User Elicitation on Single-hand Microgestures

Edwin Chan[1], Teddy Seyed[1], Wolfgang Stuerzlinger[2], Xing-Dong Yang[3], Frank Maurer[1]
University of Calgary[1], Simon Fraser University[2], Dartmouth College[3]
{chane, teddy.seyed, fmaurer}@ucalgary.ca, w.s@sfu.ca, xing-dong.yang@dartmouth.edu,

ABSTRACT
Gestural interaction has become increasingly popular, as enabling technologies continue to transition from research to retail. The mobility of miniaturized (and invisible) technologies introduces new uses for gesture recognition. This paper investigates single-hand microgestures (SHMGs), detailed gestures in a small interaction space. SHMGs are suitable for the mobile and discrete nature of interactions for ubiquitous computing. However, there has been a lack of end-user input in the design of such gestures. We performed a user-elicitation study with 16 participants to determine their preferred gestures for a set of referents. We contribute an analysis of 1,632 gestures, the resulting gesture set, and prevalent conceptual themes amongst the elicited gestures. These themes provide a set of guidelines for gesture designers, while informing the designs of future studies. With the increase in hand-tracking and electronic devices in our surroundings, we see this as a starting point for designing gestures suitable to portable ubiquitous computing.

Author Keywords
Gestures; gesture recognition; usability; touch; finger tracking; hand tracking

ACM Classification Keywords
H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

INTRODUCTION
In recent years, several techniques have been proposed to detect skin-based input [5, 7, 16, 23, 26], where touching an appendage to another part of the body serves as an input modality. There is a major advantage of skin-based input over traditional methods, since there is no (a-priori) need of any apparatus acting as a medium. Skin-based touch gestures such as tapping, pinching, and swiping, have been explored on several parts of the body. Most commonly, palms and forearms were used as touch surfaces given their relatively flat anatomy and perceived accessibility [7, 16, 27]. As gesture detection technologies become more miniaturized, gestures that are subtler and more discrete become possible. Given these developments, single-hand microgestures (SHMGs) are promising and in need of further study.

CHI'16, May 07-12, 2016, San Jose, CA, USA
© 2016 ACM. ISBN 978-1-4503-3362-7/16/05…$15.00
DOI: http://dx.doi.org/10.1145/2858036.2858589

Figure 1: Participants performing a variety of SHMGs.

A SHMG is unique relative to other touch inputs or microgestures, resulting in its moniker. Examples of SHMGs are seen in Figure 1. Traditional touch-input is performed by a user, often with her hands, on some sort of detection device, such as a digitizer or a camera-driven touch sensor. In comparison, a single-hand gesture is defined here not only as performed by a single hand, but also performed on that same hand. This is significant because it allows the gesture to be performed anytime and anywhere. In addition, the gesture can easily be performed secondarily with one hand while performing another task. In a previous elicitation study, users preferred single-hand gestures over bimanual ones; users were observed mirroring gestures on either hand to adapt to different contexts [20]. The microgesture designation suggests SHMGs are commonly performed, but rarely noticeable; SHMGs tend to be subtle yet informative. These features of SHMGs allow them to be performed naturally in public contexts where large gestures may be perceived as socially awkward [19].

Given their potential, SHMGs have been discussed as both primary and secondary topics by other researchers. Some of these discussions focus on enabling technologies such as body-mounted cameras [8] or sensors [9]. One study has elicited a gesture set from experts [30]. This work offers some insights to consider in designing gestures, yet the authors did not consult end-users directly. The result is an expert-based gesture set which differs from the user-elicited gesture set we discovered. We aim to address this lack of user input in the design and implementation of SHMGs. To better realize their potential, we are interested in going beyond technical compromises to fully understand human preferences. We accomplish this by employing Wobbrock *et al.'s* elicitation methodology [28], which has been used by numerous studies to explore user preferences and the nature of symbolic inputs. We can then compare and validate our results to the existing studies. One such study by Vatavu *et al.* elicited gestures similar to our work from end-users, but only in the context of TV interfaces using a LEAP Motion input device [26].

Our contributions begin with the classification of 1,680 elicited gestures, followed by the statistical analysis of the data using Vatavu *et al.'s* revised agreement rate [25]. We conclude with a set of design guidelines that offer qualitative insight into end-user thinking when designing SHMGs. The versatility of SHMGs make them suitable to many scenarios, and we see our work as a preliminary effort to designing better SHMGs for enhanced adoption and user experience.

RELATED WORK

A main concern with designing gestures is how well users resonate with such gesture sets, and whether a set is "easy to use." Many implementations of gesture recognition exist, and an earlier work lists 37 such implementations just for 3D gesture recognition [32]. Researchers continue to present improved technologies and new implementations, and implicitly make gestures easier to adopt and use by users. Many finger-based user interfaces involve user input performed on a passive or electronic device, such as a touch display or trackpad. These systems are ubiquitous in our surroundings now. However, a major drawback to using gestures on such devices is precisely the requirement of the device itself. Without the device, gestures cannot be used. The "surface-independent" Tickle interface [31] attempted to solve this, by permitting gestures to be made on any surface. This is an improvement, yet still requires the use of a foreign object for gesticulation. The Magic Finger similarly afforded touch input anytime and anywhere through a small device attached to a finger [33].

Device-free Hand Gestures

Device-free hand gestures require no other object to interact with, making them highly portable and literally as natural as using your hands. The earliest approaches often used some sort of Data Glove equipped with a whole set of sensors, which is well documented in Dipietro *et al.'s* survey of such systems [4]. However, wearing a glove is often inconvenient or inappropriate, which does not encourage adoption. Therefore, more recent approaches use alternative recognition methods. SixthSense [9] is worn like a necklace and makes use of a camera, projector, and fiduciary markers. Another example is the wrist-mounted camera approach used by Rekimoto *et al.* in GestureWrist [18], and later by Kim *et al.'s* Digits [6]. PinchWatch is a similar device and uses a chest-mounted camera system [8]. Saponas *et al.* presented an approach based on muscle-computer interfaces [21]. One of the latest innovations is the CyclopsRing [3], an input method for gestures including SHMGs, which achieved ~84.75% accuracy. This technology can detect most of the SHMGs we studied in this paper. All these alternatives were designed with the user in mind, but only looked at the gestures afforded by the system and not the usability of the technology. We sought to remedy this with the results of our study.

User Elicitation

Gesture interfaces are often designed without fully consulting end-users, or sacrifice usability for ease of implementation and practical reasons. As motivated by previous elicitation studies, designers and developers often do not share the same conceptual models as the end-users that should be catered to [15]. In many cases, end-users blend concepts from other systems that they have previous experience with [1]. These may include common household objects, phones, handheld controllers, etc. When comparing user-elicited and expert-elicited gesture sets, Wobbrock *et al.* discovered a user preference for user-elicited gestures. Gestures proposed by both users and experts were most preferred by users [12]. In addition, Nacenta *et al.* found that user-defined gestures are easier to remember [13]. To understand end-users, Nielsen *et al.* [15] proposed a procedure for eliciting and developing user-defined gestures. User elicitations using this procedure have offered contributions towards the design of the studied gestures and the overall user design process, despite eliciting unique types of gestures. Wobbrock *et al.* provided an agreement measure [28] to analyze and interpret elicited data, which has been widely adopted by prior elicitation studies [1, 10, 23, 26, 29, 30]. Vatavu and Wobbrock later refined this measure [25] to more accurately represent findings. Morris *et al.* [11] examined the issue of legacy bias, offering methods to reduce its effects.

Wobbrock *et al.* [29] discussed several intriguing concepts including dichotomous references, reversible gestures, and simplified mental models. Seyed *et al.* [23] noted the importance of aliasing gestures as a solution to varied user preferences, while offering atomic gestures and themes to help map gestures to users' conceptual models. Angelini *et al.* [1] looked at gestures performed on a steering wheel. Albeit different from our own study, their work identified similar metrics, such as body parts used in gesticulation, and the frequency of gesture actions, such as swipe or tap.

A recent work proposed a taxonomy of microinteractions [30], defining microgestures based on ergonomic and scenario-dependent requirements. While the premise of investigating gestures performed in relation to hand grips (due to holding objects or devices) differed from our device-free gestures, their study produced a framework for some of our findings. Indeed, their consultation with four experts of hand anatomy helped to define the physical traits that limit SHMGs. One participant was a sports therapist, while the remaining three were physiotherapists. We will relate to their findings in our discussions of physical limitations and discomforts encountered while performing gestures.

Classifying Gestures

We based our classification of elicited gestures on the *Descriptive Labeling* proposed by Nielsen [15], with gestures being recorded by their actions rather than their semantic meaning. Gestures were *chunked* and *phrased* according to Buxton's work [2], with phrases delimited by periods of tensions and relaxation. Given that each referent shown to users is a task, *i.e.,* an action, it made sense to also describe the gestures for these tasks as distinct actions. These distinct actions, termed atomic gestures by Buxton, are sometimes combined by users to create compound gestures. We employ *descriptive labeling*, *chunking*, and *phrasing* to present our data for interpretation. An example of a compound gesture is shown here:

TAP on *INDEX* with *THUMB*, then
SWIPE on *INDEX* with *THUMB*

Figure 2: A proposed gesture for the referent "Zoom in".

USER STUDY

An elicitation study was conducted to identify user preferences for SHMGs.

Participants

Sixteen paid volunteers participated in the study (7 male, 9 female). Participants were recruited using email lists and word of mouth. The participants ranged in age from 16 to 39 years (Mean = 22, SD = 4.97), and came from differing backgrounds including marketing, arts, psychology, and high school students. Of the 16 participants, 4 reported having experience with microgesture devices such as the MYO armband or LEAP Motion sensor. All participants had some experience with touch gestures, along with frequent use of devices such as smartphones, computers, or gaming consoles.

Apparatus

Since SHMGs are defined as device-free, users did not interact directly with any device. Before starting, participants were shown sample SHMGs on a laptop computer to illustrate the potentials and limitations of SHMGs. For the elicitation, referents were listed on a printout, with each referent being demonstrated on the laptop computer.

Video recording was done with a 1080p webcam mounted on a tripod, and users were able to see the live recording so they could keep their hands in view. The recording captured each of the user's gestures, as well as communication between the user and author.

Referents

We wanted to form a list of common tasks which users could relate to and which they may perform frequently. To do so, we looked at Wobbrock *et al.'s* list of referents [29] as a starting point. We then added several actions that are commonly performed on devices such as phones or computers (eg. "Play", "Next", "Copy"). The gestures were grouped into the six categories used in Piumsomboon *et al.*'s elicitation study for augmented reality [17]. Table 1 lists the final set of referents that we used in the study.

Procedure

At the start of each session, participants were asked to fill out a short survey regarding prior experience with related devices. Participants were then informed of the purpose of the study, before being *primed* [11] with a short introduction to SHMGs. This included defining SHMGs as gestures performed on the surface of the hand, from the

Category	Tasks	Category	Tasks
Transform	1. Move	Editing	19. Cut
	2. Rotate		20. Copy
	3. Enlarge		21. Paste
	4. Shrink		22. Delete
	5. Minimize		23. Accept
	6. Maximize		24. Reject
Simulation	7. Volume up		25. Undo
	8. Volume down		26. Save
	9. Mute	Menu	27. Help
	10. Play		28. Open menu
	11. Pause		29. Close Menu
	12. Stop		30. On
Browsing	13. Pan		31. Off
	14. Zoom in	Selection	32. Select single
	15. Zoom out		33. Select group
	16. Scroll		34. Find
	17. Next		
	18. Previous		

Table 1: The list of 34 referents used in the study.

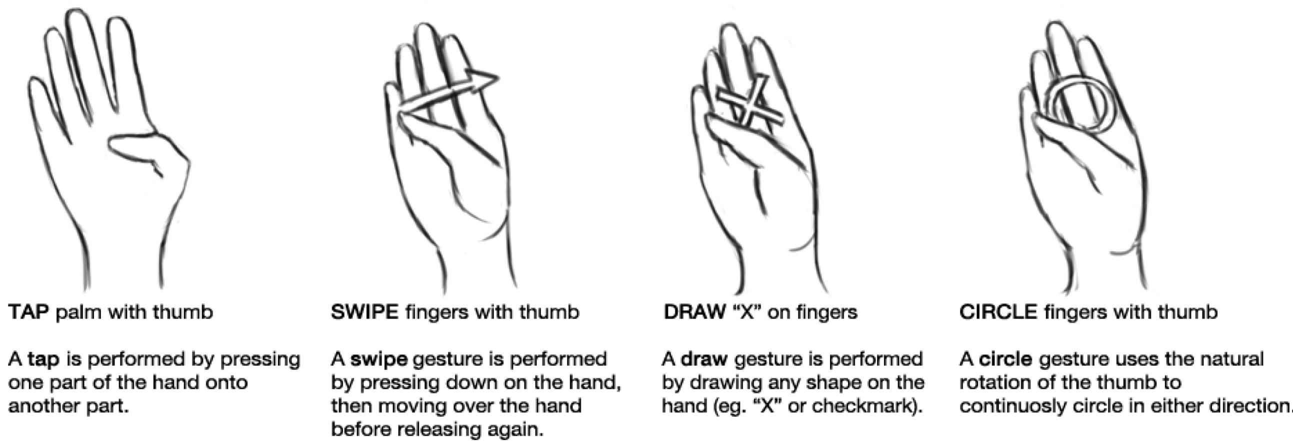

Figure 3: The four types of actions are illustrated and defined.

wrist to the fingertips, using only the fingers of the same hand, without interaction with other objects or devices. Participants were allowed to use either hand. Several types of actions, *e.g.,* tap or swipe, were shown to the participant, with variations of each explained to give them a better understanding of what gestures might be considered unique. Examples of such variations are shown in Figure 3. Participants were encouraged to design gestures based on preference without concern for implementation feasibility. However, participants were undoubtedly affected by previous associations with other input methods or implementations; this is further discussed in the results and analysis sections. We also specified gestures could be reused for different tasks if it made sense to participants.

Participants were presented with a list of 34 referents and asked to design three gestures for each task, before identifying their preferred gesture for the task. Referents were always presented in the same order to participants. By requiring three gestures, we apply the *production* technique to reduce legacy bias [11]. Each referent was demonstrated on the laptop computer, and participants were allowed to ask for clarification if required (for example, about the difference between "Move" and "Pan"). Sometimes, participants were asked to explain their choices for greater understanding of their thought process. After they completed each of the referents, we performed a semi-structured interview to elicit feedback about their experience, including potential use cases and difficulties encountered. Participants were generally enthusiastic to provide their opinion, which was encouraging for both the study and use of SHMGs.

RESULTS

From our 16 participants, we collected a total of 1,632 gestures (16 participants x 34 referents x 3 gestures). These gestures were classified with the aforementioned methodology; this process is explained in detail in the following section. From the resulting set of gestures, we calculate agreement rates between participants and interpret them. A consensus set, as defined by Wobbrock *et al.* [29], is presented for SHMGs.

Classification of Gestures

In previous work by Wobbrock *et al.* [29], their consensus set of gestures is constructed from the most common gesture elicited for each referent. In our study we followed a similar approach, but accounted for some interesting factors, which affected our results. Our approach resembles that of Piumsomboon *et al.* [17], since we group gestures that are similar rather than identical.

The primary distinction of gestures used is the type of action performed in the gesture. We were able to categorize all the gestures into four actions: *Tap, Swipe, Circle,* and *Draw.* Definitions and examples of each action are illustrated in Figure 3.

During the study, participants were asked to pick their preferred gesture after coming up with three unique gestures. In many cases, the participant would remember the action performed, but mix up the exact finger(s) used. This observation is consistent with another study, in which users expressed little concern about how many fingers were used in a gesture [30]. The confusion was also seen when comparing hand poses, where the fingers not used in the gesture would be bent in one variation but not the other. When reviewing the recordings, we were surprised by how often this happened. To account for this confusion in recalling gestures, we separated gestures that used two or less fingers from those with three or more fingers. This is less restrictive than matching the exact finger(s) used in each gesture, and seems to better represent the thought behind gestures. Some participants commented on using one or two fingers for more precise actions (such as "Select Single") while using three or more figures for tasks that seemed to need more space (such as "Select Multiple" or "Move").

From the original 1,632 gestures, we isolated the 544 preferred gestures (140 unique gestures). By following the

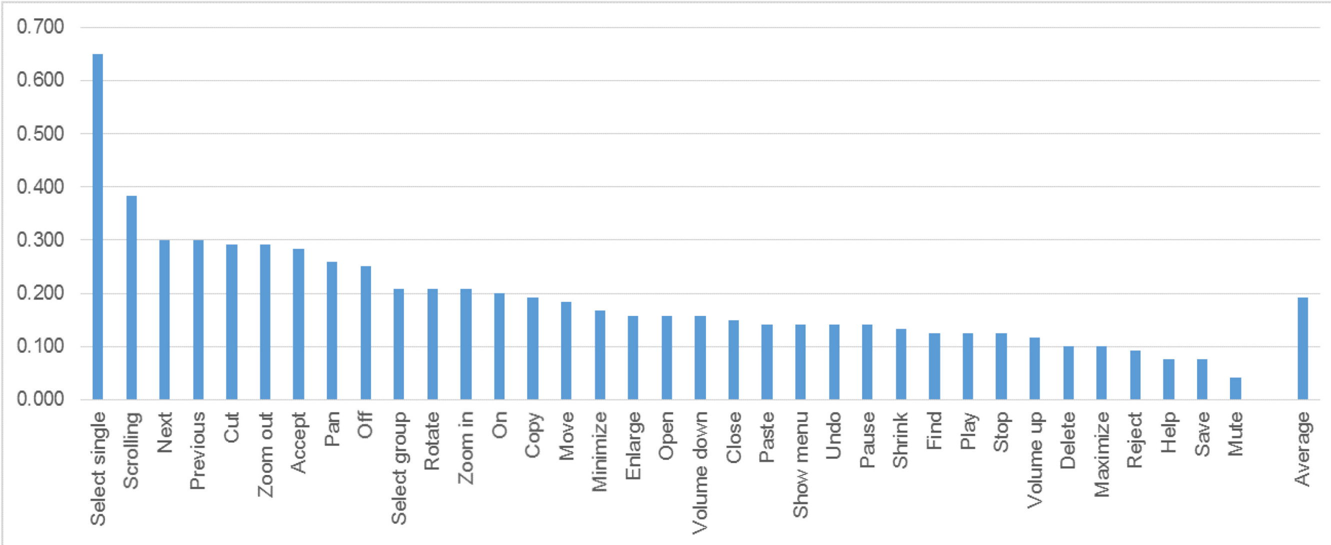

Figure 4: Agreement rates for the 34 referents.

approach defined above, we were able to reduce the set to 47 unique gestures. By taking the maximum consensus gesture for each referent in this set, we were left with 8 unique gestures, which represented 220/544 gestures or 40.4% of the entire set.

Agreement Between Participants

Previous studies [10, 17, 23, 29] used Wobbrock *et al.*'s Agreement Rate formula [28], which did not accurately represent gestures with no agreement. Gestures that had zero agreement trivially agreed with themselves. Therefore, gestures with zero agreement actually did not have an agreement rate of 0. This formula also did not account for the degrees of freedom; a gesture with 15/20 matching entries had the same agreement rate as a gesture with 30/40 matching entries, despite the latter clearly showing greater agreement for the consensus gesture [25].

A new agreement rate formula was proposed by Vatavu *et al*. [25], which accounts for the missing factors in the old formula. We measured agreement between participants using this new formula and the accompanying AGATe (AGreement Analysis Toolkit) software. The revised agreement formula is defined in Equation 1:

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

where "P is the set of all proposals for referent r, $|P|$ the size of the set, and P_i subsets of identical proposals from P" [25].

Agreement rates ranged from 0.042 (low agreement, $AR \leq 0.100$) to 0.650 (very high agreement, $AR > 0.500$). The mean AR was 0.191 (medium agreement, $0.100 < AR < 0.300$). The agreement rates of all referents are shown in Figure 4. Since the new formula calculated AR less optimistically, Vatavu *et al.* recalculated the AR of 18 previous studies [25]. In these studies, the average sample size was 19.1, while mean AR was 0.221.

Along with a new formula for agreement rate, Wobbrock *et al.* also introduced the Coagreement Rate, which looks at "how much agreement is shared between two referents r_1 and r_2." This is interesting because we can observe patterns previously left unnoticed. Existing work had already shown a significant relationship between gestures of dichotomous pairs [17, 23, 29] such as "Next"/"Previous" or "Zoom In"/"Zoom Out". Most of the pairs described were directional, where consensus gestures opposed each other directionally *(e.g.*, swipe left/right). Less focus has been placed on toggles such as "On"/"Off" or "Play"/"Pause". With the new Coagreement Rate, we not only found that "On"/"Off" (as well as "Play"/"Pause") have the same consensus gesture, but that participants who picked one gesture in r_1 often picked the same gesture in r_2. We know this since the AR for r_1 and r_2 are close to $CR(r_1,r_2)$. For example, $AR(\text{On}) = 0.200$, $AR(\text{Off}) = 0.250$, and $CR(\text{On,Off}) = 0.197$. This is different from only knowing that the same number of participants picked the consensus gesture in both referents, and suggests referents of this type should use the same gesture.

Consensus Gesture Set

As mentioned earlier, the original gesture set was reduced to 8 unique gestures. This set is rather small and even within each of the six categories of referents, there were conflicts where one gesture was preferred for several referents. This was an expected outcome, since we classified five fingers with only two types: two fingers or less, and three fingers or more. To resolve the conflicts, we looked at each instance of the consensus gesture for each referent and identified which fingers were used most. The idea behind this resolution comes from observing the participants. While participants often mixed up the exact finger they suggested for a gesture, there was a recurring theme of choosing similar gestures for seemingly related tasks. Several participants exhibited this pattern when choosing gestures for "Cut", "Copy", and "Paste", as well as "Accept" and "Reject". We observed a strong preference for keeping these gestures "close to each other" or "next to each other".

Sometimes participants arbitrarily chose different fingers for a similar gesture (such as tapping any finger and the thumb together), when they had difficulty coming up with three meaningful gestures. We tried to reduce this source of randomness by taking the most used finger(s) for each consensus gesture. Very interestingly, assigning fingers with this procedure resolved all but one conflict in the consensus gesture set.

The only remaining conflict was between "Stop", "On", and "Off". Since the top two preferred gestures for each of these referents were the same (make a fist, or tap the index/middle/ring ringers on the palm), we included both gestures for all three referents. We suggest using the same gesture for both "On" and "Off", as we previously mentioned a significant Coagreement Rate between the two. The resulting consensus set of 16 gestures representing 35 referents in 6 categories is shown in Figure 5.

Figure 5: Consensus Gesture Set

Actions

To better understand the distribution and makeup of the gestures elicited, recall our classification method which separates gestures by actions, based on Bill Buxton's work on *Chunking* and *Phrasing* [2]. When we examined the actions chosen for consensus gestures, we discovered several motifs.

Of the four action types, *Taps* were the most common (19 of 34 referents). During the think-aloud sessions, users offered some potential reasons for picking *Taps*. *Taps* were popular amongst users because of their ease with which they can be performed and their conceptual simplicity, making them easy to reproduce. Many *Tap* gestures were also preferred due to their resemblance to interaction with other devices, such as mice, trackpads, gaming devices, or remote controllers. This is apparent in the *Selection* category, where all three consensus gestures used *Taps*. A *Tap* gesture provided the precision desired when selecting a specific set of objects.

Swipes (14 of 34 referents) were frequently used when the task involved picking a value inside a continuous range, such as turning the volume up or down. In many cases they reminded users of the fluid action of sliders or radial dials. *Swipes* were also often used for tasks that were directional, such as moving something or scrolling in any direction. Of the six referents in the *Transforms* category, five made use of *Swipes.* The "Rotate" task used a *Circle* action, which was likely chosen due to the circular motion associated with rotation.

The *Draw* action appeared with six of the participants, but did not make it into the consensus set. Although drawing a question mark for "Help" or drawing an 'X' for "Close" seemed more intuitive and easier to recall, participants only resorted to the *Draw* action when experiencing difficulty devising three gestures.

Compounds gestures made up 12% of all gestures elicited, and were preferred for approximately 10% of tasks. These were mostly used for tasks that users instinctively split into smaller modules. For example, when asked to select a group of items, a participant said, "I swipe across my fingers like I am choosing the items, then I tap on my fingers to select them." In another example where a participant was asked to perform the *Save* task, the participant responded, "I have something here, then I want to make a copy here to save it."

In Figure 6, we present the distribution of action types in the preferred gestures set. The preferred gestures set represents the preferred gestures of every participant, rather than just the gestures in the consensus set. By examining the graph, we can easily tell which actions were preferred for specific gestures. For example, we can tell that *Swipes* were preferred for dichotomous pairs, which are discussed in more detail later.

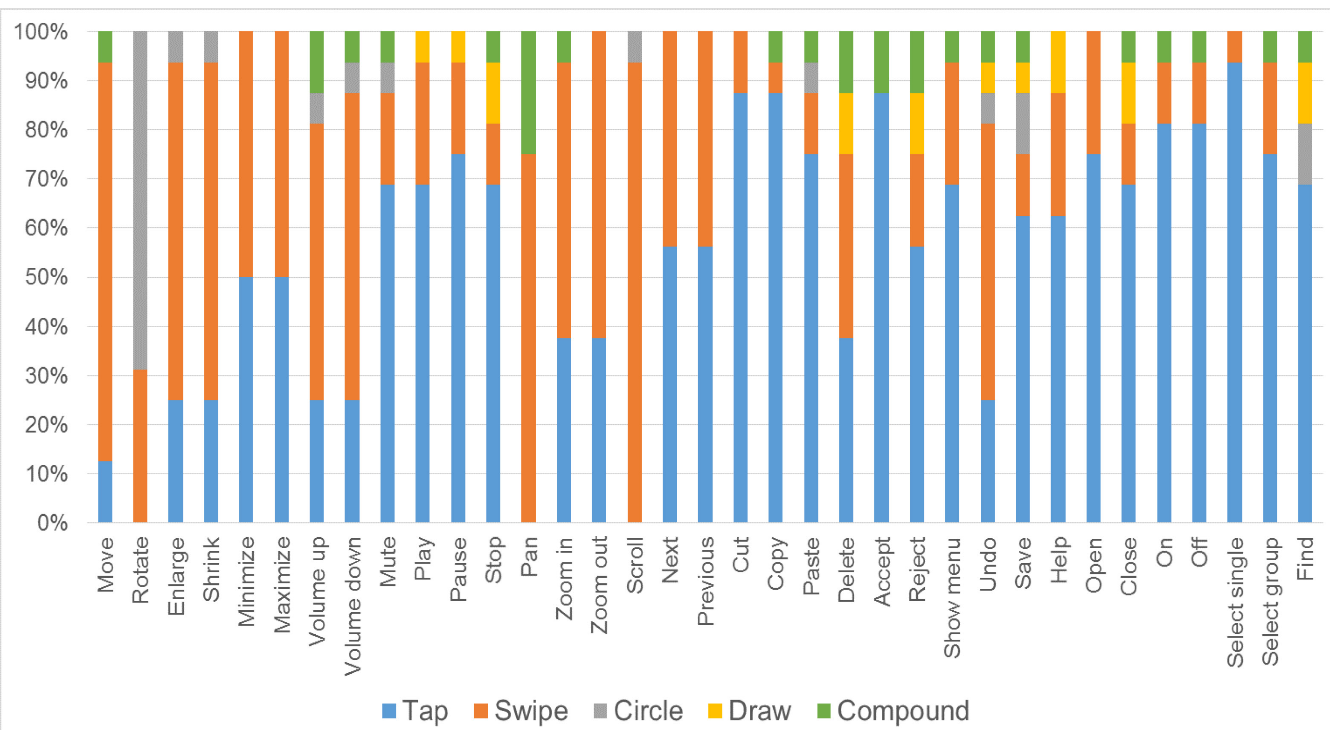

Figure 6: Distribution of action types in the preferred gesture set.

Actors

Given the physical constraints of SHMGs where gestures are performed using only a single hand, it made sense that all gestures were performed using one or more finger(s). Knowing which fingers were most common in our data helps us to quantitatively assert which actors are most suitable for SHMGs. We can then combine qualitative observations from the study with insights from existing work to suggest reasons for some of the actors standing out as most commonly used by participants. The frequencies of each finger appearing in the gestures elicited can be seen in Figure 7.

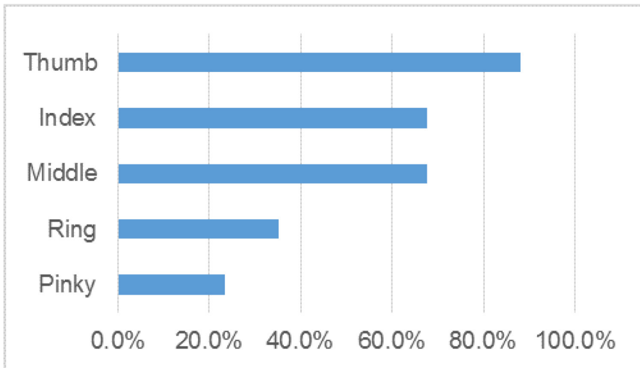

Figure 7: Frequency rates of each finger in consensus set. Sum > 100%, as multiple fingers may be used in a single gesture.

Unsurprisingly, the thumb was involved in 88% of all gestures. As explained by existing work relating hand anatomy and gestures [30], our hands are opposable through the use of our thumbs. Because of this special trait of thumbs, as well as its unique ability to rotate, the thumb can easily touch other parts of the hand, which by definition of SHMGs constitutes a gesture. Whereas other fingers have difficulty interacting with their neighbors, thumbs can touch most areas of the other fingers quite naturally. Capable of rotating, the thumb is often used for controls that involve rotation or multiple axes. For example, many gaming controls use the thumb for the D-Pad or joystick, while (two-dimensional) phone screens are often interacted with the thumb. Similarly, all the elicited *Swipe* gestures were performed with the thumb.

The preference of the index and middle fingers, when compared to the ring and pinky fingers, can also be explained by Wolf *et al.'s* summary of the anatomy of the hand [30]. Due to biomechanics and more specifically the muscles involved in moving each finger, the index finger is most suited to independent movement, followed by the middle finger. The ring finger is considered to be the least feasible, because "two muscles (*M. flexor digitorum profundus & M. flexor digitorum superficialis*) are bending synergistically the index, middle, and little finger to bring them into the palm position. In addition another muscle is responsible for stretching the ring finger (*M. extensor digitorum*), but because this muscle is also responsible for stretching the other fingers and because the ring finger has a physical connection to the middle finger (*Connexus intertendineus*), the middle finger will always move a bit in the same direction as the ring finger does."

While the pinky finger is also able to move independently like the index finger, it was seldom used in the consensus set (2 of 34 referents). Possible explanations include the greater distance between the thumb and pinky finger, as well as the reduced strength of the pinky finger compared to the index finger. Some participants avoided using the pinky finger due to potential discomfort and fatigue.

IMPLICATIONS FOR DESIGN

Combining the results above as well as the results of the interviews conducted with users at the end of each study session, we derived several guidelines for the design of SHMGs.

Previous Experience

While the agreement rate was comparable to existing studies, we believe previous experience of participants strongly affected our results. This motif has been previously documented [23, 26], and generally led to greater agreement amongst participants. However, we found that the previous experience of our participants both positively and negatively affected agreement rates. An example where it contributed positively to agreement is the "Cut" referent, which users easily associated the task with a common symbol for scissors (tapping the index and middle fingers together). In another example where previous experience may have negatively influenced agreement rates, the proposed gestures for "Mute" included using the sign language representation of the letter "m" and also simulating the action of reaching towards the back of a handheld gaming console to reduce volume. While there are physical representations for "Mute", such as a clenched fist in music performances (a gesture that participants had little prior experience with), users drew on a large variety of other previous experiences for such actions. Regardless of whether previous experiences affected agreement rates positively or negatively, the impact of these experiences was apparent in the behavior of participants.

As documented by Nebeling *et al.*, we also noticed a trend where referents which related to physical actions (such as "Cut") resulted in greater recall and agreement when metaphors were used [14]. This observation suggests that gesture designers must consider the nature of each referent, the existing metaphors, and whether these metaphors are commonly used by the expected users of the system. For referents that do not benefit from the use of metaphors, abstract gestures are more suitable as indicated by the numerous cases when users recalled specific details incorrectly.

Fingers and Postures: Their Meanings

Another topic that surfaced in other elicitation studies is the cultural meaning of various hand postures and gestures.

While symbolic hand gestures have already been discussed [23, 26], *e.g.*, "Help" with a beckoning gesture or "Mute" with a clenched fist, we found that users often chose specific fingers as well for a variety of reasons. Besides using the index finger for its dexterity or convenience, users frequently referred to the index finger as the pointer finger, which evoked a feeling of confidence or direction. A particularly interesting case is "Help", where one user used the pinky finger because "pinky is the weaker one, so you need more help."

SHMGs can be discrete and subtle, but we expect these gestures to be performed in both private and public spaces. As such, certain gestures may be less suitable than others and may need to be substituted for specific user groups.

Dichotomous Pairs and State Toggles

Another reason for choosing specific fingers was the motif of dichotomous pairings, and in some cases groupings of three or more gestures. As previously mentioned, dichotomous pairs often resulted in opposing gestures, such as swiping left to symbolize previous and swiping right to symbolize next. In Figure 6, *Swipes* are shown to be preferred for "Enlarge"/"Shrink", "Minimize"/"Maximize", "Volume up"/"Volume down", and "Zoom in"/"Zoom out". As *Swipes* were heavily preferred for dichotomous pairs in the consensus set as well, we again make the recommendation to *Swipes* for these gestures. We also recommend using identical gestures for toggles, such as "On"/"Off". Identical gestures are more suited to toggles than opposing gestures, as we identified a unique problem with hand gestures when applying certain gestures for toggles. A good example is when some users suggested closing their fist to turn the system "On", while releasing their fist to turn the system "Off". Although the gestures are unique, the hand naturally returns to a relaxed state after tension, relating to Buxton's delimitation of atomic gestures through periods of tension and relaxation [2]. As such, performing the "On" gesture results in the "Off" gesture also being performed. This difficulty was encountered for other pairs such as "Enlarge"/"Shrink" or "Volume Up"/"Volume Down", forcing users to choose other gestures.

Level of Detail

Given the variety of Actors and Actions, there are technically hundreds of possible SHMGs. However, while some users went as far as using different joints to differentiate gestures, most users settled for less detail in their gestures. Many users even complained about the lack of gestures available, as one participant described: "It's very limited, (the) amount of things you can do with one hand and touch." The difficulties participants experienced in recalling gestures in detail prompted the classification method used in our study.

One participant worried "some people would be limited in the number of hand gestures they would have based on hand mobility." This was the case for another participant who could not form a clenched fist. The dexterity of users could influence their preference of gestures.

Finally, select users were aware of variables for creating gestures but opted not to use them, as is the case when one participant used double taps instead of holds (long duration tap). The participant preferred double tapping, which felt more reassuring to them than holding a gesture for a specific duration.

Additional Variables

Due to the perceived limitation of gesture variety, users reported two interesting variables that they could potentially control in addition to the suggestions we made. First, they suggested varying the speed at which a gesture is performed. Performing a gesture slowly was perceived to offer finer adjustment, such as when performing the "Enlarge" or "Shrink" tasks. The second variable used varying forces while performing gestures such as closing a hand harder to perform "Stop" instead of "Pause". These variables may enable a larger vocabulary of natural gestures, provided the speed and force can be detected reliably. These user suggestions were made during an interview at the end of the study, and no such gestures were chosen by participants in the elicitation part.

LIMITATIONS

Here we discuss several limitations and potential extensions of our study.

Spatial Extensibility

As seen with the additional variables proposed by users, our current definition of SHMGs may not fully match the mental models of all users. This is after all the specific reason that we consulted users in our elicitation study. Participants were asked to comment on the feasibility of SHMGs as well as the study itself, and all participants commented on not being able to use mid-air spatial gestures. For example, participants asked if they could perform "Move" by tapping the thumb and index fingers together, before moving the whole hand in mid-air.

Although we defined SHMGs as gestures performed on the hand from the wrist to the fingertips, many users would have liked the option of using spatial tracking of the arm itself as well. While larger arm movements may not be suitable for discrete microgestures in public spaces, users frequently proposed small movements or rotations of the arm. This happened despite users being informed during *priming* that such spatial gestures did not fit our criteria, suggesting the desire and possible need for spatial recognition.

Elicitation Methodology

As documented by existing literature [23, 26], legacy bias may have a significant effect on results in an elicitation study. Although we applied *priming* and *production* to offset legacy bias, we encountered the same problems mentioned by Morris *et al.* [11]. That is, there is no way to determine the optimal amount of gestures each user should propose for each referent. In 55% of all cases, users did indeed choose their second or third gestures as their preferred gesture. When later asked why they did not propose gestures that seemed obvious to the researcher, users often replied, “I didn’t even think of that!” However, other users benefited less from production: “I already have a gesture in mind, so thinking of three different ones makes me start grasping for straws, because I already have a solid idea of what I would do.”

Pairing

Pairing was also proposed by Morris *et al.* [11] as another way to reduce legacy bias. With previous user experience and legacy bias having both positive and negative effects on agreement rates, *pairing* may be useful as a means to generate more optimal gestures. In the situation where a single user might run out of ideas and therefore offer arbitrary gestures as their second or third choice, having a partner may help foster additional ideas. When users pick gestures based on personal and unique experiences, a partner would be able to question the generalizability of such a gesture in a consensus set.

FUTURE WORK

To address the above-mentioned limitations, it may be interesting to perform variations of this study to note how additional variables proposed by users would affect the resulting gestures and agreement rates. While we do not expect significantly higher agreement rates when introducing greater variation, the newly available gestures may be more natural for users. Such gestures could improve recall and therefore be more preferable to users.

Specific Domains

It would also by worthy to investigate user preferences in more specific domains suited to SHMGs, such as while in public transit or while performing a primary task. While the inherent nature of SHMGs makes them less susceptible to factors which create social awkwardness [19], developing generic principles that apply universally to all contexts remains a challenge [15]. Further context-specific studies may reveal subtle factors specific to SHMGs that affect the gestures preferred by users.

CONCLUSION

We recognized the potential of single-hand microgestures (SHMGs) in ubiquitous computing amidst current technological developments. To further inform the design of SHMGs we conducted an elicitation study with end users, where we recorded a set of 1,680 gestures. We presented our findings including agreement rates, frequency statistics, and qualitative observations. Based on this we discussed several implications for the design of SHMGs. Our observations can serve both a guideline to future designers of SHMGs, as well as a reference for further studies.

REFERENCES

1. Leonardo Angelini, Francesco Carrino, Stefano Carrino, Maurizio Caon, Omar Abou Khaled, Jürgen Baumgartner, Andreas Sonderegger, Denis Lalanne, and Elena Mugellini. 2014. Gesturing on the Steering Wheel: a User-elicited taxonomy. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (AutomotiveUI '14). ACM, New York, NY, USA, Article 31, 8 pages. http://doi.acm.org/10.1145/2667317.2667414
2. William A. S. Buxton. 1995. Chunking and phrasing and the design of human-computer dialogues. In *Human-Computer Interaction*, Ronald M. Baecker, Jonathan Grudin, William A. S. Buxton, and Saul Greenberg (Eds.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA 494-499.
3. Liwei Chan, Yi-Ling Chen, Chi-Hao Hsieh, Rong-Hao Liang, and Bing-Yu Chen. 2015. CyclopsRing: Enabling Whole-Hand and Context-Aware Interactions Through a Fisheye Ring. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (UIST '15). ACM, New York, NY, USA, 549-556. http://dx.doi.org/10.1145/2807442.2807450
4. L. Dipietro, A. M. Sabatini, and P. Dario. 2008. A Survey of Glove-Based Systems and Their Applications. *Trans. Sys. Man Cyber Part C* 38, 4 (July 2008), 461-482. http://dx.doi.org/10.1109/TSMCC.2008.923862
5. Chris Harrison, Desney Tan, and Dan Morris. 2010. Skinput: appropriating the body as an input surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '10). ACM, New York, NY, USA, 453-462. http://doi.acm.org/10.1145/1753326.1753394
6. David Kim, Otmar Hilliges, Shahram Izadi, Alex D. Butler, Jiawen Chen, Iason Oikonomidis, and Patrick Olivier. 2012. 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). ACM, New York, NY, USA, 167-176. http://doi.acm.org/10.1145/2380116.2380139
7. Rong-Hao Liang, Shu-Yang Lin, Chao-Huai Su, Kai-Yin Cheng, Bing-Yu Chen, and De-Nian Yang. 2011. SonarWatch: appropriating the forearm as a slider bar. In *SIGGRAPH Asia 2011 Emerging Technologies* (SA '11). ACM, New York, NY, USA, Article 5, 1 pages. http://doi.acm.org/10.1145/2073370.2073374
8. Christian Loclair, Sean Gustafson, and Patrick Baudisch. PinchWatch: a wearable device for one-

handed microinteractions. In *Proceedings of MobileHCI*, vol. 10. 2010.

9. Pranav Mistry and Pattie Maes. 2009. SixthSense: a wearable gestural interface. In *ACM SIGGRAPH ASIA 2009 Sketches* (SIGGRAPH ASIA '09). ACM, New York, NY, USA, Article 11, 1 pages. http://doi.acm.org/10.1145/1667146.1667160

10. Meredith Ringel Morris. 2012. Web on the wall: insights from a multimodal interaction elicitation study. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces* (ITS '12). ACM, New York, NY, USA, 95-104. http://doi.acm.org/10.1145/2396636.2396651

11. Meredith Ringel Morris, Andreea Danielescu, Steven Drucker, Danyel Fisher, Bongshin Lee, m. c. schraefel, and Jacob O. Wobbrock. 2014. Reducing legacy bias in gesture elicitation studies. *interactions* 21, 3 (May 2014), 40-45. http://doi.acm.org/10.1145/2591689

12. Meredith Ringel Morris, Jacob O. Wobbrock, and Andrew D. Wilson. 2010. Understanding users' preferences for surface gestures. In *Proceedings of Graphics Interface 2010* (GI '10). Canadian Information Processing Society, Toronto, Ont., Canada, Canada, 261-268.

13. Miguel A. Nacenta, Yemliha Kamber, Yizhou Qiang, and Per Ola Kristensson. 2013. Memorability of pre-designed and user-defined gesture sets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '13). ACM, New York, NY, USA, 1099-1108. http://dx.doi.org/10.1145/2470654.2466142

14. Michael Nebeling, Alexander Huber, David Ott, and Moira C. Norrie. 2014. Web on the Wall Reloaded: Implementation, Replication and Refinement of User-Defined Interaction Sets. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces* (ITS '14). ACM, New York, NY, USA, 15-24. http://dx.doi.org/10.1145/2669485.2669497

15. Michael Nielsen, Moritz Störring, Thomas B. Moeslund, and Erik Granum. 2004. A procedure for developing intuitive and ergonomic gesture interfaces for HCI. In *Gesture-Based Communication in Human-Computer Interaction*. Springer Berlin Heidelberg, 409-420.

16. Ryosuke Ono, Shunsuke Yoshimoto, and Kosuke Sato. 2013. Palm+Act: operation by visually captured 3D force on palm. In *SIGGRAPH Asia 2013 Emerging Technologies* (SA '13). ACM, New York, NY, USA, Article 14, 3 pages. http://doi.acm.org/10.1145/2542284.2542298

17. Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn. 2013. User-defined gestures for augmented reality. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '13). ACM, New York, NY, USA, 955-960. http://doi.acm.org/10.1145/2468356.2468527

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

19. Julie Rico and Stephen Brewster. 2010. Usable gestures for mobile interfaces: evaluating social acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '10). ACM, New York, NY, USA, 887-896. http://dx.doi.org/10.1145/1753326.1753458

20. Gustavo Alberto Rovelo Ruiz, Davy Vanacken, Kris Luyten, Francisco Abad, and Emilio Camahort. 2014. Multi-viewer gesture-based interaction for omni-directional video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14). ACM, New York, NY, USA, 4077-4086. http://dx.doi.org/10.1145/2556288.2557113

21. T. Scott Saponas, Desney S. Tan, Dan Morris, Ravin Balakrishnan, Jim Turner, and James A. Landay. 2009. Enabling always-available input with muscle-computer interfaces. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology* (UIST '09). ACM, New York, NY, USA, 167-176. http://doi.acm.org/10.1145/1622176.1622208

22. Richard A. Schmidt and Timothy D. Lee. 2011. *Motor Control and Learning: A Behavioral Emphasis. (5th. ed.)* Human Kinetics, Champaign, Illions, 2.

23. Teddy Seyed, Chris Burns, Mario Costa Sousa, Frank Maurer, and Anthony Tang. 2012. Eliciting usable gestures for multi-display environments. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces* (ITS '12). ACM, New York, NY, USA, 41-50. http://doi.acm.org/10.1145/2396636.2396643

24. Kentaro Takemura, Akihiro Ito, Jun Takamatsu, and Tsukasa Ogasawara. 2011. Active bone-conducted sound sensing for wearable interfaces. In *Proceedings of the 24th Annual ACM Symposium Adjunct on User Interface Software and Technology* (UIST '11 Adjunct). ACM, New York, NY, USA, 53-54. http://doi.acm.org/10.1145/2046396.2046419

25. Radu-Daniel Vatavu and Jacob O. Wobbrock. 2015. 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). ACM, New York, NY, USA, 1325-1334. http://doi.acm.org/10.1145/2702123.2702223

26. Radu-Daniel Vatavu and Ionut-Alexandru Zaiti. 2014. 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). ACM, New York, NY, USA, 131-138. http://doi.acm.org/10.1145/2602299.2602316

27. Martin Weigel, Vikram Mehta, and Jürgen Steimle. 2014. More than touch: understanding how people use skin as an input surface for mobile computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14). ACM, New York, NY, USA, 179-188. http://doi.acm.org/10.1145/2556288.2557239
28. Jacob O. Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A. Myers. 2005. Maximizing the guessability of symbolic input. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '05). ACM, New York, NY, USA, 1869-1872. http://doi.acm.org/10.1145/1056808.1057043
29. Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '09). ACM, New York, NY, USA, 1083-1092. http://doi.acm.org/10.1145/1518701.1518866
30. Katrin Wolf, Anja Naumann, Michael Rohs, and Jörg Müller. 2011. Taxonomy of microinteractions: defining microgestures based on ergonomic and scenario-dependent requirements. In *Proceedings of the 13th IFIP TC 13 International Conference on Human-Computer Interaction - Volume Part I* (INTERACT'11), Pedro Campos, Nuno Nunes, Nicholas Graham, Joaquim Jorge, and Philippe Palanque (Eds.), Vol. Part I. Springer-Verlag, Berlin, Heidelberg, 559-575.
31. Katrin Wolf, Robert Schleicher, Sven Kratz, and Michael Rohs. 2013. Tickle: a surface-independent interaction technique for grasp interfaces. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction* (TEI '13). ACM, New York, NY, USA, 185-192. http://doi.acm.org/10.1145/2460625.2460654
32. Ying Wu and Thomas S. Huang. 1999. Vision-based gesture recognition: A review. In *Gesture-based Communication in Human-Computer Interaction*. Springer Berlin Heidelberg, 103-115.
33. Xing-Dong Yang, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2012. Magic finger: always-available input through finger instrumentation. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology* (UIST '12). ACM, New York, NY, USA, 147-156. http://doi.acm.org/10.1145/2380116.2380137