WristLens: Enabling Single-Handed Surface Gesture Interaction for Wrist-Worn Devices Using Optical Motion Sensor

Author:
Yeo, HS; Lee, J; Bianchi, A; Samboy, A; Koike, H; Woo, W; Quigley, A

Publication details:
ACM International Conference Proceeding Series
pp. 1 - 8
9781450376037 (ISBN)
2153-1633 (ISSN)

Event details:
Proceedings of the Augmented Humans International Conference

Publication Date:
2020-03-16

Publisher DOI:
https://doi.org/10.1145/3384657.3384797

WristLens: Enabling Single-Handed Surface Gesture Interaction for Wrist-Worn Devices Using Optical Motion Sensor

Hui-Shyong Yeo
University of St Andrews
Fife, Scotland, UK
hsy@st-andrews.ac.uk

Juyoung Lee
KAIST
Daejeon, Republic of Korea
ejuyoung@kaist.ac.kr

Andrea Bianchi
KAIST
Daejeon, Republic of Korea
andrea@kaist.ac.kr

Alejandro Samboy
Soluciones GBH
Santo Domingo, Dominican Republic
a.samboy@gbh.com.do

Hideki Koike
Tokyo Institute of Technology
Tokyo, Japan
koike@c.titech.ac.jp

Woontack Woo
KAIST
Daejeon, Republic of Korea
wwoo@kaist.ac.kr

Aaron Quigley
University of St Andrews
Fife, Scotland, UK
aquigley@st-andrews.ac.uk

ABSTRACT

WristLens is a system for surface interaction from wrist-worn wearable devices such as smartwatches and fitness trackers. It enables eyes-free, single-handed gestures on surfaces, using an optical motion sensor embedded in a wrist-strap. This allows the user to leverage any proximate surface, including their own body, for input and interaction. An experimental study was conducted to measure the performance of gesture interaction on three different body parts. Our results show that directional gestures are accurately recognized but less so for shape gestures. Finally, we explore the interaction design space enabled by WristLens, and demonstrate novel use cases and applications, such as on-body interaction, bimanual interaction, cursor control and 3D measurement.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**; **Pointing devices**; **Pointing**.

KEYWORDS

Single-handed, gesture, on-body interaction, smartwatch

ACM Reference Format:
Hui-Shyong Yeo, Juyoung Lee, Andrea Bianchi, Alejandro Samboy, Hideki Koike, Woontack Woo, and Aaron Quigley. 2020. WristLens: Enabling Single-Handed Surface Gesture Interaction for Wrist-Worn Devices Using Optical Motion Sensor. In *AHs '20: Augmented Humans International Conference (AHs '20), March 16–17, 2020, Kaiserslautern, Germany*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3384657.3384797

AHs '20, March 16–17, 2020, Kaiserslautern, Germany

ACM ISBN 978-1-4503-7603-7/20/03…$15.00
https://doi.org/10.1145/3384657.3384797

1 INTRODUCTION

WristLens aims to enhance and augment human capabilities for input and interaction with arbitrary surfaces, including our own body. Instead of instrumenting the environments with sensors, here we leverage natural uses of existing wearable devices to amplify the human body and proximate surfaces for interaction.

Wrist-worn devices such as smartwatches and fitness trackers are now commonplace. They provide useful features such as notifications, health monitoring and map navigation, which sees them in regular use. A prior evaluation, in-the-wild, measured an average usage of 5.4 times per hour [32] for such devices, demonstrating the potential of smartwatches as companion devices.

However, interacting with such wearables typically necessitates both hands, as one arm is wearing the device while the other hand is required to operate the touch screen or physical buttons. This can be difficult or not possible, especially if the other hand is not available, such as when holding or operating something (coffee, umbrella, remote controller) or if the user has diminished physical capabilities. While alternative interaction methods based on voice input (e.g., "Ok Google") or midair gestures (e.g., WearOS's [6] tilt to wake or scroll) are certainly possible, these techniques are often error prone and can be socially awkward [39] to perform in public.

This work takes advantage of the fact that the wrist, and hence the device strap, is often close to one's body or other surfaces one is interacting with. We introduce WristLens, a system to enable surface gesture recognition for wrist-worn devices, based on optical motion sensing. It allows the user to leverage arbitrary nearby surfaces for input and interaction, including the user's own body, thus allowing single-handed and eyes-free interaction. With an optical sensor embedded in the device-strap, the user can glide on any surface as if using a computer mouse.

In addition, unlike midair gestures, leveraging nearby surfaces for interaction provides passive tactile feedback and can help avoid unintended gestures (e.g., Midas touch). The surface also acts as a support for the hand, thus reducing hand fatigue caused by gorilla arm effects [17]. Finally, WristLens also enables novel use cases such as eyes-free interaction, on-body interaction, bimanual interaction, cursor control and object measurement — scenarios that we will describe in the applications section.

2 RELATED WORK

Our related work spans multiple research areas within HCI. Here we focus on on-body, on-surface and single-handed interaction which are most relevant for this form of device-strap to surface input. We also discuss social acceptability of such interfaces.

2.1 On-body Interaction

Body-based interaction naturally offers a suitable surface for gestural interaction, providing passive tactile feedback [44]. It also affords proprioception and supports eyes-free input techniques for mobile and wearable scenarios. As an example, the work by Wagner et al. [45] presents an overview of the usability and social acceptance implications of using the body as a canvas for input for body-centric interaction techniques.

To achieve on-body input, several technologies have been explored. A common approach is to augment the clothes worn by a user with sensors such as RFID [9], capacitive array [36] or conductive threads [19, 33]. However, the interactive capability is then only available when the user is wearing that particular piece of cloth. Alternative technologies such as a wearable electronic skin [31, 46] are also possible, but are limited to the worn area only.

A different approach requires the user to wear a wearable device. Skinput [14] detects the location of finger taps on the arm and hand by analyzing mechanical vibrations that propagate through the body using a novel array of sensors worn as an armband, whereas OmniTouch [11] achieves similar detection using a depth camera mounted on the shoulder. Gesture such as rubbing the face [25] or touching the nose [22] can also be detected with wearable sensors.

A final approach relies on external (grounded) sensors. For example, Run&Tap [10] investigates on-body tapping as a potential input technique for runners. Belly gestures [44] supports unistroke gesture patterns on the abdomen. However, these systems were studied with bulky hardware or external sensors such as the Kinect and Optitrack, which are incompatible with the wearable application scenarios envisaged here.

2.2 On-surface Interaction

While the previous section focused on interaction techniques that use the body as an input canvas, this section describes methods to detect input on surfaces other than the body. A common approach is to augment the surface [37], along with, for example, a projector-camera (pro-cam) system [47, 48]. However, this approach requires a bulky apparatus and is not portable for wearable applications. Anywhere Surface Touch [29] consists of a camera system worn below the wrist, pointing towards the finger area to detect finger taps on any surface it is resting on. LightRing [20] consists of a gyroscope and infrared sensor that detects finger flexion and palm rotation when the wrist is resting on any surface.

By contrast, our approach employs a small and wearable optical motion sensor. Indeed, this sensor has been used extensively to enable novel input techniques [13, 27, 30, 49, 50] in the HCI community. Much of the prior work leverages the optical motion sensing combined with other sensors (such as macro camera, accelerometer and proximity sensor) for augmenting input directly on the fingers. The most similar work to ours is Magic Finger [49], which allows a user to control a mouse using a finger and a rich set of interactions based on the relative motion of a finger on any surface. However, this technique requires a specialized micro camera and a relatively large sensing box, hindering practical applications in the wild. Heo et al. [16] embedded an optical motion sensor below the watch strap, but it requires a second hand to perform swipe gestures from below. Finally, Ni and Baudisch [28] demonstrated how it is possible to repurpose an optical mouse to simulate a motion scanner device.

2.3 Single-Handed Interaction

To address the shortcoming of smartwatch interactions typically requiring two hands for input, several researchers have proposed methods that enable single-hand interaction based on mid-air gesture [1, 24, 53] and static posture recognition [54]. Mid-air motion gestures can cause arm fatigue [17], as well as being error-prone [4] and socially awkward [38], whereas static posture lacks expressiveness (e.g., no support for touch and swipe gestures). In contrast, our approach is based on gesture input on the surface itself, hence providing a physical support for the arm and limiting excessive fatigue [44].

2.4 Social Acceptability

A common issue with on-body interfaces is their social acceptability, as these devices often require the user to touch their own body parts or to perform awkward midair gesture in public in order to operate them. Profita et al. [34] and Harrison et al. [12] studied the societal perceptions of textile or projected interfaces at different on-body locations. Whack Gestures [18] aims to allow inexact and inattentive interaction with mobile devices, thus minimizing social acceptability concerns. In determining social acceptance of such interfaces, Montero et al. [26] show that the user's perception of others' ability to notice them is an important factor. Rico et al. [38] studied social acceptability with respect to location and audience, and further provide design recommendations and evaluation guidelines. In reality, what is considered socially acceptable is evolving as new devices and forms of interaction enter day to day life.

3 DESIGN AND PROTOTYPE

Given the ubiquity of surfaces on and around us, our overall goal is to broaden the range of surfaces (including the user's body) and hence the nature of input available in single-handed interactions. Hence, the motivation is to design and implement an input technique for wrist-worn devices that can support (i) eyes-free input (ii) and can operate with only one hand. Furthermore, it can iii) leverage a nearby surface as a canvas such as the user's body or surrounding flat objects (e.g., a table or wall), which implicitly also provides (iv) passive haptic feedback. Put simply, the user can control the spatial location of the input (relative x-y coordinates), consequently enabling gestures (touch, swipe and draw shapes).

3.1 Hardware

Our system employs the optical motion sensor commonly found in modern computer mice. This sensoor not only fulfills the above

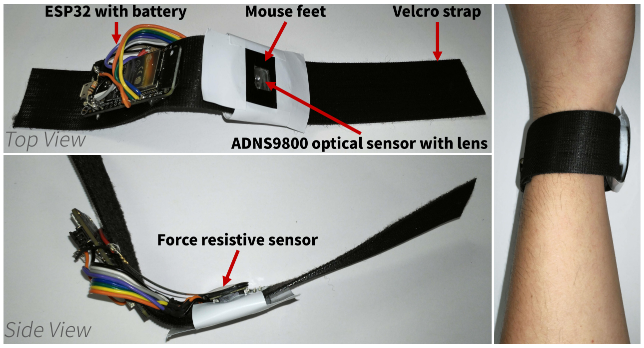

Figure 1: WristLens consists of a wrist band with optical sensor attached below. The sensor is connected to an ESP32 micro controller that is able to process and send data wirelessly.

requirements, but it is also very low-cost, readily available, robust and a proven technology. This sensor measures changes in position by obtaining sequential surface images and mathematically determining the direction and magnitude of the movement. Typical laser optical sensors work on most surfaces except highly reflective ones such as glass, although specific high-end laser optical sensors can also handle glass and other reflective materials (e.g., Logitech's Darkfield technology).

Our setup here includes an Avago ADNS-9800 [42] sensor fitted with an ADNS-6190-002 [41] lens, which is attached to the bottom part of a Velcro wrist band (Figure 1), while an Android smartwatch can be attached on the top (not shown in the figure). The sensor is connected to an ESP32 micro controller through SPI, along with a small lithium battery. A force sensitive resistor (FSR) was fitted to enable hard press for selections through clicking. The sensor readings are polled and sent wirelessly to a PC or directly to the smartwatch through WiFi using UDP to minimize latency.

3.2 Software

We created the experimental software (written in Java on a PC) and several demo applications (written in the Unity game engine). We used a modified $P Point-Cloud Recognizer [43] that can recognize both directional gestures and shape gestures. In fact, the original $P recognizer is direction invariant by design, and therefore cannot recognize the direction of unistroke gestures. For the optical sensor, we fixed the resolution at 400 CPI with a refresh rate of 50Hz. We also used the smartwatch's IMU data to account for wrist rotation, and hence achieve 3D surface tracking.

4 EXPERIMENTAL STUDY

We conducted an experiment to evaluate the accuracy of the gesture recognition software on three distinct body parts, following similar studies in prior work [13, 34, 44, 45]: *THIGH*, *ABDOMEN* and *OTHER ARM*. Specifically, we selected the upper *THIGH* around the hip area because the arm is just beside this area when in a relaxed posture [23]. Inattentive "whack gestures" [18] were also performed in this area, which were also previously explored by Profita et al. [34]. For *ABDOMEN*, Vo et al. [44] suggested the abdomen is especially

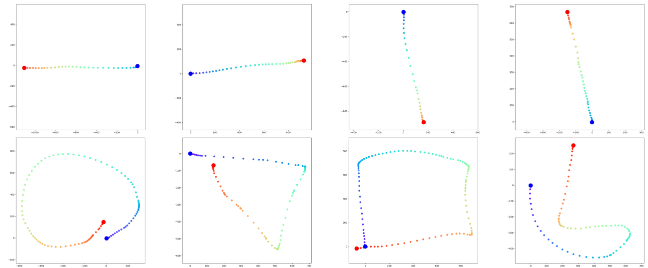

Figure 2: Selected sample gestures drawn by participants.

appropriate for gesture interaction as it offers a fairly large surface that can be easily reached with any hand in any circumstance. Lastly, Profita et al. [34] also demonstrated that the *OTHER ARM* is one of the preferred areas for interaction as it is unobtrusive, can be easily accessed, and appears the least "awkward" or the most "normal" to participants.

4.1 Participants

We recruited 12 participants (3 females, 1 left-handed) from University of St Andrews. Their ages ranged from 19 to 30 (M: 24.1, SD: 3.9). They were compensated with a £10 voucher. Four participants had experience with wearable devices (fitness tracker or smartwatch), and two of them wore the device daily. Participants were asked to stand throughout the experiment while facing a monitor approximately 1 meter away. They were asked to wear the device on the same hand they would normally wear a watch. As a result, ten participants chose to wear the device on the non-dominant hand but two preferred to wear it on their dominant hand.

4.2 Procedure

Before the experiment started, participants were instructed to try out all the gestures and conditions to gain familiarity with the system. During this time, gesture results and trails were shown on a display. Then, 2 blocks of data were collected and used as a template for the $P recognizer for real-time recognition. Next, 10 blocks of real-time data were collected. For the first 4 blocks (considered as training phase and hence discarded in the analysis of the results), the recognized gesture was shown on screen after each trial. There was no visual feedback instead for the remaining 6 blocks, participants performed the gestures using their imagination and muscle memory.

For each trial, the participants attempted the gesture by approaching the wrist to the appropriate body surface with the sensor facing the surface, then they drew the gesture, and finally left the surface. This process was repeated for 3 different conditions. In total, we collected 2880 valid data points = 12 participants x 3 body parts (*THIGH*, *ABDOMEN*, *OTHER ARM*) x 8 gestures (4 directional and 4 shape) x 10 blocks. Conditions were fully counterbalanced and the order of the gestures within each block was randomized. The gestures are motivated by previous work [8, 53]. Examples by participants can be seen in Figure 2 and 3. At the end of each condition, participants completed a NASA TLX [15] survey to assess the perceived workloads, level of comfort and social acceptability in front of colleagues or strangers. The study took approximately 60 minutes in total.

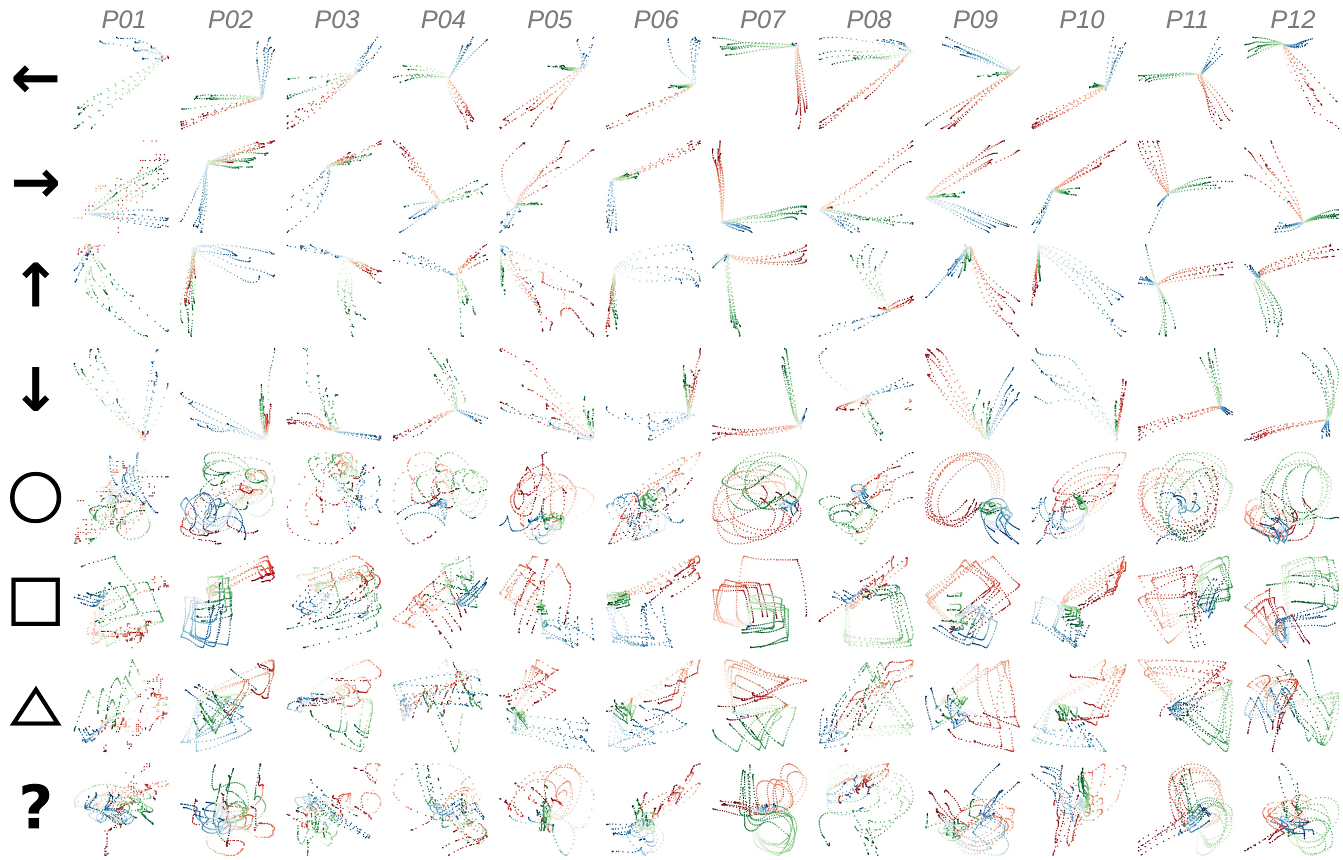

Figure 3: Participant data points from block 5 to 10, excluding practice blocks, each column is a participant, from left to right: P1 to P12. Gestures from top to bottom: left swipe, right swipe, up swipe, down swipe, rectangle, circle, triangle and question mark. Red: THIGH, Green: ABDOMEN, Blue: OTHER ARM.

THIGH

	←	→	↑	↓	↓	□	△	?
←|Left swipe	93.1% (67)	0%	0%	1.4% (1)	4.2% (3)	1.4% (1)	0%	0%
→|Right swipe	11.1% (8)	86.1% (62)	0%	0%	0%	1.4% (1)	1.4% (1)	0%
↑|Up swipe	5.6% (4)	0%	94.4% (68)	0%	0%	0%	0%	0%
↓|Down swipe	4.2% (3)	2.8% (2)	2.8% (2)	88.9% (64)	0%	0%	0%	1.4% (1)
↓|Circle	4.2% (3)	5.6% (4)	0%	2.8% (2)	55.6% (40)	6.9% (5)	16.7% (12)	8.3% (6)
□|Rectangle	5.6% (4)	2.8% (2)	0%	0%	9.7% (7)	55.6% (40)	19.4% (14)	6.9% (5)
△|Triangle	5.6% (4)	6.9% (5)	0%	1.4% (1)	2.8% (2)	12.5% (9)	69.4% (50)	1.4% (1)
?|Question	4.2% (3)	1.4% (1)	4.2% (3)	1.4% (1)	12.5% (9)	9.7% (7)	8.3% (6)	58.3% (42)

ABDOMEN

	←	→	↑	↓	↓	□	△	?
←|Left swipe	98.6% (71)	0%	0%	0%	0%	0%	0%	1.4% (1)
→|Right swipe	1.4% (1)	97.2% (70)	0%	0%	0%	0%	1.4% (1)	0%
↑|Up swipe	0%	0%	97.2% (70)	1.4% (1)	0%	0%	1.4% (1)	0%
↓|Down swipe	0%	0%	1.4% (1)	98.6% (71)	0%	0%	0%	0%
↓|Circle	5.6% (4)	2.8% (2)	0%	1.4% (1)	65.3% (47)	2.8% (2)	11.1% (8)	11.1% (8)
□|Rectangle	2.8% (2)	1.4% (1)	1.4% (1)	1.4% (1)	13.9% (10)	65.3% (47)	9.7% (7)	4.2% (3)
△|Triangle	4.2% (3)	0%	0%	0%	2.8% (2)	6.9% (5)	79.2% (57)	6.9% (5)
?|Question	0%	0%	5.6% (4)	2.8% (2)	2.8% (2)	9.7% (7)	4.2% (3)	75.0% (54)

OTHER ARM

	←	→	↑	↓	↓	□	△	?
←|Left swipe	90.3% (65)	0%	1.4% (1)	6.9% (5)	0%	1.4% (1)	0%	0%
→|Right swipe	0%	95.8% (69)	0%	4.2% (3)	0%	0%	0%	0%
↑|Up swipe	5.6% (4)	1.4% (1)	86.1% (62)	4.2% (3)	0%	0%	2.8% (2)	0%
↓|Down swipe	0%	0%	0%	98.6% (71)	1.4% (1)	0%	0%	0%
↓|Circle	8.3% (6)	6.9% (5)	9.7% (7)	6.9% (5)	37.5% (27)	13.9% (10)	12.5% (9)	4.2% (3)
□|Rectangle	4.2% (3)	1.4% (1)	6.9% (5)	4.2% (3)	12.5% (9)	58.3% (42)	6.9% (5)	5.6% (4)
△|Triangle	1.4% (1)	2.8% (2)	6.9% (5)	2.8% (2)	6.9% (5)	9.7% (7)	65.3% (47)	4.2% (3)
?|Question	6.9% (5)	2.8% (2)	12.5% (9)	9.7% (7)	6.9% (5)	6.9% (5)	6.9% (5)	47.2% (34)

Figure 4: Confusion matrix for condition THIGH, ABDOMEN and OTHER ARM.

4.3 Results

The first 4 blocks were considered as training data and discarded from the analysis. Results were analyzed using one- and two-way ANOVA tests followed by Bonferroni correction post-hoc analysis with $\alpha = 0.01$. Average time and accuracy were not statistically different across the three conditions and average results per condition are reported in Figure 5. A deeper analysis of the errors, reveals instead that errors for different gestures were statistically different across conditions ($F_{(2,264)} = 6.5$, $p < 0.01$) and gestures ($F_{(7,264)} = 43$, $p < 0.01$) but not their interaction. Post-hoc comparison reveals that the directional gestures (Up, Down, Left and Right) were statistically different from shape gestures (Question mark, Triangle, Circle and Rectangle), with the second group causing many more errors than the simple gestures. Confusion matrices for gestures in

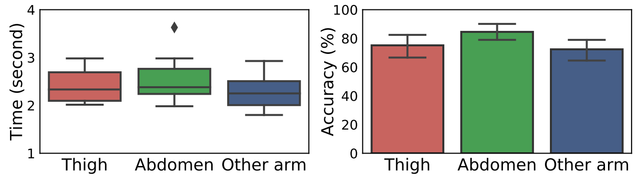

Figure 5: Completion time and accuracy across 3 conditions.

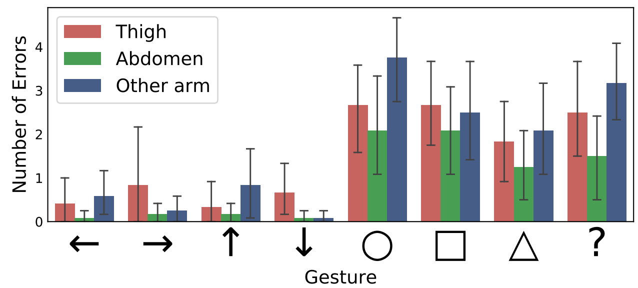

Figure 6: Errors for each type of gesture.

each condition are shown in Figure 4. Finally, we found significant differences for the workload, measured with the NASA TLX questionnaire ($F_{(2,33)} = 3.3$, $p < 0.05$), with *OTHER ARM* significantly more demanding than *ABDOMEN*.

5 DISCUSSION

Overall, directional gestures are accurately recognized but less so for shape gestures. There are several factors that contribute to this result. Naturally, shape gestures are more difficult to draw consistently than simple directional line gestures. This issue becomes more prominent when gestures are performed on the body, due to surface friction and different spatial orientations of users, as we explain below. Fortunately, in many applications for wearable devices, directional gestures combined with and tap detection are often sufficient for basic interaction, such as swipe to scroll through a list and tap to select an item within the list.

Users spatial orientation - From the visualized gesture trails (Figure 3), we can observe that different participants (indicated by columns) performed the same gestures with inconsistent directions/orientations even for the same body parts (indicated by colors). In fact, Vo et al. [44] noted that users employ different mental spatial orientations depending on the complexity of the gesture they have to draw. In addition, when no visual orientation cues are provided users often draw gestures following symmetries relative to the current view and their perception of the horizontal and vertical axis.

This means that users have a different perception of what is left and right and what is top and bottom on their body, i.e., they may invert one or both axes when prompted to perform a gesture. For example, some users related the lower part of their torso to the upper part of the screen, versus the upper part of the torso to the lower part of the screen, i.e., some moved their hand down to move the cursor up, and vice-versa. Indeed, during the practice sessions, some participants were confused and kept changing the manner in which they would perform certain gestures even on the same surface. To mitigate this, we asked participants to decide on one way and to be consistent throughout the experiment. Yet, we can still notice that P8 and P11 performed a mistake at the gesture DOWN and UP, respectively.

Non-dominant hand - The majority of participants performed gestures with their non-dominant hand, which, by nature, does not possess the same fine motor skills as the dominant hand. The experiment could have been performed with the dominant hand, and therefore we might expect better results but less ecological validity as the majority of people wear watches on their non-dominant hand. From the visualized gesture trails (Figure 3), we can observe that the data of majority participants were noisy, except for participant 7, 9 and 11, which look particularly clean. The cleaner data allows for better recognition rate, suggesting that the recognition is highly user-dependent. Using more advanced machine learning technique could improve the recognition rate.

Surface properties - The sensor readings also depend on the surface properties such as the size of the area, its flatness and material type. In particular, performing shape gestures on the OTHER ARM was difficult due to lack of surface area. In fact, it is difficult to draw a gesture with an appropriate size in the absence of a large enough canvas, as is visualized with the blue lines in Figure 3. Clearly those lines are shorter than those drawn on the other two body parts. Furthermore, on some occasions, some participants did not realize that the sensor was actually outside the boundaries of the surface, and hence not touching the surface of the forearm, resulting in missing data points. To compensate for this issue, it was also observed that some participants would slightly rotate their arm during gesture input, as a way to create more surface area.

Loose clothing can also cause noisy readings, since these tend to form folds which makes sliding movements difficult. Specifically, sliding movements that stretch the cloth are easier than movements that fold the cloth. Some participants also had to empty their pants pocket (smartphone and car key) as they create uneven surfaces. As indicated before, the laser optical sensor works well on many different surface types but not all types such as reflective surfaces (e.g., glass). Through our 12 participants study, we found that our sensor works well on almost all types of clothes worn by our participants, except on one type of khakis pant that appears reflective to the sensor, as observed in the noisy data from P1 in Figure 3 (red)).

Gesture delimiter - Initially we utilized a force sensitive resistor to detect a hard press as the gesture delimiter. However, during a pilot test we observed it was unnatural and cumbersome to maintain a consistent pressure when drawing the gesture. Therefore we removed this requirement, and used only the sensor's built-in lift detection threshold. Yet, it was noted during the experiment that participants sometimes accidentally performed two consecutive gestures, even when they lift the device slowly for adjustments.

User feedback - From the questionnaire results, it can be seen that ABDOMEN has the lowest workload, followed by THIGH and then OTHER ARM. However, ABDOMEN was also voted the least socially acceptable location to perform these gestures, either in front of friends or strangers. Therefore, THIGH may be the

Figure 7: Left: 3D object measurement tool. Middle: Bimanual interaction for controlling stylus properties. Right: Anywhere cursor control for inputting text using gesture keyboard technique.

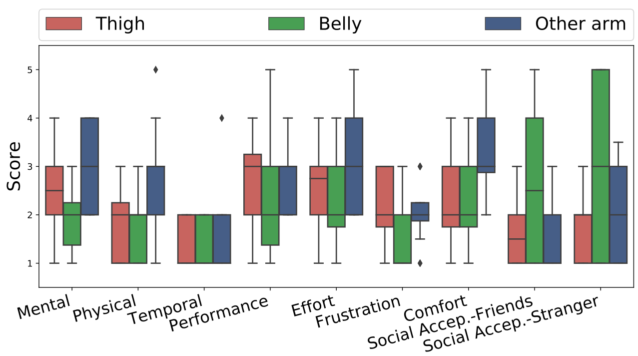

Figure 8: Standard NASA TLX score (range from 1 to 5), along with comfortability and social acceptability scores, which are inverted to match plot. Lower scores are better.

participants' most preferred body location, maintaining a balance between perceived workload and social acceptability.

6 APPLICATIONS AND USE CASES

In this section, we demonstrate how WristLens can enable several applications and scenarios leveraging both gesture input and 2D position tracking, as shown in Figure 7, 9 and the video figure.

On-body Interaction: WristLens enables surface gesture interaction on body parts while on the go (Figure 9). This can be useful to control mobile or wearable devices, such as dismissing notifications, skipping to the next song, changing volume, etc. This can be performed single-handedly and an in eyes-free manner, effortlessly near the upper thigh or abdomen area, without requiring a user to take out the phone from pockets, which might causes higher cognitive load, especially while on the go. Extending input to simple body interaction opens up the space of what is possible for interactive spaces, without requiring the instrumentation of new surfaces, or bodies with other sensors.

Bimanual Interaction: Since smartwatches are often worn on the non-dominant hand, WristLens can be leveraged to enable bimanual interaction in desktop or tablet environments. For example, while holding a mouse on the right hand and wearing WristLens on the left hand, standard two-handed gestures such as zoom, rotate, panning can be realized (see video figure). While using a stylus on the right hand to paint, the left hand can intuitively swipe or rotate

Figure 9: Left: taking out the phone to control the music player while walking on the street is dangerous, but mitigated by WristLens. Right: User can draw gestures on the body to perform simple commands without losing focus.

on a surface to alter the properties of the painting brush, such as color, size and tilt angle (Figure 7 middle).

Anywhere Cursor Control: WristLens also supports cursor control, for example in situations where a mouse is not available, such as in a conference meeting space, or when using a laptop in a coffee shop and users have only access to the computers' trackpad. In addition, combining cursor control with gesture keyboard [21] allows users to input text with only one hand (Figure 7 right). Indeed, WristLens works almost like a computer mouse since it uses the same optical sensor. The difference is that the center point is at the wrist instead of the palm, so it might require more arm movement instead of subtle wrist movements. Nonetheless, this can potentially avoid repetitive stress injuries in the wrist joint, since movement from the elbow causes less wrist strain.

Dirty or Covered Hand In some scenarios where the hand is dirty (e.g., kitchen, operation room, factory), the user might be reluctant to touch the screen or hold a pointing device. With WristLens, the user can use the smartwatch strap on the wrist area to glide over any nearby surface to achieve point and click input.

3D Measuring Tool: By measuring the distance travelled by the optical motion sensor when sliding on a surface, WristLens can be used to achieve 2D measurements. Furthermore, by combining this functionality with the rotational tracking in the smartwatch, objects can be measured in 3D (Figure 7 left and video figure).

Future Applications We also envision future applications that can be enabled with WristLens. For one, text can be scanned [3] and digitized by sliding through WristLens as if using a marker pen. Different textures (clothing or desk surfaces) can be recognized

[40, 52], enabling placement awareness and shortcut commands. This technique further enables cross-device interaction such as picking and dropping data [35]. Albeit not the focus of this work, it could also allow under side or back-of-device [2, 50, 51] interaction, avoiding occlusion and the fat-finger problem [5]. Finally, we believe that WristLens can be useful for virtual reality environment because proprioception allows users to accurately touch their body parts without looking at them while wearing VR display.

7 LIMITATIONS AND FUTURE WORK

This work presents several limitations related to the current prototype, which we aim to overcome in future work.

Form factor: In the current prototype, the embedded optical sensor is extruded from the watch-strap of 5mm, because the lens itself is 3mm and it requires a 2-3mm gap for the focal length. Therefore, it is not as flat as a standard watch-strap. Future work will need to improve the form factor and shrink the thickness (smaller sensor and lens) of the device.

Limited evaluation: Our experimental study was conducted with participants standing still. Thus, we cannot generalize our results for situations in which the user is on the move, such as walking or running. We also did not evaluate the performance of 2D targeting and how it compares to a computer mouse. In future work, a complete evaluation with these conditions should be conducted.

Surface recognition: Our current prototype does not recognize which surface/body part it is touching, thus it could not enable context-aware interaction such as body-shortcuts. We could utilize the IMU data to infer the touched body part [7] but this is very user dependent, because, as seen in the results, different users orient their wrist differently over time. In future work, we propose to use the raw image captured by the optical sensor for surface/texture recognition [40] using machine learning techniques.

8 CONCLUSION

We presented WristLens — a system that enables surface gesture interaction using a wrist-worn wearable device. The system is based on a low-cost optical motion sensor commonly found in computer mice. Our evaluation results show that directional gestures can be accurately recognized on different body parts, but less so for shape gestures. Participants also feel it's comfortable and socially acceptable to perform these gestures on the thighs and on the other arm in public, but less acceptable on the abdomen area. While more work needs to be done to fully evaluate the system in real-world conditions, we envision such type of sensor, when included in future wearable device, unlocks a high potential interaction modality.

ACKNOWLEDGMENTS

Andrea Bianchi was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICT Consilience Creative program (IITP-2020-2015-0-00742) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation). Juyoung Lee and Woontack Woo were supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2019-0-01270, WISE AR UI/UX Platform Development for Smartglasses).

REFERENCES

[1] Teo Babic, Harald Reiterer, and Michael Haller. 2017. GestureDrawer: One-handed Interaction Technique for Spatial User-defined Imaginary Interfaces. In *Proceedings of the 5th Symposium on Spatial User Interaction (SUI '17)*. ACM, New York, NY, USA, 128–137. https://doi.org/10.1145/3131277.3132185

[2] Patrick Baudisch and Gerry Chu. 2009. Back-of-device Interaction Allows Creating Very Small Touch Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 1923–1932. https://doi.org/10.1145/1518701.1518995

[3] bidouille. 2019. Optical Mouse Cam. https://www.bidouille.org/hack/mousecam.

[4] Doug A. Bowman, Ernst Kruijff, Joseph J. LaViola, and Ivan Poupyrev. 2004. *3D User Interfaces: Theory and Practice.* Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA.

[5] Bill Buxton. 2007. *Extending Interaction.* 331–336. https://doi.org/10.1016/B978-012374037-3/50077-8

[6] Wear OS by Google. 2019. Navigate your watch with wrist gestures. https://support.google.com/wearos/answer/6312406?hl=en.

[7] Xiang 'Anthony' Chen and Yang Li. 2016. Bootstrapping User-Defined Body Tapping Recognition with Offline-Learned Probabilistic Representation. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 359–364. https://doi.org/10.1145/2984511.2984541

[8] 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 (UIST '16)*. Association for Computing Machinery, New York, NY, USA, 861–872. https://doi.org/10.1145/2984511.2984563

[9] Tiago Guerreiro, Ricardo Gamboa, and Joaquim Jorge. 2009. Mnemonical Body Shortcuts for Interacting with Mobile Devices. In *Gesture-Based Human-Computer Interaction and Simulation*, Miguel Sales Dias, Sylvie Gibet, Marcelo M. Wanderley, and Rafael Bastos (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 261–271.

[10] Nur Al-huda Hamdan, Ravi Kanth Kosuru, Christian Corsten, and Jan Borchers. 2017. Run&Tap: Investigation of On-Body Tapping for Runners. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS '17)*. ACM, New York, NY, USA, 280–286. https://doi.org/10.1145/3132272.3134140

[11] Chris Harrison, Hrvoje Benko, and Andrew D. Wilson. 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, 441–450. https://doi.org/10.1145/2047196.2047255

[12] Chris Harrison and Haakon Faste. 2014. Implications of Location and Touch for On-body Projected Interfaces. In *Proceedings of the 2014 Conference on Designing Interactive Systems (DIS '14)*. ACM, New York, NY, USA, 543–552. https://doi.org/10.1145/2598510.2598587

[13] Chris Harrison and Scott E. Hudson. 2010. Minput: Enabling Interaction on Small Mobile Devices with High-precision, Low-cost, Multipoint Optical Tracking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 1661–1664. https://doi.org/10.1145/1753326.1753574

[14] Chris Harrison, Desney Tan, and Dan Morris. 2010. Skinput: Appropriating the Body As an Input Surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 453–462. https://doi.org/10.1145/1753326.1753394

[15] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. https://doi.org/10.1177/154193120605000909 arXiv:https://doi.org/10.1177/154193120605000909

[16] Seongkook Heo, Michelle Annett, Benjamin Lafreniere, Tovi Grossman, and George Fitzmaurice. 2017. No Need to Stop What You'Re Doing: Exploring No-Handed Smartwatch Interaction. In *Proceedings of the 43rd Graphics Interface Conference (GI '17)*. Canadian Human-Computer Communications Society, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 107–114. https://doi.org/10.20380/GI2017.14

[17] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-air Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1063–1072. https://doi.org/10.1145/2556288.2557130

[18] Scott E. Hudson, Chris Harrison, Beverly L. Harrison, and Anthony LaMarca. 2010. Whack Gestures: Inexact and Inattentive Interaction with Mobile Devices. In *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '10)*. ACM, New York, NY, USA, 109–112. https://doi.org/10.1145/1709886.1709906

[19] Thorsten Karrer, Moritz Wittenhagen, Leonhard Lichtschlag, Florian Heller, and Jan Borchers. 2011. Pinstripe: Eyes-Free Continuous Input on Interactive Clothing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1313–1322. https://doi.org/10.1145/1978942.1979137

[20] Wolf Kienzle and Ken Hinckley. 2014. LightRing: Always-Available 2D Input on Any Surface. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. Association for Computing Machinery, New York, NY, USA, 157–160. https://doi.org/10.1145/2642918.2647376

[21] Per-Ola Kristensson and Shumin Zhai. 2004. SHARK2: A Large Vocabulary Shorthand Writing System for Pen-based Computers. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology (UIST '04)*. ACM, New York, NY, USA, 43–52. https://doi.org/10.1145/1029632.1029640

[22] Juyoung Lee, Hui-Shyong Yeo, Murtaza Dhuliawala, Jedidiah Akano, Junichi Shimizu, Thad Starner, Aaron Quigley, Woontack Woo, and Kai Kunze. 2017. Itchy Nose: Discreet Gesture Interaction Using EOG Sensors in Smart Eyewear. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers (ISWC '17)*. Association for Computing Machinery, New York, NY, USA, 94–97. https://doi.org/10.1145/3123021.3123060

[23] Mingyu Liu, Mathieu Nancel, and Daniel Vogel. 2015. Gunslinger: Subtle Arms-down Mid-air Interaction. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, 63–71. https://doi.org/10.1145/2807442.2807489

[24] Anders Markussen, Mikkel Rønne Jakobsen, and Kasper Hornbæk. 2014. Vulture: A Mid-air Word-gesture Keyboard. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1073–1082. https://doi.org/10.1145/2556288.2556964

[25] Katsutoshi Masai, Yuta Sugiura, and Maki Sugimoto. 2018. FaceRubbing: Input Technique by Rubbing Face Using Optical Sensors on Smart Eyewear for Facial Expression Recognition. In *Proceedings of the 9th Augmented Human International Conference (AH '18)*. Association for Computing Machinery, New York, NY, USA, Article Article 23, 5 pages. https://doi.org/10.1145/3174910.3174924

[26] Calkin S. Montero, Jason Alexander, Mark T. Marshall, and Sriram Subramanian. 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)*. ACM, New York, NY, USA, 275–278. https://doi.org/10.1145/1851600.1851647

[27] A. Nguyen and A. Banic. 2015. 3DTouch: A wearable 3D input device for 3D applications. In *2015 IEEE Virtual Reality (VR)*. 55–61. https://doi.org/10.1109/VR.2015.7223324

[28] Tao Ni and Patrick Baudisch. 2009. Disappearing 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, 101–110. https://doi.org/10.1145/1622176.1622197

[29] Takehiro Niikura, Yoshihiro Watanabe, and Masatoshi Ishikawa. 2014. Anywhere Surface Touch: Utilizing Any Surface as an Input Area. In *Proceedings of the 5th Augmented Human International Conference (AH '14)*. Association for Computing Machinery, New York, NY, USA, Article Article 39, 8 pages. https://doi.org/10.1145/2582051.2582090

[30] Shahriar Nirjon, Jeremy Gummeson, Dan Gelb, and Kyu-Han Kim. 2015. TypingRing: A Wearable Ring Platform for Text Input. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '15)*. ACM, New York, NY, USA, 227–239. https://doi.org/10.1145/2742647.2742665

[31] Aditya Shekhar Nittala, Anusha Withana, Narjes Pourjafarian, and Jürgen Steimle. 2018. Multi-Touch Skin: A Thin and Flexible Multi-Touch Sensor for On-Skin Input. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 33, 12 pages. https://doi.org/10.1145/3173574.3173607

[32] Stefania Pizza, Barry Brown, Donald McMillan, and Airi Lampinen. 2016. Smartwatch in Vivo. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5456–5469. https://doi.org/10.1145/2858036.2858522

[33] Ivan Poupyrev, Nan-Wei Gong, Shiho Fukuhara, Mustafa Emre Karagozler, Carsten Schwesig, and Karen E. Robinson. 2016. Project Jacquard: Interactive Digital Textiles at Scale. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 4216–4227. https://doi.org/10.1145/2858036.2858176

[34] Halley P. Profita, James Clawson, Scott Gilliland, Clint Zeagler, Thad Starner, Jim Budd, and Ellen Yi-Luen Do. 2013. Don'T Mind Me Touching My Wrist: A Case Study of Interacting with On-body Technology in Public. In *Proceedings of the 2013 International Symposium on Wearable Computers (ISWC '13)*. ACM, New York, NY, USA, 89–96. https://doi.org/10.1145/2493988.2494331

[35] Jun Rekimoto. 1997. Pick-and-Drop: A Direct Manipulation Technique for Multiple Computer Environments. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology (UIST '97)*. Association for Computing Machinery, New York, NY, USA, 31–39. https://doi.org/10.1145/263407.263505

[36] Jun Rekimoto. 2001. GestureWrist and GesturePad: Unobtrusive Wearable Interaction Devices. In *Proceedings of the 5th IEEE International Symposium on Wearable Computers (ISWC '01)*. IEEE Computer Society, USA, 21.

[37] Jun Rekimoto and Masanori Saitoh. 1999. Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. Association for Computing Machinery, New York, NY, USA, 378–385. https://doi.org/10.1145/302979.303113

[38] Julie Rico and Stephen Brewster. 2010. Usable Gestures for Mobile Interfaces: Evaluating Social Acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 887–896. https://doi.org/10.1145/1753326.1753458

[39] Steven Schirra and Frank R. Bentley. 2015. "It's Kind of Like an Extra Screen for My Phone": Understanding Everyday Uses of Consumer Smart Watches. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. ACM, New York, NY, USA, 2151–2156. https://doi.org/10.1145/2702613.2732931

[40] Dennis Schüsselbauer, Andreas Schmid, Raphael Wimmer, and Laurin Muth. 2018. Spatially-Aware Tangibles Using Mouse Sensors. In *Proceedings of the Symposium on Spatial User Interaction (SUI '18)*. ACM, New York, NY, USA, 173–173. https://doi.org/10.1145/3267782.3274690

[41] Avago Technologies. 2019. ADNS-6190-002 LaserStream Gaming Round Lens. https://datasheet.octopart.com/ADNS-6190-002-Avago-datasheet-8326645.pdf.

[42] Avago Technologies. 2019. ADNS-9800 LaserStream Gaming Sensor. https://datasheet.octopart.com/ADNS-9800-Avago-datasheet-10666463.pdf.

[43] Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. 2012. Gestures As Point Clouds: A $P Recognizer for User Interface Prototypes. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction (ICMI '12)*. ACM, New York, NY, USA, 273–280. https://doi.org/10.1145/2388676.2388732

[44] Dong-Bach Vo, Eric Lecolinet, and Yves Guiard. 2014. Belly Gestures: Body Centric Gestures on the Abdomen. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI '14)*. ACM, New York, NY, USA, 687–696. https://doi.org/10.1145/2639189.2639210

[45] Julie Wagner, Mathieu Nancel, Sean G. Gustafson, Stephane Huot, and Wendy E. Mackay. 2013. Body-centric Design Space for Multi-surface Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1299–1308. https://doi.org/10.1145/2470654.2466170

[46] Martin Weigel, Tong Lu, Gilles Bailly, Antti Oulasvirta, Carmel Majidi, and Jürgen Steimle. 2015. iSkin: Flexible, Stretchable and Visually Customizable On-Body Touch Sensors for Mobile Computing. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2991–3000. https://doi.org/10.1145/2702123.2702391

[47] Andrew D. Wilson. 2005. PlayAnywhere: A Compact Interactive Tabletop Projection-Vision System. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology (UIST '05)*. Association for Computing Machinery, New York, NY, USA, 83–92. https://doi.org/10.1145/1095034.1095047

[48] Robert Xiao, Scott Hudson, and Chris Harrison. 2016. DIRECT: Making Touch Tracking on Ordinary Surfaces Practical with Hybrid Depth-Infrared Sensing. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces (ISS '16)*. Association for Computing Machinery, New York, NY, USA, 85–94. https://doi.org/10.1145/2992154.2992173

[49] Xing-Dong Yang, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2012. Magic Finger: Always-available Input Through Finger Instrumentation. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. ACM, New York, NY, USA, 147–156. https://doi.org/10.1145/2380116.2380137

[50] Xing-Dong Yang, Pourang Irani, Pierre Boulanger, and Walter Bischof. 2009. One-handed Behind-the-display Cursor Input on Mobile Devices. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)*. ACM, New York, NY, USA, 4501–4506. https://doi.org/10.1145/1520340.1520690

[51] Xing-Dong Yang, Pourang Irani, Pierre Boulanger, and Walter Bischof. 2009. One-Handed behind-the-Display Cursor Input on Mobile Devices. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)*. Association for Computing Machinery, New York, NY, USA, 4501–4506. https://doi.org/10.1145/1520340.1520690

[52] Hui-Shyong Yeo, Juyoung Lee, Andrea Bianchi, David Harris-Birtill, and Aaron Quigley. 2017. SpeCam: Sensing Surface Color and Material with the Front-facing Camera of a Mobile Device. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17)*. ACM, New York, NY, USA, Article 25, 9 pages. https://doi.org/10.1145/3098279.3098541

[53] Hui-Shyong Yeo, Juyoung Lee, Hyung-il Kim, Aakar Gupta, Andrea Bianchi, Daniel Vogel, Hideki Koike, Woontack Woo, and Aaron Quigley. 2019. WRIST: Watch-Ring Interaction and Sensing Technique for Wrist Gestures and Macro-Micro Pointing. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '19)*. Association for Computing Machinery, New York, NY, USA, Article Article 19, 15 pages. https://doi.org/10.1145/3338286.3340130

[54] Yang Zhang and Chris Harrison. 2015. Tomo: Wearable, Low-Cost Electrical Impedance Tomography for Hand Gesture Recognition. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, 167–173. https://doi.org/10.1145/2807442.2807480