# Eliciting Wrist and Finger Gestures to Guide Recognizer Design

Qi Feng Liu*
Cheriton School of Computer Science
University of Waterloo

Keiko Katsuragawa†
National Research Council of Canada
Cheriton School of Computer Science
University of Waterloo

Edward Lank‡
Cheriton School of Computer Science
University of Waterloo

## Abstract

While hand gestures, i.e. movements of the fingers and wrist, are a low-effort input modality, sensing and recognition of these small-scale gestures is challenging. In particular, while many authors have explored varying designs of hardware to support hand gesture input, each systems recognize their own gesture set, rendering challenging comparisons between different capture and recognition systems. In this paper, we explore the design of hand and finger gesture input by conducting an elicitation study to understand the trade-offs between hand, wrist, and arm gestures. Alongside this, to evaluate the overall potential of wrist-worn recognition, we explore the design of hardware to recognize gestures by contrasting an IMU-only recognizer with a simple low-cost wrist-flex sensor. We discuss the implications of our work both to the comparative evaluation of systems and to the design of enhanced hardware sensing.

**Keywords:** Gesture, Sensors, Recognition, Hand, Finger

**Index Terms:** H.5.2 [User Interfaces]: User Interfaces—Graphical user interfaces (GUI); H.5.m [Information Interfaces and Presentation]: Miscellaneous

## 1 Introduction

In human-computer interaction, a significant body of work exists in gestural interaction, including work in free-space gesture [11,34,37], surface gesture [37,38], motion gestures [27], and hand gestures [5]. Our interest is specifically in free space hand gestures, which can be characterized as hand movements performed by the wrist and fingers (versus, for example, movements of the arm [1,30]).

Hand gesture interaction is attractive both because of the dexterity of the hand and the expressivity of pointing and gesturing as a communication modality. While gestural input is natural and expressive, gesture design is a significant challenge: gestures are not self-revealing [38], so gestural interfaces typically stress recall over recognition [37]. To address these challenges, researchers aim to create gesture sets that "make sense" to end-users [4,6,11,27,38].

We are not the first researchers to study hand gesture input. However, when one examines past work on hand gestures, one challenge in assessing technologies proposed for capturing and interpreting input is the wide variety of gesture sets tested. Many commercial and/or research systems recognize different small sets of finger and hand gestures [10, 11, 24, 29]. For example, the Myo arm band recognizes five hand gestures, a recent system by Zhang *et al.* [39] recognizes two gesture sets (one set of eight gestures and a second set of five gestures), and a third recent system by Wen *et al.* [36] uses only a smartwatch accelerometer and gyroscope to recognize a set of five repeated-movement gestures. While some overlap exists in these gesture sets, each of the three gesture sets differs significantly from the other sets.

*e-mail: qfliu@uwaterloo.ca
†e-mail: Keiko.Katsuragawa@nrc-cnrc.gc.ca
‡e-mail: lank@uwaterloo.ca

To support input and recognition of hand gestures, one question we address in this paper is what types of gestures are good candidates for recognition. To accomplish this, we present an elicitation study that explores hand, wrist, and forearm manipulations to perform a series of interactions with ubiquitous computing artifacts and with personal devices, heavily influenced by work by Shimon [2] and Ruiz [27] on elicitation. We synthesize the results of our study, analyze the gestures created, and discuss the similarities and differences between the elicited gestures and the gesture sets used in past work on wrist-worn gesture sensing [5,19,36,39,40]. We also leverage our reference gesture set in a simplified recognition study to evaluate the efficacy of wrist-worn [5, 39] against IMU-based recognizer strategies [36]. We explore this question by carefully replicating state of the art IMU-recognition [36] and wrist-worn recognition [5].

Overall, our contributions are two-fold. First, we highlight characteristics of hand gesture input as elicited from end users. Contrasting this gesture set with the gestures proposed by past research, we note the existence of a new class of gestures, *compound gestures*, which have not typically been evaluated by wrist-worn gesture recognizers. Second, we build a simple wrist-worn recognizer based upon past work by Dementyev and Pardiso [5] and contrast its efficacy against IMU-only recognition on both the types of gestures recognized by past work in wrist-worn gesture capture (what we term component gestures) and on our new class of compound gestures. This contrast highlights the complementary nature of IMU-only and wrist-worn sensor based recognition and also provides evidence that wrist-worn gesture capture systems can be used to recognize both component and compound gestures.

## 2 Related Work

This paper focuses specifically on wrist- and hand-gestures of the kind typically sensed by wrist-worn sensors [5,36,39,40]. As our interest is in this subtle gesture space, we do not consider larger-scale arm motions.

Researchers have explored various forms of arm- and wrist-worn sensors to sense input. Commercially, the Leap motion is a camera-based system that captures hand and finger motion. It provides accurate tracking for fixed-location input (e.g. input to a personal computer), but is poorly suited to mobile contexts because of the limitation of placement flexibility [33]. Alongside camera-based systems, the MYO armband [9] reliably senses five wrist- and hand-gestures. Researchers have used IMUs in smartphones [22] and smartwatches [14] to detect arm movements and point-and-click actions. Smartwatches have also been used to detect hand gestures performed repeatedly [36] and to detect vibrations transmitted through the body [16]. Finally, a variety of wrist-worn sensors have been tested to capture wrist and hand movements [5,7,8,10,13,16,18,19,31,36,39].

Each of these research systems exhibits an ability to achieve a very high recognition rate on gestures analyzed during evaluation. However, one thing that we have noted in our analysis of past work is that every system has used a gesture set that, while exhibiting broad similarities, differ in significant ways. For example, some examine finger tapping [16], bending [19] or pinching [39], while others examine static pose [24,40] versus dynamic movements [16,36]. Gesture sets vary extensively in size as well: the Serendipity system of Wen *et al.* [36] was evaluated on a set of five gestures, McIntosh *et al.*'s EMPress system [19] was evaluated on fifteen gestures, and Zhang and Harrison [39], for reasons that are unclear, evaluated the Tomo system on two different gesture sets – a set of 8 wrist gestures and a partially overlapping set of 5 finger gestures in 2015 [39].

Despite on-going interest in wrist-worn gesture capture, we posit that the use of unique gesture sets to evaluate competing technologies presents two drawbacks. First, from the perspective of past systems, it is difficult to analyze the benefits and drawbacks of any one recognizer technology that exists in the literature [39]. If everyone generates over 90% recognition accuracy, then why consider, as one example, electrical impedance tomography with all of its complexities [39] when researchers have already obtained 90% recognition using only the IMU on a smartwatch [36] on gesture sets of approximately the same size? Second, and related, from the perspective of new system design and evaluation, we are left suspicious of each new system. Is the new system really necessary? What does the new system add in terms of class of gesture?

The initial goal of this paper is to understand what classes of gestures need to be sensed. It is true that researchers could simply pick any one gesture set used in past research, but the question becomes which of the myriad sets of gestures should one select? Should the gestures by dynamic, i.e. require movement [16, 36]? If so, then this biases against sophisticated techniques that have been shown to detect static pose with high accuracy [5, 40]. Should all gestures be little more than point-and-click operations [14, 17], or are more abstract poses necessary [3, 15, 24, 33, 35]? Perhaps the gold standard for gesture recognition should be various forms of signed letter alphabets [21] or template-based alphabets [30]?

## 3 USER-DEFINED GESTURE SET

In this section, we describe the results of an elicitation study to understand gestural input. Elicitation studies [32, 38] are a common technique for understanding potential end-users' concepts of gesture parameters and of how gestural manipulations should map onto commands [4, 27, 33, 38]. An initial question we asked, given the myriad set of elicitation studies available, is whether past gesture elicitation studies might provide information to us on appropriate gestural interaction. Overall, we believe that the class of gestures captured by wrist-worn recognizers [5, 19, 36, 39, 40] has not been studied directly via elicitation. Our analysis found that past elicitation studies of arm and hand movement gestures either allowed large-scale full arm gestures [28], allowed any body part gestures [6], or strictly limited studies to a very small number of joint movement (i.e., microgestures) [4]. Each of these past elicitation studies addresses a class of wrist and finger gestures, but not the class of gestures recognizable by wrist-worn sensors.

It should be noted that there is nothing untoward in the experimental design in past elicitation studies: the assumption past researchers make is that a specific recognition technology – camera-based tracking [28], smartphone IMUs [27], or a Leap Motion [4] – will be used to capture gestures for recognition; researchers wish to determine gestures that make sense given specific gesture-capture technology. In contrast, wrist-worn recognizers move with the participant, relaxing spatial constraints imposed by localized sensing (e.g. Leap Motion), while still capturing wrist and finger movements. This means that, while we cannot allow any-body-part gestures [28], we do not need to limit to a small number of joint movements [4] or to a spatially constrained input region [33].

In early pilot studies eliciting gestural input, we noted that, to perform gestures with hands, there are three joints on a human arm that define these gestural manipulations: fingers, wrist, and elbow. Free-space gestures also comprise arm gestures which also involve shoulder movement, but we explicitly focus our elicitation study toward elbow, wrist, and finger movement because of the gorilla arm effect [12] and the potential reticence of users to perform large-scale gestures in public contexts [25].

Table 1: Tasks for Elicitation Study.

| Application | Task |
|---|---|
| 1. Map Application on Smartwatch Display | Pan Left<br>Pan Right<br>Zoom In<br>Zoom Out<br>Activate Help<br>Go to Home<br>Scroll Up<br>Scroll Down |
| 2. Map Application on External Display | Pan Left<br>Pan Right<br>Zoom In<br>Zoom Out<br>Activate Help<br>Go to Home<br>Scroll Up<br>Scroll Down |
| 3. Phone Application (audio only) | Answer Call<br>Hangup Call<br>Ignore Call |
| 4. Music Application (audio only) | Play<br>Stop<br>Play Next<br>Play Previous |

Given that we wish to explore the class of gestures recognizable by wrist-worn sensors, our approach to elicitation is identical to Chan *et al.* [4], Vatavu *et al.* [33] and others: We constrain gestural input to movements of the fingers, wrist, and elbow by asking participants to keep their elbow next to their body while performing gestures. This constraint, similar to the use of wrist and ankle weights [28], is designed to overcome the legacy bias of large scale arm movements [20, 28] by discouraging energetic whole-arm movements. Participants are free to perform either micro-gestures (i.e. small scale movements of the fingers) [4] or larger scale wrist and finger movements [19].

### 3.1 Experiment Design

#### 3.1.1 Elicited Tasks

While it is important to consider both task and context in elicitation studies, it is also the case that participants frequently overload gestures by considering context-sensitive interpretations (e.g., the phone-to-mouth gesture in smartphone motion gestures varies depending on whether the phone is ringing [27]; drag with one finger in surface gestures varies depending on whether the gesture starts on the object or not [38]). We chose three contexts for our elicitation study: interacting with an external display, interacting with a wearable device such as a smartwatch, and controlling a smartphone while the phone is not in the user's hand. To limit the length of the study, we restricted our application domain to map navigation for the first two contexts; for the third context, a smartphone while the phone is not in the user's hand, we considered audio-only interactions and selected two applications to control: the phone application and a music application. Our tasks are shown in Table 1. Our task selection was heavily influenced by common tasks in Shimon's and Ruiz's papers [2, 27]. There were a total of 23 task commands elicited from the study.

### 3.1.2 Participants

We recruited 16 participants (10 male) from the general student body of our institution to participate in our study.

### 3.1.3 Procedure

At the beginning of the experiment, the researcher asked the participants to wear an LG G smartwatch on the hand on which he/she usually wears a watch. We used the smartwatch as a placebo, thus encouraging participants to bias toward gestures that could, conceivably, be sensed by treating the watch as a quasi "magic brick" [27].

All participants were shown a video demonstrating three possible joints which they could use (finger joints, wrist joint, and elbow joint) and two example gestures for each joint movement. They were shown "make fist" and "telephone hand sign" for finger joint, "hand sway left" and "hand sway right" for wrist joint, "forearm sway left" and "forearm sway right" for elbow joint. We also showed two combination movements, "make fist then forearm sway to left" and "make fist then hand sway right". The participants were informed they would be asked to come up with a preferred gesture for each task using any of three joints. They were also told that the video was merely an illustration of the possible joints to be used, and they could create any gesture they wish using finger, hand, or forearm movements. We also asked that within an application domain (e.g., a map app on the smartwatch) they do not reuse gestures, but noted that they could re-use gestures after switching application domains. To simulate the map application, a map was shown both on the smartwatch and on a nearby projection screen. All application domains and the task within an application domain were counter-balanced.

For each task, we asked participants to come up with two gestures. If they could not come up with a second gesture, they were allowed to skip the second gesture. We did this to ensure that we elicited a large set of possible gestures within the one hour allocated for each participant balanced against the need to avoid frustrating participants if they could not come up with a second reasonable gesture. We also encouraged participants to vary the joint used in each gesture elicited to encourage participants to fully explore the richness of finger, hand, and forearm movement. After identifying gestures for each task, participants were asked to perform the gesture and to describe the gesture and the joints involved.

## 3.2 Elicitation Study Results

In this section, we highlight agreement scores, joint analysis, and gestural parameters used to discriminate gestures.

### 3.2.1 Agreement

Because participants could choose from different joints and a total of 7 different possible combinations of joints to perform gestures, gestures exhibited very low consensus [33].

As a result, we grouped similar gestures together in a task, then calculated the agreement score following guidelines from Chan *et al.*, Morris *et al.*, and Piumsomboon *et al.* [4, 20, 23]. Specifically, we used the following guidelines from related work to group the gestures:

- If gestures differ only by direction, i.e. "turn wrist to left" and "turn wrist to right", we grouped them together [20].

- As per Chan *et al.* [4], gestures that use two or fewer fingers (i.e., hand faces down, index finger moves up and index plus middle finger moves up) were grouped together and gestures that use three or more fingers (i.e. all five fingers touching together and first three fingers touching together) were grouped together.

- We leveraged Piumsomboon's definition of gesture similarity, i.e., path gestures that have consistent directionality although the gesture is performed with different static hand poses are grouped together [23], for example fist-moves-right and open-hand-moves-right.

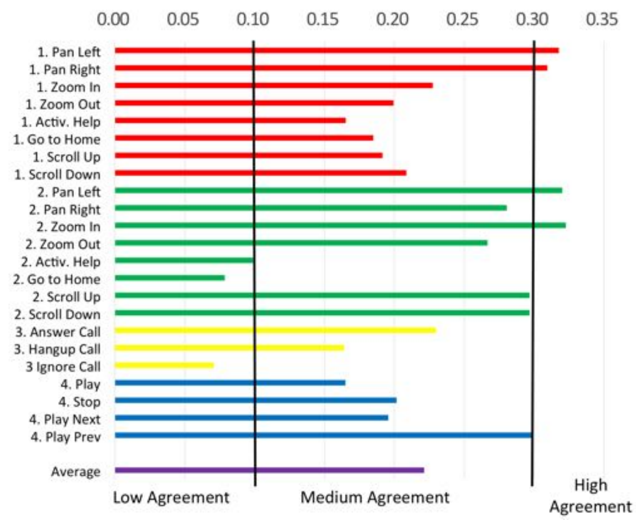


Figure 1: Agreement score for all 23 tasks. 1-4 represent the applications as shown in Table 1: Map Application on Smartwatch Display (1), Map Application on External Display (2), Phone Application (3), and Music Application (4).

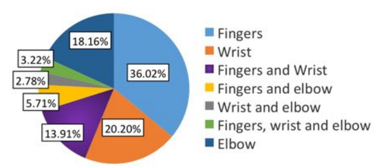


Figure 2: Distribution of joint usage for all gestures.

Grouping the above three gesture types together helps us to understand whether moderate or high agreement exists on, respectively, wrist movements, finger movements or poses, and hand poses. This, then, can guide the design of gesture sensing by targeting types of movement.

We used the formula proposed by Vatavu *et al.* [32] to calculate agreement score of our grouped gestures:

$$AR(r) = \frac{|P|}{|P|-1} \sum_{P_i \subseteq P} (\frac{|P_i|}{|P|})^2 - \frac{1}{|P|-1}$$

where $P$ is the set of all proposals for referent, $|P|$ is the size of the set, and $P_i$ are subsets of identical gestures from $P$.

Using Vatavu's interpretation of agreement values, our results ranged from 0.071 (low agreement, $AR \leq 0.1$) to 0.323 (high agreement, $0.3 < AR \leq 0.5$) as shown in Figure 1. Most gestures exhibited moderate agreement, $0.1 \leq AR \leq 0.3$, an expected result [32]. Gestures with moderate or high agreement included all of wrist movements, hand poses and finger movements.

### 3.2.2 Joint Analysis

A question we asked during analysis was whether there was a bias toward one joint or another. The distribution of joints used for each gesture are shown in Figure 2. We can see from this figure that 74.38% of the gestures involve only one joint, 22.40% of the

Table 2: Our set of 15 wrist and finger consensus gestures collected from elicitation study. See Figure 3 for visual depiction of gesture components.

| No. | Gesture Type | Components |
|---|---|---|
| 1. | Component | b. Fist |
| 2. | Component | d. Phone |
| 3. | Component | e. Thumb |
| 4. | Component | h. Flex |
| 5. | Component | i. Extend |
| 6. | Component | g. Close |
| 7. | Compound | g. Close + e. Spread |
| 8. | Compound | a. Point + m. Pro + h. Flex |
| 9. | Compound | a. Point + l. Sup + h. Flex |
| 10. | Compound | m. Pro + h. Flex |
| 11. | Compound | l. Sup + l. Sup |
| 12. | Compound | m. Pro + j. Adduct |
| 13. | Compound | f. Pinch + c. Spread |
| 14. | Compound | m. Pro + k. Abduct |
| 15. | Compound | m. Pro + i. Extend |

gestures involve two joints, and 3.22% of the gestures involve three joints. Of these single joint gestures, fingers are the most common, then wrist, and elbow is least common. Fingers and wrist is the most common multi-joint combination. Elicited joint preference during exit interviews triangulates well with this data: the most preferred joints are fingers (12/16 participants, followed by wrist 5/16, and none chose elbow as their first choice). Note that one participant chose both wrist and finger as his first choice. Finger and wrist movements accounted for over 70% of all elicited gestures and all elicited consensus set gestures were finger and wrist gestures..

### 3.3 Consensus Gesture Set

Recall that our primary goal involved understanding what types of gestures should be recognized by wrist-worn recognizers. To explore this question, we captured all consensus gestures for each task (e.g. the gestures with the highest agreement score for each task). Participants in our study prefer gestures that leverage finger or wrist manipulations over forearm movement. Table 2 shows our consensus hand gesture set. Some of these gestures (e.g. Flex, Close, Fist) were used in multiple contexts. As noted, some gestures represent combinations of more than one simpler gesture, i.e. compound gestures.

### 3.4 Discussion

A first question we asked was whether and how our consensus set differs from gesture sets leveraged during wrist-worn recognizer studies [5, 19, 36, 39, 40]. To address this question, we took the union of gesture sets analyzed by every paper on wrist-worn gesture recognizers that we could identify published between 2014 and 2017 [5,19,36,39,40]. While acknowledging that the above set of papers may not be exhaustive, they produce a useful cross section of related work. The union of past gesture sets is presented in Figure 3.

Contrasting Table 2 with the gestures from past work, Figure 3, we see that many of the component gestures represented in Table 2 are also found in work on wrist-worn gesture recognizers. However, what is missing from every paper are the compound gestures highlighted in Table 2.

Past research has focused primarily on simple component gestures, rather than on the combination of pose+wrist or multiple wrist gestures in sequence. One question that this observation of the existence of compound gestures poses is how effectively past recognizers perform on compound gestures, i.e. are current wrist-worn approaches sufficient to recognize gestures like the mix of


Figure 3: 27 gestures collected from related work and 13 component gesture. The single letter acronyms represent related systems: E:Empress, T:Tomo, S:Serendipity, F:WristFlex and M:MYO.

component and compound gestures elicited from participants and presented in Table 2.

## 4 Recognizing Compound Gestures: A Feasibility Study

Given our understanding of component and compound gestures, in this section we explore the technical requirements for capturing hand-gesture input. The assumption made by wrist-worn gesture recognizers is that hand and finger movement – enacted by tendons in the wrist and hand – can be accurately sensed by wrist-worn recognizers. However, because past evaluations have focused on component gestures, not compound gestures that combine wrist + finger movements or multiple wrist movements in sequence, it is unclear whether combining multiple signals – either simultaneously or sequentially – can be accurately captured by wrist-worn sensors. It may be that the signals are sufficiently differentiable that wrist-worn recognizers can either separate out individual component signals or can be trained on an overall gesture set comprised of both component and compound gestures. It may, however, also be the case that certain signals (e.g. wrist movements) might be so strongly sensed at the wrist that they overwhelm more subtle signals (e.g. hand poses).

To address these questions, in this section:

- We describe the design of a simple wrist-band recognizer that includes a smartwatch's IMU and four bend sensors, unimaginatively dubbed *WristRec*.
- We evaluate recognition of bend sensor alone, IMU-alone, and bend+IMU recognition using WristRec.

To re-iterate the goal of our design and analysis is not to build the best capture system for hand and wrist gesture but, instead, to evaluate sensing strategies. Given past research, it remains unclear how well wirst-worn recognizers will perform on gesture sets that include compound gestures [39].

### 4.1 Designing A Simple Wrist-Worn Recognizer

Consider the gestures captured from our elicitation study. A gesture recognizer tuned to wrist and hand-gestures in our consensus set must discriminate direction of movement, hand pose, joint combinations, timing, and movement scale. IMU sensing can capture many aspects of wrist and forearm movement – direction, timing, and scale, for example, but only if the wrist is displaced during movement. As we analyzed IMU signals from manipulations of the hand and fingers that result in more subtle spatial movements consisting of skin stretching at the wrist, angular deformations between forearm and hand, and movement of the carpal and ulnar tendons along the inner surface of the wrist, we found that these movement were not detected by an IMU.

To analyze more subtle angular deformations and finger movements, we examined wrist-worn signals more carefully. Looking carefully at the wrist, we noted that both finger and wrist deformations resulted in subtle movements of the wrist, a result leveraged by Dementyev and Paradiso [5]. To examine this deformation in more detail, we attached an elasticized strap around a researchers wrist and performed movements from our gesture set. As gestures were performed, we noted that the strap was deformed in various ways due to finger movement. When the hand was flexed, the sides of the wrist (near thumb and pinky) were stretched, pushing the band into a more elongated oval. As fingers were moved, disruption along the carpal and ulnar tendons resulted in deformations of the band along the palm-facing side of the wrist and stretching along the back-of-hand side of the wrist. As well, pinky and ring finger movement caused twisting of the band along the pinky-facing side of the wrist as the skin was tightened. Essentially, we note that, along with displacement of the wrist due to larger-scale movement, small scale movements caused stretching, pulling, and deformation around the wrist, and the capture of these deformations might support gesture recognition.

While many technologies can be used to measure how movements of fingers and wrist affect components of the wrist [5, 7, 8, 10, 13, 16, 18, 31, 36, 39], our observation of strap deformation during subtle movements motivated the use of sensors that can capture this strap deformation, i.e. flex sensors to sense band movement. Flex sensors are inexpensive, low-powered, low-noise, rapidly adapting, variable resistors in which the deflection of the sensor or pressure on the sensor varies the resistance such that subtle surface deformations can be measured dynamically via a simple programmable micro-controller. To date, flex sensor gesture sensing work has been primarily work by Dementyev and Paradiso [5], who only measure the ability of flex sensors to capture five finger bend gestures. However, the characteristics of flex sensors recommend them as a sensing technology worthy of further exploration. Contrasting them directly with alternative sensing, we note that camera-based systems require higher power consumption during capture. Camera-based systems also require computer vision algorithms to interpret input, and these algorithms are processing intensive, further consuming power. Techniques such as electrical impedence tomography result in increased latency, a constraint noted by Zhang and Harrison in their system design [39]. Finally, our experience with Myo input signals indicates that, to capture signals, the medical-grade, stainless steel EMG sensors are necessary to maximize signal-to-noise ratio, whereas flex sensors provide low-noise input less expensively. In summary, while flex sensors have much to recommend them as a capture technology, Dementyev and Paradiso only evaluate their flex-sensor based system on finger bend gestures [5]. It is unclear how a system based on flex sensors would perform on compound gestures or a larger gesture set.

One open question when designing a bend-sensor based wrist recognizer is where to place bend sensors. Dementyev and Paradiso simply placed an array of 15 flex sensors around the inner surface of the wrist (i.e. the palm-facing side of the wrist) with the flex sensors aligned along the elbow to wrist direction (axial alignment). However, to maximize input signals, bend sensor positioning should be adjusted to maximize input signal, such that signal-to-noise ratio for movement is maximized. As well, if signals are clearly linked (i.e. two adjacent bend sensors produce identical signals during movement differing only in magnitude) there is little need to redundantly measure the same signal. To test signal strength available from wrist deformation and to optimize bend sensor placement, we conducted a pilot study with 5 participants. We glued one axial and one tangential bend sensor to a Fitbit sleep band. We then rotated the sleep band around each participant's wrist in increments of 15 degrees, so that all 48 positions (24 locations for two orientation) were tested. Participants performed wrist and pose gestures for each location drawn from our consensus gesture set. We found four locations of maximum signal strength. Considering the top of the wrist (center of the back of the hand) as 0°, these positions were at 0°(axial orientation), 90°(tilt), 180°(axial), and 270°(tilt).

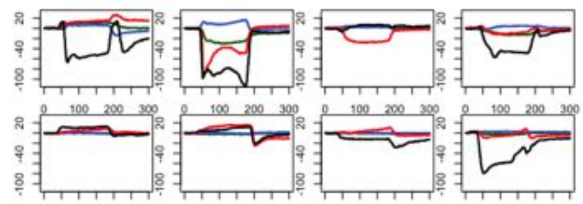

Figure 4: Bend sensor data for one instance of one participant's logged performance of each of the gestures . From top left, Flex, Extend, Spread and Fist. From bottom left, Index, Middle, Ring and Pinky finger bend gestures.

Figure 4 depicts the four bend sensor traces from each of these locations for a simplified set of eight gestures. The top row includes wrist gestures: wrist flex and a set of hand pose gestures – wrist extend, fingers spread, fist. The bottom row presents data from bends of the index, middle, ring, and pinky fingers. Flex and extend have high signal amplitude compared to other gestures. Index, middle, and ring finger bends have smaller but recognizable distinct sensor shapes. Finally, pinky finger bend gestures (bottom right) exhibits an interesting shape. Bending the pinky results in ring finger bending and significant axial deflection of the skin on the wrist at the 270°position. We capture the dynamics of this interaction in our bend sensor data, where a larger signal (from two finger bend) is gradually attenuated as the participant straightens his or her ring finger.

### 4.1.1 System Design

We constructed our prototype WristRec system using the results of the pilot study analyzing flex sensor performance. Our prototype recognition system was constructed from four bi-directional flex sensors at four sides of the wrist and a 5V Arduino board.

We first cut the Fitbit sleep band into 1 inch width to match the width of the smartwatch band. Four flex sensors are placed on the bottom surface of the Fitbit band against participants skin. We used adhesive bandages to provide custom positioning for bend sensors along the length of the Fitbit sleep band. On top of the Fitbit One sleep band, participants wore an LG G-watch R. The LG G-watch R allowed us to contrast IMU-based recognition measured by the smartwatch, wrist-deformation recognition measured by bend sensors, and combined recognition via both IMU and deformation. Figure 5 shows our prototype.

## 4.2 Recognition Test

In this experiment, we used both bend sensor, accelerometer and gyroscope data as captured by a smartwatch IMU, and the combination of data sources to evaluate recognition on our consensus gesture set.

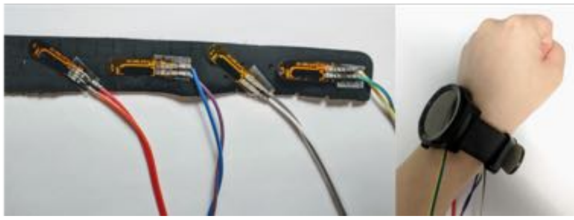

Figure 5: Left: Bend sensors mounted on a Fitbit One sleep band. Right: Experiment setup with a smartwatch on top of the band with sensors.

#### 4.2.1 Participants

12 paid participants were recruited, 9 male and 3 female. All participants used the hand they would typically wear a watch on to perform gestures. For 10 participants, this was their left hand; for 2 participants, this was their right hand.

#### 4.2.2 Apparatus and Data Collection

Using our hardware configuration described previously, we captured data from both the bend sensors (via an Arduino board) and the IMU on the LG smartwatch. The Arduino board was connected to an external computer via USB cable and the smartwatch streamed IMU sensor data to the same computer via Bluetooth. The sampling rate for the bend sensors was 100Hz, and the sampling rate for the accelerometer and gyroscope exceeded 200Hz. Data were synchronized on the external computer. We used two 2-inch tilt sensors at 0°and 180°and two 1-inch bend sensors oriented tilt at 90°and 270°for this study.

#### 4.2.3 Method

Our experimental methodology in both number of participants and in procedure replicates the experimental methodology of Zhang and Harrison [39].

Our recognizer was placed on participants' wrists and participants were seated at a table facing the computer used to coordinate experimental data.

Our gesture set consisted of the 13 component gestures and 9 compound gestures (the compound gestures from Table 2) plus their gesture components). This creates an overall gesture set of 22 gestures and ensures that each component gesture that comprises a compound gesture is included in our gesture set, specifically to evaluate whether compound gestures can be discriminated reliably from their constituent component gestures.

Gestures were captured in two blocks. First, participants completed a training block. A video of the gesture was shown before each gesture, and the participants performed all gestures in the gesture set once to familiarize the gesture set and system interface. Second, participants completed ten data collection blocks. For the component gestures, a picture of the gesture was shown in the computer display to cue the gesture. For the compound gestures, because the multi-component movement was difficult to understand using icons, we showed participants the same video we used for the training.

During each trial, participants start from a resting position with their elbow on the armrest. A gesture is presented on a computer screen and participants press the space key on a keyboard to start the gesture. The participants hear two auditory cues for each gesture: the first one is a signal at the time they press the key for them to start performing the gesture; the second one marks the end of the transitioning move of the gesture and participants were asked to finish their gesture before this second cue. Audio cues were spaced at 3.5 seconds apart for the compound gestures (due to combined movement requirements) and at 3 seconds apart for the component gesture set. Participants performed each gesture set one time each in a block and they repeated this for 10 blocks. The order of the gestures within the block was randomized. In total, we collected 220 gestures per participant.

#### 4.2.4 Data Cleaning and Classification

We analyze recognition accuracy in two ways, replicating the analysis presented by Dementyev and Paradiso [5] with their WristFlex system. We first use 10-fold cross validation to assess recognizer performance, which yields an optimistic upper bound. Next, we use the first 5 blocks as training data and the second 5 blocks as test data for analysis, a more traditional approach with separated training and test data.

We also analyze recognition accuracy using three different inputs, flex sensor only, IMU only, and combined flex sensor plus IMU with data cleaning and classification proceeding as follows:

1. We use flex sensor data alone to perform recognition. To clean data, we first normalized the data, then used low pass filter to eliminate noise. To reduce feature space, we down-sampled the data to 30 data points spaced equidistant over the input. From 3 seconds of component gestures and 3.5 seconds of compounds gestures, this process produced 10 Hz data for the component gesture and 11.7 Hz data for the compound gesture. Overall, our final data was a vector of 248 entries: the bend sensor data down sampled to 30 data points from each location (4x30), the deltas for each sensor between adjacent samples (4x30), and the overall minimum and maximum of each sensor input (8 values).
2. We replicate the same data processing algorithm described in the Serendipity system [36] which uses only the accelerometer and gyroscope data from IMU to recognize gestural input. Instead of taking the lower 10 bands, we took 30 power bands, yielding a total of 37 x 7 features x 2 sensors = 518 features.
3. We combine the flex sensor data and IMU watch sensor data to form an enlarged vector feature space for recognition. We used the expanded 766 feature vector combining the 248 bend sensor features described above with the 518 IMU vectors we used for Serendipity algorithm.

Given the above 248-point vector for bend sensor data and 518-point vector for IMU data, we analyzed data using both a SVM with polynomial kernel in the same manner as Serendipity [36] and a random forest algorithm. For the SVM, the data was fed into a multi-class classification SVM as an n-dimensional vector (time was a dimension) and cut with hyperplanes (Polynomial kernel). We set the parameters for the Random Forest algorithm to the following values: the number of features to consider in feature selection is calculated by $1+\log_2$(total number of features), the max depth of a tree is 100, and the number of trees is 100.

### 4.3 Results

The random forest algorithm outperformed SVM for each configuration. Because this is a feasibility study evaluating the potential of wrist-worn recognition, we focus on random forest performance; we note that more complex machine learning algorithms and larger data sets should significantly enhance performance.

Figure 6A shows 10-fold cross validation accuracy and Figure 6B shows test-training set accuracy for the random forest algorithm for the 13 component gestures gesture set, the 9 compound gestures gesture set, and the overall 22-gesture set including both component and compound gestures for IMU, Bend, and IMU+Bend. Our recognition accuracy on the overall 22-gesture gesture set with Bend sensor and IMU was 88.1% for 10-fold cross validation and 78.9% for separate training and test data.

For comparison, Dementyev and Paradiso [5] report 96.3% accuracy for cross validation and 80.5% (with recognizer feedback)

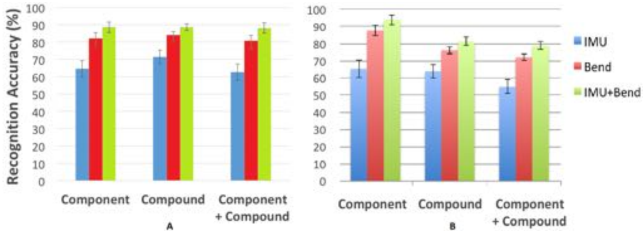


Figure 6: Recognition accuracy with IMU-only, Bend sensor only and IMU + Bend sensor for 13 component gestures, 9 compound gestures and 22 component and compound gestures using (A) 10-fold cross validation and (B) separated training and test data. Error bars show std error.

and 69.3% (no recognizer feedback) accuracy via separated training and test data. Note, however, that Dementyev and Paradiso's cross-validation accuracy value is somewhat difficult to interpret, as their recognizer provided feedback during the collection of their test data and this feedback condition may have been included in cross validation analysis. Considering separated training and test data, our approach is most comparable to their no feedback condition (we obtain 78.9% on 22-gestures versus 69.3% for Dementyev and Paradiso on 5-gestures).

Figures 6 also shows IMU only recognition, a replication of Wen et al.'s Serrendipity system. Note that our gesture sets are all significantly larger than their 5-gesture set. We find accuracies of 65.3% for component 63.9% for compound and 54.9% for the combined gesture set using only the IMU data. A repeated measure one-way ANOVA (3 sensor type; IMU only, Bend sensor only and IMU+Bend) on recognition accuracy indicates that significant differences exist for sensor type ($F_{2,22} = 32.125$, $p < .001$). Post hoc paired t tests revealed significant differences between IMU only and IMU+Bend ($p < .001$), Bend only and IMU+Bend ($p < .001$).

Finally, Figure 7 presents the confusion matrix for our 22 gestures with IMU+Bend data together using using separated training and test data. Note the area of higher confusion on similar gestures. In particular, we see higher levels of confusion between a.Point, f.Pinch and e.Thumb. We also observed higher confusion on compound gestures partially overlapped. In particular, f.Pinch + c.Spread and g.Close + c.Spread, g.Close and g.Close + c.Spread, m.Pro + k.abduct and m.Pro + k.adduct had higher confusion.

### 4.4 Discussion

As noted, the primary goal of this section was to determine whether wrist-worn gesture recognition technologies could successfully discriminate compound gesture sets of the form elicited by our participants. Overall, we find that the combination of wrist-worn IMU+bend sensors provides the highest recognition rate. Furthermore, as shown in Figure 6, the combined sensors have very good discriminatory power for compound gestures, preserving overall accuracy even when both compound gestures and their constituent component gestures are represented in the same gesture set. Compound gestures and overall recognition rate when both compound and component are included remain stable, even with a limited training set of five gestures per participant.

One thing that does seem obvious is that wrist-worn recognition should be more effective for gestures that involve wrist movement versus gestures that involve finger movement (i.e. hand pose gestures). To evaluate this, we separate gestures into two categories: the top row identifies hand pose gestures where changes to wrist-hand alignment and hand orientation are not required to perform the gestures (i.e. gestures are simply finger movement); and wrist gestures where significant wrist flex or rotation occurs either because

| | a. Point | b. Fist | c. Spread | d. Phone | e. Thumb | f. Pinch | g. Close | h. Flex | i. Extend | j. Adduct | k. Abduct | l. Sup | m. Pro | g. Close + e. Spread | a. Point + m. Pro + h. Flex | a. Point + l. sup + h. Flex | m. Pro + h. Flex | l. Sup + l. Sup | m. Pro + j. Adduct | f. Pinch + c. Spread | m. Pro + k. Abduct | m. Pro + l. Extend |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | 33 | 8 | 4 | 0 | 3 | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| b | 8 | 40 | 1 | 0 | 4 | 5 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| c | 4 | 0 | 43 | 0 | 10 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| d | 0 | 0 | 0 | 45 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 2 | 0 | 0 | 1 | 1 | 0 | 0 |
| e | 4 | 2 | 0 | 0 | 49 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| f | 8 | 3 | 2 | 0 | 10 | 37 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| g | 0 | 0 | 0 | 3 | 0 | 0 | 35 | 0 | 0 | 0 | 0 | 0 | 0 | 15 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 0 |
| h | 0 | 1 | 5 | 0 | 0 | 1 | 0 | 50 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| i | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 60 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| j | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 50 | 3 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| k | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 55 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| l | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 56 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| m | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 54 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ge | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 55 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 |
| amh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 46 | 8 | 2 | 0 | 0 | 0 | 3 | 0 |
| alh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 56 | 1 | 0 | 0 | 0 | 0 | 0 |
| mh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 49 | 0 | 4 | 0 | 3 | 1 |
| ll | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 60 | 0 | 0 | 0 | 0 |
| mj | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 47 | 0 | 9 | 2 |
| fc | 0 | 0 | 0 | 3 | 0 | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 17 | 0 | 0 | 0 | 0 | 0 | 32 | 0 | 0 |
| mk | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 1 | 12 | 0 | 40 | 3 |
| mi | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 1 | 3 | 0 | 1 | 50 |

Figure 7: Confusion Matrix for 22 gestures with IMU + Bend sensor using separated training and test data.

of movement at the wrist joint or at the elbow joint. Using our more traditional training versus test data analysis, for hand pose gestures, IMU accuracy is 42.9% versus 69.7% for wrist gestures. Bend + IMU is significantly more accurate – 70.5% for hand pose gestures, and 89.7% for wrist gestures.

## 5 Synthesis

Hand gesture input, and, in particular, input involving wrist and finger movements, has a number of attractive attributes. It is a low-effort form of input, thus limiting the fatigue experienced by users when performing larger scale arm movements [12]. Alongside issues of fatigue, it is also a more subtle form of input, thus limiting potential issues of performance anxiety [26].

One challenge with wrist and hand gestures has been inconsistency in gesture sets. While multiple elicitation studies have been performed for camera-based capture systems [4, 33], we note that it is less clear that wrist-worn systems have benefited from the same elicitation-based analysis of gesture sets. Zhang and Harrison highlight the challenges of this lack of gesture set consistency: the end result has been a lack of clarity in wrist-worn gesture recognizer evaluation specifically because each system has been evaluated on gesture sets which, while largely similar, are all distinct one from the other. Zhang and Harrison address this by evaluating their Tomo system on finger bend gestures similar to the gesture set used by Dementyev and Paradiso [5].

In this paper, our approach is slightly different: we conduct an elicitation study and contrast the results of that elicitation study with gesture sets that have been recently used in system evaluation. We posit two take-aways from our elicitation study. First, in terms of strength, past research has selected gesture sets that are similar to component gestures extracted from elicitation. As a result, past gestures are representative in scale and form to the one class of gestures we find in elicitation. Second, in terms of weakness, past research has focused primarily on simple component gestures, rather than on the combination of pose+wrist or of multiple wrist gestures in sequence, i.e. compound gestures.

Once identified, a related question around the existence of compound gestures arises. Specifically, while these compound gestures exist, can wrist-worn gesture recognizers discriminate compound gestures? Or will individual components within the compound gesture exhibit such significant signal strength that subtle compound gestures provide too similar to reliably discriminate? This is a particular concern because of the nature of component gestures elicited in our study. The component gestures include both movements of the fingers (pinch, point, first) and movements of the wrist (flex, extend, adduct). A wrist-worn sensor co-localized with wrist movement might be perturbed so significantly by a large wrist movement that smaller deformations from finger movements such as pointing might be too subtle to discriminate.

To answer this question, note, again, Figure 6. In this Figure, we break recognition down into recognition rates on component versus compound gestures and contrast this with a gesture set that combines all component and compound gestures into one overall data set to understand overall behavior. In each case we see that, while IMU-based recognition performs well – significantly better than chance – wrist-worn deformations contain significant additional information not fully captured by the IMU alone. Overall, given the similarity in form between compound and overall gesture recognition accuracy in Figure 6, we observe that at least one technology, flex sensors seem equally effective on compound gestures. This result was not immediately obvious to us, particularly as flex sensors have only been evaluated on finger bend gestures, but was a reassuring validation of the efficacy of wrist-worn gesture capture and recognition.

Finally, we note one additional minor component to this work. Given the success of Wen et al.'s system in recognizing five hand gestures using only a smartwatch-based IMU [36], we wished to explore whether additional wrist-mounted sensors (beyond the generic IMU) would significantly enhance performance. We find that it does, i.e. that the IMU combined with a wrist-worn sensor suite that senses wrist deformations provides complementary information that significantly enhances recognition rates over IMU-only. The addition of the IMU also improves recognition in comparison to flex-sensors-only capture, further highlighting the complementary nature of wrist movements and wrist deformations.

## 6 Conclusion

In this paper, we describe an exploration of the space of wrist and hand gestures. To accomplish this goal, we present results of an elicitation study and a synthesis of past work in wrist and hand gesture input. One significant result of the elictation study is the identification of compound gestures – gestures comprised of combinations of wrist and finger movements – as a class of gestures that has been, based on our analysis of past work, rarely explored during the evaluation of gesture capture systems.

Alongside questions of gesture form, we also explore wrist-based recognition and its sufficiency within the space of hand and wrist gestures gleaned from our elicitation study. While our initial concern was that wrist-worn recognition might be insufficient for compound gestures (due to larger signals from wrist movement overwhelming hand-pose and finger-movement gestures), an early evaluation of wrist-worn flex sensors coupled with IMU signals demonstrates the feasibility of wrist-worn recognition for both component gestures (consisting of either wrist or hand movement) and compound gestures that combine wrist and hand movement.

## References


[1] R. Aigner, D. Wigdor, H. Benko, M. Haller, D. Lindbauer, A. Ion, S. Zhao, and J. Tzu Kwan Valino Koh. Understanding Mid-Air Hand Gestures: A Study of Human Preferences in Usage of Gesture Types for HCI. Technical report, 11 2012.

[2] S. S. Arefin Shimon, C. Lutton, Z. Xu, S. Morrison-Smith, C. Boucher, and J. Ruiz. Exploring Non-touchscreen Gestures for Smartwatches. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, CHI '16, pp. 3822–3833. ACM, New York, NY, USA, 2016. doi: 10.1145/2858036.2858385

[3] G. Bailly, J. Müller, M. Rohs, D. Wigdor, and S. Kratz. ShoeSense: A New Perspective on Gestural Interaction and Wearable Applications. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, CHI '12, pp. 1239–1248. ACM, New York, NY, USA, 2012. doi: 10.1145/2207676.2208576

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

[5] A. Dementyev and J. A. Paradiso. WristFlex: Low-power Gesture Input with Wrist-worn Pressure Sensors. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, pp. 161–166. ACM, New York, NY, USA, 2014. doi: 10.1145/2642918.2647396

[6] H. Dong, A. Danesh, N. Figueroa, and A. E. Saddik. An Elicitation Study on Gesture Preferences and Memorability Toward a Practical Hand-Gesture Vocabulary for Smart Televisions. *IEEE Access*, 3:543–555, 2015. doi: 10.1109/access.2015.2432679

[7] A. Ferrone, X. Jiang, L. Maiolo, A. Pecora, L. Colace, and C. Menon. A fabric-based wearable band for hand gesture recognition based on filament strain sensors: A preliminary investigation. In *2016 IEEE Healthcare Innovation Point-Of-Care Technologies Conference*, HI-POCT, pp. 113–116, 11 2016. doi: 10.1109/HIC.2016.7797710

[8] R. Fukui, M. Watanabe, T. Gyota, M. Shimosaka, and T. Sato. Hand Shape Classification with a Wrist Contour Sensor: Development of a Prototype Device. In *Proceedings of the 13th International Conference on Ubiquitous Computing*, UbiComp '11, pp. 311–314. ACM, New York, NY, USA, 2011. doi: 10.1145/2030112.2030154

[9] F. Haque, M. Nancel, and D. Vogel. Myopoint: Pointing and Clicking Using Forearm Mounted Electromyography and Inertial Motion Sensors. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pp. 3653–3656. ACM, 2015. doi: 10.1145/2702123.2702133

[10] C. Harrison, D. Tan, and D. Morris. Skinput: Appropriating the Body As an Input Surface. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, CHI '10, pp. 453–462. ACM, New York, NY, USA, 2010. doi: 10.1145/1753326.1753394

[11] L. Hespanhol, M. Tomitsch, K. Grace, A. Collins, and J. Kay. Investigating Intuitiveness and Effectiveness of Gestures for Free Spatial Interaction with Large Displays. In *Proceedings of the 2012 International Symposium on Pervasive Displays*, PerDis '12, pp. 6:1–6:6. ACM, New York, NY, USA, 2012. doi: 10.1145/2307798.2307804

[12] J. D. Hincapié-Ramos, X. Guo, P. Moghadasian, and P. Irani. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-air Interactions. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, CHI '14, pp. 1063–1072. ACM, New York, NY, USA, 2014.

[13] P.-G. Jung, G. Lim, S. Kim, and K. Kong. A Wearable Gesture Recognition Device for Detecting Muscular Activities Based on Air-Pressure Sensors. *IEEE Transactions on Industrial Informatics*, 11(2):485–494, 4 2015. doi: 10.1109/TII.2015.2405413

[14] K. Katsuragawa, K. Pietroszek, J. R. Wallace, and E. Lank. Watchpoint: Freehand Pointing with a Smartwatch in a Ubiquitous Display Environment. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, AVI '16, pp. 128–135. ACM, 2016. doi: 10.1145/2909132.2909263

[15] D. Kim, O. Hilliges, S. Izadi, A. D. Butler, J. Chen, I. Oikonomidis, and P. Olivier. Digits: Freehand 3D Interactions Anywhere Using a Wrist-worn Gloveless Sensor. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, UIST '12, pp. 167–176. ACM, 2012. doi: 10.1145/2380116.2380139

[16] G. Laput, R. Xiao, and C. Harrison. ViBand: High-Fidelity Bio-Acoustic Sensing Using Commodity Smartwatch Accelerometers. In *Proceedings of the 29th Annual ACM Symposium on User Interface*

*Software and Technology*, UIST '16, pp. 321–333. ACM, New York, NY, USA, 2016. doi: 10.1145/2984511.2984582

[17] M. Liu, M. Nancel, and D. Vogel. Gunslinger: Subtle Arms-down Mid-air Interaction. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology*, UIST '15, pp. 63–71. ACM, 2015. doi: 10.1145/2807442.2807489

[18] Z. Lu, X. Chen, Q. Li, X. Zhang, and P. Zhou. A Hand Gesture Recognition Framework and Wearable Gesture-Based Interaction Prototype for Mobile Devices. *IEEE Transactions on Human-Machine Systems*, 44(2):293–299, 4 2014. doi: 10.1109/THMS.2014.2302794

[19] J. McIntosh, C. McNeill, M. Fraser, F. Kerber, M. Löchtefeld, and A. Krüger. EMPress: Practical Hand Gesture Classification with Wrist-Mounted EMG and Pressure Sensing. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, CHI '16, pp. 2332–2342. ACM, New York, NY, USA, 2016. doi: 10.1145/2858036.2858093

[20] M. R. Morris, A. Danielescu, S. Drucker, D. Fisher, B. Lee, m. c. schraefel, and J. O. Wobbrock. Reducing Legacy Bias in Gesture Elicitation Studies. *interactions*, 21(3):40–45, 5 2014. doi: 10.1145/2591689

[21] P. Paudyal, J. Lee, A. Banerjee, and S. K. S. Gupta. DyFAV: Dynamic Feature Selection and Voting for Real-time Recognition of Fingerspelled Alphabet Using Wearables. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces*, IUI '17, pp. 457–467. ACM, New York, NY, USA, 2017. doi: 10.1145/3025171.3025216

[22] K. Pietroszek, A. Kuzminykh, J. R. Wallace, and E. Lank. Smartcasting: A Discount 3D Interaction Technique for Public Displays. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design*, OzCHI '14, pp. 119–128. ACM, New York, NY, USA, 2014. doi: 10.1145/2686612.2686629

[23] T. Piumsomboon, A. Clark, M. Billinghurst, and A. Cockburn. User-defined Gestures for Augmented Reality. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pp. 955–960. ACM, New York, NY, USA, 2013. doi: 10.1145/2468356.2468527

[24] Z. Ren, J. Meng, J. Yuan, and Z. Zhang. Robust Hand Gesture Recognition with Kinect Sensor. In *Proceedings of the 19th ACM International Conference on Multimedia*, MM '11, pp. 759–760. ACM, New York, NY, USA, 2011. doi: 10.1145/2072298.2072443

[25] J. Rico and S. Brewster. Gestures All Around Us: User Differences in Social Acceptability Perceptions of Gesture Based Interfaces. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '09, pp. 64:1–64:2. ACM, New York, NY, USA, 2009. doi: 10.1145/1613858.1613936

[26] J. Rico and S. Brewster. Usable gestures for mobile interfaces: Evaluating social acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pp. 887–896. ACM, New York, NY, USA, 2010. doi: 10.1145/1753326.1753458

[27] J. Ruiz, Y. Li, and E. Lank. User-defined motion gestures for mobile interaction. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, CHI '11, pp. 197–206. ACM, 2011. doi: 10.1145/1978942.1978971

[28] J. Ruiz and D. Vogel. Soft-constraints to reduce legacy and performance bias to elicit whole-body gestures with low arm fatigue. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 3347–3350. ACM, 2015.

[29] T. S. Saponas, D. S. Tan, D. Morris, R. Balakrishnan, J. Turner, and J. A. Landay. Enabling Always-available Input with Muscle-computer Interfaces. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology*, UIST '09, pp. 167–176. ACM, New York, NY, USA, 2009. doi: 10.1145/1622176.1622208

[30] S. Siddhpuria, K. Katsuragawa, J. R. Wallace, and E. Lank. Exploring At-Your-Side Gestural Interaction for Ubiquitous Environments. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, DIS '17, pp. 1111–1122. ACM, New York, NY, USA, 2017. doi: 10.1145/3064663.3064695

[31] R. Tidwell, R. Akl, S. Akumalla, S. Karlaputi, D. Struble, and K. Kavi. Evaluating the Feasibility of EMG and Bend Sensors for Classifying Hand Gestures. In *Proceedings of the International Conference on Multimedia and Human Computer Interaction*, MHCI '13, pp. 1–8. ASET, Toronto, ON, Canada, 2013.

[32] R.-D. Vatavu and J. O. Wobbrock. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pp. 1325–1334. ACM, New York, NY, USA, 2015. doi: 10.1145/2702123.2702223

[33] R.-D. Vatavu and I.-A. Zaiti. Leap Gestures for TV: Insights from an Elicitation Study. In *Proceedings of the 2014 ACM International Conference on Interactive Experiences for TV and Online Video*, TVX '14, pp. 131–138. ACM, New York, NY, USA, 2014. doi: 10.1145/2602299.2602316

[34] D. Vogel and R. Balakrishnan. Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*, UIST '05, pp. 33–42. ACM, 2005. doi: 10.1145/1095034.1095041

[35] D. Way and J. Paradiso. A Usability User Study Concerning Free-Hand Microgesture and Wrist-Worn Sensors. In *Proceedings of the 2014 11th International Conference on Wearable and Implantable Body Sensor Networks*, BSN '14, pp. 138–142. IEEE Computer Society, Washington, DC, USA, 2014. doi: 10.1109/BSN.2014.32

[36] H. Wen, J. Ramos Rojas, and A. K. Dey. Serendipity: Finger Gesture Recognition Using an Off-the-Shelf Smartwatch. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, CHI '16, pp. 3847–3851. ACM, New York, NY, USA, 2016. doi: 10.1145/2858036.2858466

[37] D. Wigdor and D. Wixon. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1st ed., 2011.

[38] J. O. Wobbrock, M. R. Morris, and A. D. Wilson. User-defined Gestures for Surface Computing. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, CHI '09, pp. 1083–1092. ACM, 2009. doi: 10.1145/1518701.1518866

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

[40] Y. Zhang, R. Xiao, and C. Harrison. Advancing Hand Gesture Recognition with High Resolution Electrical Impedance Tomography. In *Proceedings of the 29th Annual ACM Symposium on User Interface Software and Technology*, UIST '16, pp. 843–850. ACM, New York, NY, USA, 2016. doi: 10.1145/2984511.2984574