Understanding Same-Side Interactions with Wrist-Worn Devices

Frederic Kerber
DFKI, Computer Science
Campus Saarland
Saarbrücken, Germany
frederic.kerber@dfki.de

Markus Löchtefeld
Aalborg University
Aalborg, Denmark
mloc@create.aau.dk

Antonio Krüger
DFKI, Computer Science
Campus Saarland
Saarbrücken, Germany
krueger@dfki.de

Jess McIntosh
Bristol Interaction Group,
University of Bristol
Bristol, United Kingdom
jm0152@my.bristol.ac.uk

Charlie McNeill
Bristol Interaction Group,
University of Bristol
Bristol, United Kingdom
cm0763@my.bristol.ac.uk

Mike Fraser
Bristol Interaction Group,
University of Bristol
Bristol, United Kingdom
mike.fraser@bristol.ac.uk

ABSTRACT

We investigate one-handed, same-side gestural interactions with wrist-worn devices. We contribute results of an elicitation study with 26 participants from various backgrounds to learn about gestures people would like to do when only able to interact using the arm on which they wear the device, e.g. while carrying something in the opposite hand. Based on the analysis of 1,196 video-taped gestures, 145 atomic gestures could be identified, which in turn were used to create a set of 296 unique gesture combinations. From these, we identified a conflict-free set of 43 gestures to trigger 46 common smartwatch tasks. The results show that symbolic gestures such as drawing a question mark for activating a help function are consistently used across participants. We further found symbolic and continuous gestures to be used significantly more often by men. Based on the results, we derived guidelines that should be considered when designing gestures for SSI.

Author Keywords

Same-side interaction; one-handed interaction; wrist-worn devices; smartwatches

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Since the development of the first Linux-based smartwatch in 2001 [23], the advancing miniaturization of technical components has made it possible to create more and more powerful wrist-worn devices with smart features such as touchscreens, customizable watchfaces and extensions through the already well-established app store concept. Despite their technical sophistication, the small screen size, usually only about 1.5" to 2.5", generates new challenges, in particular with respect to touch input. Consequently, recent research investigated alternative input possibilities such as utilizing the wristband [1, 6, 26], adding additional sensors [25, 38] or using gestural input [29]. Although they all provide viable alternatives to classic touch input and solve the problem of occlusion [3], most of them cannot solve another common problem of device interactions on the go: namely, they still require the opposite hand for interaction. However, the opposite hand may not be easily available, e.g. when walking in the city and carrying purchases, or in an industrial context while doing factory work. As shown by Ng et al. [24], such an encumbrance negatively affects standard gestures commonly performed on touchscreens. Hands-free alternatives such as speech input may also not always be appropriate, e.g. due to noisy environments or privacy concerns. We therefore follow Rekimoto's approach of one-handed and thereby same-side interactions. In contrast to other approaches (e.g. [29, 33]), we follow a participatory design approach instead of targeting the topic from a technical perspective, i.e. creating an input technology and analyzing its features. In an elicitation study with 26 participants, we let people invent gestures for a set of 46 common smartwatch interaction tasks. We thereby completely abstracted from any technical restrictions and were only interested in the way people desired to interact. Based on the findings in our study, we contribute (1) a characterization of user-defined gestures for same-side interaction (SSI) with wrist-worn devices, (2) a conflict-free set of SSI gestures for 46 common tasks that allow interacting with wrist-worn devices, (3) an overview about the technical feasibility of these gestures with respect to currently available recognition methods from the related work, and (4) guidelines that should be kept in mind when selecting or designing gestures for SSI.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

NordiCHI '16, October 23–27, 2016, Gothenburg, Sweden

© 2016 ACM. ISBN 978-1-4503-4763-1/16/10...$15.00

DOI: http://dx.doi.org/10.1145/2971485.2971519

The rest of the paper is organized as follows: We first give an overview on related work with respect to different input possibilities both for opposite-side interaction (OSI) and SSI, as well as corresponding sensing technologies. We then present the results of our elicitation study in which we analyzed the characteristics of user-generated SSI gestures for a set of 46 common smartwatch interaction tasks we created based on examining popular apps as well as related work. Based on the results of the study, we provide a conflict-free set of SSI gestures for these tasks and an analysis of the technical feasibility of sensing the contained gestures with currently available approaches. We conclude by providing guidelines that should be considered when designing or selecting gestures for SSI.

RELATED WORK

We will first provide an overview on related work with respect to different input possibilities as well as corresponding sensing technologies. After providing the results of our elicitation study, we will come back to the technical approaches and relate the user-generated gestures to them.

In terms of interaction with wearable devices, Profita et al. found the wrist and the forearm to be the most socially acceptable area to position such devices [27]. Comparable to a normal watch, such wearable computing devices will change body position as well as size and interface over time, following cultural concepts and fashion [19]. The increasing variety and popularity of smartwatches underpins the trend towards a wrist-worn computing device, but as of now, only a limited set of interaction techniques exists. Most of the related work in this field focuses on Opposite-Side Interaction (OSI) techniques, meaning they require the hand that is not wearing the wearable device to operate it. In this paper, we focus on leveraging the capabilities of such a device using the arm that is wearing it, so-called Same-Side Interaction (SSI) [12]. We split the related work into four parts, the first two being about interaction with wrist-worn devices, the third covering possible sensing techniques to enable SSI with a wrist-worn device and the last part considering prior work in terms of eliciting user interactions.

Opposite-Side Interaction (OSI)

Early touch input on smartwatches was investigated by Raghunath and Narayanaswami [28]. Most of their design guidelines have been used in Android Wear. Blasko et al. [4] used bidirectional strokes that were segmented by tactile landmarks, which enabled eyes-free input on rectangular [5] and circular watches. In contrast, the work of Ashbrook et al. investigated touch interaction on a circular wristwatch [2] without tactile landmarks. They developed a model based on empirical data that allows determining the error rate for variously-sized buttons placed around the rim. Another suitable interaction technique for smartwatches would be touch input on the back of the device [3, 26].

Instead of relying on touch input, Xiao et al. developed a multi-degree-of-freedom, mechanical interface for smartwatches [38] which allowed for continuous 2D panning and twist as well as binary tilt and click. Pasquero et al. developed a turnable bezel for such mechanical continuous input [25]. Similar to this, Kim et al. presented a prototype of a wrist-worn device that utilized an array of infrared proximity sensors to interpret hand gestures made over it [15]. In [10], Kerber et al. compared different OSI techniques that can be found in today's commercially available smartwatches.

Same-Side Interaction (SSI)

One of the first SSI-operated wrist-worn devices was presented by Rekimoto with the GestureWrist [29]. The authors used capacitive sensors and an accelerometer to sense wrist-shape changes and forearm movements for input. This allows the user to input commands using only one arm, but requires additional sensors which are typically not available in today's wrist-worn devices.

Kerber et al. presented the results of a user study comparing the performance of static and dynamic peephole interactions for a navigation task on a smartwatch [11]. This is, to the best of our knowledge, the first direct comparison of OSI and SSI. While they found the touch interaction of the static peephole (OSI) to be on average 12% faster, it was marginal compared to the advantage of using only one arm to interact (SSI).

In [12], Kerber et al. used electromyography (EMG) to control a smartwatch. Their initial pilot study compared SSI using a Myo Wristband against traditional OSI touch to control a music player on a smartwatch. While their preliminary study did not find any significant difference in terms of task completion time, keeping in mind the early state of commercial EMG devices, their results demonstrated the general feasibility of such a SSI approach.

As can be seen, the majority of prior work examined OSI and the approaches that relied on SSI have not been convincingly successful enough to present an alternative to touch input. On the other hand, an effective SSI approach would provide advantages in many everyday situations in which OSI is not easily applicable. In terms of touch input, gloves (e.g. in winter or in industrial contexts) could be a hindrance. At least for situations that do not require visual feedback (e.g. skipping tracks or changing the volume in a music app), input possibilities that do not require interacting with the device itself or its direct surroundings are advantageous if the device is covered by clothes such as shirts, jackets or protective clothes. It should also not be forgotten that OSI often requires both hands – the one that carries the device and the other to execute the interaction. Hence, situations requiring carrying or holding something (e.g. during shopping or factory work) are problematic for OSI. We strongly believe that gestural SSI can be an alternative in these situations. Consequently, we aim to understand how participants would like to use their wrist-worn devices with SSI.

Sensing Technologies

In this paper, we focus on SSI with a wrist-worn device. Generally speaking, a variety of different interactions are possible with just one arm, ranging from arm gestures to single-finger movements. Even though a variety of sensors exist that allow sensing this type of gesture from a distance, such as the Microsoft Kinect or Leap Motion, we focus only on sensing technologies that can be integrated into a wearable device.

Inertial measurement units (IMU) using accelerometers, gyroscopes and sometimes magnetometers have become common in today's wearable devices to detect gestures or for activity recognition. Early work by Rekimoto explored an accelerometer to detect arm orientation for gestures [29]. Ward et al. [33] used a combination of microphones and accelerometers to detect gestures and activities. Lately, the IMU inside smartwatches has even been used to detect and support Cardiopulmonary Resuscitation (CPR) movements [9].

With Digits, Kim et al. [14] presented a system that is based on a combination of an infrared emitter and camera as well as a laser line to reconstruct 3D hand poses. In contrast to this, the use of ultrasonic sensors has been explored in [18]. While the authors used them to realize a slidebar on the arm that would be operated in an OSI manner, the approach could be used to detect the same-side hand approaching the watch as well. These techniques work very similar to a depth sensing camera (such as the time of flight camera DUO mini lv1[1]). Such could be integrated in to the smartwatch and sense the orientation of the fingers of the hand.

Sensing the muscle activation inside the arm is another possible technology to detect arm and hand gestures. Electromyography (EMG) has been successfully used to detect sign language that contains some of the most complex hand gestures [17]. In a recent paper, Nagar and Zhu used a wrist-based EMG system to detect the three gestures involved in playing rock, paper, scissors [22]. Nevertheless, EMG has the drawback that it is often used only at the upper part of the lower arm, which is not really convenient for a wrist-worn device. Rekimoto used capacitive sensors inside the wristband to sense simple hand gestures [29].

Eliciting user interactions

Our approach to understanding the users' needs and wishes in terms of SSI is to employ an elicitation study. Incorporating users in the design process has been successfully conducted on many prior occasions as part of participatory design [30]. Confronting users with target actions and having them perform interactions that they think should cause them, has already been done several times in the past. Good et al. [8] used it to develop a command-line email interface. It was also used by Wobbrock et al. [35] to design EdgeWrite. Furthermore, Wobbrock et al. [36] used an elicitation technique to design surface computing gestures. More related to our use-case, Weigel et al. [34] explored the possibility of skin-based interactions in such a manner, whereas Vatavu and Zaiti investigated user-defined gestures for interactive TV [32].

The works mentioned above showed, for different use-cases and with different approaches, the applicability of such elicitation studies to bring up meaningful interactions through participatory design. As smartwatches and other wrist-worn devices are becoming more and more ubiquitous in people's every day lives, we expect participants of our study to have an even deeper understanding of their needs and wishes. Therefore, we are fairly certain that such an elicitation study will generate meaningful results.

ELICITATION STUDY

To get an understanding of SSI gestures participants are willing to do, we utilized a guessability study methodology [35] that presents desired effects (e.g. "The volume is increased.") and elicits their causes, i.e. the gestures, to invoke them.

Tasks and Procedure

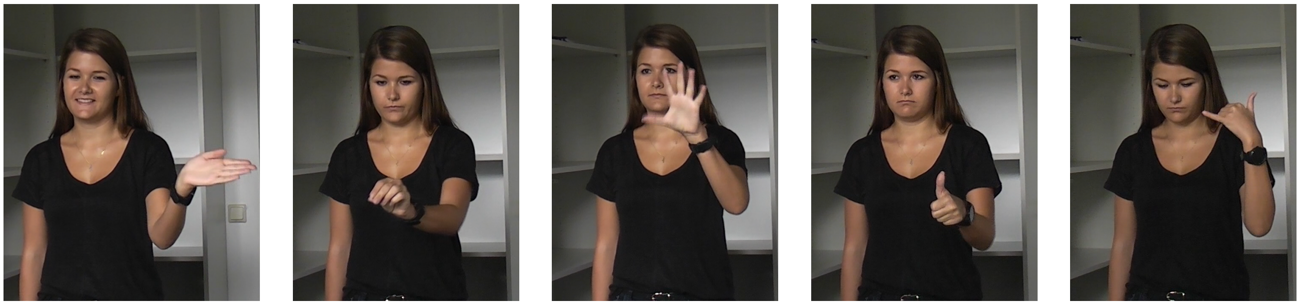
Before executing the actual elicitation study, we first conducted a set of tasks (effects) representing typical situations when interacting with a smartwatch. To compile this set, we considered related work ([16, 34, 36]) on the one hand and examined popular app types in the respective app stores on the other. We selected 46 different tasks (cf. Figure 2) and presented them in random order to every participant and for each, the participants were instructed to invent a gesture while using only the arm where the smartwatch was worn. As current sensing technology could potentially capture only a subset of the gestures participants are willing to do, we opted for not using any sensing at all. Instead, the participants were filmed from two different angles to ensure that every gesture is clearly visible on the taped material for a manual analysis afterwards. This approach opened up the possibility to observe a completely uninfluenced behavior, free of possible restrictions due to current technology, and has been used in similar settings in the past [16, 34, 36]. To avoid introducing other factors that could potentially influence participants, we also refrained from providing any output or reaction to the participants' gestures. However, people were equipped with a switched-off smartwatch (LG G Watch R) for reference (e.g. when executing gestures like bringing the smartwatch in the direct view area). We instructed the participants that they could execute any gesture they can think of as long as it contains a movement of either shoulder, elbow, wrist, thumb or fingers of the arm where the smartwatch was worn.

To mimic a realistic use case, the participants were told to be walking through the city after buying some goods. Consequently, they were standing during the study and carried a bag with a book in the opposite hand. For each of the 46 tasks, the experimenter read out aloud a description of the desired effect to achieve, e.g. "Consider that you are listening to some music. Now increase the volume." and was available in case the participants had any requests. After executing their desired gesture, the participants were presented two questions on a nearby display to rate the goodness and perceived ease of their invented gesture each on a 7-point scale. In total, we collected 1,196 gestures (46 gestures per participant) during single-user sessions with a median duration of 21.5 minutes (min=16 minutes, max=29 minutes). Figure 1 shows five examples of these gestures.

Participants

For the elicitation study, we recruited 26 voluntary participants (aged 15-38 years, median age 26 years, 8 females) from various cultural backgrounds (Europe, Middle East, India). 24 of them were right-handed, whereas only two

[1] `https://duo3d.com/product/duo-mini-lv1`, last accessed 28/07/2016



(a) Previous track **(b) Select single item** **(c) Take a picture** **(d) Approve payment** **(e) Emergency call**

Figure 1. Five different gestures performed during the elicitation study. Captions name the effect that should be achieved with the invented gestures.

were left-handed; five wore their watch on their dominant side. Their occupations included child care worker; chimney sweeper; accountant; engineer; pupil; student in computer science, English, marketing, physics, product design or psychology; researcher in computer science, secretary and software engineer. In total, 14 of them had a background related to computer science. All but one participant had owned a touch-enabled smartphone for at least one year and only one participant owned a device comparable to a current smartwatch (Striiv Fusion). None of the them received any compensation for participating.

Analysis Methodology

In the following, we describe our approach to analyzing the collected video material. We first manually described every user-defined gesture (e.g. "Rotate lower arm inwards.") and classified it according to the *involved body parts*: fingers (including thumb), wrist, elbow and shoulder. We further distinguished between *continuous* and *discrete* gestures as previously done by e.g. [36, 37]. Continuous gestures describe movements that directly relate to the impact they have, e.g. rotating the lower arm to scroll in a longer text, whereas discrete gestures describe closed movements such as pointing somewhere. Another dimension we investigate is the *size of the movement* or in other words the required space to execute it. A movement of the index finger alone is considered as a small movement whereas for example waving with the complete arm is considered a large movement. We do not further distinguish intermediate steps, i.e. a gesture is considered either small or large. We also annotated whether the gesture is inspired by *touch-based interaction*, such as the pinch gesture typically used when zooming out, and lastly, we also analyzed whether the gesture is mimicking an interaction in the physical world (either specifically or symbolically). An example of the former is changing the value of a virtual spin-control by rotating the lower arm with thumb and index expanded as they would grasp the physical control; the latter can be illustrated with a gesture such as drawing a question mark when calling a help function.

Results

An analysis of the 1,196 videotaped gestures showed that they consisted of 296 unique gestures. We then decomposed combined movements into their atomic parts as far as practicable, e.g. "Rotate lower arm while clenching into a fist" was split into "Rotate lower arm" and "Clench into a fist". After this step, 145 atomic base gestures were identified which were further analyzed. From the 296 unique gestures, 121 (40.9%) consisted of only an atomic base gesture, whereas the other 175 gestures (59.1%) were combinations of two or more atomic base gestures, the largest one consisting of six parts. We then analyzed the distribution of the gestures per participant. Each participant used 17 to 37 unique gestures (median=30.5) to trigger the given 46 tasks. The repetition count for the most often utilized gesture varied between 3 and 18 (median=4.5). A Dixon's Q test with 99.5% confidence level revealed two outliers (Q=$0.53 > 0.517$ and Q=$0.54 > 0.517$) which we excluded from further analysis.

Based on the remaining 24 participants (7 female, 13 with a background related to computer science), we re-analyzed the set of 1,104 gestures and identified 284 unique ones, of which 121 (42.6%) were atomic base gestures, whereas the other 163 gestures (57.4%) were combinations of two or more atomic base gestures. Each of the remaining participants used 17 to 37 unique gestures (median=31) to trigger the given 46 tasks. The repetition count for the most often utilized gesture now varied between 3 and 10 (median=4). The most frequently used gesture was applied 36 times, whereas 140 gestures were only used once and additional 47 only twice.

None of the participants had any problems finding gestures for the given tasks, which is also supported by an average goodness rating of 5.29 for the 1,104 gestures. An average rating of 6.27 for ease shows that the participants were able to create gestures they deemed to be easily executable. Mann-Whitney tests revealed that male participants, based on their self-assessment, considered their gestures significantly better suited ($U = 149,323, p < 0.001$) and easier to execute ($U = 141,857, p < 0.001$) than women did for their gestures. Furthermore, men used significantly more symbolic gestures ($U = 136,735, p < 0.01$) and more continuous gestures ($U = 131,031, p < 0.05$) than women did. Table 1 gives an overview of the gestures along the dimensions presented earlier. We report our findings separately for the 145 atomic base gestures and the resulting 284 unique gestures.

Gesture Composition

As already reported, 163 of the gestures did not consist of only a single movement, but were a combination of movements. We thereby distinguish gestures that are executed suc-

Dimension	Base Gestures	Unique Gestures
Involved body parts		
Fingers (F) only	39	56
Wrist (W) only	9	6
Elbow (E) only	18	19
Shoulder (S) only	5	5
F and W	1	16
F and E	3	30
F and S	2	15
W and E	2	4
W and S	1	-
E and S	53	54
F, W and E	3	5
F, W and S	-	1
F, E and S	7	64
W, E and S	2	8
F, W, E and S	-	1
Continuous	17	36
Discrete	128	248
Large	75	142
Small	70	142
Touch-inspired	12	20
Real-world-like	16	35
Symbolic	76	147

Table 1. Results from the analysis of the 145 atomic base gestures and the 284 unique gestures built from them.

cessively, e.g. "Extend index finger, then retract index finger" and those that can be seen as variants of other gestures, as is the case for "Rotate lower arm with extended index finger", in which the main gesture is "Rotate lower arm" and the extended index finger only results in a modification. From the 145 atomic base gestures, 13 have also been used to create variants with "Extend index finger" (45×) and "Clench into a fist" (26×) being the most used modifiers. In total, 61 main gestures have been used with modifiers to create 106 variants.

Agreement Score

We further analyzed the gestures in relation to the tasks they were executed for. Therefore, we grouped the 1,104 gestures on a per-task basis into 46 groups and analyzed these groups individually. Based on the agreement concept presented in [35, 36] and refined in [31], we computed the corresponding score. The overall agreement score $\mathcal{AR}$ is thereby defined as

$$\mathcal{AR} = \frac{\sum_{t \in T} \left(\frac{|P_t|}{|P_t|-1} \sum_{P_i \subseteq P_t} \left(\frac{|P_i|}{|P_t|} \right)^2 - \frac{1}{|P_t|-1} \right)}{|T|}$$

In the equation above, t refers to a task in the set of all tasks T, P_t is the set of gestures executed for task t, and P_i is a subset of P_t consisting of identical gestures. A higher score thereby refers to a higher agreement, i.e. a larger number of people that consistently chose an identical gesture. To give a specific example, consider a task like "Pause" for which three different gestures have been created by the participants. The three gestures were repeated 7, 5 and 3 times respectively. Based on this gesture distribution, the following agreement score is computed:

$$\mathcal{AR}_{\text{Pause}} = \tfrac{15}{14} \left(\left(\tfrac{7}{15}\right)^2 + \left(\tfrac{5}{15}\right)^2 + \left(\tfrac{3}{15}\right)^2 \right) - \left(\tfrac{1}{14}\right) = 0.32$$

In Figure 2, the agreement scores for the 46 tasks are illustrated in blue. The overall agreement score is $A = 0.07$. An unpaired t-test revealed a no significant difference in the agreement score between male (M=0.08, SD=0.1) and female (M=0.06, SD=0.09) participants, $t(90) = -0.983$, $p = 0.328$. The same analysis was conducted without considering different variants as separate gestures. The corresponding agreement scores are depicted in Figure 2 in orange. The overall agreement score is slightly increased to $A = 0.09$. Again, no significant difference in the agreement score between men (M=0.1, SD=0.11) and women (M=0.08, SD=0.11) could be observed; $t(90) = -0.89, p = 0.376$.

Continuous Tasks/Gestures

From the set of 46 tasks, seven tasks can be seen as continuous, e.g. "Decrease volume" or "Zoom out". As outlined above, 36 unique gestures were considered continuous, resulting in 118 gesture executions containing one of them.

Considering all gestures, we have a consensus of 89.5%, i.e. 89.5% of the continuous (discrete) tasks are executed with a continuous (discrete) gesture. A Mann-Whitney-Test revealed that consensus of gestures for discrete tasks is significantly higher than for continuous tasks, $U = 114,696, p < 0.001$. A Rosner's Extreme Studentized Deviate test with a significance level of 0.05 revealed eight outliers. Consistently, all of them were continuous tasks with a higher number of discrete gestures chosen. However, for four of these tasks, the majority chose a continuous gesture, and in the other four cases, the most often repeated gesture was a continuous one.

Negatively Connoted Tasks/Gestures

From the set of tasks we defined for the elicitation study, several can be seen as having a negative connotation, for example "Delete" or "Reject call". To get a better understanding of what people consider as "negative", we let five independent researchers rate the 46 tasks as positively or negatively. We considered a task as negative if at least four of the five ratings supported this. In total, 11 tasks were classified as negative. On the other side, the analysis revealed 19 atomic base gestures that contained a movement away from the body – something we hypothesize is connected to unwanted or negative actions. Furthermore, nine additional gestures exist that have a clear negative connotation, e.g. "Thumb down".

Considering all executed gestures, we have a consensus of 80.9%, i.e. 80.9% of the tasks that are considered to have

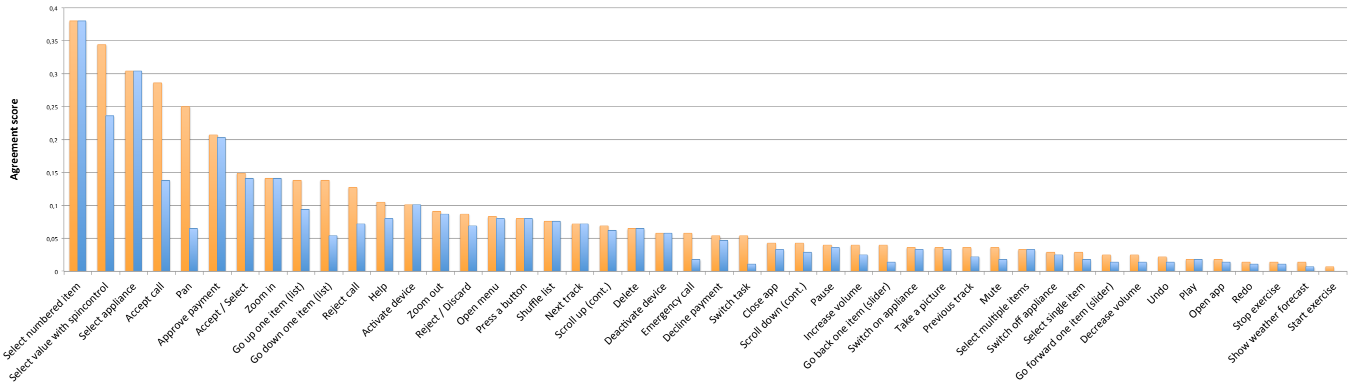


Figure 2. Agreement scores, representing the amount of people that consistently chose an identical gesture, for the 46 tasks without (orange) and with (blue) considering gesture variants as separate gestures, ordered by score values for the former.

a negative connotation are executed with a gesture that contained a movement away from the body or had a negative connotation, whereas the tasks without a negative connotation were executed with a gesture that did not contain a movement away from the body. A Mann-Whitney test revealed that consensus on gestures for negatively connoted tasks is significantly lower than for non-negatively connoted tasks, $U = 144,888, p < 0.001$. A Rosner's Extreme Studentized Deviate test with a significance level of 0.05 revealed seven tasks as outliers, which we further investigated.

In line with the results from comparing negatively and non-negatively connoted tasks, all seven tasks were rated as negatively connoted, but at most half of the participants in the elicitation study chose a gesture that includes a movement away from the body or with a negative connotation. An in-depth analysis of the used gestures showed an overlap with the task's opposite in five of the seven cases, e.g. for "De-activate device", participants used the same gestures they used for "Activate device". In this sense, participants did not regard the negative aspect of the specific task but invented gestures they deemed suitable for both the positive and the negative task. A similar effect could be observed for the sixth task, namely "Mute", which showed overlaps with "Increase volume" – a task that could be seen as its opposite. In case of the last task, "Delete", two aspects could be observed. One the one hand, 50% of the invented gestures were in fact negatively connoted. Regarding the remaining twelve gestures, on the other hand, we often see a combination of two unbiased gestures that together make up a negative impression, e.g. "Clench into a fist" and "Spread fingers" are not negatively connoted by themselves, but if both are combined consecutively, they create the impression that something is grabbed and dropped vigorously, which can be seen as a negatively connoted action.

Conflict-free Gesture Set for SSI

For the next part of this paper, we considered the most frequently used gesture(s) for each of the 46 tasks to create a gesture set suitable to be used for controlling a smartwatch solely based on same-side interactions. For 35 tasks, a unique most frequently used gesture could be found with repetition counts between 3 and 15 (median=5.5). 14 of the chosen gestures were conflict-free, i.e. the gesture was only used to trigger one task. The other gestures formed nine sets with potential conflicts, i.e. the same gesture was used to trigger multiple tasks. Based on the nature of the task set, not every overlap is in fact a conflict. To give a specific example, we consider the two tasks "Decline payment" and "Switch off appliance in smart home". Both have "Thumb down" associated as the most frequently used gesture. When considering the context, there is no actual overlap as both situations are not very likely to occur in parallel without having a clear indication for the user which of the two tasks is the currently active one, e.g. a full-screen dialog could request approval for the payment. Based on this approach, only two of the nine sets actually contain a conflict that needs to be resolved. For both sets, we identified the candidate with the lowest repetition count to replace it with (one of) the second most used gesture(s). For tasks with more than one most frequently used gesture, we consider as suitable all gestures that do not conflict (in the same sense as above) with other tasks. The concept of having more than one triggering action, called aliasing, is known to increase input guessability [7, 35]. With the same idea in mind, the distinct gesture sets could also be extended by the second or third most frequently used gesture as long as no conflicts are introduced. Table 2 (page 8) shows the gestures that are used in the conflict-free gesture set for our 46 tasks (Table 3, on the same page).

As outlined above, we also wanted to check which of the gestures could potentially be sensed by one or several of the technologies presented in the related work section. Our estimate with respect to the three types of sensing, electromyography, inertial measurement unit and infrared/depth sensing, is also included in Table 2. Although this looks promising at first glance as every gesture could potentially be sensed, some limitations have to be taken into account. First of all, a mechanism to distinguish involuntary motion from gestures done on purpose is required. As using the touchscreen or pressing a button would again require the opposite hand, this is not a suitable approach, but a motion delimiter such as the one presented in [13] could be used. Furthermore, it has to be considered that many different gestures have to be detectable and reliably distinguishable in parallel, as high recognition rates with low rates of false-positives are crucial for such a system. However, recent work also combining different types of sensors such as electromyography and pressure sensors [20] seems promising in this respect. Nonetheless, only an evaluation with corresponding implementations of the systems mentioned can provide reliable insights, but this is beyond the scope of this paper.

Discussion

The analysis showed that the participants in general did not have any problems generating gestures for SSI that are both suitable to invoke the given tasks and easy to perform.

An overall agreement rate of below 0.1 is remarkable compared to related elicitation studies that typically reached higher values (see [31] for the results of 18 studies). To assess this result, we should keep in mind that all but one participant did not have prior experience with smartwatches or comparable devices. Consequently, no interaction model (legacy bias [21]) was present as a fallback (as is for example the case with typically well-known mouse interaction when eliciting surface gestures). Although the result indicates that a personalized interaction approach might be necessary, we consider the compiled conflict-free gesture set as meaningful for two reasons: (1) Due to aliasing as explained above, the practical applicability of the gesture set is further expanded and (2) the gesture set can be seen as an informed starting point, thereby reducing the need for the user as well as the developer to first define gestures for every interaction.

Based on the results from our elicitation study, we also derived guidelines that should be kept in mind when designing gestures for SSI. Although they are not surprisingly different from what is known from similar research for gestural input, we see it as additional confirmation for the field of smartwatch interaction, and especially SSI, which has not been investigated before.

Similar tasks are triggered by similar gestures

If we consider similar tasks, e.g. “Go up one item in a list”, “Go back one item in a slider”, “Scroll up in a text” and “Undo” which all four are concerned with going back to something that has been seen before, we also see similar gestures in the conflict-free gesture set, namely “Move arm from down to up”, “Stroke upwards” (used for two tasks) and “Rotate lower arm outwards”. In total, seven groups of similar tasks could be defined – five of these also show similar gestures, whereas the other two groups are dominated by symbolic gestures.

Opposing tasks are triggered by opposing gestures

Considering opposing tasks such as “Go up one item in a list” and “Go down one item in a list”, we also see opposing gestures in the final gesture set, i.e. “Move arm from down to up” and “Move arm from up to down”. In total, 15 sets with opposing gestures could be defined – six of these contain opposing gestures, whereas six additional ones are made up of symbolic gestures. Only three sets, namely “Undo/Redo”, “Go back/forward in a slider” and “Go to previous/next track” did not directly contain opposing gestures. In these three cases, the “forward” direction is consistently associated with a movement towards the body. However, the countermotion is not expressed by a movement away from the body. Instead, another movement towards the body is chosen, but, a completely different one. As going back to a previously seen or heard item is rather positive (people want to experience the same thing again), this could serve as an explanation.

Atomic gestures have higher agreement than combined ones

Although the set of all user-generated gestures contains a large number of gestures that are combinations of more than one atomic gesture, only two gestures in the final gesture set are combined ones. As combinations provide more variability, it is not surprising that the set that is made of the most-often chosen, identical gestures mainly contains unambiguous, atomic gestures. A positive side-effect of this trend is, however, that the gestures are potentially easier to remember and faster to perform.

Symbolic gestures have a high agreement

The sets of the most-frequently used gestures for the 46 tasks contain symbolic gestures in 31 cases (67.4%), which indicates on the one hand that people have similar symbolic depictions in mind, and on the other hand that these symbolic depictions could be expressed appropriately by means of one-handed gestural interactions. To give a prominent example, consider the task “Accept a call”, for which more than 50% of the participants executed a movement of their hand towards their ear, which symbolizes moving a phone to their ear.

CONCLUSION AND FUTURE WORK

In this paper, we investigated one-handed, same-side interactions with wrist-worn devices such as smartwatches. We presented the results of an elicitation study with 26 participants from various backgrounds to learn about SSI gestures that are preferred by the users. We analyzed a set of 46 common smartwatch tasks and identified a conflict-free set consisting of 43 SSI gestures. We provide further insights with respect to involved body parts, the composition of gestures and their symbolic nature in the specific context. We also found symbolic as well as continuous gestures to be used significantly more often by men.

To the best of our knowledge, the presented results are the first in-depth analysis of SSI with wrist-worn devices. The fact that users had no difficulties finding gestures for all use-cases while being limited to using only one arm, demonstrates that SSI is a real alternative for smartwatch interaction while on the go or in other situations that hinder OSI.

For future work we will investigate possible sensing techniques in more depth. While our initial survey suggests that most gestures should be possible to sense, an in-depth analysis, especially of the capabilities of a wrist-worn EMG, needs to be conducted. Furthermore, we will follow up with a social acceptability study of these gestures and see how practical they are in everyday life. Last but not least, we have to consider the connection of user input (what we investigated here) and interaction feedback or output, respectively – especially when mixed with the next input interaction (e.g. open a message, read a bit, scroll in the message, etc.).

ID	Description	Detectable via
1	Shake wrist	IMU
2	Point with index finger	EMG, IMU/IR&DC
3	Wave out to in	EMG, IMU
4	Show number 1-5	EMG/IR&DC
5	Slow stroke 90° downwards	EMG, IMU
7	Thumb up	EMG, IMU/IR&DC
11	Thumb down	EMG, IMU/IR&DC
12	Make a fist	EMG/IR&DC
16	Wave up to down	EMG, IMU
21	Draw a question mark	IMU
22	Pinch	EMG/IR&DC
28	Unpinch/zoom	EMG/IR&DC
33	Move hand to ear	IMU
35	Move hand (dynamic peephole)	IMU
36	Move arm in parallel to body to the inner side	IMU
40	Rotate lower arm outwards (continuous)	IMU
41	Move hand to mouth	IMU
43	Push flat hand forward	EMG, IMU
45	Spread fingers	EMG/IR&DC
52	Rotate lower arm outwards	IMU
54	Move arm down to up	IMU
55	Move arm up to down	IMU
56	Rotate lower arm inwards (continuous)	IMU
62	Rotate lower arm 180°	IMU
63	Move arm to rest position (hanging)	IMU
66	Move hand to direct view area	IMU
69	Wave up to down with index finger only	EMG
72	Stroke upwards	EMG, IMU
74	Stroke downwards	EMG, IMU
88	Wave arm away from the body	IMU
89	Circular movement of lower arm inwards	IMU
98	Shake lower arm left and right	IMU
104	Move arm further away from head / eyes	IMU
117	Extend index finger	EMG/IR&DC
119	Stretch arm above head	IMU
139	Rotate lower arm	IMU
140	Tap	EMG/IR&DC
142	Grab virtual item, throw away to the outside	EMG, IMU/IR&DC
146	Snap with middle and thumb	EMG/IR&DC
156	Extend middle finger	EMG/IR&DC
166	Clap open hand on leg	IMU
178	Move arm up to down (continuous)	IMU
218	Retract index finger	EMG/IR&DC

Table 2. Gestures used in the conflict-free gesture set. Directional indications refer to the body of the performer, e.g. "out to in" means a movement from the outside to the inner side.

The last column indicates the sensing technology that could potentially detect the gesture. A comma-separated list indicates that a combination of technologies is required; alternatives are indicated by "/".
EMG = Electromyography
IMU = Inertial Measurement Unit
IR&DC = Infrared Sensing or Depth Camera

Task	Utilized Gesture IDs	Conformity
Activate device	66	21%
De-activate device	1, 63	33%
Open app	7, 12, 45, 62	29%
Close app	88	17%
Select single item	4, 12, 69, 140	33%
Select multiple items	4	17%
Select numbered item	4	63%
Zoom in	28	33%
Zoom out	22, 104	42%
Pan	35	50%
Accept/Select	7	38%
Reject/Discard	88	25%
Open menu	45	29%
Help	21	33%
Switch task	52	13%
Undo	72	13%
Redo	36, 89	17%
Delete	142	21%
Go up one item (list)	54	29%
Go down one item (list)	55	29%
Go back one item (slider)	72	17%
Go forward one item (slider)	3	13%
Scroll up (cont.)	40	21%
Scroll down (cont.)	16, 178, 56	38%
Shuffle list	98	21%
Accept call	33	54%
Reject call	88	33%
Emergency call	33	25%
Mute	41	17%
Increase volume	40	13%
Decrease volume	56	13%
Next track	3	25%
Previous track	72	17%
Play	117 + 156	13%
Pause	43	17%
Approve payment	7	46%
Decline payment	11	21%
Select appliance	2	54%
Switch on appliance	2, 146	25%
Switch off appliance	11	13%
Press a button	2	29%
Select value with spin control	139	50%
Start exercise	5	8%
Stop exercise	166, 74, 12, 88	33%
Show weather forecast	119	13%
Take a picture	117 → 218, 12	25%

Table 3. Conflict-free gesture set for 46 common smartwatch interaction tasks based on the most-frequently used user-defined SSI gestures. Alternatives are indicated by a comma-separated list, simultaneous executed gestures are indicated by a +, gestures that are executed successively are connected by a →. The third column indicates the percentage of participants that chose the exact gesture during the elicitation study.

REFERENCES

1. Ahn, Y., Hwang, S., Yoon, H., Gim, J., and Ryu, J.-h. BandSense: Pressure-Sensitive Multi-Touch Interaction on a Wristband. In *Proc. CHI EA '15*, ACM (2015), 251–254.
2. Ashbrook, D., Lyons, K., and Starner, T. An Investigation into Round Touchscreen Wristwatch Interaction. In *Proc. MobileHCI '08*, ACM (2008), 311–314.
3. Baudisch, P., and Chu, G. Back-of-Device Interaction Allows Creating Very Small Touch Devices. In *Proc. CHI '09*, ACM (2009), 1923–1932.
4. Blasko, G., and Feiner, S. An Interaction System for Watch Computers Using Tactile Guidance and Bidirectional Segmented Strokes. In *Proc. ISWC '04*, IEEE Computer Society (2004), 120–123.
5. Blasko, G., and Feiner, S. Evaluation of an Eyes-Free Cursorless Numeric Entry System for Wearable Computers. In *Proc. ISWC '10* (2006), 21–28.
6. Funk, M., Sahami, A., Henze, N., and Schmidt, A. Using a Touch-Sensitive Wristband for Text Entry on Smart Watches. In *Proc. CHI EA '14*, ACM (2014), 2305–2310.
7. Furnas, G. W., Landauer, T. K., Gomez, L. M., and Dumais, S. T. The Vocabulary Problem in Human-System Communication. *Commun. ACM 30*, 11 (1987), 964–971.
8. Good, M. D., Whiteside, J. A., Wixon, D. R., and Jones, S. J. Building a User-Derived Interface. *Commun. ACM 27*, 10 (1984), 1032–1043.
9. Gruenerbl, A., Pirkl, G., Monger, E., Gobbi, M., and Lukowicz, P. Smart-Watch Life Saver: Smart-Watch Interactive-Feedback System for Improving Bystander CPR. In *Proc. ISWC '15*, ACM (2015), 19–26.
10. Kerber, F., Kiefer, T., and Löchtefeld, M. Investigating Interaction Techniques for State-of-the-Art Smartwatches. In *Proc. CHI EA '16* (2016), 2540–2547.
11. Kerber, F., Krüger, A., and Löchtefeld, M. Investigating the Effectiveness of Peephole Interaction for Smartwatches in a Map Navigation Task. In *Proc. MobileHCI '14*, ACM (2014), 291–294.
12. Kerber, F., Lessel, P., and Krüger, A. Same-Side Hand Interactions with Arm-placed Devices Using EMG. In *Proc. CHI EA '15*, ACM (2015), 1367–1372.
13. Kerber, F., Schardt, P., and Löchtefeld, M. WristRotate: A Personalized Motion Gesture Delimiter for Wrist-Worn Devices. In *Proc. MUM '15*, ACM (2015), 218–222.
14. Kim, D., Hilliges, O., Izadi, S., Butler, A. D., Chen, J., Oikonomidis, I., and Olivier, P. Digits: Freehand 3D Interactions Anywhere Using a Wrist-Worn Gloveless Sensor. In *Proc. UIST '12*, ACM (2012), 167–176.
15. Kim, J., He, J., Lyons, K., and Starner, T. The Gesture Watch: A Wireless Contact-Free Gesture Based Wrist Interface. In *Proc. ISWC '07*, IEEE Computer Society (2007), 1–8.
16. Lee, S.-S., Kim, S., Jin, B., Choi, E., Kim, B., Jia, X., Kim, D., and Lee, K.-p. How Users Manipulate Deformable Displays as Input Devices. In *Proc. CHI '10*, ACM (2010), 1647–1656.
17. Li, Y., Chen, X., Tian, J., Zhang, X., Wang, K., and Yang, J. Automatic Recognition of Sign Language Subwords Based on Portable Accelerometer and EMG Sensors. In *Proc. ICMI-MLMI '10*, ACM (2010), 17:1–17:7.
18. Liang, R.-H., Lin, S.-Y., Su, C.-H., Cheng, K.-Y., Chen, B.-Y., and Yang, D.-N. SonarWatch: Appropriating the Forearm as a Slider Bar. In *Proc. SIGGRAPH Asia '11*, ACM (2011), 5:1–5:1.
19. Martin, T. L. Time and Time Again: Parallels in the Development of the Watch and the Wearable Computer. In *Proc. ISWC '02*, IEEE Computer Society (2002), 5–11.
20. McIntosh, J., McNeill, C., Fraser, M., Kerber, F., Löchtefeld, M., and Krüger, A. EMPress: Practical Hand Gesture Classification with Wrist-Mounted EMG and Pressure Sensing. In *Proc. CHI '16* (2016), 2332–2342.
21. Morris, M. R., Danielescu, A., Drucker, S., Fisher, D., Lee, B., schraefel, m. c., and Wobbrock, J. O. Reducing Legacy Bias in Gesture Elicitation Studies. *interactions 21*, 3 (2014), 40–45.
22. Nagar, A., and Zhu, X. Gesture Control by Wrist Surface Electromyography. In *Proc. PerCom Workshops '15* (2015), 556–561.
23. Narayanaswami, C., Kamijoh, N., Raghunath, M., Inoue, T., Cipolla, T., Sanford, J., Schlig, E., Venkiteswaran, S., Guniguntala, D., Kulkarni, V., and Yamazaki, K. IBM's Linux Watch: the Challenge of Miniaturization. *Computer 35*, 1 (2002), 33–41.
24. Ng, A., Williamson, J., and Brewster, S. The Effects of Encumbrance and Mobility on Touch-Based Gesture Interactions for Mobile Phones. In *Proc. MobileHCI '15*, ACM (2015), 536–546.
25. Pasquero, J., Stobbe, S. J., and Stonehouse, N. A Haptic Wristwatch for Eyes-Free Interactions. In *Proc. CHI '11*, ACM (2011), 3257–3266.
26. Perrault, S. T., Lecolinet, E., Eagan, J., and Guiard, Y. Watchit: Simple Gestures and Eyes-Free Interaction for Wristwatches and Bracelets. In *Proc. CHI '13*, ACM (2013), 1451–1460.
27. Profita, H. P., Clawson, J., Gilliland, S., Zeagler, C., Starner, T., Budd, J., and Do, E. Y.-L. Don't Mind Me Touching My Wrist: A Case Study of Interacting with On-Body Technology in Public. In *Proc. ISWC '13*, ACM (2013), 89–96.

28. Raghunath, M. T., and Narayanaswami, C. User Interfaces for Applications on a Wrist Watch. *Personal Ubiquitous Comput. 6*, 1 (Jan. 2002), 17–30.

29. Rekimoto, J. GestureWrist and GesturePad: Unobtrusive Wearable Interaction Devices. In *Proc. ISWC '01*, IEEE Computer Society (2001), 21–27.

30. Schuler, D., and Namioka, A., Eds. *Participatory Design: Principles and Practices*. L. Erlbaum Associates Inc., 1993.

31. Vatavu, R.-D., and Wobbrock, J. O. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In *Proc. CHI '15*, ACM (2015), 1325–1334.

32. Vatavu, R.-D., and Zaiti, I.-A. Leap Gestures for TV: Insights from an Elicitation Study. In *Proc. TVX '14*, ACM (2014), 131–138.

33. Ward, J. A., Lukowicz, P., and Tröster, G. Gesture Spotting Using Wrist Worn Microphone and 3-Axis Accelerometer. In *Proc. sOc-EUSAI '05*, ACM (2005), 99–104.

34. Weigel, M., Mehta, V., and Steimle, J. More Than Touch: Understanding How People Use Skin as an Input Surface for Mobile Computing. In *Proc. CHI '14*, ACM (2014), 179–188.

35. Wobbrock, J. O., Aung, H. H., Rothrock, B., and Myers, B. A. Maximizing the Guessability of Symbolic Input. In *Proc. CHI EA '05*, ACM (2005), 1869–1872.

36. Wobbrock, J. O., Morris, M. R., and Wilson, A. D. User-Defined Gestures for Surface Computing. In *Proc. CHI '09*, ACM (2009), 1083–1092.

37. Wu, M., Shen, C., Ryall, K., Forlines, C., and Balakrishnan, R. Gesture Registration, Relaxation, and Reuse for Multi-Point Direct-Touch Surfaces. In *Proc. TableTop Workshop '06*, IEEE (2006), 183–190.

38. Xiao, R., Laput, G., and Harrison, C. Expanding the Input Expressivity of Smartwatches with Mechanical Pan, Twist, Tilt and Click. In *Proc. CHI '14*, ACM (2014), 193–196.