell as a selection trigger. In addition, participants expressed that they would not feel uncomfortable if their faces were visible as long as their interactions were not recorded.

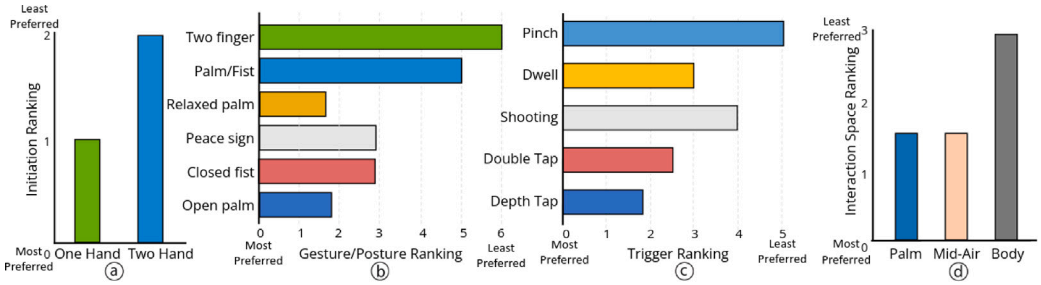

Fig. 4. Median User Preferences for (a) number of hands and (b) gestures for initiating input space, (c) gestures for triggering selections, and (d) overall preferences for input space. For ranks, the lower, the better (i.e., 1 better than 5).

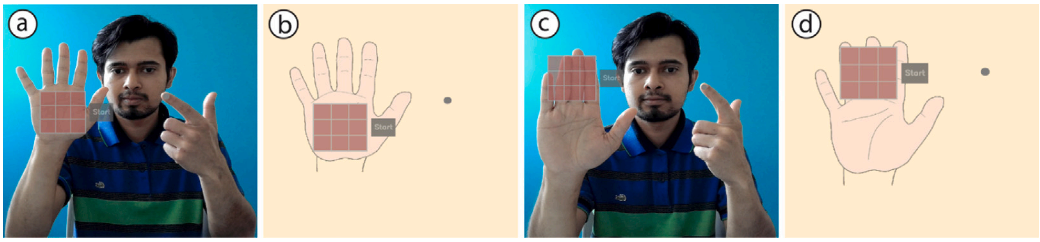

Fig. 5. Techniques: (a) S2H Palm (b) H2S Palm (c) S2H Finger (d) H2S Finger.

4. User Study 1

Our survey reveals preferences regarding the interaction initialization, anchoring style, and selection trigger. To the best of our knowledge, no previous work related to hand or palm as an interaction space investigated factors that could influence users' performance when using the hand space. Thus, we conduct our first study to quantify users' performance on remediation techniques, the number of items that can be placed on the hand, and preferred items location (i.e., palm vs. finger). Our research questions include:

- RQ1: How do the grid size and the item placement factors affect the performance of palm-based interaction?
- RQ2: What remediation technique (i.e., H2S or S2H) is suitable for palm-based interaction?

4.1. Grid size and placement

We keep the grid size bounded inside the palm or the finger area. We ran a pilot study (N = 6) where we asked participants to place 2, 4, and 6 items on the palm and on the finger area, aligning elements either vertically, horizontally and in a grid layout. Results reveal that placing 6 items on the palm or the finger area is difficult. We hence use a maximum of 4×4 grid size — resulting in grids of 2×2, 3×3, and 4×4 items. Each cell (i.e., item) is visually scaled once the hand moves closer or further from the camera.

The grid can be located in either the palm or in the finger area. The palm area covers the space situated just under the fingers and over the wrist. The finger area covers the space from the start of the fingers up to the fingertips. For consistency, we use the same grid size in both palm and fingers area, i.e., approximately the palm width and height.

4.2. Remediation technique

The survey did not reveal any clear preferences between Screen-to-Hand (S2H) and Hand-to-Screen (H2S) remediation techniques. We hypothesize that the senses of body ownership and agency might differ when actually using the techniques, not when only visible via a form. Consequently, we investigate actual users' performances with both options, S2H and H2S. In our survey, we found that participants prefer using a virtual UI that covers the entire palm and also moves with their hands. Thus, for both remediation techniques, we dynamically move and resize the UI and its element according to the user's hand movement. For instance, if a user moves their hand close to the screen, the attached UI will also get larger. However, only the on-screen representation of the palm-attached UI is larger, i.e., in the display space. UI elements remain the same size relative to the palm, i.e., in the motor space. With the S2H technique, users can directly see themselves via the camera feed, along with the palm-attached UI (non-dominant hand) and a cursor (dominant hand) to interact with the items (Fig. 5a and d). With the H2S technique, users can see an abstract hand representation along with the palm-attached UI (non-dominant hand) and a cursor (dominant hand) (Fig. 5b and c).

4.3. Participants and apparatus

We recruited 12 right-handed participants (8 males, age range between 20 and 40 years, mean age 26.83, s.d. 5.48 years) via on-campus flyers and word-of-mouth. For all three studies, we only recruit participants who have access to a laptop with a 14-inch screen. We added this, as otherwise, UI elements may not be consistent across participants' devices. All participants participated in this study had prior experience using public touchscreen devices. They received $15 for participation.

Due to restrictions regarding in-person studies, we conducted the study remotely via Zoom. We developed a web application that uses participants' device cameras to track their hand and finger movements. The front end of the web application was developed using Javascript, HTML, and CSS. The back end was developed with NodeJS. The front-end uses hand tracking models provided by MediaPipe (MediaPipe, 2021a,b), a framework with trained classifier models for detecting different body parts such as faces, hands, eyes, etc. from video streams (Zhang et al., 2020). We used their solution as it provides a reliable hand-tracking solution (e.g., an average precision of 95.7% in detecting palm (Bazarevsky and Zhang, 2023)) with minimum noise. Note that the solution has some limitations for extreme use cases where hand tracking is unreliable due to complete overlaps between both hands — which are common for any vision-based solutions. To avoid complete overlap of both hands, we instructed participants to use their left hand to point to the camera and their right-hand index finger to interact with the virtual windows projected on the right hand. We employed this approach based on the Kinematic chain model (Guiard, 1987), in which the dominant hand was responsible for precision and timing, while the non-dominant hand was used for less precise tasks (e.g., anchoring virtual windows). In addition, we asked them to close other fingers than the index finger (i.e., a closed posture with the other fingers). This approach was specifically designed to minimize hand occlusion throughout the interactions, fostering a seamless and unobstructed user experience. The models are written in C++ and compiled into WebAssembly using the Bazel build system. The compiled WebAssembly binaries are available online as NPM packages. The backend server is used to host the web application and to log study-related data in a database.

We provided participants with a link to access the web application. The application ran on participants' web browser, received the image frames from their device camera, and leveraged MediaPipe's tracking algorithms to detect hand movements. On trial completion, the web application sent trial-related data (e.g., trial time) to the database.

4.4. Procedure

During the study, participants were asked to stay connected via Zoom. They were then given the link to the web application, and the researcher ensured that participants selected the correct experimental conditions in the app options through screen share.

To start interacting with the system, participants raised their non-dominant hand and face it toward the camera while standing in front of the laptop. Then, the web application displayed a start button and a grid of items on the hand. An author running the study ensured that the device cameras were capturing users' hands when they were interacting with the content using their palms. The start button was always placed to the left-center, outside the grid (Fig. 5a–d). The goal is to keep the moving direction of the dominant hand entering the palm area from the thumb side while controlling the trial distance. To begin a trial, participants moved the cursor (index on the dominant hand) on top of the start button and performed a Dwell action to validate their selection. Next, the app highlighted the random trial target cell with a red background. Participants then moved the cursor on top of the target cell to complete the trial. The system highlighted hovered cells with a blue color. Previous research showed that dwell time of approximately 600 ms performed the best for item selection tasks (Isomoto et al., 2021). In addition, they showed that users preferred a dwell time between 500 ms and 600 ms. Therefore, in this study, we used a dwell time of 600 ms to trigger a selection. Once the trial was completed, the system highlighted the 'start' button again to launch the next trial.

We used a 2 × 2 × 3 within-subject design, with *Remediation* (H2S and S2H), *Grid Location* (Palm, Finger), and *Grid Size* (2 × 2, 3 × 3, 4 × 4) as factors. Participants had to select every cells 3 times, resulting in 12 selections for the 2 × 2 *Grid Size*, 27 selections for the 3 × 3 *Grid Size*, and 48 selections for the 4 × 4 *Grid Size*. Thus, each participant had a total of 87 selections for each *Remediation* × *Grid Location* combinations, for a total of 348 trials per participants. Participants were given practice trials to get familiarized with the techniques.

The presentation order of the *Remediation* was counterbalanced across participants. The order of *Grid Location* and *Grid Size* were randomized for each *Remediation*. Trials ended on successful selection. Participants were instructed to complete the trials as quickly and accurately as possible. We collected participants' preferences at the end of the session. Participation lasted approximately 60 min (including breaks and practice).

4.5. Results

We break down the trial completion time (from trial start to the correct selection) into two parts: (i) *Target visit time*, from the trial start to the final selection attempt, and (ii) *Selection time*, the dwell time. As the selection can be made with other trigger options (e.g., depth tap), we did our data analysis on *Target visit time*. We used a repeated measures ANOVA and post-hoc pairwise comparisons (Bonferroni α-level = 0.05) to analyze *Target visit time*, and Friedman and Wilcoxon tests with Bonferroni correction to analyze user preferences.

Target visit time (Fig. 6a): *Grid Location* has a significant impact on *Target visit time* ($F_{1,10} = 9.32$, $p < 0.05$, $\eta^2 = 0.48$). Participants are significantly faster in the Palm area (mean 1997 ms) than in the Finger area (mean 2525 ms). We also observe a significant effect of *Grid Size* on *Target visit time* ($F_{2,20} = 17.68$, $p < 0.001$, $\eta^2 = 0.64$): The 4 × 4 (mean 2495 ms) grid is significantly slower than both the 3 × 3 (mean 2100 ms) and the 2 × 2 (mean 1601 ms) grids. There is no significant effect of *Remediation* ($F_{1,10} = 1.36$, $p = 0.27$, $\eta^2 = 0.12$). In general, participants are faster with S2H (mean 2235 ms) than with H2S (mean 2270 ms). Results from the study also showed that participants take longer time if targets are located near the little finger area (Fig. 6b–d), i.e., the furthest from the 'start' button. This result confirms findings from previous work, showing that pointing performances degrade as fingers move away from a reference point when indirect feedback about the finger position is provided (Gustafson et al., 2010).

Preference rating: We collected participants' preferences using 5-point Likert scale questions. Participants rated four *Remediation* × *Grid Location* combinations (i.e., H2S-Palm, H2S-Finger, S2H-Palm and S2H-Finger) and three *Grid Size*. We found differences between the four combinations ($\chi^2(3, N = 12) = 25.49, p < 0.0001$): Between H2S-palm and S2H-palm, H2S-Finger and S2H-Palm, and H2S-Finger and S2H-finger. S2H-Palm is the most preferred technique (mean 4.75), followed by S2H-finger (mean rating 3.83), H2S-Palm (mean rating 3.08), and H2S-Finger (mean rating 2.08). We also observe significant differences among the three *Grid Size* ($\chi^2(2, N = 12) = 21.83, p < 0.0001$). We found differences between all pairs where 2 × 2 is the most preferred grid size (mean rating 4.92), followed by 3 × 3 (mean rating 3.67) and 4 × 4 (mean rating 1.83).

4.6. Discussion

Answering RQ1: Our results indicate that using the palm area is faster than using the finger area. Target visit time increases significantly for palm with the 4 × 4 grid size as a large number of items have to be accommodated on a small space. Thus, we suggest avoiding using this grid size on Palm. In addition, item selection was slower toward the pinky's side. Indeed, besides longer movement times from start to target, the hand detector seldom failed to detect both hands in case of strong overlap. Thus, items further away should be larger to accommodate both complications.

Answering RQ2: We did not see any significant difference in the Remediation. However, participants expressed their preference for using S2H-Palm and S2H-finger. Note that it could be a side effect of the novelty factor — augmented reality with S2H versus web app with H2S. However, this could also indicate that the sense of body ownership is indeed appealing to users, even if it does not have any visible effect on performances. Thus, we continue our exploration with these two options, namely S2H and Palm.

5. User Study 2

Study 1 shows that more than three items in a row affects selection time. However, we surveyed UIs for 8 public touch interactive devices (4 ATM machines, 3 ticket vending systems, and 1 food ordering self-checkout system) and observed that several UIs include four or five items in a row or in a column (e.g., CIBC ATM (CIBC, 2021), or City

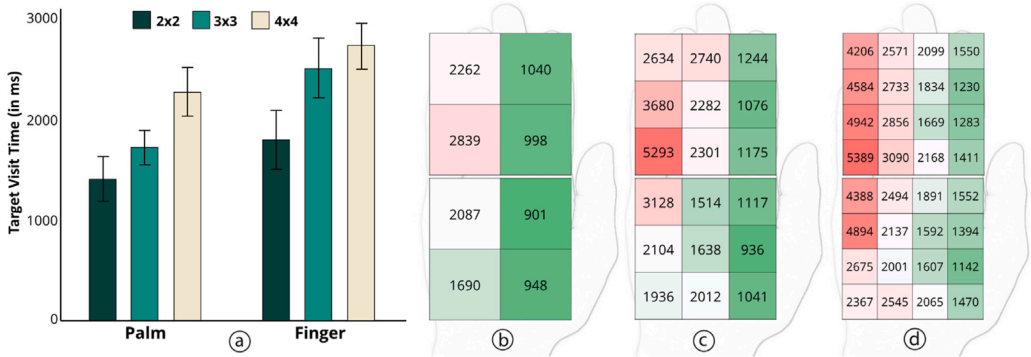

Fig. 6. (a) Mean target visit time for Palm and Finger across different Grid Size. Error bars: ±2 S.E. Target visit time in (b) 2 × 2, (c) 3 × 3 and (d) 4 × 4 grid cells.

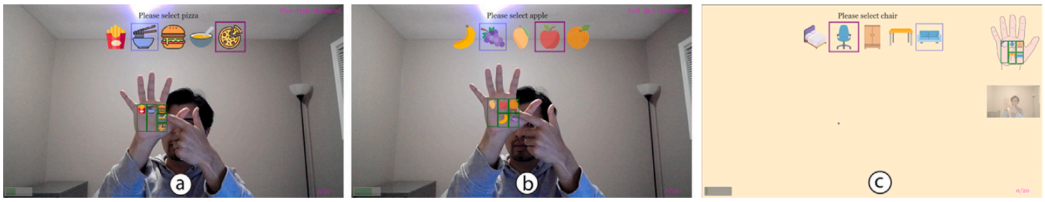

Fig. 7. Full integration (a) Left-to-Right with Flow Layout and (b) Right-to-Left with Grid Layout. Partial integration (c) Left-to-Right with Grid.

Bank ATM (CityBank, 2018)). Hence, further exploration is needed to map items from the on-screen UI onto the palm. Our research questions include:

- RQ3: How does the layout factor affect user performance in a palm-based interaction?
- RQ4: Which integration approach (i.e., Full Integration vs. other approaches) is suitable for palm-based interaction?

5.1. *Design factors*

We explore solutions to map five items onto the palm area. We consider the following design factors:

5.1.1. *Layout*

We consider two layout strategies: *Grid* and *Flow*. With the *Grid* (Fig. 7b), the system positions up to 3 items in the first row and continues placing the remaining items in rows below. The *Flow* strategy (Fig. 7a) is similar to writing on a piece of paper without correctly planning the space needed and continuing writing vertically near the edge. The system positions up to 3 items in the first row and continues placing the remaining items in the last column.

5.1.2. *Scanning direction*

The UI items can be scanned from *Left-to-Right* (Fig. 7a), i.e., the Scanning Direction in our geographic area, or *Right-to-Left* (Fig. 7b), i.e. the movement direction of the pointing finger. Indeed, for right-handed users, the right index finger enters the palm via the right side, hence creating a right-to-left scan of items.

5.1.3. *Integration*

We explore two options to integrate PalmSpace in an existing interface: *Full* and *Partial*. In the *Full* version, the original interface has to be modified to let users see themselves via Augmented Reality on the full screen (similar to S2H in our previous studies), as shown in Fig. 7a. In the *Partial* version, a static palm icon with the palm-attached UI is positioned in an empty space of the original UI (Fig. 7b). We also include the raw camera feed without any visual augmentation. We are primarily interested in the degree of efforts deploying PalmSpace would require.

5.2. *Participants and apparatus*

We recruited 12 right-handed participants (4 females, age range between 21 and 40 years, mean age 25.67, SD 5.92) via flyers and word-of-mouth. Two participants from Study 1 participated in Study 2. All participants had prior experience using public touchscreen devices (mean 9.17 years) and received $15 for their participation. The web application was similar to the one used in study 1.

5.3. *Procedure*

We conducted the study remotely via Zoom. Participants were first given information about the study and provided with the web application link. They were asked to share their device screen to ensure they selected the experimental options in the correct order on the web application.

The task consisted in selecting an item from a set of items displayed at the top of the screen (Fig. 7a). The participants were asked to stand and raise their non-dominant hand and to face it toward the camera. Then, a start button appeared near the thumb of the non-dominant hand in order to keep the movement direction of the dominant hand while controlling the trial distance. They were asked to adjust the camera in such a way that left hands were always visible in the middle of the screen. The system indicated the target item with a purple rectangular border on the original UI and a textual prompt (e.g., *Please select apple*). Participants used the live augmented camera feed with *Full* or the icon image with *Partial* to locate the item and to select it by positioning the index finger in the corresponding area for 600 ms. Similar to study 1, a cursor indicated the index position. A green rectangle indicated the hoovered item based on the cursor position. A correctly performed selection ended the trial, and showed the next trial.

5.4. *Design*

We used a 2 × 2 × 2 within-subject design with *Layout* (Grid, Flow), *Scanning Direction* (Left-to-Right, Right-to-Left), and *Integration* (Full and Partial) as factors. The order for *Integration* was counter-balanced across participants. *Layout* and *Scanning Direction* were randomized across participants for each *Integration*. Each trial involved a target

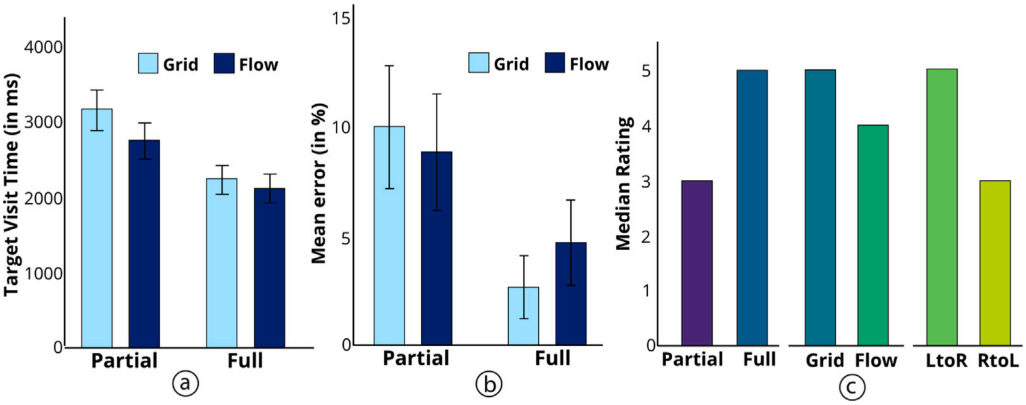

Fig. 8. (a) Mean target visit time and (b) Mean error rates. Error bars: ±2 S.E. (c) Median user ratings.

randomly selected among available items. Participants performed 20 repetitions of each *Layout* – *Scanning Direction* – *Integration* combinations, for a total of 160 trials per participant. We used eight different sets of grocery icons – one for each combination condition (Fig. 7a–c) – to prevent learning and memorizing items' positions across conditions. Participants completed 5 random practice trials before starting a new combination. We also collected subjective feedback via 5-point Likert scale questions. A session lasted approximately 45 min, including short breaks and practice.

5.5. Results

We removed 39 trials (2.03%) trials with completion times exceeding the mean time by more than 3 standard deviations.

5.5.1. Target visit time

We define the *Target visit time* (Fig. 8a) as the elapsed time between the trial start and the final attempt to make a successful selection, i.e., just before the dwell time action. We use a repeated measures ANOVA and Bonferroni adjusted (α-level = 0.05) post-hoc pairwise comparisons. *Integration* has a significant effect on target visit time ($F_{1,11} = 28.36$, $p < 0.001$, $\eta^2 = 0.72$). Full Integration (mean 2169 ms) is significantly faster than Partial Integration (mean 2963 ms). We did not find main effects of *Layout* (mean target visit time for Grid 2705 ms and Flow 2428 ms) or *Scanning Direction* (mean target visit time for LtoR 2416 ms and RtoL 2717 ms).

5.5.2. Error rate

(Fig. 8b): Error rate corresponds to the number of trials with at least one selection attempt in another cell than the target, divided by the total number of trials. We use Friedman tests with Wilcoxon tests for post-hoc pairwise comparisons to analyze error rates. The overall error rate is 6.43% (121 trials with error). *Integration* has an significant effect on error rate ($\chi^2(1, N = 12) = 12.00, p < 0.001$). Pairwise comparisons show that Full (mean 4%) has significantly fewer errors than Partial (mean 9%) ($z = -3.05, p < .05, r = 0.72$). There is no significant difference in error rates regarding *Layout* and *Scanning Direction* options.

5.5.3. Preferences

(Fig. 8c): There is a significant effect of *Integration* on users preference ($z = -2.05, p < .05, r = 0.59$): Full (mean rating 4.42) is preferred over Partial (3.08). We also observe a significant effect of *Scanning Direction* ($z = -2.57$, $p < 0.05, r = 0.74$), with participants rating LtoR (mean rating 4.50) higher than RtoL (mean rating 3.33). We did not find any significant difference between *Layout* options ($z = -0.91, p = 0.36, r = 0.26$).

5.6. Summary

Answering RQ3: Though we did not see any significant difference of *Scanning Direction* on users' performance, we saw a strong preference for using LtoR than RtoL *Scanning Direction*. We believe this is due to the fact that all our participants were native left-to-right readers. Indeed, participants commonly express a strong preference for tasks with a left-to-right direction than right-to-left (Friedrich and Elias, 2016). The Scanning Direction is hence more important than the motor actions flow used by PalmSpace.

Answering RQ4: Results reveal that Full *Integration* has a better overall performance across target visit time, error rate, and preference. This suggests that modifying the original interface to include users to see themselves on the display via Augmented Reality leads to improved users performance. We hence continue our following study with this Full Integration.

6. User Study 3

We accommodate different UIs commonly seen on self-service ticketing vending machines and compare PalmSpace performances with mid-air inputs (i.e., MidAir). Our primary research questions in this study is: (RQ5) Does PalmSpace yield improved user performance compared to MidAir techniques?

6.1. Participants and apparatus

We recruited 12 right-handed participants (4 females, age range between 20 and 35 years, mean age 26.0 years, st dev 6.26 years). All participants had prior experience using public touchscreen interfaces (mean 8 years of experience). Four participants from Study 1 and 2 participated in Study 3. Note that all the studies are conducted at least three months apart from each other. We anticipate no learning transfer of the techniques from one study to another. Participants were compensated with $15 for their time. We used the same platforms (e.g., MediaPipe) as previous studies. We leveraged MediaPipe's algorithms to detect hand movements captured via participant's device camera for both palm and mid-air inputs.

6.2. Procedure

As for previous studies, we met participants via Zoom to provide information about the study and the application link and to ensure they ran the study properly via screen sharing. We also verified that the device cameras captured their hands when they were interacting with the content using their palms. Participants were required to perform selection tasks to purchase train tickets using both palm and mid-air. We designed a set of four UIs (i.e., Main UI, Ticket Type, Payment

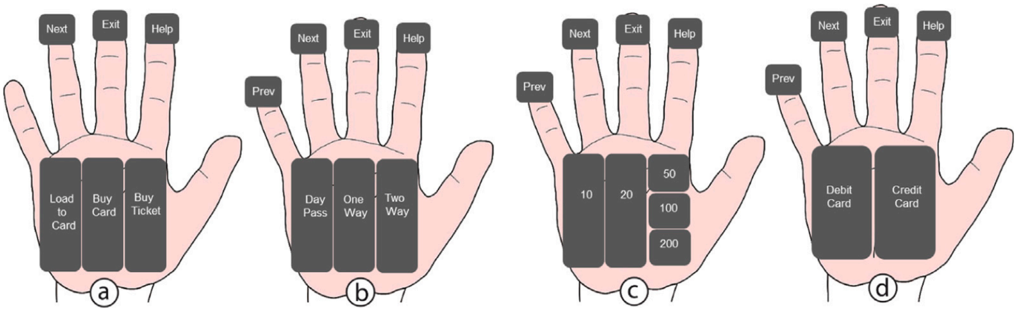

Fig. 9. PalmSpace UIs: (a) Main, (b) Ticket Type, (c) Payment Amount, and (d) Payment Method.

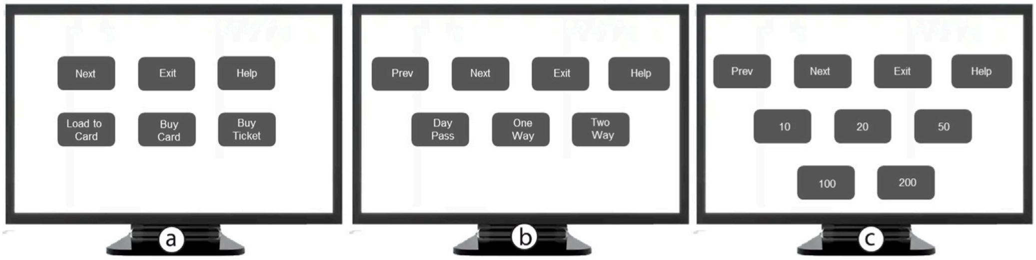

Fig. 10. UIs for MidAir Input. (a) Main, (b) Ticket Type, and (c) Payment Amount UIs.

Amount, and Payment Method) to mimic a standard ticketing vending machine application (Fig. 9). The Main UI contains the start button and the current trial instruction (e.g., *Buy Ticket for One Way Travel costing $20 using Credit Card*). The instruction is placed at the top of all UI screens throughout the trial. A trial starts by selecting the start button, which also triggers a timer and displays the contents of the Main UI. The Start button is located on the thumb for the PalmSpace and at the middle of the right side of the screen for the MidAir. The Main UI contains options, such as Buy Ticket, Buy Card, and Load to a card (Fig. 9a). Selecting an option from the Main UI opens the Ticket Type window to purchase Day Pass, One-Way, or Two-Way tickets (Fig. 9b). After the option selection, the application navigates users to the Payment amount with options $10, $20, $50, $100 and $200 (Fig. 9c), and finally to the Payment method (i.e., Credit card and Debit card) UIs (Fig. 9d). If the participant makes a wrong selection (e.g., select $10 instead of $20 in our example), the application ends the trial while marking it as an error and prompts participants to redo the trial straight away. Sequentially selecting correct items from all UIs ends the trial, stops the timer, and shows the prompt for the next trial information.

With PalmSpace, users are required to raise their non-dominant hand and face it toward the camera. Once the application detects the hand, it displays the UI onto the palm via the embedded video feed. Similar to our previous studies, a gray-colored circular cursor of radius 10 pixels (≈ 0.25 cm) was mapped to users' right-hand index fingertip. Participants selected a UI element by hovering the cursor on the palm's corresponding area for a predefined dwell time. We placed a set of standard UI elements (e.g., next or previous buttons) that are commonly used for navigation on the fingertips (Fig. 9). Note that the presentation of UI elements on the palm was different for each UI as we followed the best practices for arranging the items identified in our previous studies. For instance, with a UI containing more than 3 elements, we used the *Flow* layout to accommodate elements on the palm. Another example is the increased size allocated to elements further away from the start button to assist in faster selections (Fig. 9c). With MidAir, the UIs were placed in a fixed location at the center of the screen, mirroring the expected UI arrangement in touchscreen interfaces (Fig. 10). To leverage mid-air dwell selections, we used larger items — each item with a dimension of 160 × 80 pixels or 4.2 × 2.1 cm in a 22-inch display. Similar to PalmSpace, a cursor was mapped to users' right-hand index fingertip. Participants selected the UI elements by moving the cursor on top of an element and hovering for a predefined dwell time. For both techniques, we observed that a dwell time of 600 ms resulted in higher errors when switching between multiple windows. With a pilot study, we found that a dwell time of 1000 ms is suitable for triggering selections for both PalmSpace and MidAir techniques. Both techniques hence use a 1000 ms dwell time. For all the techniques, participants selected the UI elements by moving the cursor on top of an element and hovering for a predefined dwell time. Indeed, as for PalmSpace, we could not reliably detect the mid-air tap action via a simple webcam.

6.3. Design

We used a within-subject design with *Technique* (PalmSpace, MidAir) counterbalanced across participants. For each trial, the study application automatically created the instruction by randomly selecting an option from each UI. After 10 random practice trials, participants performed 30 repetitions with each *Technique*, for a total of 60 trials per participant. After completion of all trials, participants were asked to provide subjective feedback on the techniques using 5-point Likert scale questions. A session lasted approximately 40 min, including short breaks and practice.

6.4. Data log

We recorded the trial time, which was measured from the time users selected the start button to the time they successfully selected the target element in the last UI (i.e., complete a task such as *Buy Ticket for One Way Travel costing $20 using Credit Card* successfully by navigating though multiple UIs). An error was recorded when the participants made a wrong selection. As mentioned previously, participants were instructed to repeat a trial if they made any mistakes while using any of the user interfaces during that specific trial. Cursor movement was calculated by measuring the total on-screen path covered by the cursor during a trial. Additionally, we gathered subjective feedback from participants regarding the techniques using a 5-point Likert scale.

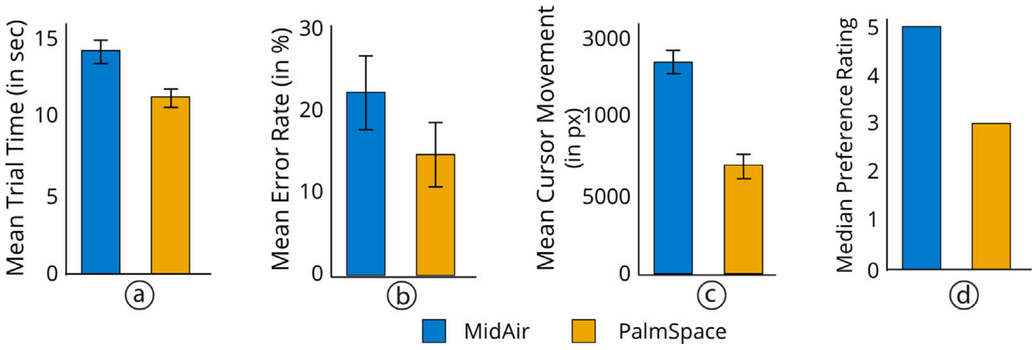

Fig. 11. Mean (a) trial time, (b) error rate, and (c) cursor movement and (c) error rate for two techniques. Error bars: ±2 S.E. (d) Median preference rating for MidAir and PalmSpace.

6.5. Results

We used a t-test to analyze the trial time. We found a significant effect of *Technique* on trial time ($t(11) = 3.04, p < 0.05$). PalmSpace (mean time 11,295 ms) is significantly faster than MidAir (mean time 14,170 ms) (Fig. 11a). We further analyzed the error rate defined by the number of error trials divided by the total number of trials (Fig. 11b). *Technique* had no effect on error rate (Wilcoxon $Z = -0.98, p = 0.32$, Palm 15% and MidAir 22% across 5 interfaces). During a trial, we also recorded on-screen cursor movement length to determine finger movement by participants for both techniques (Fig. 11c). A t-test showed a significant difference in the cursor movement length for *Technique* ($t(11) = 6.10, p < 0.001$). Participants moved the cursor more on the screen with MidAir (2670 px) than with Palm (1378 px). Participants rated PalmSpace (mean 4.33) significantly higher than MidAir (mean 3.00) (Wilcoxon $Z = -2.00, p < 0.05$) (Fig. 11d).

Subjective feedback: We also received participants' feedback on the investigated techniques. They expressed that the palm offers haptic feedback, which is useful for mapping and accessing items. Furthermore, they stated that interacting with MidAir techniques by poking in the air occasionally resulted in a sense of unnaturalness. Although our results suggest the palm is an efficient space for accessing on-screen content, two participants expressed their concerns regarding arm fatigue due to the two-handed interaction concept with PalmSpace. One participant mentioned "I do not like using two hands up at all times". Another participant stated "It was easier to use just one hand rather than two, as used in PalmSpace".

6.6. Summary

Answering RQ5: Our results show that, compared to MidAir techniques, PalmSpace significantly reduces trial completion time. With PalmSpace, items can be accessed easily by moving the index finger within a small palm area. This makes the palm an efficient area to accommodate UIs, which does not require large finger movements to access on-screen items. In addition, the palm offers haptic feedback, which is found useful for mapping and accessing the items.

7. In the wild demonstration

We set up PalmSpace in actual public spaces to collect feedback and observations in ecologically valid settings.

7.1. Conditions

We tested PalmSpace at a university's main entrance and in a restaurant. The setup consisted of a regular Dell XPS laptop with a WiFi connection, a table, and pieces of paper with "TRY ME" inscriptions (Fig. 12). We included inscriptions since there was no interactive system to replace, and perhaps a laptop alone would not have had engaged passers-by. The main screen consisted of textual inscriptions "Show your left palm toward the camera. Touch the buttons on the left hand's palm using your right hand's index finger" at the top, and the camera feed in full screen. Once a left hand is detected, the system would display a menu on the palm with four items: View Map (of the local area), Food Menu (of a nearby restaurant), Bus Routes (with timetables of the university bus stop), and Weather Today (weather summary of the current week). Once an item was selected, the system displayed the content on the right side of the screen. We did not intervene or encourage anyone to interact with our system. The program automatically logged users' interactions with PalmSpace (e.g., the number of selection events). Observations were made from a distance, and we collected feedback if users actively looked around to discuss the concept they had just used. We ran this observational study for 2 h (morning) at the university main entrance and 2 h (lunchtime) in the restaurant. Note that we did not simulate a touch alternative that could have put participants at risk.

7.2. Feedback and observations

We have 166 successful palm detection events, lasting around 8s on average. On these 166 palm detection events, we have 118 item selections ($\approx$ 0.7 selection/palm detection). This indicates that most of the users selected an item once the menu was triggered. Several participants were holding a beverage. In this case, participants left as there was no solution to put down the beverage required for our two-handed interaction.

From our observations and informal discussions, we gather that several participants successfully used PalmSpace by combining palm for menu appearance and dwell for item selection while facing the screen. Interestingly, several participants tried to perform a click action at first but quickly realized the presence of a timer at the bottom of the screen for the dwell action. We received positive feedback from users regarding PalmSpace. One participant commented that it "works really well". Several participants joined in groups, often starting by waving at the camera as they saw themselves in the live video feed. Users interested in the prototype often asked about the goals of the research project and were impressed by the fact that the application was simply running on a standard laptop in a web browser. They also expressed their interest in using such a system in public spaces. Lastly, the restaurant owner appreciated seeing so many customers engaged in this atypical activity in his establishment — which indicates a favorable reception of PalmSpace deployment in public spaces.

7.3. Discussion

Interestingly, users showed us a new way to use PalmSpace: Using the left hand to display the menu and the right hand to point in front of the palm but toward the screen. This solution is less comfortable than actually touching the palm. We believe that, instead of textual

Fig. 12. In the wild demonstration in a (a) university main hall and (b) in a restaurant. (c) A user interacting with PalmSpace.

instructions, other forms of guidance could have led to a better understanding of the pointing action (e.g., pictures or figures, or even animations). Overall, users expressed interest in using the PalmSpace-like system in public spaces. We believe the reasons include curiosity and the novelty effect of a palm-based augmented reality application. Yet, the system was engaging users without a required or necessary task. This is encouraging, and we conjecture that PalmSpace can be used in real situations with actual needs (e.g., buying a transit ticket). However, even if PalmSpace is engaging, the two-handed input can prevent interaction as soon as one hand is busy (e.g., holding a cup of tea).

8. Design guidelines

Based on the study results, we present our findings as design guidelines for palm interaction.

- *Palm vs. Midair*: We observe that participants are significantly faster with palm than with mid-air. This can be attributed to several reasons. Firstly, PalmSpace provides easy access to items by simply moving the index finger within a confined space on the palm. Secondly, users can benefit from proprioception effects. Lastly, PalmSpace incorporates haptic feedback for users during the process of accessing items, allowing them to receive feedback on their actions through tactile sensations on the palm. This reduces the reliance solely on visual feedback, which is the case with the MidAir technique. We suggest designers consider the palm as an alternate input space as selecting items in the palm area is faster than in mid-air.
- *Palm vs. finger area*: Our findings from Study 1 indicate that accessing items in the palm area is faster compared to the finger area. Touching the palm may provide increased stability compared to using fingertips, as the palm offers a larger continuous surface area that is conducive to fine motor control and precision in both discrete and continuous interactions. As a result, we recommend that designers consider placing more items on the palm area instead of the finger area.
- *Grid size and item placement*: We have observed that when using a 4 × 4 grid size on the palm and finger area, the target visit time increases. Therefore, we recommend avoiding placing four or more items in horizontal or vertical directions on the palm or finger area. For more items, designers should consider rearranging items in multiple rows. Furthermore, item selection tends to be slower toward the side of the pinky finger, which can be attributed to the limited available space on that finger compared to others. As a result, we recommend that designers avoid placing multiple items on the pinky finger in any camera-based hand and finger tracking solutions that might suffer from occlusion.
- *Presentation and Remediation*: Although no significant differences were observed in remediation, our participants showed a preference for S2H remediation technique for both Palm and finger interfaces. With S2H, user interfaces are virtually augmented onto the users' palm, offering a more immersive and interactive experience, and providing a better sens of body ownership and sense of embodiment than a palm icon on the screen (H2S). Additionally, the interfaces follow the user's hand, allowing for freedom of movement and flexibility in using the virtual interface, also increasing the sense of embodiment via the sense of agency. This enhances the overall user experience by making the interaction more intuitive and engaging. Consequently, we suggest considering S2H remediation technique for palm-based interfaces.
- *UI integration*: We observed that full integration outperforms partial integration in terms of target visit time, error rate, and preference. With full integration, users can see themselves onscreen via Augmented Reality, which creates an immersive experience that enhances usability and engagement of the interface. Therefore, we recommend utilizing this UI integration technique for palm-based interfaces.
- *Scanning direction*: Our study 2 results indicate that there is no significant difference in users' performance based on scanning direction. However, participants expressed a strong preference for left-to-right (LtoR) scanning direction over right-to-left (RtoL) scanning direction. It is important to note that our participants are native left-to-right readers, which we believe to be the primary reason for their preference. As a result, we recommend that designers consider scanning direction based on the native reading direction of their users, and not necessarily based on the actual motor movements of the fingers.
- *Discoverability*: Designers should consider providing instructions (e.g., animations or videos) to demonstrate how to initiate the interface with their palms. Once users initiate the interface, they can intuitively figure out how to use it.
- *Privacy*: PalmSpace superimposes UIs directly on the palm via the camera feed (i.e., S2H) over superimposing a virtual palm in UIs (i.e., H2S). When superimposing UIs on the palm space, and hence using the camera feed, privacy should be considered, which was also raised by a few participants during the study. Indeed, while others might not be able to see the screen and its content, any mid-air or palm-based inputs can be visible by passerby. The interface could randomize items so that inputs are not guessable via motor movements alone for instance.

9. Limitations and future work

We intentionally designed PalmSpace to be compatible with any public displays equipped with cameras, without requiring any hardware modifications. The performance and reliability of PalmSpace depend on several factors such as camera resolution, environmental parameters where the system is being used, and also the quality of the hand-tracking system. Currently, PalmSpace uses conventional webcams with hand-detection modules provided by MediaPipe. The hand-tracking modules may face limitations in implementing tap gestures with complete reliability, as they can be affected by hand and finger occlusion issues, which are commonly encountered in computer vision approaches. Future research could focus on developing techniques that are robust to environmental parameters, and accurate for hands and fingers detection. These advancements have the potential to enhance the performance and usability of palm-based techniques in real-world scenarios.

We explore options using a screen to provide visual feedback, i.e., for novice users. With a robust hand-detection mechanism, future work can also explore the novice to expert transition and the relations between memorability, items location, and palm landmarks (e.g., joints). Further research could be focused on investigating whether the need for visual guidance can be replaced by memorizing target placement patterns using fixed landmarks on the palm. The current prototype uses the left hand for anchoring the input space and the index finger of the right hand to control the cursor. Future work can explore how to include multiple cursors using multiple fingers to allow more expressive gestures like touchpads on laptops.

With the survey, we mainly focused on exploring various design factors for PalmSpace. During this exploration, we received participant feedback on specific techniques such as Depth Tap that necessitate depth cameras, while alternative approaches can be designed using traditional RGB cameras. The decision to use RGB cameras for our studies was influenced by the practicality of PalmSpace in real-world scenarios, such as self-checkout machines at stores, where any hardware modifications or changes would not be required to use PalmSpace. Consequently, all user studies were conducted using the RGB cameras available rather than employing depth cameras. Nevertheless, we acknowledge the potential for future iterations of PalmSpace machines to incorporate different cameras, including depth cameras. This consideration aligns with our anticipation that integrating depth cameras could facilitate the implementation of novel techniques like Depth Tap. Furthermore, the introduction of depth cameras has the potential to expand PalmSpace features by enabling the transition from 2D interfaces to 3D interfaces, thus broadening input and output capabilities. However, it is crucial to emphasize that delving into these possibilities would necessitate further research.

We recruited a limited number of participants for our user studies (e.g., 12 participants for each study except the initial survey). We acknowledge that larger participant pools from diverse backgrounds may better represent the variances in user performance and elucidate more diverse preferences and observations. In addition, it can lead to new data analysis regarding gender, users' palm size, and other related factors.

We would like to caution that prolonged use of PalmSpace may potentially result in arm fatigue. This can be primarily attributed to the dual-hand usage of PalmSpace, which requires increased physical effort and may cause discomfort when using both hands. Mitigation strategies such as exploring more ergonomic designs for PalmSpace, allowing users to customize the placement of items, incorporating short interactions, and enforcing breaks during prolonged use should be considered to reduce the arm fatigue associated with PalmSpace.

Further research could investigate the potential of multi-layer UIs on the palm and compare their performance with on-screen UIs in different tasks. Additionally, researchers could investigate the potential of adaptive UIs that dynamically adjust different UI elements based on the user's interaction patterns. For example, adaptive UIs could update the location and size of items based on the frequency of their selection by users, or adjust item locations based on users' interaction styles, in order to enhance usability and user experience.

From our knowledge so far, our research is among the first to explore the palm area as an input space using only a commercially available camera with no depth-sensing capabilities. We intentionally limited our prototype's capabilities in order to become relevant for a wide range of public interfaces and also to conduct remote user studies. Our research objective, in a way, perfectly aligned with our obligation to conduct user studies without any physical interactions. Hopefully, our work will help future work to further explore the interaction space of the palm.

10. Conclusion

In this paper, we took an initial step to explore PalmSpace, a viable input technique that leverages users hand palm to interact with public touchscreen devices. First, we conduct design studies to ascertain the preferences regarding different design factors for interfaces located in the palm area and expectations for alternatives to public touchscreen interfaces. Based on the design studies' conclusions, we conduct three user studies to determine which factors impact target selection times and user preferences. We also run a demo in the wild – in a restaurant – to collect users' feedback. Finally, we discuss design guidelines for palm interaction, limitations of our current work, and future work ideas.

Previous research on mid-air and on-body interaction has often utilized specialized and wearable sensors that are not feasible for regular use in public interfaces due to factors such as high costs and the need to wear additional devices. In contrast, PalmSpace aims to minimize hardware requirements, relying primarily on software-based solutions and utilizing only a camera for sensing. As advancements continue to be made in camera technology, wearable devices, and sensing technologies, it is expected that high-resolution information, such as per-pixel depth information, will become more readily available for integration into commercial products. This will eventually enable such PalmSpace-like solutions to robustly and accurately track users' activities and map them with public UIs. Furthermore, in light of the health risks associated with interacting with public interfaces, there is a growing incentive to develop solutions that mitigate and isolate risk factors. We anticipate that this trend will continue to gain momentum, driving further research and development in this area. Therefore, the research and design of interfaces like PalmSpace, which leverage emerging advanced sensing technologies, will be crucial in the near future. As a stepping stone, this work aims to accelerate the development of such interfaces, providing a foundation for future researchers to build upon.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Mohammad Khalad Hasan reports financial support was provided by Natural Sciences and Engineering Research Council of Canada.

Data availability

Data will be made available on request.

Acknowledgments

We thank Saniya Alam and Sabrina Alam for their assistance in preparing the images in the paper. This work was made possible by NSERC Discovery Grant (#RGPIN-2019-05211), the Canada Foundation for Innovation Infrastructure Fund (#40440).

References

Ahlström, D., Hasan, K., Irani, P., 2014. Are you comfortable doing that? Acceptance studies of around-device gestures in and for public settings. In: Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services. MobileHCI '14, Association for Computing Machinery, New York, NY, USA, pp. 193–202. http://dx.doi.org/10.1145/2628363.2628381.

Aigner, R., Wigdor, D.J., Benko, H., Haller, M., Lindlbauer, D., Ion, A., Zhao, S., Tzu, J., Valino, K., Electronica, A., 2012. Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for HCI. In: Microsoft Research Technical Report MSR-TR-2012-111. Microsoft, Online, pp. 1–10.

Alallah, F., Neshati, A., Sakamoto, Y., Hasan, K., Lank, E., Bunt, A., Irani, P., 2018. Performer vs. Observer: Whose comfort level should we consider when examining the social acceptability of input modalities for head-worn display? In: Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology. VRST '18, Association for Computing Machinery, New York, NY, USA, http://dx.doi.org/10.1145/3281505.3281541.

Argelaguet, F., Hoyet, L., Trico, M., Lecuyer, A., 2016. The role of interaction in virtual embodiment: Effects of the virtual hand representation. In: 2016 IEEE Virtual Reality. VR, pp. 3–10. http://dx.doi.org/10.1109/VR.2016.7504682.

Bachmann, D., Weichert, F., Rinkenauer, G., 2018. Review of three-dimensional human-computer interaction with focus on the leap motion controller. Sensors 18 (7), 2194.

Baudisch, P., Xie, X., Wang, C., Ma, W.-Y., 2004. Collapse-to-zoom: Viewing web pages on small screen devices by interactively removing irrelevant content. In: Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology. UIST '04, Association for Computing Machinery, New York, NY, USA, pp. 91–94. http://dx.doi.org/10.1145/1029632.1029647.

Bazarevsky, V., Zhang, F., 2023. On-device, real-time hand tracking with MediaPipe. Retrieved March 15, 2023 from https://ai.googleblog.com/2019/08/on-device-real-time-hand-tracking-with.html.

Breslauer, N., Galić, I., Kukec, M., Samardžić, I., 2019. Leap motion sensor for natural user interface. Tehnički vjesnik 26 (2), 560–565.

Burigat, S., Chittaro, L., Vianello, A., 2012. Dynamic visualization of large numbers of off-screen objects on mobile devices: An experimental comparison of wedge and overview+detail. In: Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services. MobileHCI '12, Association for Computing Machinery, New York, NY, USA, pp. 93–102. http://dx.doi.org/10.1145/2371574.2371590.

Butler, A., Izadi, S., Hodges, S., 2008. SideSight: Multi-touch interaction around small devices. In: Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology. UIST '08, Association for Computing Machinery, New York, NY, USA, pp. 201–204. http://dx.doi.org/10.1145/1449715.1449746.

Chen, X.A., Schwarz, J., Harrison, C., Mankoff, J., Hudson, S.E., 2014a. Air+touch: Interweaving touch & in-air gestures. In: Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology. UIST '14, Association for Computing Machinery, New York, NY, USA, pp. 519–525. http://dx.doi.org/10.1145/2642918.2647392.

Chen, X.A., Schwarz, J., Harrison, C., Mankoff, J., Hudson, S., 2014b. Around-body interaction: Sensing and interaction techniques for proprioception-enhanced input with mobile devices. In: Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices and Services. MobileHCI '14, Association for Computing Machinery, New York, NY, USA, pp. 287–290. http://dx.doi.org/10.1145/2628363.2628402.

CIBC, 2021. CIBCOnline.CIBC.com. Retrieved September 9, 2021 from https://www.cibconline.cibc.com.

CityBank, 2018. City.bank. Retrieved September 9, 2022 from https://www.city.bank.

Cockburn, A., Quinn, P., Gutwin, C., Ramos, G., Looser, J., 2011. Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback. Int. J. Hum.-Comput. Stud. 69 (6), 401–414. http://dx.doi.org/10.1016/j.ijhcs.2011.02.005, URL https://www.sciencedirect.com/science/article/pii/S107158191100022X.

Delamare, W., Silpasuwanchai, C., Sarcar, S., Shiraki, T., Ren, X., 2019. On gesture combination: An exploration of a solution to augment gesture interaction. In: Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces. ISS '19, Association for Computing Machinery, New York, NY, USA, pp. 135–146. http://dx.doi.org/10.1145/3343055.3359706.

Dezfuli, N., Khalilbeigi, M., Huber, J., Özkorkmaz, M., Mühlhäuser, M., 2014. PalmRC: Leveraging the palm surface as an imaginary eyes-free television remote control. Behav. Inf. Technol. 33 (8), 829–843. http://dx.doi.org/10.1080/0144929X.2013.810781.

Freeman, E., Brewster, S., Lantz, V., 2014. Tactile feedback for above-device gesture interfaces: Adding touch to touchless interactions. In: Proceedings of the 16th International Conference on Multimodal Interaction. ICMI '14, Association for Computing Machinery, New York, NY, USA, pp. 419–426. http://dx.doi.org/10.1145/2663204.2663280.

Freeman, E., Brewster, S., Lantz, V., 2016. Do that, there: An interaction technique for addressing in-air gesture systems. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. CHI '16, Association for Computing Machinery, New York, NY, USA, pp. 2319–2331. http://dx.doi.org/10.1145/2858036.2858308, URL https://doi-org.ezproxy.library.ubc.ca/10.1145/2858036.2858308.

Friedrich, T.E., Elias, L.J., 2016. The write bias: The influence of native writing direction on aesthetic preference biases. Psychol. Aesthetics Creat. Arts 10 (2), 128.

Guiard, Y., 1987. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. J. Motor Behav. 19 (4), 486–517.

Gupta, S., Morris, D., Patel, S.N., Tan, D., 2013. AirWave: Non-contact haptic feedback using air vortex rings. In: Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing. UbiComp '13, Association for Computing Machinery, New York, NY, USA, pp. 419–428. http://dx.doi.org/10.1145/2493432.2493463.

Gustafson, S., Baudisch, P., Gutwin, C., Irani, P., 2008. Wedge: Clutter-free visualization of off-screen locations. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '08, Association for Computing Machinery, New York, NY, USA, pp. 787–796. http://dx.doi.org/10.1145/1357054.1357179, URL https://doi-org.ezproxy.library.ubc.ca/10.1145/1357054.1357179.

Gustafson, S., Bierwirth, D., Baudisch, P., 2010. Imaginary interfaces: Spatial interaction with empty hands and without visual feedback. In: Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology. UIST '10, Association for Computing Machinery, New York, NY, USA, pp. 3–12. http://dx.doi.org/10.1145/1866029.1866033.

Gustafson, S., Holz, C., Baudisch, P., 2011. Imaginary phone: Learning imaginary interfaces by transferring spatial memory from a familiar device. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology. UIST '11, Association for Computing Machinery, New York, NY, USA, pp. 283–292. http://dx.doi.org/10.1145/2047196.2047233.

Gustafson, S.G., Irani, P.P., 2007. Comparing visualizations for tracking off-screen moving targets. In: CHI '07 Extended Abstracts on Human Factors in Computing Systems. In: CHI EA '07, Association for Computing Machinery, New York, NY, USA, pp. 2399–2404. http://dx.doi.org/10.1145/1240866.1241014.

Gustafson, S.G., Rabe, B., Baudisch, P.M., 2013. 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, Association for Computing Machinery, New York, NY, USA, pp. 889–898. http://dx.doi.org/10.1145/2470654.2466114.

Han, T., Hasan, K., Nakamura, K., Gomez, R., Irani, P., 2017. SoundCraft: Enabling spatial interactions on smartwatches using hand generated acoustics. In: Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology. UIST '17, Association for Computing Machinery, New York, NY, USA, pp. 579–591. http://dx.doi.org/10.1145/3126594.3126612.

Harrison, C., Benko, H., Wilson, A.D., 2011. OmniTouch: Wearable multitouch interaction everywhere. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology. UIST '11, Association for Computing Machinery, New York, NY, USA, pp. 441–450. http://dx.doi.org/10.1145/2047196.2047255.

Harrison, C., Hudson, S.E., 2009. Abracadabra: Wireless, high-precision, and unpowered finger input for very small mobile devices. In: Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology. UIST '09, Association for Computing Machinery, New York, NY, USA, pp. 121–124. http://dx.doi.org/10.1145/1622176.1622199.

Harrison, C., Ramamurthy, S., Hudson, S.E., 2012. On-body interaction: Armed and dangerous. In: Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction. TEI '12, Association for Computing Machinery, New York, NY, USA, pp. 69–76. http://dx.doi.org/10.1145/2148131.2148148.

Harrison, C., Tan, D., Morris, D., 2010. Skinput: Appropriating the body as an input surface. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '10, Association for Computing Machinery, New York, NY, USA, pp. 453–462. http://dx.doi.org/10.1145/1753326.1753394.

Hasan, K., Ahlström, D., Irani, P., 2013. Ad-binning: Leveraging around device space for storing, browsing and retrieving mobile device content. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '13, Association for Computing Machinery, New York, NY, USA, pp. 899–908. http://dx.doi.org/10.1145/2470654.2466115.

Hasan, K., Ahlström, D., Irani, P.P., 2015a. Comparing direct off-screen pointing, peephole, and flick & pinch interaction for map navigation. In: Proceedings of the 3rd ACM Symposium on Spatial User Interaction. SUI '15, Association for Computing Machinery, New York, NY, USA, pp. 99–102. http://dx.doi.org/10.1145/2788940.2788957.

Hasan, K., Ahlström, D., Irani, P., 2015c. SAMMI: A spatially-aware multi-mobile interface for analytic map navigation tasks. In: Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services. MobileHCI '15, Association for Computing Machinery, New York, NY, USA, pp. 36–45. http://dx.doi.org/10.1145/2785830.2785850, URL https://doi-org.ezproxy.library.ubc.ca/10.1145/2785830.2785850.

Hasan, K., Ahlström, D., Kim, J., Irani, P., 2017. AirPanes: Two-handed around-device interaction for pane switching on smartphones. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. CHI '17, Association for Computing Machinery, New York, NY, USA, pp. 679–691. http://dx.doi.org/10.1145/3025453.3026029.

Hasan, K., Kim, J., Ahlström, D., Irani, P., 2016. Thumbs-up: 3D spatial thumb-reachable space for one-handed thumb interaction on smartphones. In: Proceedings of the 2016 Symposium on Spatial User Interaction. SUI '16, Association for Computing Machinery, New York, NY, USA, pp. 103–106. http://dx.doi.org/10.1145/2983310.2985755.

Hausen, D., Boring, S., Greenberg, S., 2013. The unadorned desk: Exploiting the physical space around a display as an input canvas. In: IFIP Conference on Human-Computer Interaction. Springer, South Africa, pp. 140–158.

HoloLens, 2021. Microsoft hololens. Retrieved September 9, 2022 from https://www.microsoft.com/en-us/hololens.

Hoshi, T., 2011. Development of aerial-input and aerial-tactile-feedback system. In: 2011 IEEE World Haptics Conference. pp. 569–573. http://dx.doi.org/10.1109/WHC.2011.5945548.

Hoshi, T., Iwamoto, T., Shinoda, H., 2009. Non-contact tactile sensation synthesized by ultrasound transducers. In: World Haptics 2009 - Third Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. pp. 256–260. http://dx.doi.org/10.1109/WHC.2009.4810900.

Hoshi, T., Takahashi, M., Iwamoto, T., Shinoda, H., 2010. Noncontact tactile display based on radiation pressure of airborne ultrasound. IEEE Trans. Haptics 3 (3), 155–165. http://dx.doi.org/10.1109/TOH.2010.4.

Hossain, Z., Hasan, K., Liang, H.-N., Irani, P., 2012. EdgeSplit: Facilitating the selection of off-screen objects. In: Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services. MobileHCI '12, Association for Computing Machinery, New York, NY, USA, pp. 79–82. http://dx.doi.org/10.1145/2371574.2371588.

Irani, P., Gutwin, C., Yang, X.D., 2006. Improving selection of off-screen targets with hopping. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '06, Association for Computing Machinery, New York, NY, USA, pp. 299–308. http://dx.doi.org/10.1145/1124772.1124818.

Isomoto, T., Yamanaka, S., Shizuki, B., 2021. Relationship between dwell-time and model human processor for dwell-based image selection. In: ACM Symposium on Applied Perception 2021. SAP '21, Association for Computing Machinery, New York, NY, USA, http://dx.doi.org/10.1145/3474451.3476240.

Jones, B., Sodhi, R., Forsyth, D., Bailey, B., Maciocci, G., 2012. Around device interaction for multiscale navigation. In: Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services. MobileHCI '12, Association for Computing Machinery, New York, NY, USA, pp. 83–92. http://dx.doi.org/10.1145/2371574.2371589, URL https://doi-org.ezproxy.library.ubc.ca/10.1145/2371574.2371589.

Ketabdar, H., Roshandel, M., Yüksel, K.A., 2010. MagiWrite: Towards touchless digit entry using 3D space around mobile devices. In: Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services. MobileHCI '10, Association for Computing Machinery, New York, NY, USA, pp. 443–446. http://dx.doi.org/10.1145/1851600.1851701, URL https://doi-org.ezproxy.library.ubc.ca/10.1145/1851600.1851701.

Kilteni, K., Groten, R., Slater, M., 2012. The sense of embodiment in virtual reality. Presence: Teleoperators Virtual Environ. 21 (4), 373–387. http://dx.doi.org/10.1162/PRES_a_00124, arXiv:https://direct.mit.edu/pvar/article-pdf/21/4/373/1625283/pres_a_00124.pdf.

Kratz, S., Rohs, M., 2009. Hoverflow: Exploring around-device interaction with IR distance sensors. In: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services. MobileHCI '09, Association for Computing Machinery, New York, NY, USA, http://dx.doi.org/10.1145/1613858.1613912.

Kratz, S., Rohs, M., Guse, D., Müller, J., Bailly, G., Nischt, M., 2012. PalmSpace: Continuous around-device gestures vs. Multitouch for 3D rotation tasks on mobile devices. In: Proceedings of the International Working Conference on Advanced Visual Interfaces. AVI '12, Association for Computing Machinery, New York, NY, USA, pp. 181–188. http://dx.doi.org/10.1145/2254556.2254590, URL https://doi-org.ezproxy.library.ubc.ca/10.1145/2254556.2254590.

Laput, G., Xiao, R., Chen, X.A., Hudson, S.E., Harrison, C., 2014. Skin buttons: Cheap, small, low-powered and clickable fixed-icon laser projectors. In: Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology. UIST '14, Association for Computing Machinery, New York, NY, USA, pp. 389–394. http://dx.doi.org/10.1145/2642918.2647356.

Le, H.V., Kosch, T., Bader, P., Mayer, S., Henze, N., 2018. PalmTouch: Using the palm as an additional input modality on commodity smartphones. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18, Association for Computing Machinery, New York, NY, USA, pp. 1–13. http://dx.doi.org/10.1145/3173574.3173934.

LeapMotion, 2021. UltraLeap.com. Retrieved September 9, 2021 from https://www.ultraleap.com.

Lien, J., Gillian, N., Karagozler, M.E., Amihood, P., Schwesig, C., Olson, E., Raja, H., Poupyrev, I., 2016. Soli: Ubiquitous gesture sensing with millimeter wave radar. ACM Trans. Graph. 35 (4), http://dx.doi.org/10.1145/2897824.2925953.

Martinez Plasencia, D., Joyce, E., Subramanian, S., 2014. MisTable: Reach-through personal screens for tabletops. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '14, Association for Computing Machinery, New York, NY, USA, pp. 3493–3502. http://dx.doi.org/10.1145/2556288.2557325.

MediaPipe, 2021a. MediaPipe - Google LLC. Retrieved September 9, 2022 from https://mediapipe.dev.

MediaPipe, 2021b. MediaPipe code repository - Google LLC. Retrieved September 9, 2022 from https://github.com/google/mediapipe.

Montero, C.S., Alexander, J., Marshall, M.T., Subramanian, S., 2010. Would you do that? Understanding social acceptance of gestural interfaces. In: Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services. MobileHCI '10, Association for Computing Machinery, New York, NY, USA, pp. 275–278. http://dx.doi.org/10.1145/1851600.1851647.

Müller, H., Löcken, A., Heuten, W., Boll, S., 2014. Sparkle: An ambient light display for dynamic off-screen points of interest. In: Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational. NordiCHI '14, Association for Computing Machinery, New York, NY, USA, pp. 51–60. http://dx.doi.org/10.1145/2639189.2639205.

NIH, 2021. New coronavirus stable for hours on surfaces. Retrieved May 20, 2023 from https://www.nih.gov/news-events/news-releases/new-coronavirus-stable-hours-surfaces.

Niikura, T., Hirobe, Y., Cassinelli, A., Watanabe, Y., Komuro, T., Ishikawa, M., 2010b. In-air typing interface for mobile devices with vibration feedback. In: ACM SIGGRAPH 2010 Emerging Technologies. SIGGRAPH '10, Association for Computing Machinery, New York, NY, USA, http://dx.doi.org/10.1145/1836821.1836836.

Oh, S.Y., Yoon, B., Kim, H.-i., Woo, W., 2020. Finger contact in gesture interaction improves time-domain input accuracy in HMD-based augmented reality. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. In: CHI EA '20, Association for Computing Machinery, New York, NY, USA, pp. 1–8. http://dx.doi.org/10.1145/3334480.3383098, URL https://doi-org.ezproxy.library.ubc.ca/10.1145/3334480.3383098.

Qin, Q., Rohs, M., Kratz, S., 2011. Dynamic ambient lighting for mobile devices. In: Proceedings of the 24th Annual ACM Symposium Adjunct on User Interface Software and Technology. In: UIST '11 Adjunct, Association for Computing Machinery, New York, NY, USA, pp. 51–52. http://dx.doi.org/10.1145/2046396.2046418.

Qualtrics, 2021. Qualtrics XM - experience management software. Retrieved September 9, 2022 from https://www.qualtrics.com.

Rakkolainen, I., 2007. How feasible are star wars mid-air displays. In: 2007 11th International Conference Information Visualization. IV '07, pp. 935–942. http://dx.doi.org/10.1109/IV.2007.64.

Rico, J., Brewster, S., 2010b. Usable gestures for mobile interfaces: Evaluating social acceptability. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '10, Association for Computing Machinery, New York, NY, USA, pp. 887–896. http://dx.doi.org/10.1145/1753326.1753458.

Sodhi, R., Poupyrev, I., Glisson, M., Israr, A., 2013. AIREAL: Interactive tactile experiences in free air. ACM Trans. Graph. 32 (4), http://dx.doi.org/10.1145/2461912.2462007.

Song, J., Sörös, G., Pece, F., Fanello, S.R., Izadi, S., Keskin, C., Hilliges, O., 2014. In-air gestures around unmodified mobile devices. In: Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology. UIST '14, Association for Computing Machinery, New York, NY, USA, pp. 319–329. http://dx.doi.org/10.1145/2642918.2647373, URL https://doi-org.ezproxy.library.ubc.ca/10.1145/2642918.2647373.

Spindler, M., Schuessler, M., Martsch, M., Dachselt, R., 2014. Pinch-drag-flick vs. Spatial input: Rethinking zoom & pan on mobile displays. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '14, Association for Computing Machinery, New York, NY, USA, pp. 1113–1122. http://dx.doi.org/10.1145/2556288.2557028, URL https://doi-org.ezproxy.library.ubc.ca/10.1145/2556288.2557028.

Sridhar, S., Markussen, A., Oulasvirta, A., Theobalt, C., Boring, S., 2017. WatchSense: On- and above-skin input sensing through a wearable depth sensor. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. CHI '17, ACM, New York, NY, USA, pp. 3891–3902. http://dx.doi.org/10.1145/3025453.3026005, URL http://doi.acm.org/10.1145/3025453.3026005.

Surale, H.B., Matulic, F., Vogel, D., 2019. Experimental analysis of barehand mid-air mode-switching techniques in virtual reality. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 1–14.

Takashima, K., Shinshi, N., Kitamura, Y., 2015. Exploring boundless scroll by extending motor space. In: Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services. MobileHCI '15, Association for Computing Machinery, New York, NY, USA, pp. 557–566. http://dx.doi.org/10.1145/2785830.2785884.

Uddin, M.S., Gutwin, C., Lafreniere, B., 2016. HandMark menus: Rapid command selection and large command sets on multi-touch displays. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. CHI '16, Association for Computing Machinery, New York, NY, USA, pp. 5836–5848. http://dx.doi.org/10.1145/2858036.2858211.

Wang, C.-Y., Chu, W.-C., Chiu, P.-T., Hsiu, M.-C., Chiang, Y.-H., Chen, M.Y., 2015. 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, Association for Computing Machinery, New York, NY, USA, pp. 153–160. http://dx.doi.org/10.1145/2785830.2785886, URL https://doi-org.ezproxy.library.ubc.ca/10.1145/2785830.2785886.

Weigel, M., Mehta, V., Steimle, J., 2014. 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, Association for Computing Machinery, New York, NY, USA, pp. 179–188. http://dx.doi.org/10.1145/2556288.2557239.

Xiao, R., Cao, T., Guo, N., Zhuo, J., Zhang, Y., Harrison, C., 2018. LumiWatch: On-arm projected graphics and touch input. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18, Association for Computing Machinery, New York, NY, USA, pp. 1–11. http://dx.doi.org/10.1145/3173574.3173669.

Yang, X.-D., Hasan, K., Bruce, N., Irani, P., 2013. Surround-see: Enabling peripheral vision on smartphones during active use. In: Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology. UIST '13, Association for Computing Machinery, New York, NY, USA, pp. 291–300. http://dx.doi.org/10.1145/2501988.2502049.

Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C.-L., Grundmann, M., 2020. MediaPipe hands: On-device real-time hand tracking. pp. 1–5, arXiv, arXiv:2006.10214.

Zhang, Y., Zhou, J., Laput, G., Harrison, C., 2016. SkinTrack: Using the body as an electrical waveguide for continuous finger tracking on the skin. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. CHI '16, Association for Computing Machinery, New York, NY, USA, pp. 1491–1503. http://dx.doi.org/10.1145/2858036.2858082.

Zhou, J., Zhang, Y., Laput, G., Harrison, C., 2016. AuraSense: Enabling expressive around-smartwatch interactions with electric field sensing. In: Proceedings of the 29th Annual Symposium on User Interface Software and Technology. UIST '16, Association for Computing Machinery, New York, NY, USA, pp. 81–86. http://dx.doi.org/10.1145/2984511.2984568.