Evaluating Position and Orientation of Smartwatch Camera for Real-World Object Selection

Kaoru Shirane
shirane@iplab.cs.tsukuba.ac.jp
University of Tsukuba
Tsukuba, Ibaraki, Japan

Myungguen Choi
choi@iplab.cs.tsukuba.ac.jp
University of Tsukuba
Tsukuba, Ibaraki, Japan

Kousei Nagayama
nagayama@iplab.cs.tsukuba.ac.jp
University of Tsukuba
Tsukuba, Ibaraki, Japan

Buntarou Shizuki
shizuki@cs.tsukuba.ac.jp
University of Tsukuba
Tsukuba, Ibaraki, Japan

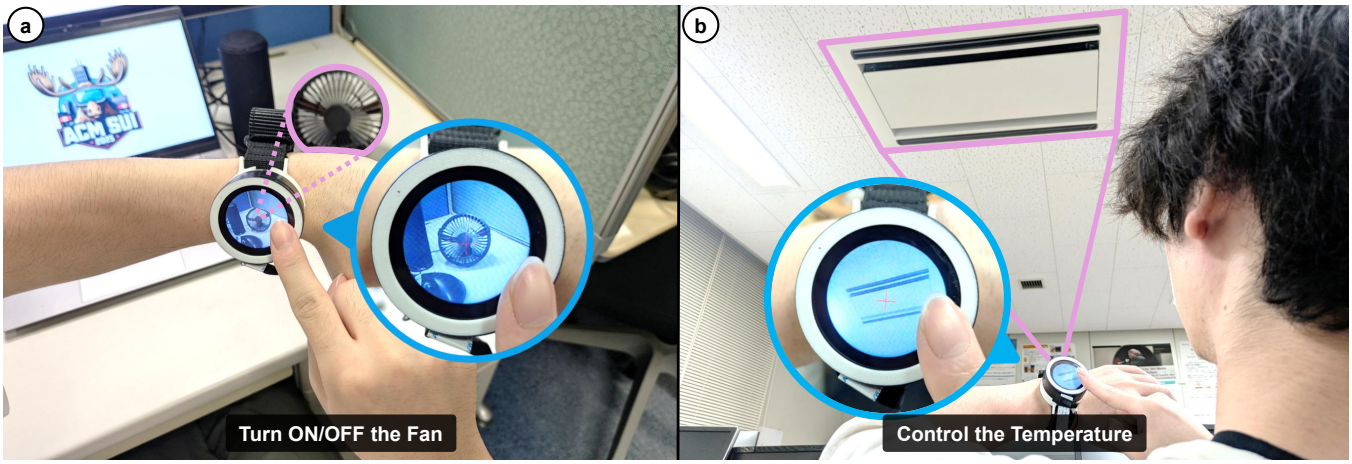

Figure 1: Examples of real-world object selection using a smartwatch-based pointing method. (a) The user selects a fan by aligning it within the view of a downward-facing camera mounted on the underside of the watch band. (b) The user selects an air conditioner using a side-facing camera embedded in the smartwatch display.

Abstract

This study investigated the pointing performance of a method for selecting real-world objects using a smartwatch equipped with a camera mounted at various positions and orientations. In this method, users point at targets by aligning these targets with the center of the smartwatch display while viewing the camera feed. Although this approach is promising for enabling interaction with real-world objects using only a smartwatch, the impact of camera position and orientation on pointing performance has not been explored. Therefore, we conducted a study to evaluate pointing performance across different combinations of camera positions and orientations. The results showed that configurations with the camera mounted on the band, rather than the bezel, led to faster pointing speeds. Additionally, camera orientations facing downward or sideways resulted in better pointing performance compared with forward orientations.

CCS Concepts

• **Human-centered computing** → **Pointing**.

Keywords

Smartwatch, Camera, Pointing

ACM Reference Format:
Kaoru Shirane, Myungguen Choi, Kousei Nagayama, and Buntarou Shizuki. 2025. Evaluating Position and Orientation of Smartwatch Camera for Real-World Object Selection. In *ACM Symposium on Spatial User Interaction (SUI '25), November 10–11, 2025, Montreal, QC, Canada.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3694907.3765935

SUI '25, November 10–11, 2025, Montreal, QC, Canada

ACM ISBN 979-8-4007-1259-3/2025/11
https://doi.org/10.1145/3694907.3765935

1 Introduction

Interacting with real-world objects through computers is an important research area within the field of Human-Computer Interaction (HCI). In addition to electronic devices, even objects not connected

to a computer can serve as interaction targets in various scenarios, such as when referencing real-world objects to voice agents [31] or large language model (LLM) agents [10, 29]. A common approach for selecting these objects is to position them within the camera's field of view (FOV) and use techniques such as ray casting [29, 31, 35] to indicate the target. This operation helps clarify the specific real-world object being referenced, which is essential for preventing selection errors (e.g., accurately turning on one of several lights) and enabling context-aware interactions (e.g., asking "when does this close?" while pointing at a physical store) [31]. Numerous researchers have proposed methods for selecting real-world objects using head-mounted displays (HMDs) [10, 29] and mobile devices, such as smartphones [24, 31, 35] and smartwatches [12].

Most research on real-world object selection using mobile devices has focused on smartphones, leaving the potential of smartwatches relatively underexplored. When selecting real-world objects with a smartphone, users must hold the device so that the target objects appear within the FOV. This requirement introduces challenges, such as increased selection time and reduced usability in hands-busy situations (e.g., while cooking or assembling objects). By contrast, smartwatches are worn on the wrist and are always readily accessible, making them promising platforms for fast and usable interaction with real-world objects.

In this study, we investigate the pointing performance of a real-world object selection method using a camera mounted on a smartwatch. In this method, users point at real-world targets—such as a fan or an air conditioner—by aligning the target within the camera view shown on the smartwatch display (Figure 1). The target is then selected by either touching the smartwatch display or performing a hand gesture with the hand wearing the smartwatch. Although this technique is a straightforward extension of smartphone-based real-world object selection methods, its pointing performance has not been evaluated in the context of smartwatch interaction.

Additionally, the optimal positional relationship between a smartwatch and its camera has not been investigated. Some smartwatches are equipped with cameras positioned to capture the back of the hand (Neptune Pine[1]) or oriented perpendicular to the forearm (Samsung Gear 2[2]). As a unique example, the WristCam[3] features a camera mounted on the band. These cameras are primarily intended for purposes such as video recording or face recognition rather than pointing interactions, and their suitability for pointing-based object selection tasks has not been evaluated. Such camera positions and orientations may be suboptimal for object selection, as the camera's perspective may not align intuitively with the user's expectations. Moreover, depending on the target's location, this misalignment can cause the smartwatch display to tilt at awkward angles, making it difficult for users to clearly view the target.

Our goal is to identify the optimal camera position and orientation for selecting real-world objects using a smartwatch-based pointing method. To evaluate the pointing performance of this method across different camera positions and orientation, we conducted a user study in which participants performed a pointing task targeting real-world objects. The study compared selection performance across five camera configurations, combining two camera positions (on the smartwatch display bezel and on the underside of the watch band) with three camera orientations (forward, side, and downward). The results revealed that forward-facing camera configurations resulted in low pointing performance. By contrast, the Band-downward configuration enabled fast and intuitive pointing but imposed a higher physical load. Side-oriented camera configurations offered the best balance, enabling faster pointing with reduced physical effort.

2 Related Work

This section reviews prior work in three areas: (1) methods for selecting real-world objects, (2) object selection techniques using smartwatches, and (3) smartwatches equipped with cameras.

2.1 Methods for Selecting Real-World Objects

In the field of HCI, selecting real-world objects—such as electronic devices or physical objects not connected to computers—is an important task. Electronic devices are typically selected for the purpose of direct operation, while other objects may be selected by LLMs as reference information [10, 29]. Recognition of real-world objects is commonly performed using image recognition techniques, with cameras embedded in HMDs [10, 29], rings [25], smartphones [5, 8, 24, 31, 35], or smartwatches [12] used to capture the surrounding environment.

Most real-world object selection methods using mobile devices employ smartphones [24, 31, 35], while only a few have explored the use of smartwatches. Fang et al. [12] proposed *WatchThis*, which enables interaction with LLMs using camera images displayed on a smartwatch. In this method, the smartwatch display is flipped upward to show images captured from the back of the device, effectively creating a see-through experience. Users point to real-world objects with their index finger and issue voice commands to the LLM, allowing it to reference the indicated objects during interaction.

While *WatchThis* relies on a flip-up display mechanism, to the best of our knowledge, no prior research has investigated the performance of real-world object selection using smartwatches with a standard (non-flipped) display orientation. In this work, we address this gap by evaluating selection performance using a smartwatch in a conventional display position.

2.2 Object Selection Techniques Using Smartwatches

Most research on interaction techniques with smartwatches has focused on interaction through the watch's on-screen UI. Because of their small display size compared to smartphones or PCs, various techniques have been proposed for the selection of small UI elements [6, 15, 28, 33, 47, 48, 50]. In addition, finger occlusion is a significant issue ("fat finger problem" [39]), and several methods have been developed to reduce or eliminate it. The most common approach involves one-handed input using the hand wearing the

[1]https://www.kickstarter.com/projects/neptune/neptune-pine-smartwatch-reinvented (accessed: 2025-06-21)
[2]https://www.samsung.com/us/support/downloads/?model=N0021373 (accessed: 2025-06-21)
[3]https://wristcam.com/products/wristcam-original-edition-noir-42-44-45mm (accessed: 2025-06-21)

smartwatch, enabling interaction primarily through finger gestures [13, 14, 20, 32, 40, 42].

Smartwatches have also been used to select objects beyond the smartwatch UI, such as objects on large displays [22, 34], objects in Augmented and Virtual Reality (AR/VR) environments [7, 18, 23, 45], and real-world objects [1, 12]. In these studies, objects are selected by using the smartwatch as a touchpad [7, 18, 45], or employing ray casting techniques based on the orientation of the smartwatch [7, 22, 23]. Most smartwatch-based ray casting methods project rays in the direction of the user's forearm, similar to hand-based ray casting. While this direction can be effective in scenarios in which the targets are located on external surfaces (e.g., large displays or surrounding environments), it may not be suitable in situations in which users interact while looking at the smartwatch screen—such as in the *WatchThis* scenario. In this setting, users select real-world objects by viewing the camera feed on the smartwatch display.

2.3 Smartwatches Equipped with Cameras

Various applications using cameras placed on smartwatches or wristbands have been proposed, including diet monitoring [37, 38], body pose estimation [30, 46, 49], AR applications [41], and novel input methods [43]. Sen et al. [38] proposed a diet monitoring method that automatically captures images of meals while eating using a camera placed at the front position of the smartwatch band. Thomas et al. [41] implemented a mobile AR prototype that displays images from a camera located diagonally beneath the smartwatch band. Jannat et al. [21] proposed a face-centered spatial user interface for smartwatches that detects the user's face position using a front-facing camera and uses the spatial relationship between the face and the smartwatch.

The appropriate position and orientation of a smartwatch camera vary depending on the application's use. Sen et al. [37] compared seven camera positions to investigate the most suitable placement for capturing images of meals during eating. The results showed that two cameras positioned on the outer side of the wrist and one positioned directly beneath the band were effective. However, the optimal camera position and orientation for pointing using a smartwatch camera have not been investigated. In this work, we evaluate pointing speed and accuracy across different camera positions and orientations to identify the most appropriate configuration for pointing tasks.

3 Implementation

In this research, we evaluated the pointing performance of five camera configurations, each defined by a combination of camera position and orientation (Figure 2). We consider two camera positions: the smartwatch display (Display) and the underside of the band (Band). Additionally, we define three camera orientations: forward, side, and downward. The forward orientation refers to the orientation of the forearm wearing the watch, i.e. the 3 o'clock on the smartwatch. The side orientation refers to the 12 o'clock on the smartwatch. The downward orientation refers to the orientation normal to the smartwatch display and away from the user's face when they look at the watch. Among the six possible combinations, the Display-downward configuration was excluded from the study because the watch band obstructed the camera's view. Consequently, we adopted the following five camera configurations: Display-forward (Figure 2a), Display-side (Figure 2b), Band-forward, Band-side, and Band-downward (Figure 2c).

Cameras on commercial off-the-shelf (COTS) smartwatches are typically located on the display, and in some models, they are mounted on the band. Based on these designs, we selected Display and Band as the two camera positions in our study. Although COTS smartwatches with band-mounted cameras typically place them on the side (Figure 2c), we adopted the underside of the band as the Band position in our study. This is because orienting the camera in the forward or downward orientation from the side position would result in a misalignment with the display orientation, making interaction less intuitive. Therefore, we positioned the camera on the underside of the band for all Band configurations.

We selected forward and side camera orientations, which are commonly used in COTS smartwatches (Figure 2a, b), as well as a downward orientation, which is rarely adopted in existing products. We used a downward-facing camera based on the assumption that pointing movements would become more intuitive and efficient, as the camera and display face the same direction. While other camera positions and orientations are also conceivable, we limited the study to the five configurations to ensure feasibility within the allotted experimental time.

We developed a custom prototype smartwatch to evaluate the pointing performance of configurations with various camera positions and orientations. This was necessary because (1) no existing commercial smartwatches support the required camera configurations, and (2) displaying images from an external camera on COTS smartwatches is technically difficult. Additionally, we designed a wearable switch that can be attached to the user's fingertip, enabling one-handed interaction using the same hand that wears the prototype smartwatch.

3.1 Hardware

3.1.1 Smartwatch Hardware. Our prototype smartwatch consists of two main components: the watch module and the camera module. The watch module integrates a Seeed Studio XIAO ESP32S3 Sense microcontroller, a 1.28-inch round TFT display (240 × 240 pixels; effective display area: 32.40 mm), and a microSD card for local image storage. It is powered by a 3.7 V, 120 mAh lithium-polymer battery, allowing standalone operation without the need for a tethered host device. These components are illustrated in Figure 3a. The camera module used was an OV2640, which was mounted using additional attachments or embedded components depending on the specific camera position and orientation. The total weight of the prototype smartwatch, including all components, was approximately 50 g.

In the Display conditions, the camera was embedded within a smartwatch case fabricated using a 3D printer. The mounting position varied depending on the camera's orientation: In the forward condition, it was placed on the front (Figure 2a), whereas in the side condition, the camera was embedded on the lateral side of the case(Figure 2b). In the Band conditions, a modular attachment was used to mount the camera onto the wristband. Custom attachments were designed for each orientation and secured with a Velcro strap, allowing for flexible positional adjustment (Figure 2c).

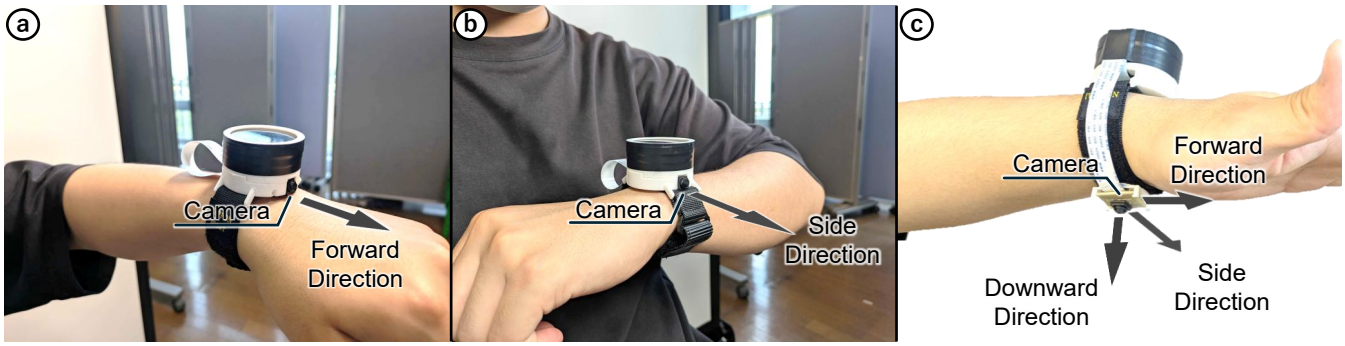

Figure 2: Camera configurations used in our study. (a) Display-forward configuration. (b) Display-side configuration. (c) Band-forward, Band-side, and Band-downward configurations.

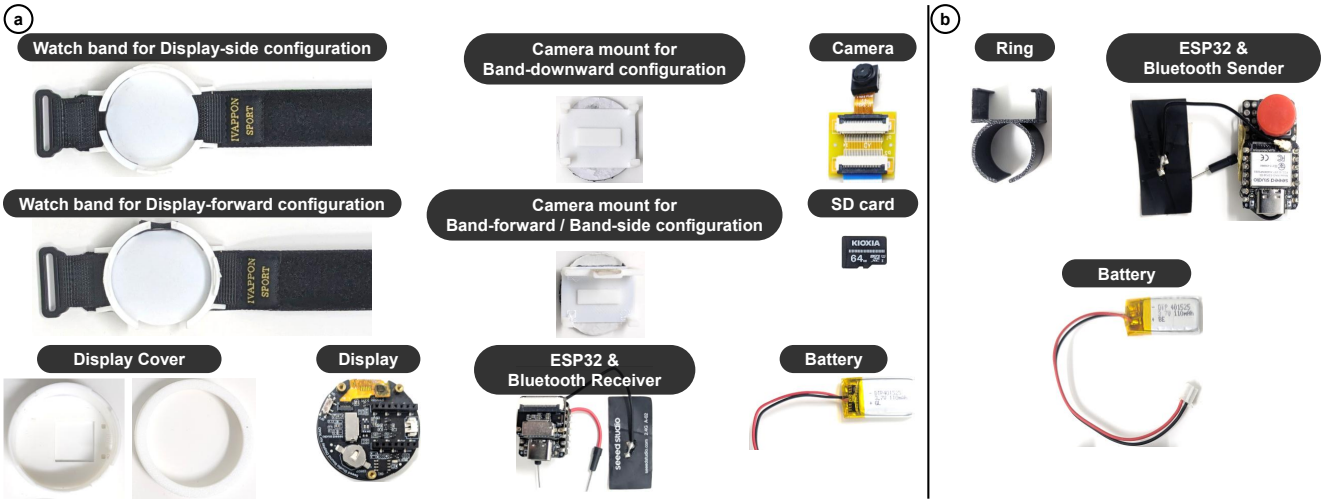

Figure 3: Overview of the hardware components used in the smartwatch and the wearable switch. (a) Smartwatch components. (b) Wearable switch components.

3.1.2 Wearable Switch Hardware. The wearable switch consists of a microcontroller (Seeed Studio XIAO ESP32S3 Sense), a 3.7 V, 120 mAh lithium-polymer battery, and a mounting ring. We designed the ring using a 3D printer to securely attach the device to the user's finger. These components are illustrated in Figure 3b.

3.2 Software

The smartwatch and the wearable switch communicate via Bluetooth Low Energy. The smartwatch display shows the live camera feed along with a red crosshair cursor indicating the center of the screen. When the camera orientation is set to forward, the camera feed is rotated by 90° to account for the difference between the physical camera orientation and the standard display alignment.

In this study, users could perform selection actions either by tapping the smartwatch display or by pressing the wearable switch. After a selection action, the word "Capture" briefly appears on the display as visual feedback, and the current camera frame is saved to the microSD card connected to the smartwatch's microcontroller.

4 Study

In this study, we investigated the pointing performance of a smartwatch-based object selection method across various camera configurations.

4.1 Participants

Fifteen participants (3 female, 12 male; mean age = 22.5 years, SD = 1.1 years; IDs: P1–P15) from a local university were recruited for the study. Fourteen participants were right handed, and one was left handed. Five participants reported the daily use of smartwatches, one had prior experience, and nine had never owned one.

4.2 Design

We employed a within-participant design. The study included three independent variables:

- *Configuration*: Display-forward (DF), Band-forward (BF), Display-side (DS), Band-side (BS), Band-downward (BD)
- *Vertical Angle*: −22.5°, 0°, 22.5°
- *Horizontal Angle*: −45.0°, −22.5°, 0°, 22.5°, 45.0°

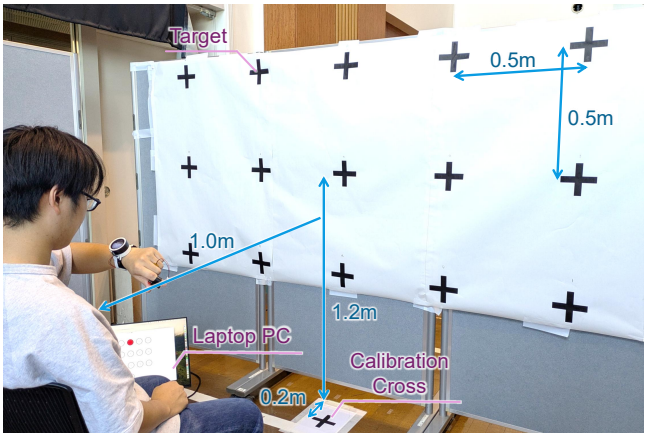

Figure 4: Experimental setup.

Configuration refers to the combination of the camera's position and its orientation. *Vertical Angle* and *Horizontal Angle* refer to the position of a target relative to the participants. A positive *Vertical Angle* indicates an upward position, while a negative value indicates a downward position. Similarly, a positive *Horizontal Angle* indicates a position to the right, while a negative value indicates a position to the left.

The order of the *Configuration* was counterbalanced using a Latin square design. The participants completed 15 trials (3 *Vertical Angles* × 5 *Horizontal Angles*) for each *Configuration*. The order of the 15 trials was randomized. One set of 15 trials was defined as a session, and the participants completed two sessions. Considering the physical load associated with the pointing task, we limited the number of sessions to two. This allowed the entire study to be completed within approximately 60 minutes while maximizing data collection and improving data reliability by mitigating fatigue-related effects. In total, we collected 5 configurations × 2 sessions × 15 trials = 150 valid datasets per participant, resulting in 2,250 datasets across all 15 participants.

4.3 Procedure and Task

Upon arriving at the designated space at the local university, the participants read and signed an informed consent form. They then received detailed instructions about the task. The participants wore the smartwatch on their left wrist and the wearable switch on their left index finger. The experimental setup is shown in Figure 4. Participants were seated on chairs positioned 1.0 m away from partitions with paper targets attached. Each target was a 10 cm cross, and a total of 15 targets were printed on a single sheet of paper, arranged according to combinations of *Vertical Angle* and *Horizontal Angle*. The paper was affixed to the partition such that the center target was positioned 1.2 m above the floor.

The study task involved pointing to a target using the smartwatch under each *Configuration*. The participants first confirmed the target location by referring to a laptop PC positioned on the floor in front of them, slightly to the left. The PC displayed 15 circles representing the possible target positions, and the circle corresponding to the current target was highlighted in red. After confirming the target, the participants pointed at a calibration cross placed on the floor 0.2 m away from the partition, then tapped the smartwatch display to align the initial position for the pointing task (initial phase). Next, the participants performed the pointing task using the specified *Configuration* (pointing phase). When the target was aligned with the center of the smartwatch's display view, the participants pressed the wearable switch. The participants were instructed to perform the pointing actions as quickly and accurately as possible, aiming to align the target at the center of the smartwatch display. The target indicator on the laptop PC was then updated to show the next target position. This sequence constituted one trial, and the participants repeated the process to complete the selection of all 15 targets.

Each participant completed 15 trials in one session, with each trial corresponding to a unique combination of *Vertical Angle* and *Horizontal Angle*. Each participant completed two consecutive sessions. After completing the two sessions for a given *Configuration*, the participants removed the smartwatch and the wearable switch and then completed the System Usability Scale (SUS) [4], the NASA Task Load Index (NASA-TLX) [16], and the Borg CR-10 [3]. This procedure was repeated five times, corresponding to the five different *Configurations*. After completing all trials, the participants were interviewed and asked to indicate their preferred *Configuration*. Each study took approximately 60 minutes.

4.4 Metrics

We analyzed the following metrics:

Pointing Time: This refers to the total time elapsed from touching the smartwatch display during the initial phase to pressing the button during the pointing phase.

Pointing Error: This refers to the distance (in millimeters) between the center of the smartwatch display and the target's center on the smartwatch display. This metric was recorded only when the target was visible on the smartwatch display at the moment the wearable switch was pressed.

Questionnaire: Subjective measures of usability, workload, and physical strain, collected using the SUS, NASA-TLX, and Borg CR-10 questionnaires, were used to evaluate the usability and physical demands of each *Configuration*.

4.5 Analysis

The pointing time and pointing error did not follow a normal distribution. Therefore, we applied the aligned rank transform (ART) [17, 36, 44] to each dataset, followed by a three-way repeated measures ANOVA with *Configuration*, *Vertical Angle*, and *Horizontal Angle* as factors. Within-factor post-hoc analyses were conducted using ART-C [11] with Holm correction [19]. For these metrics, effect sizes are reported using η_p^2.

We used the Friedman test to evaluate the SUS, NASA-TLX, and Borg CR-10 scores with *Configuration* as a factor, and cross-factor pairwise comparisons were performed using Wilcoxon signed-rank tests with Holm correction.

5 Results

In the Band-forward configuration, the target was not visible in 2.89% of the captured images (13 out of 450) due to occlusion by the

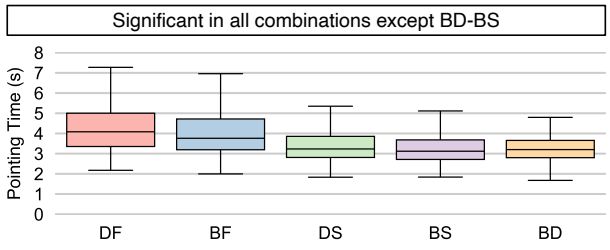

Figure 5: Pointing time for each *Configuration*.

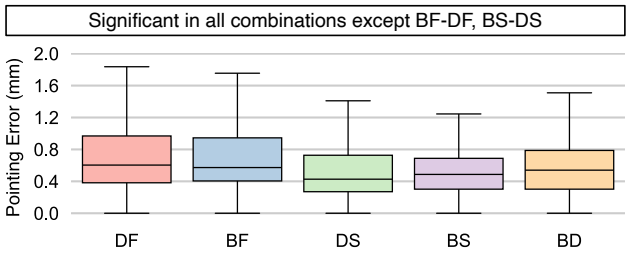

Figure 6: Pointing error for each *Configuration*.

participants' fingers. As a result, these 13 data points were excluded from the analysis, yielding a total of 2,237 valid pointing error data points.

5.1 Pointing Time

The median pointing times (lower is better) for each *Configuration* were as follows: Display-forward = 4.08 s, Band-forward = 3.76 s, Display-side = 3.23 s, Band-side = 3.12 s, and Band-downward = 3.20 s. We observed significant effects of *Configuration* ($F_{4,2161} = 139.23, p < .01, \eta_p^2 = .205$), *Vertical Angle* ($F_{2,2161} = 55.81, p < .01, \eta_p^2 = .049$), and *Horizontal Angle* ($F_{4,2161} = 12.79, p < .01, \eta_p^2 = .023$) on the pointing time. A significant interaction was also found between *Vertical Angle* and *Horizontal Angle* ($F_{8,2161} = 2.10, p < .05, \eta_p^2 = .008$). Figure 5 presents the pointing times for all *Configuration*.

Post hoc tests revealed that the Band-side and Band-downward configurations resulted in significantly faster pointing times than the other configurations ($p < .01$). Notably, configurations with the camera mounted on the band were significantly faster than those with the camera on the display, even when the camera orientation was the same. Specifically, Band-forward was significantly faster than Display-forward ($t_{2161} = -4.23, p < .01$), and Band-side was significantly faster than Display-side ($t_{2161} = -3.06, p < .01$).

5.2 Pointing Error

The median pointing errors (lower is better) for each *Configuration* were as follows: Display-forward = 0.60 mm, Band-forward = 0.57 mm, Display-side = 0.43 mm, Band-side = 0.49 mm, and Band-downward = 0.54 mm. We observed significant effects of *Configuration* ($F_{4,2148} = 19.87, p < .05, \eta_p^2 = .036$) on the pointing error. However, there were no significant effects of *Vertical Angle* ($F_{2,2148} = 1.43, p = .24, \eta_p^2 = .001$) or *Horizontal Angle* ($F_{4,2148} = 0.29, p = .73, \eta_p^2 = .001$). Figure 6 presents the pointing errors for all *Configuration*. Post hoc tests revealed that the Display-side and Band-side configurations resulted in significantly lower pointing errors than the other configurations ($p < .05$).

5.3 System Usability Scale (SUS)

The average overall SUS scores (higher is better) for each *Configuration* were as follows: Display-forward = 55.83, Band-forward = 47.83, Display-side = 74.67, Band-side = 79.00, and Band-downward = 71.83. A Friedman test was conducted on the overall SUS scores across *Configuration*, and significant differences were found ($\chi^2_{4,N=15} = 21.62, p < .01$). Figure 7 presents the SUS scores for all *Configuration*.

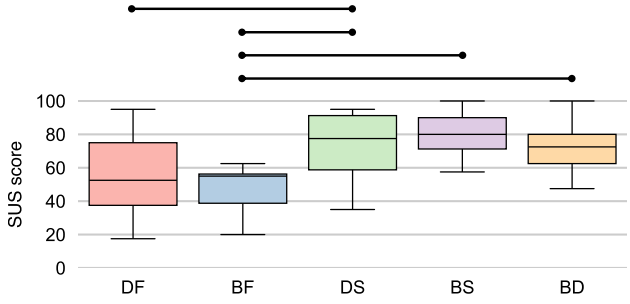

Figure 7: SUS score for each *Configuration*. Significant differences ($p < .05$) between conditions are indicated by horizontal bars.

5.4 NASA Task Load Index (NASA-TLX)

The average overall workload scores (lower is better) for each *Configuration* were as follows: Display-forward = 55.27, Band-forward = 63.67, Display-side = 40.98, Band-side = 40.49, and Band-downward = 52.53. We performed the Friedman test on the responses for each of the six questions and the overall score. *Configuration* was significantly affected by physical demand ($\chi^2_{4,N=15} = 21.23, p < .01$), effort ($\chi^2_{4,N=15} = 11.79, p < .05$), frustration ($\chi^2_{4,N=15} = 24.23, p < .01$), and overall workload ($\chi^2_{4,N=15} = 23.93, p < .01$). Figure 8 presents the NASA-TLX scores for all *Configuration*.

5.5 Borg CR-10

The average Borg scores (lower is better) for each *Configuration* were as follows: Display-forward = 4.27, Band-forward = 4.77, Display-side = 2.90, Band-side = 2.10, and Band-downward = 4.57. A Friedman test was conducted on the Borg scores with *Configuration*, and significant differences were found ($\chi^2_{4,N=15} = 28.15, p < .01$). Figure 9 presents the BORG scores for all *Configuration*. Post hoc tests revealed that the Display-side and Band-side configurations resulted in significantly lower Borg scores than the other configurations ($p < .05$).

5.6 Preference of Configuration

The average preference ranks (lower is better) for each *Configuration* were as follows: Display-forward = 3.47, Band-forward = 4.77, Display-side = 2.13, Band-side = 1.87, and Band-downward = 3.07. Thus, Band-side, Display-side, Band-downward, Display-forward,

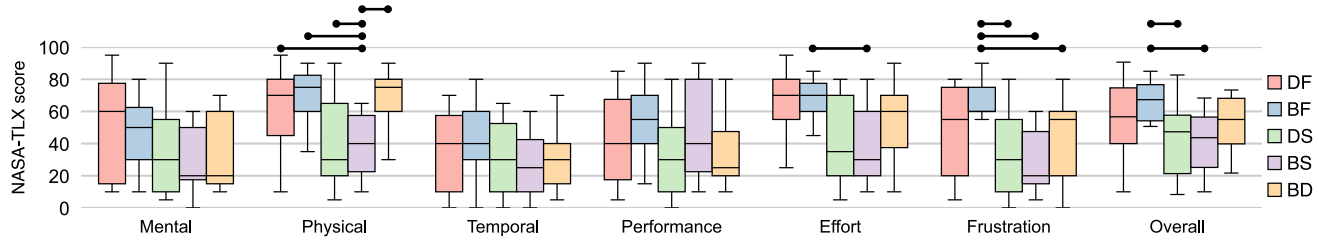

Figure 8: NASA-TLX score for each *Configuration*. Significant differences ($p < .05$) between conditions are indicated by horizontal bars.

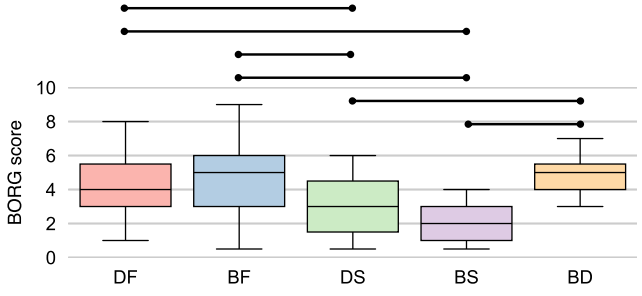

Figure 9: BORG score for each *Configuration*. Significant differences ($p < .05$) between conditions are indicated by horizontal bars.

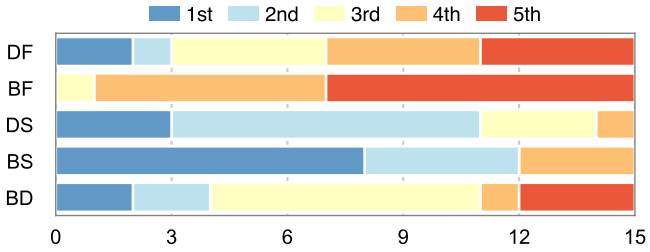

Figure 10: Results of preference for *Configuration*.

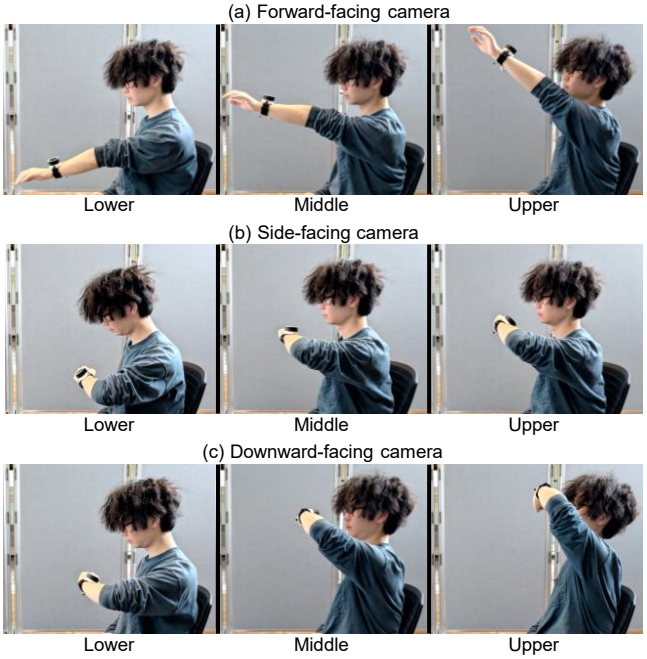

Figure 11: Examples of postures when pointing at targets located at different heights: (a) forward-facing camera configurations, (b) side-facing camera configurations, and (c) downward-facing camera configurations.

and Band-forward were preferred in this order. Figure 10 presents the participants' preference for *Configuration*.

6 Discussion

6.1 Pointing Performance of Each Configuration

6.1.1 Forward orientation. The configurations with forward-facing cameras (Display-forward and Band-forward) exhibited lower pointing performance than the other configurations. Specifically, these configurations resulted in significantly longer pointing times and higher pointing errors. Additionally, BORG scores were significantly higher than those of the Display-side and Band-side configurations, indicating greater physical demand. This increased demand can be attributed to the need for users to raise their arms high when using these configurations (Figure 11a); most participants moved their entire arm, while P3 moved only their forearm. These findings suggest that forward-facing camera configurations are not suitable for smartwatch-based object selection tasks. Although using the arm's orientation for pointing has been adopted in smartwatch-based interactions with large displays and AR/VR environments [7, 22, 23], our findings indicate that this method is not effective when users need to view the smartwatch screen during object selection.

Although the Band-forward configuration resulted in significantly faster pointing times than the Display-forward configuration, the participants preferred the latter. Six participants reported

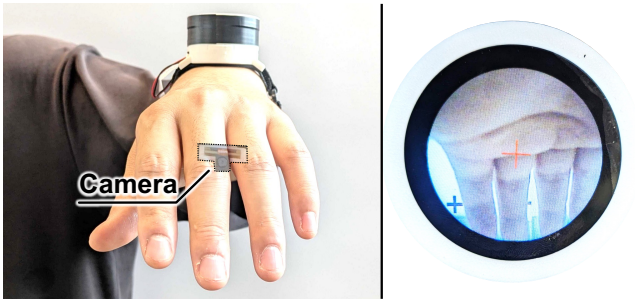

Figure 12: Illustrations of camera occlusion by the user's finger in the Band-forward configuration.

that the Band-forward configuration was physically demanding because they had to dorsiflex their wrists to prevent their hands from appearing in the camera view. Indeed, the Band-forward was the only configuration in which the target was occasionally obscured by the participants' fingers in the captured images (Figure 12). Additionally, when using the Band-forward configuration in everyday scenarios, there is a possibility that the camera could be blocked by items carried, such as bags or luggage. Therefore, if a forward-facing camera is to be installed on a smartwatch, it should be positioned on the display rather than on the band.

6.1.2 Side orientation. The configurations with side-facing cameras (Display-side and Band-side) demonstrated superior pointing performance compared with the other configurations. They yielded significantly shorter pointing times, lower pointing errors, and lower BORG scores than the Display-forward and Band-forward configurations. Furthermore, the participants reported the highest preference for these configurations. This is likely because the target could be aligned with the display simply by rotating the wrist, thereby reducing physical load (Figure 11b). These results suggest that side-facing cameras are preferable for smartwatch-based pointing tasks.

6.1.3 Downward orientation. The Band-downward configuration demonstrated significantly faster pointing times compared with the Display-forward, Band-forward, and Display-side configurations. Six participants rated this configuration as intuitive and easy to use for pointing tasks. This advantage likely stems from the alignment of the camera orientation with the user's line of sight, enabling intuitive pointing by simply positioning the smartwatch to overlap with the target visually.

However, the Band-downward configuration resulted in significantly higher physical demand than the Band-side configuration, as indicated by a higher Borg score compared with both the Display-side and Band-side configurations. These findings suggest that pointing using the Band-downward configuration imposes a greater physical load. Nine participants reported that this configuration was physically demanding, especially when selecting targets positioned at the top. This is because users must raise their arms high to align the camera with targets above (Figure 11c). For instance, P10 commented, "I had to lean backward to select upper targets, which was physically demanding." Therefore, the Band-downward configuration presents a trade-off between a high pointing speed—stemming from its intuitiveness—and increased physical load.

6.1.4 Summary and design implications. In summary, concerning pointing performance and physical load, the camera orientations can be ranked as follows: side-facing, downward-facing, and forward-facing. Regarding camera position, the Band-forward configuration occasionally suffered from occlusion caused by the user's finger, suggesting that the Display-forward configuration is preferable when employing a forward-facing camera. Thus, when designing a smartwatch intended for pointing-based interaction, a side-facing camera mounted on the band proves to be the most promising option. Moreover, incorporating both side-facing cameras and cameras oriented in other directions—as seen in COTS devices, such as the VTech KidiZoom Smartwatch DX3[4] and the LOKMAT APPLLP 5 Max[5]—makes it is possible to enable comfortable smartwatch-based pointing without compromising camera configurations that may be useful for other applications.

6.2 Effect of Target Position

Although there was no significant interaction between target position and camera configuration in terms of pointing time or pointing error, participant feedback suggested that target position influenced physical load and display visibility. First, the participants reported increased physical strain when pointing upward with the Band-downward configuration, as described in Section 6.1.3. By contrast, the Display-side and Band-side configurations enabled users to select upper targets with simple wrist rotation, resulting in lower physical demand. Therefore, for tasks that involve frequent selection of upper targets, side-oriented camera configurations are more suitable.

However, the side-facing camera configurations may pose challenges when pointing toward the side opposite the hand wearing the smartwatch (in this study, the right side). Three participants reported difficulty when selecting targets on the right using these configurations. For example, P6 commented, "I found it difficult to select the target on the right because I had to put my elbow forward." This difficulty likely arises from the natural posture of viewing a watch, which requires users to rotate their torsos to aim the camera toward the opposite side, thereby increasing physical strain. Although not explicitly reported, similar challenges may also occur with the Band-downward configuration, as it involves a comparable body posture when pointing toward the right.

Finally, the Band-downward configuration is particularly well suited for selecting targets located directly below the user, as it allows interaction from a natural smartwatch viewing posture. By contrast, side-faced camera configurations tend to tilt the smartwatch display at an angle that makes it difficult to view when pointing downward. Therefore, the Band-downward configuration is especially effective for selecting objects placed on the floor or on a desk while the user is standing.

[4]https://www.vtechkids.com/product/detail/20103/KidiZoom_Smartwatch_DX3 (accessed: 2025-08-16)
[5]https://www.lokmat.cn/en-ca/products/lokmat-appllp-max?srsltid=AfmBOooZdzLrqeGqFSopXPytISsrprxrkXR1hqvYnfgqSbhuO_kgjPvE (accessed: 2025-08-16)

6.3 Effectiveness of Smartwatch-based Pointing

Our study investigated the pointing performance of a smartwatch-based method using a mounted camera and demonstrated its potential for real-world object selection. The results showed that pointing errors were consistently small across all camera configurations. In particular, the Displaye-side configuration achieved the lowest median pointing error of 0.43 mm, corresponding to 0.08°, calculated based on a user-to-display-center distance of 0.3 m. This level of accuracy exceeds that reported for gaze-based and head-based pointing [27]. These findings suggest that smartwatch-based pointing is a viable method for selecting even considerably small real-world objects with high accuracy.

Additionally, smartwatch-based pointing methods offer several advantages over those using other mobile devices. Unlike smartphones, smartwatches do not need to be held in the hand, allowing for faster and more seamless object selection, particularly in hands-busy situations. While wearable rings also provide similar benefits in terms of one-handed interaction, smartwatches offer the added advantage of displaying live camera feeds. This enables users to select objects and menu items on the smartwatch display simultaneously, facilitating more intuitive and efficient interaction with real-world objects. Furthermore, smartwatches support a wide range of finger gestures using the same hand that performs pointing [42], allowing for diverse diverse inputs without disrupting the pointing posture. These advantages highlight the promise of smartwatches as a platform for real-world interaction and underscore the need for further research in this area.

6.4 Contextual and Privacy Considerations for Smartwatch-based Pointing

Smartwatch-based pointing methods can be applied to a variety of contexts, such as controlling appliances or retrieving information from real-world objects; however, this technique is not optimal for all situations. For example, in tasks such as playing music or interacting with voice agents that do not rely on the user's physical environment, voice input alone is often more appropriate and efficient. Similarly, when using HMDs, head-based or gaze-based pointing methods are preferable, as they allow for hands-free interaction. Conversely, smartwatch-based pointing is particularly suitable when voice input is inappropriate—such as in noisy environments or in cases in which it is necessary to explicitly specify the target object— or when the user is not wearing an HMD. Therefore, selecting the interaction method that best matches the specific context and task requirements is important.

Cameras built into smartwatches may raise privacy concerns when using smartwatch-based pointing methods in public spaces. Unlike smartphones, which are often carried in pockets with their cameras concealed, smartwatch cameras remain continuously exposed, posing privacy concerns. Similar concerns have been raised in prior studies on smart glasses, in which always-visible cameras have led to issues of perceived discomfort and social tension [2, 9, 26]. These studies recommend strategies such as using explicit hand gestures to activate the camera or providing visual or auditory indicators (e.g., LED lights or sound cues) to signal when the camera is in use, thereby increasing transparency for bystanders. Incorporating privacy-aware design considerations is essential to enable the safe and socially acceptable use of smartwatch-based pointing methods in public spaces.

6.5 Limitations

This study has several limitations. First, there are constraints related to the apparatus. The prototype smartwatch used in this study differs from COTS smartwatches, particularly in its thickness, which results in a shorter distance between the display and the user's face. Additionally, the camera used in this study has a lower resolution (240 × 240 pixels) compared with the cameras found in COTS smartwatches. Although the participants did not report task impacts specifically attributable to these issues, they may nonetheless have influenced pointing performance.

Second, this study investigated only two camera positions and three camera orientations. For instance, the WristCam, a COTS smartwatch, features a camera mounted on the side of the band, a configuration whose pointing performance remains unexamined in our research. Additionally, one participant (P10) suggested that a 45° camera angle on the band might offer the most effective configuration. These insights indicate the need for further investigation into alternative camera positions and orientations.

Third, the evaluation was conducted in a controlled, static environment. Since this study represents an initial investigation into the pointing performance of a smartwatch-based pointing method, we adopted a simple pointing task performed in a seated posture. As a result, the study does not address usability, physical fatigue, or pointing performance of this method in real-world contexts, such as walking, multitasking, or extended use. Future work should explore these scenarios to better understand the practical applicability of smartwatch-based pointing in everyday settings.

Finally, only the pointing performance of a single object selection method was investigated. In this study, we used a method in which the object at the center of the smartwatch display was selected when a button was pressed. However, methods such as selecting objects by touching the display or using voice input can also be employed, and the pointing performance of these methods was not examined in this study. In particular, since the touch input method requires the use of both hands, the input posture may differ from that used here, and the appropriate camera position and orientation may also vary.

7 Conclusion

In this study, we investigated the pointing performance of a smartwatch-based method using a mounted camera. Our results showed that side-facing camera configurations enabled faster and more accurate pointing with reduced physical load compared with forward-facing configurations. While the Band-downward configuration achieved the fastest pointing performance by aligning naturally with the user's gaze, it also introduced higher physical demand, particularly when selecting upper targets. Additionally, our findings indicate that real-world object pointing using a smartwatch can be performed with high accuracy. This work is the first to evaluate how camera position and orientation affect pointing performance with smartwatches, and it establishes a foundation for designing more effective smartwatch-based interaction techniques.

References

[1] Günter Alce, Andreas Espinoza, Ted Hartzell, Staffan Olsson, Dennis Samuelsson, Mattias Wallergård, and Thomas Mandl. 2018. UbiCompass: An IoT Interaction Concept. *Adv. in Hum.-Comp. Int.* 2018 (Jan. 2018), 12 pages. https://doi.org/10.1155/2018/5781363

[2] Divyanshu Bhardwaj, Alexander Ponticello, Shreya Tomar, Adrian Dabrowski, and Katharina Krombholz. 2024. In Focus, Out of Privacy: The Wearer's Perspective on the Privacy Dilemma of Camera Glasses. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 577, 18 pages. https://doi.org/10.1145/3613904.3642242

[3] Gunnar A Borg. 1982. Psychophysical bases of perceived exertion. *Medicine and science in sports and exercise* 14, 5 (1982), 377–381.

[4] John Brooke. 1995. SUS: A quick and dirty usability scale. *Usability Eval. Ind.* 189 (11 1995).

[5] Kaifei Chen, Jonathan Fürst, John Kolb, Hyung-Sin Kim, Xin Jin, David E. Culler, and Randy H. Katz. 2018. SnapLink: Fast and Accurate Vision-Based Appliance Control in Large Commercial Buildings. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4, Article 129 (Jan. 2018), 27 pages. https://doi.org/10.1145/3161173

[6] Xiang 'Anthony' Chen, Tovi Grossman, and George Fitzmaurice. 2014. Swipeboard: a text entry technique for ultra-small interfaces that supports novice to expert transitions. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST '14)*. Association for Computing Machinery, New York, NY, USA, 615–620. https://doi.org/10.1145/2642918.2647354

[7] Yuan Chen, Keiko Katsuragawa, and Edward Lank. 2020. Understanding Viewport- and World-based Pointing with Everyday Smart Devices in Immersive Augmented Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376592

[8] Adrian A. de Freitas, Michael Nebeling, Xiang 'Anthony' Chen, Junrui Yang, Akshaye Shreenithi Kirupa Karthikeyan Ranithangam, and Anind K. Dey. 2016. Snap-To-It: A User-Inspired Platform for Opportunistic Device Interactions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 5909–5920. https://doi.org/10.1145/2858036.2858177

[9] Tamara Denning, Zakariya Dehlawi, and Tadayoshi Kohno. 2014. In Situ with Bystanders of Augmented Reality Glasses:Perspectives on Recording and Privacy-MediatingTechnologies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2377–2386. https://doi.org/10.1145/2556288.2557352

[10] Mustafa Doga Dogan, Eric J Gonzalez, Karan Ahuja, Ruofei Du, Andrea Colaço, Johnny Lee, Mar Gonzalez-Franco, and David Kim. 2024. Augmented Object Intelligence with XR-Objects. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Association for Computing Machinery, New York, NY, USA, Article 19, 15 pages. https://doi.org/10.1145/3654777.3676379

[11] Lisa A. Elkin, Matthew Kay, James J. Higgins, and Jacob O. Wobbrock. 2021. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21)*. Association for Computing Machinery, New York, NY, USA, 754–768. https://doi.org/10.1145/3472749.3474784

[12] Cathy Mengying Fang, Patrick Chwalek, Quincy Kuang, and Pattie Maes. 2024. WatchThis: A Wearable Point-and-Ask Interface powered by Vision-Language Models for Contextual Queries. In *Adjunct Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST Adjunct '24)*. Association for Computing Machinery, New York, NY, USA, Article 54, 4 pages. https://doi.org/10.1145/3672539.3686776

[13] Jun Gong, Xing-Dong Yang, and Pourang Irani. 2016. WristWhirl: One-handed Continuous Smartwatch Input using Wrist Gestures. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 861–872. https://doi.org/10.1145/2984511.2984563

[14] Anhong Guo and Tim Paek. 2016. Exploring tilt for no-touch, wrist-only interactions on smartwatches *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 17–28. https://doi.org/10.1145/2935334.2935345

[15] Chris Harrison and Scott E. Hudson. 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* (Victoria, BC, Canada) *(UIST '09)*. Association for Computing Machinery, New York, NY, USA, 121–124. https://doi.org/10.1145/1622176.1622199

[16] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, Netherlands, 139 – 183. https://doi.org/10.1016/S0166-4115(08)62386-9

[17] James J. Higgins and Suleiman Tashtoush. 1994. An aligned rank transform test for interaction. *Nonlinear World* 1, 2 (1994), 201 – 211.

[18] Teresa Hirzle, Jan Rixen, Jan Gugenheimer, and Enrico Rukzio. 2018. WatchVR: Exploring the Usage of a Smartwatch for Interaction in Mobile Virtual Reality. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188629

[19] Sture Holm. 1979. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics* 6, 2 (1979), 65 – 70.

[20] Da-Yuan Huang, Liwei Chan, Shuo Yang, Fan Wang, Rong-Hao Liang, De-Nian Yang, Yi-Ping Hung, and Bing-Yu Chen. 2016. DigitSpace: Designing Thumb-to-Fingers Touch Interfaces for One-Handed and Eyes-Free Interactions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1526–1537. https://doi.org/10.1145/2858036.2858483

[21] Marium-E Jannat, Thuan T Vo, and Khalad Hasan. 2022. Face-Centered Spatial User Interfaces on Smartwatches. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 393, 7 pages. https://doi.org/10.1145/3491101.3519720

[22] Keiko Katsuragawa, Krzysztof Pietroszek, James R. Wallace, and Edward Lank. 2016. Watchpoint: Freehand Pointing with a Smartwatch in a Ubiquitous Display Environment. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (Bari, Italy) *(AVI '16)*. Association for Computing Machinery, New York, NY, USA, 128–135. https://doi.org/10.1145/2909132.2909263

[23] Daniel Kharlamov, Brandon Woodard, Liudmila Tahai, and Krzysztof Pietroszek. 2016. TickTockRay: smartwatch-based 3D pointing for smartphone-based virtual reality. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology* (Munich, Germany) *(VRST '16)*. Association for Computing Machinery, New York, NY, USA, 365–366. https://doi.org/10.1145/2993369.2996311

[24] Daehwa Kim, Vimal Mollyn, and Chris Harrison. 2023. WorldPoint: Finger Pointing as a Rapid and Natural Trigger for In-the-Wild Mobile Interactions. *Proc. ACM Hum.-Comput. Interact.* 7, ISS, Article 442 (Nov. 2023), 19 pages. https://doi.org/10.1145/3626478

[25] Maruchi Kim, Antonio Glenn, Bandhav Veluri, Yunseo Lee, Eyoel Gebre, Aditya Bagaria, Shwetak Patel, and Shyamnath Gollakota. 2024. IRIS: Wireless ring for vision-based smart home interaction. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Association for Computing Machinery, New York, NY, USA, Article 114, 16 pages. https://doi.org/10.1145/3654777.3676327

[26] Marion Koelle, Katrin Wolf, and Susanne Boll. 2018. Beyond LED Status Lights - Design Requirements of Privacy Notices for Body-worn Cameras. In *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction* (Stockholm, Sweden) *(TEI '18)*. Association for Computing Machinery, New York, NY, USA, 177–187. https://doi.org/10.1145/3173225.3173234

[27] Mikko Kytö, Barrett Ens, Thammathip Piumsomboon, Gun A. Lee, and Mark Billinghurst. 2018. Pinpointing: Precise Head- and Eye-Based Target Selection for Augmented Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173655

[28] Gierad Laput, Robert Xiao, Xiang 'Anthony' Chen, Scott E. Hudson, and Chris Harrison. 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* (Honolulu, Hawaii, USA) *(UIST '14)*. Association for Computing Machinery, New York, NY, USA, 389–394. https://doi.org/10.1145/2642918.2647356

[29] Jaewook Lee, Jun Wang, Elizabeth Brown, Liam Chu, Sebastian S. Rodriguez, and Jon E. Froehlich. 2024. GazePointAR: A Context-Aware Multimodal Voice Assistant for Pronoun Disambiguation in Wearable Augmented Reality. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 408, 20 pages. https://doi.org/10.1145/3613904.3642230

[30] Hyunchul Lim, Yaxuan Li, Matthew Dressa, Fang Hu, Jae Hoon Kim, Ruidong Zhang, and Cheng Zhang. 2022. BodyTrak: Inferring Full-body Poses from Body Silhouettes Using a Miniature Camera on a Wristband. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 3, Article 154 (Sept. 2022), 21 pages. https://doi.org/10.1145/3552312

[31] Sven Mayer, Gierad Laput, and Chris Harrison. 2020. Enhancing Mobile Voice Assistants with WorldGaze. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3313831.3376479

[32] Florian Müller, Sebastian Günther, Niloofar Dezfuli, Mohammadreza Khalilbeigi, and Max Mühlhäuser. 2016. ProxiWatch: Enhancing Smartwatch Interaction through Proximity-based Hand Input. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA,

2617–2624. https://doi.org/10.1145/2851581.2892450
[33] Stephen Oney and Brad Myers. 2012. ZoomBoard: a diminutive QWERTY soft keyboard using iterative zooming for ultra-small devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2733–2736.
[34] Krzysztof Pietroszek, Liudmila Tahai, James R. Wallace, and Edward Lank. 2017. Watchcasting: Freehand 3D interaction with off-the-shelf smartwatch. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. 172–175. https://doi.org/10.1109/3DUI.2017.7893335
[35] Yue Qin, Chun Yu, Wentao Yao, Jiachen Yao, Chen Liang, Yueting Weng, Yukang Yan, and Yuanchun Shi. 2023. Selecting Real-World Objects via User-Perspective Phone Occlusion. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 531, 13 pages. https://doi.org/10.1145/3544548.3580696
[36] K. C. Salter and R. F Fawcett. 1993. The art test of interaction: a robust and powerful rank test of interaction in factorial models. *Communications in Statistics - Simulation and Computation* 22, 1 (1993), 137–153. https://doi.org/10.1080/03610919308813085
[37] Sougata Sen, Vigneshwaran Subbaraju, Archan Misra, Rajesh Balan, and Youngki Lee. 2020. Annapurna: An automated smartwatch-based eating detection and food journaling system. *Pervasive and Mobile Computing* 68 (2020), 101259. https://doi.org/10.1016/j.pmcj.2020.101259
[38] Sougata Sen, Vigneshwaran Subbaraju, Archan Misra, Rajesh Krishna Balan, and Youngki Lee. 2015. The case for smartwatch-based diet monitoring. In *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*. 585–590. https://doi.org/10.1109/PERCOMW.2015.7134103
[39] Katie A Siek, Yvonne Rogers, and Kay H Connelly. 2005. Fat finger worries: how older and younger users physically interact with PDAs. In *Human-Computer Interaction-INTERACT 2005: IFIP TC13 International Conference, Rome, Italy, September 12-16, 2005. Proceedings 10*. Springer, 267–280.
[40] Ke Sun, Yuntao Wang, Chun Yu, Yukang Yan, Hongyi Wen, and Yuanchun Shi. 2017. Float: One-Handed and Touch-Free Target Selection on Smartwatches *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 692–704. https://doi.org/10.1145/3025453.3026027
[41] Derianna Thomas and Lars Erik Holmquist. 2020. WristAR: A Wrist-Mounted Augmented Reality Interface. In *Proceedings of the 19th International Conference on Mobile and Ubiquitous Multimedia* (Essen, Germany) *(MUM '20)*. Association for Computing Machinery, New York, NY, USA, 312–314. https://doi.org/10.1145/3428361.3432077
[42] Riku Tsunoda, Myungguen Choi, and Buntarou Shizuki. 2024. Thumb-to-Finger Gesture Recognition Using COTS Smartwatch Accelerometers. In *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia (MUM '24)*. Association for Computing Machinery, New York, NY, USA, 184–195. https://doi.org/10.1145/3701571.3701600
[43] Wouter Van Vlaenderen, Jens Brulmans, Jo Vermeulen, and Johannes Schöning. 2015. WatchMe: A Novel Input Method Combining a Smartwatch and Bimanual Interaction. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI EA '15)*. Association for Computing Machinery, New York, NY, USA, 2091–2095. https://doi.org/10.1145/2702613.2732789
[44] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963
[45] Dennis Wolf, John J. Dudley, and Per Ola Kristensson. 2018. Performance Envelopes of in-Air Direct and Smartwatch Indirect Control for Head-Mounted Augmented Reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 347–354. https://doi.org/10.1109/VR.2018.8448289
[46] Erwin Wu, Ye Yuan, Hui-Shyong Yeo, Aaron Quigley, Hideki Koike, and Kris M. Kitani. 2020. Back-Hand-Pose: 3D Hand Pose Estimation for a Wrist-worn Camera via Dorsum Deformation Network. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20)*. Association for Computing Machinery, New York, NY, USA, 1147–1160. https://doi.org/10.1145/3379337.3415897
[47] Robert Xiao, Gierad Laput, and Chris Harrison. 2014. Expanding the input expressivity of smartwatches with mechanical pan, twist, tilt and click. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 193–196. https://doi.org/10.1145/2556288.2557017
[48] Ryo Yamada and Pattie Maes. 2019. A Non-Touchscreen Tactile Wearable Interface as an Alternative to Touchscreen-Based Wearable Devices. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.
[49] Hui-Shyong Yeo, Erwin Wu, Juyoung Lee, Aaron Quigley, and Hideki Koike. 2019. Opisthenar: Hand Poses and Finger Tapping Recognition by Observing Back of Hand Using Embedded Wrist Camera. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 963–971. https://doi.org/10.1145/3332165.3347867
[50] Yang Zhang, Junhan Zhou, Gierad Laput, and Chris Harrison. 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* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1491–1503. https://doi.org/10.1145/2858036.2858082