Eliciting User-defined Zenithal Gestures for Privacy Preferences

Francisco J. Martínez-Ruiz[1] [a] and Santiago Villarreal-Narvaez[2] [b]

[1] *Universidad Autónoma de Zacatecas, Centro Histórico C.P. 98000, Zacatecas, Mexico*

[2] *Université Catholique de Louvain (UCLouvain), LouRIM, B-1348 Louvain-la-Neuve, Belgium*

Keywords: Ambient Intelligence, Gesture-based User Interfaces, Gesture Elicitation Study, Privacy, Public Space, Zenithal Camera, Zenithal Gestures.

Abstract: Common spaces are full of cameras recording our pictures purposely or unintentionally, which causes privacy concerns. Instead of specifying our privacy preferences on one device or sensor at a time, we want to capture them once for an entire building through zenithal gestures in order to notify all devices in this building. For this purpose, we present an elicitation study of gestures elicited from thirty participants to notify reactions, acceptance or refusal of actions, via gestures recognized by a zenithal camera placed on the ceiling at the entrance. This perspective is different from the tradition frontal or lateral perspective found in other studies. After classifying the results into forty-six gesture classes, we suggest a consensus set of ten user-defined zenithal gestures to be used in a common space inside a building.

1 INTRODUCTION

In Ambient Intelligence (AmI) (Cook et al., 2009), common spaces, like public places, shopping malls, railway stations, are full of cameras and sensors where our image and data could be recorded on purpose or unintentionally (Ashok et al., 2014). For example, a person can pass by a public display and be captured for personalizing or customizing the user interface, which is perceived as a good benefit, or for suggesting some products in a shop close by, which is perhaps not considered as a benefit, depending on privacy preferences of the end user. These preferences can be communicated explicitly from one device to another (Kray et al., 2010), but this poses an additional problem of preferences transfer.

We investigate an alternative approach: instead of informing device by device, privacy preferences of users (Lopez et al., 2014) are transmitted to an entire building through zenithal gestures so as to notify all concerned devices. We define *zenithal gestures* as end-users body gestures (Kendon, 2004) acquired by a sensor located on top of the end user, hence its name borrowing the term “zenith”. Although zenithal gestures could cover in principle any human limb, zenithal gestures typically consist of upper-body gestures (*e.g.*, head and shoulders (Vanderdonckt et al., 2019), arms (Liu et al., 2015) until hands (Bostan et al., 2017)) acquired by a zenithal (usually spherical) camera hanging from the ceiling.

A zenithal camera offers the following advantages: it enables capturing an entire scene in a snapshot instead of a part of the environment, which is particularly appropriate when this environment dynamically changes (Cook et al., 2009); algorithms for gesture recognition (Coyette et al., 2007) are more simple than for gestures acquired with a frontal or a lateral camera, such as a MS Kinect Azure because the physical space should not be scaled, translated, or rotated in pre-processing (Vanderdonckt et al., 2018); it provides precise support for person detection (Li and Ma, 2018), localization (Grd et al., 2018), and counting (Antić et al., 2009) in various environments (e.g., counting the customers coming in or out a shop), three important functions in global human activity recognition. However, the image captured by a zenithal camera may suffer from irregularities of head movements and partial limb occlusion (e.g., neck movements cannot be differentiated as they are hidden by the head).

To support our initial scenario, we conduct a gesture elicitation study (Wobbrock et al., 2009) of zenithal gestures that notifies reactions and acceptance of gestures recognized from a zenithal perspective. We collect and propose a set of gestures to be used in a common space inside a building. For this purpose, we simulate the environment by equipping

[a] https://orcid.org/0000-0002-8842-7556
[b] https://orcid.org/0000-0001-7195-1637

Martínez-Ruiz, F. and Villarreal-Narvaez, S.
Eliciting User-defined Zenithal Gestures for Privacy Preferences.
DOI: 10.5220/0010259802050213
In *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2021) - Volume 2: HUCAPP*, pages 205-213
ISBN: 978-989-758-488-6

it with a collection of sensors capable of recognizing the presence of human beings and responding to their presence (Sadri, 2011). This paper aims at exploring gestures that users associate with privacy preferences and identifying the reaction of users while defining privacy settings in common spaces.

The rest of the paper is organized as follows: Section 2 presents work related to end user interaction, concentrating on privacy preferences, ambient intelligence, and gestures. Section 3 describes and conducts the research method used for eliciting zenithal gestures from participants of an experiment and Section 4 presents its main results and discuss them. Finally, Section 5 provides conclusion and discusses some future avenue to this work.

2 RELATED WORK

This section presents an overview of the concepts and work about privacy and gesture elicitation studies in general, then specifically in common spaces.

2.1 Privacy Preferences and Ambient Intelligence

Privacy is a particularly important issue for users and governments (Lopez et al., 2014). In a survey documented in the RFID Position Statement of Consumer Privacy and Civil Liberties Organizations[1], privacy is considered more important than the benefits of AmI itself. The phantom of the "Big Brother" is present for users (Cook et al., 2009). Although an always watching system could be useful at home, museums, business, health care facilities, and in public buildings (Sadri, 2011), it is also useful in a personal space. For instance, St. Pierre *et al.* (2014) propose a recording system that helps users to recover information from previous sessions to make it available only in an individual context. Nevertheless, end users tolerate a certain reduction in privacy in the social network realms. A photo or a tagging is acceptable if shared with friends (Besmer and Richter Lipford, 2010).

However, this is a vastly different story in public spaces. How to provide users with specific configurations and services in a public area? Grd *et al.* (2018) proposed a system to encode biometric information and deliver a hash id, so identity and biometric data of users remain safe. Similarly, Ashok *et al.* (2014) introduced a way to inform cameras privacy preferences of users with IR LEDs an approach that can be used also in buildings. Finally, the myriad of cameras

[1]See https://www.aclu.org/other/rfid-position-statement

Figure 1: Gesture variations for a "X" symbol to express a Decline action (Erazo et al., 2017).

available nowadays are informed with gestures about privacy concerns (Koelle et al., 2018), but with frontal cameras and hand gestures.

A zenithal camera can be effectively used as a indoor GPS whose data can be merged with additional data (Valera et al., 2007). In our exploratory study, end users have an active role as they are engaged in participatory design (Bergold and Thomas, 2012) because privacy preferences are explicitly expressed and then broadcasted to cameras and sensors in the public space. Contrarily to the "display blindness" of public interfaces where users avoid participating because information is considered not relevant (Müller et al., 2009), this scenario foster end users to participate by communicating their preferences.

2.2 The Design and Elicitation of Gestures

Gestures can be described using multiple notations and classifications (Grijincu et al., 2014). More technology and more reasonably priced becomes available now to identify and register gestures, such as a Microsoft Kinect Azure or radar-based devices (Magrofuoco et al., 2019). Also, there is an effort from open source communities to make available libraries of these sensors to a wider audience to implement an AmI environment and delivering a series of services. The research method consists of applying a gesture elicitation study (GES) (Wobbrock et al., 2009; Morris et al., 2010): a set of voluntary participants is presented with a series of referents, which materializes a task or an action to be executed by system according to the end user's gesture. Then, gestures are elicited from participants and analysed to identify a consensus set of gestures. This kind of study allows users to propose gestures that they feel more acceptable for a specific task (Rodriguez and Marquardt, 2017). A Systematic Literature Review (SLR) of existing GES reveals that they are primarily conducted depending on several conditions (Villarreal-Narvaez et al., 2020), depending on the type of users (*e.g.*, able-bodied people vs. people with disabilities), depending on their tasks (*e.g.*, actions performed in an AmI environment (Grijincu et al., 2014), sketching actions (Kieffer et al., 2010; Sangiorgi et al., 2012)), depending on the device (*e.g.*, a tabletop (Morris et al., 2010), a radar-based device (Magrofuoco et al., 2019), or a combination of devices (Kray et al., 2010)), depending on which human limbs are concerned (*e.g.*, hands (Bostan et al., 2017), arms (Liu et al., 2015)), and depending on the physical environment where end users are interacting (*e.g.*, in public settings (Rico and Brewster, 2010)).

Since this study elicits gestures from participants in order to express their privacy settings, the closest studies are related to opt-in/out actions (Erazo et al., 2017; Koelle et al., 2018; Rodriguez and Marquardt, 2017), which are different from general privacy concerns. For example, different gesture patterns can be used to express a same action, like variations for a "X" symbol to express a Decline action (Fig. 1).

While the elicitation considers multiple factors from participants, such as naturalness, intuitiveness, memorability, etc., the proposal can be affected by the legacy bias (Morris et al., 2010): most participants, who are often users of WIMP interfaces (windows, icons, menus and pointing mechanisms), are prone to an unintentional reuse of familiar metaphors and tend to propose the same gestures they already know, even if the conditions are quite different. We apply different techniques to minimize this bias, such as kinesthetic priming (Hoff et al., 2016) and soft constraints (Ruiz and Vogel, 2015). Since our study is specifically tailored to privacy concerns, social acceptability plays an important role (Rico and Brewster, 2010): people may elicit different gestures when they are in a public environment, such as discrete gestures, and less constrained gestures when they are in a private environment.

3 ZENITHAL GESTURE ELICATION STUDY

We conducted a GES following the methodology originally defined from the literature (Wobbrock et al., 2009; Vatavu and Wobbrock, 2015) to collect users' preferences for zenithal gestures associated to privacy actions.

3.1 Participants

Thirty voluntary participants (15 Females, 15 Males; aged from 18 to 55 years, M=26.50, SD=11.43) were recruited for the study via a contact list broadcasted in our organization. Most participants were students (25), the the rest being employees (4 professors and 1 administrative assistant). The age groups were intended to be as representative as possible for adopters of wearable computing, a technology that the percentage of individuals using it is the highest for the age group 25-34 years. All the participants were frequent users of computers and smartphones. Most of the users have experience with videogames (25 users, with 5 among them reporting to be frequent users) and gestural interface knowledge was extremely limited.

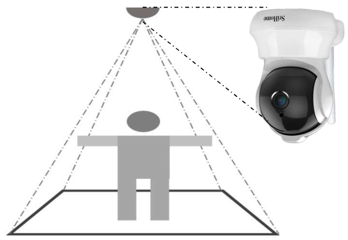

Figure 2: Apparatus used in the setup.

3.2 Apparatus

The controlled experiment took place in the main entrance of the building, where some space was devoted to dispose the experiment settings and materials. A space of two square meters space was marked with an "X" to indicate the interaction position to participants captured by a SriHome 3 MegaPixel AI IP camera[2] (Fig. 2). This location was selected to simulate a crowded space (the space was used by other people and passersby) to test gestures under real conditions.

3.3 Procedure

During the **introduction to the experimental setting**, each session was initiated by an explanation of the privacy settings to be tested, the position of the camera, the signing of a consent, form and the filling of a socio-demographic questionnaire about their experience with different devices and demographic information (based on a 7-point Likert scale (Likert, 1932) ranging from 1=strongly disagree to 7=strongly agree). The experimenter explained to participants the following tasks that they had to perform and the allowed types of zenithal gestures they could propose to be compliant with our definition.

During the **test phase**, each session applied the original GES protocol (Wobbrock et al., 2009). Each participant was verbally submitted to 10 *referents* (Table 1) in a sequential order: S1=Accept Sharing, S2=Decline Sharing, T1=Accept Tagging, T2=Decline Tagging, L1=Accept Locating, L2=Decline Locating, B1=Accept Blurring, B2=Decline Blurring, and OkAll=Accept all general actions, NotOkAll=Decline all general actions. For each referent, the participant proposed one full-body gesture (they could use the full range of gestures such as head, hands, arms,

[2]See http://www.sricam.com/

chest, legs, and feet) with the following conditions: each gesture should fit well with the referent, it should be easy to produce and to remember, it should not involve any vocal or touch interaction. Participants were instructed to remain as natural at all times as possible. They operated with the belief that no technological constraint (*e.g.*, no constraint due to gesture recognition) was imposed to preserve the natural and intuitive character of the elicitation. Proposed gestures were captured by the zenithal camera in a video sequence, one per participant, and saved in a file for analysis.

During the **post-test phase**, participants filled in the IBM Post-Study System Usability Questionnaire (PSSUQ) (Lewis, 2002). While this questionnaire is traditionally used for testing software interaction, it was selected here among other questionnaires to enable participants to express their level of satisfaction with the perceived setup usability, the convenience of the zenithal camera, and the testing process. This questionnaire has been empirically validated with a large number of participants on a significant set of stimuli (Lewis, 2006), it is widely applicable for any system, and it benefits from a proved α=.89 reliability coefficient between its results and the perceived system usability (Lewis, 1995). Each IBM PSSUQ closed question is measured using a 7-point Likert scale (1=strongly disagree, 2=largely disagree, 3=disagree, 4=neutral, 5=agree, 6=largely agree, 7=strongly agree) and four measures are computed: system usefulness (SysUse: Items 1-5), quality of the information (InfoQual: Items 6-11), quality of the interaction (InterQual: Items 12-15), and system quality (Overall: Item 16). A whole session in average took 12 minutes: 5 minutes for each questionnaire and 2 minutes for eliciting zenithal gestures.

3.4 Design

Our study was within-subjects with two groups (female vs. male) and with one independent variable: REFERENT, a nominal variable with 10 conditions, representing privacy preferences, *i.e.*, Accept or Decline actions (Table 1). Most systems reacted late at preventing the sharing or tagging photos and included no tagging, no taking photos, and blurring options (Besmer and Richter Lipford, 2010). Sharing and tagging have been treated in social networks. So, we proposed a list of privacy preferences to be informed a priori at the entrance of the public space or building. In contrast to (Koelle et al., 2018) who elicited frontal gestures for the opt-in and opt-out actions, we considered the Accept (Ok) and DECLINE (NOK) symmetric options. Lastly, we decided to include global preferences that should be valid in any case. These selectors will inform the public space that we accept or decline all preferences in a single command. Two groups were considered to investigate whether the gestures elicited by participants are sensitive to gender: would male and female participants prefer zenithal gestures differently or consistently since their respective preference is subject to social acceptance?

Table 1: List of referents by action on various objects.

Referents	Accept (Ok)	Decline (NOk)
Sharing	S1	S2
Tagging	T1	T2
Locating	L1	L2
Blurring face	B1	B2
General	OkAll	NotOkAll

4 RESULTS AND DISCUSSION

A total amount of 300 gestures were elicited from 2 groups × 15 participants × 10 referents, which we clustered into groups of similar types according to the pattern criteria devised in (Erazo et al., 2017) (see Fig. 1 for some examples of variations clustered in the same class/pattern). For this purpose, all gestures were labeled with the terms described in GestureML[3] to maintain a consistent vocabulary, thus resulting into 46 classes of zenithal gestures (Fig. 4).

Fig. 3 shows the agreement rates (Vatavu and Wobbrock, 2016) computed by AGATe (Vatavu and Wobbrock, 2015) for each REFERENT condition sorted by decreasing order of the agreement rate. The values are aggregated for both groups. Below are the two most preferred gestures elicited by participants or each referent. No GES exists today that covers the same referents. The closest GES elicited gestures for opt-in/out in both commercial/private setups (Rodriguez and Marquardt, 2017) (Fig. 3).

Overall, these values belong to the range of medium agreement (M=.129, SD=.085) according to the magnitude table (Vatavu and Wobbrock, 2015). This is aligned with classical results of GES but can also be explained by the size of the classification: the more classes are included in the classification, the more choices participants invented, and the less the agreement rates become. The Accept blurring face referent stands out as it is the only one benefiting from a high agreement (M=.363), thus suggesting that the most frequent gesture class elicited

[3]See http://www.gestureml.org

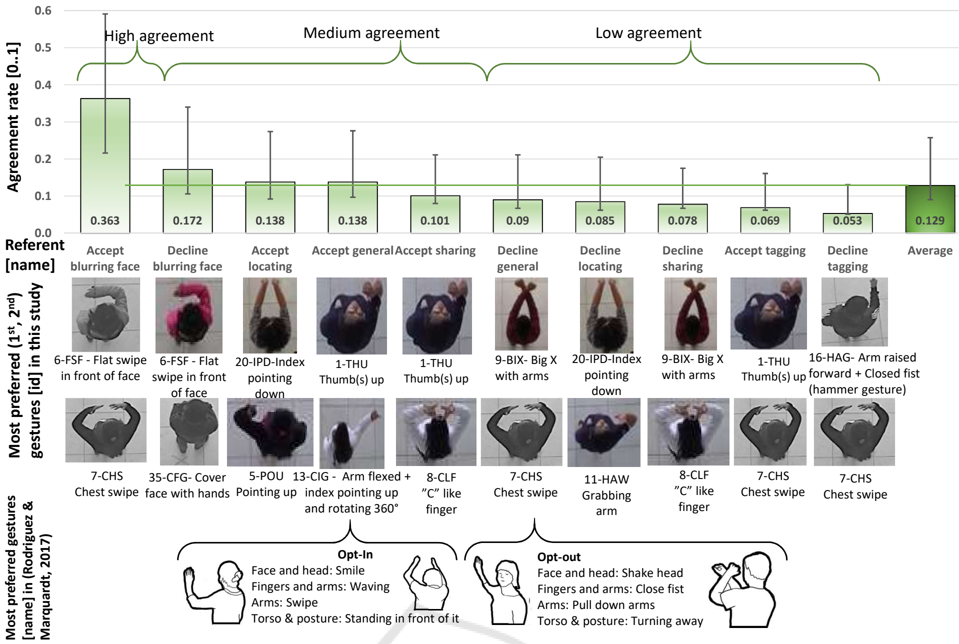

Figure 3: Agreement rates among gestures for all referents, sorted in decreasing order of their value with the most most preferred gestures and (below) comparison with gestures elicited for Opt-in and Opt-out in (Rodriguez and Marquardt, 2017).

for this referent should be selected. Surprisingly, its opposed referent, *i.e.*, Decline blurring face, received the second highest agreement rate among participants (M=.172), although it belongs to the medium agreement range. This is the case for the next four referents until Accept sharing (M=.101). The paired referents Accept tagging (M=.069) and Decline tagging (M=.053) received the lowest agreement rates, perhaps because participants were not familiar with any metaphor to represent these actions, as opposed to other ones which are more familiar for them.

AGATe (Vatavu and Wobbrock, 2015) also computes a statistical test to determine whether these rates are correlated. Since our sampling contains five pairs of related referents, each pair with a Accept and a Decline action, we discovered four correlations: a statistically significant effect ($p \leq .050^{*}$) between Accept sharing and Decline sharing ($V_{rd(1,N=60)} = 5.188$) and between Accept general and Decline general ($V_{rd(1,N=60)} = 5.188$), a highly significant one ($p \leq .010^{**}$) between Accept locating and Decline locating ($V_{rd(1,N=60)} = 9.981$), and a very highly significant one ($p \leq .001^{***}$) between Accept blurring face and Decline blurring face ($V_{rd(1,N=60)} = 49.561$), which confirms the two high rates for this pair of referents. All Accept (Ok) referents ($V_{rd(4,N=150)} = 204.373$) taken together as well as all Decline (NOk) referents ($V_{rd(4,N=150)} = 31.680$) were also related in a very highly significant manner (p^{***}=.001), thus suggesting that the gesture classes respectively proposed by participants exhibit a high internal cohesion and are consistently chosen. We did not find any statistically significant difference between the two groups (female *vs.* male): $V_{bg}(2,N{=}30)$ always returned a value $p \geq .239$, *n.s.*.

In order to establish a basic mapping between categories, we opened the pool of options to the second most frequent gestures selected by participants. The two most preferred gesture classes for Accept Sharing are classes coded as THD/1 and CLF/8 respectively. We remark that most of the users avoid a mirror gesture (*e.g.*, right hand and then left hand). The Accept tagging referent was related to CHS/7, a gesture that includes the chest (upper body to indicate that privacy setting affected the whole person). For the Decline tagging, the strong gesture coded HAG/16 represents a hammer gesture, which could be also suitable. The Accept locating referent used a space gesture IPD/20 (Index Pointing Down) and for the Decline locating, hand(s) waving coded HAW/11 in order to alert the system to avoid record user location (this is an ambiguous use of this gesture that could be used for opt-in commands). The Accept Blurring face referent is marked with FSF/6 (Face Swipe). Participants reproduced this legacy behaviour, probably because of their previous experience with drawing GUI applications. Next, for the Decline Blurring face, covering the face was the second option but it is a good fit because, instead of a motion gesture, users have proposed a static one CFG/35. Finally, the Accept General ref-

Figure 4: Classification of zenithal gestures resulting from the study.

erent was associated by participants in second term with the circle gesture CIG/13, a motion gesture that involves users and their environment. For the Decline General, the big "X" was selected as the first term. Again, it is a good fit to express the decline of all options because the gesture is clear in a zenithal view and also can be understood and remembered by participants in subsequent interactions with the system. The big "X" was also elicited for Opt-out referent in a commercial setup (Rodriguez and Marquardt, 2017).

Fig. 5 graphically depicts the mean PSSUQ score per participant. Nine participants in the 30 sampling assessed zenithal gestures as they elicited them in the setup as "excellent" proposals suitable for expressing the referents ($M \geq 6$). Fourteen of them assessed these gestures with a "good" averaged score ($5 \leq M \leq 6$), which represents the vast majority of them. This is also depicted in Fig. 6, where the first bag of the Pareto distribution covers the interval $[5,6[$, while the second corresponds to the "excellent" region with the interval $[6,7]$. These two bags already cover 80% of the population sampling, which means a good representativeness. Only three participants out of thirty were not confident in the gestures they proposed by reporting they were not suitable. *P*22 may be an outlier: the lowest score was answered for all questions.

Fig. 7 reports the results of the IBM PSSUQ questionnaire. On the one hand, most questions were an-

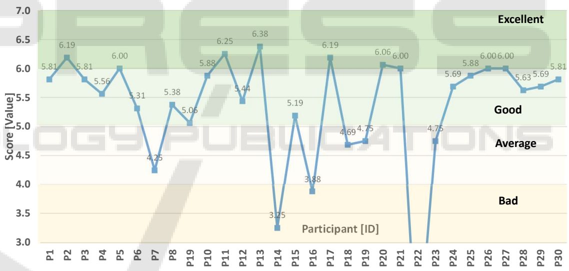

Figure 5: Results of the IBM PSSUQ.

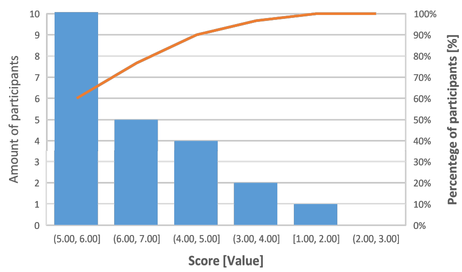

Figure 6: Pareto distribution of IBM PSSUQ.

swered positively, which directly impacted the first metric SysUse (M=5.54, SD=1.25). PSSUQ metrics having a value greater or equal than five are considered assessed positively, even if 5 is not the median value of the 7-point Likert scale. This happens for SysUse, InterQual (M=5.58, SD=1.28) and Overall satisfaction (M=5.33, SD=1.51), but not the infor-

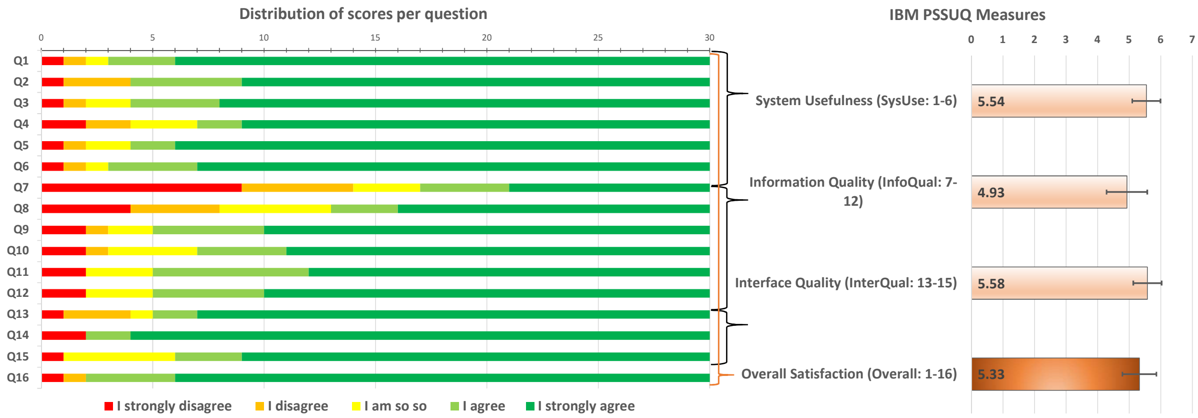

Figure 7: Results of the IBM PSSUQ.

mation quality InfoQual metric (M=4.93, SD=1.80). This is mainly due to questions Q7 and Q8 which were assessed rather negatively and with a wide distribution. Q7 assesses the quality of error messages, which is an inappropriate criteria since gestures, as they were presented, do not include any feedback in case of error. All elicited gestures were accepted. Similarly, Q8 assesses the error recovery, which is not really appropriate in this case.

Regarding the location, the selected space was a real entrance of the Computer Engineering faculty building. The conditions of the interaction were real. Students and passersby at any time. However, the fortuity audience was influenced in the time of response of participants and in the thinking time. In future studies, we must avoid a space with the public until we have a more developed set of gestures. Most of the participants were students of the computer science engineering faculty. So, they have some experience with GUI interactions, but could be biased by their background. The consequence was that some motion hand gestures were part of the gestures proposed as participants were familiar with these gestures in video games. However, other types of background influenced the results. For instance, one participant is a folkloric dancer, so the swirl gesture naturally emerged and another participant, who was part of the university soccer team, proposed leg and feet gestures.

5 CONCLUSION AND FUTURE WORK

In this paper, we gathered and studied the gestures that a collective of users has suggested in order to select different privacy settings. Most of the gestures have been reduced to simpler versions and loosely gathered in 46 categories. Since we have avoided any limitation in the gesture proposal. Participants have proposed complex gestures including arms, legs, and hand motions. Participants started to create a vocabulary, especially with the use of gesture number 9 (Big "X") as prefix to indicate declination of specific privacy options. The Legacy bias was present in some of the gestures and interaction suggestions. An interesting discovery was that instead of producing mirror gestures to express the dichotomy of allow/decline an option, users changed to a different suggestion. Finally, this is an exploratory exercise that gave us a repertory of gestures that can be used in forthcoming studies. Also, we can introduce some extra guidelines (Vanderdonckt, 1995) for the participants taking into account the current results and avoid the occlusion in the zenithal plane and small gestures. Besides that, testing Welcome and Ending gestures to open and close the privacy preferences configuration session should be considered. Since this was a first exploration of the zenithal gestures, we must ask participants to take more time. Also, we must avoid the rush times in the entrance of the building to see if we get an alternative group of categories. In addition, in this study most of the participants have a computer science background. So, we should include a more diverse pool of users. Finally, we should include feedback in the mock-up interface. For instance, adding lights in future iterations of the elicitation study to help participants to test their proposals.

REFERENCES

Antić, B., Letić, D., Ćulibrk, D., and Crnojević, V. (2009). K-means based segmentation for real-time zenithal people counting. In *Proc. ICIP '09*, pages 2565–2568.

Ashok, A., Nguyen, V., Gruteser, M., Mandayam, N., Yuan, W., and Dana, K. (2014). Do Not Share! Invisible

Light Beacons for Signaling Preferences to Privacy-Respecting Cameras. In *Proc. of VLCS '14*, page 39–44, New York. ACM.

Bergold, J. and Thomas, S. (2012). Participatory research methods: A methodological approach in motion. *Historical Social Research/Historische Sozialforschung*, pages 191–222.

Besmer, A. and Richter Lipford, H. (2010). Moving beyond untagging: Photo privacy in a tagged world. In *Proc. of CHI '10*, page 1563–1572, New York. ACM.

Bostan, I., Buruk, O. T., Canat, M., Tezcan, M. O., Yurdakul, C., Göksun, T., and Özcan, O. (2017). Hands as a controller: User preferences for hand specific on-skin gestures. In *Proc. of DIS '17*, pages 1123–1134.

Cook, D. J., Augusto, J. C., and Jakkula, V. R. (2009). Ambient intelligence: Technologies, applications, and opportunities. *Pervasive and Mobile Computing*, 5(4):277 – 298.

Coyette, A., Schimke, S., Vanderdonckt, J., and Vielhauer, C. (2007). Trainable sketch recognizer for graphical user interface design. In *Proc. of INTERACT '07*, volume 4662, pages 124–135. Springer.

Erazo, O., Rekik, Y., Grisoni, L., and Pino, J. A. (2017). Understanding gesture articulations variability. In *Proc. of INTERACT '17*, pages 293–314, Cham. Springer.

Grd, P., Tomičić, I., and Baca, M. (2018). Privacy improvement model for biometric person recognition in ambient intelligence using perceptual hashing. In *Proc. of CECC '18*, New York. ACM.

Grijincu, D., Nacenta, M. A., and Kristensson, P. O. (2014). User-defined interface gestures: Dataset and analysis. In *Proc. of ITS '14*, page 25–34, New York. ACM.

Hoff, L., Hornecker, E., and Bertel, S. (2016). Modifying gesture elicitation: Do kinaesthetic priming and increased production reduce legacy bias? In *Proc. of TEI '16*, page 86–91, New York. ACM.

Kendon, A. (2004). Gesture: Visible action as utterance.

Kieffer, S., Coyette, A., and Vanderdonckt, J. (2010). User interface design by sketching: A complexity analysis of widget representations. In *Proc. of EICS '10*, page 57–66, New York. ACM.

Koelle, M., Ananthanarayan, S., Czupalla, S., Heuten, W., and Boll, S. (2018). Your smart glasses' camera bothers me! exploring opt-in and opt-out gestures for privacy mediation. In *Proc. of the NordiCHI '18*, page 473–481, New York. ACM.

Kray, C., Nesbitt, D., Dawson, J., and Rohs, M. (2010). User-defined gestures for connecting mobile phones, public displays, and tabletops. In *Proc. of MobileHCI '10*, pages 239–248, New York. ACM.

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *Int. J. of Human-Computer Interaction*, 7(1):57–78.

Lewis, J. R. (2002). Psychometric evaluation of the PSSUQ using data from five years of usability studies. *Int. J. of Human-Comp. Int.*, 14(3-4):463–488.

Lewis, J. R. (2006). Sample sizes for usability tests: Mostly math, not magic. *interactions*, 13(6):29–33.

Li, L. and Ma, J. (2018). Zenithal people detection based on improved faster r-cnn. In *IEEE 4th Int. Conf. on Computer and Com.*, ICCC '18, pages 1503–1508.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22(140):55–.

Liu, M., Nancel, M., and Vogel, D. (2015). Gunslinger: Subtle arms-down mid-air interaction. In *Proc. of UIST '15*, pages 63–71. ACM.

Lopez, M., Pedraza, J. P., Carbó, J., and Molina, J. M. (2014). Ambient intelligence: Applications and privacy policies. In *Proc. of PAAMS '14*, volume 430 of *Communications in Computer and Information Science*, pages 191–201. Springer.

Magrofuoco, N., Pérez-Medina, J. L., Roselli, P., Vanderdonckt, J., and Villarreal, S. (2019). Eliciting contact-based and contactless gestures with radar-based sensors. *IEEE Access*, 7:176982–176997.

Morris, M. R., Wobbrock, J. O., and Wilson, A. D. (2010). Understanding users' preferences for surface gestures. In *Proc. of Graphics Interface 2010*, GI '10, pages 261–268. Canadian Information Processing Society.

Müller, J., Wilmsmann, D., Exeler, J., Buzeck, M., Schmidt, A., Jay, T., and Krüger, A. (2009). Display blindness: The effect of expectations on attention towards digital signage. In *Pervasive Computing*, pages 1–8. Springer.

Rico, J. and Brewster, S. (2010). Usable gestures for mobile interfaces: evaluating social acceptability. In *Proc. of CHI '10*, pages 887–896, New York. ACM.

Rodriguez, I. B. and Marquardt, N. (2017). Gesture elicitation study on how to opt-in & opt-out from interactions with public displays. In *Proc. of ISS '17*, pages 32–41, New York. ACM.

Ruiz, J. and Vogel, D. (2015). Soft-constraints to reduce legacy and performance bias to elicit whole-body gestures with low arm fatigue. In *Proc. of CHI '15*, pages 3347–3350, New York. ACM.

Sadri, F. (2011). Ambient intelligence: A survey. *ACM Comput. Surv.*, 43(4).

Sangiorgi, U. B., Beuvens, F., and Vanderdonckt, J. (2012). User interface design by collaborative sketching. In *Proc. of DIS '12*, pages 378–387. ACM.

Valera, A., Weiss, M., Vallés, M., and Díez, J. (2007). Bluetooth-networked trajectory control of autonomous vehicles. *IFAC Proc. Vol.*, 40(1):198–203. 8th IFAC Symposium on Cost Oriented Automation.

Vanderdonckt, J. (1995). Accessing guidelines information with sierra. In *Proc. of INTERACT '95*, pages 311–316. Chapman & Hall.

Vanderdonckt, J., Magrofuoco, N., Kieffer, S., Pérez, J., Rase, Y., Roselli, P., and Villarreal, S. (2019). Head and shoulders gestures: Exploring user-defined gestures with upper body. In *Design, User Experience, and Usability. User Experience in Advanced Technological Environments*, pages 192–213, Cham. Springer.

Vanderdonckt, J., Roselli, P., and Pérez-Medina, J. L. (2018). !FTL, an articulation-invariant stroke gesture recognizer with controllable position, scale, and rotation invariances. In *Proc. of ICMI '18*, page 125–134, New York. ACM.

Vatavu, R.-D. and Wobbrock, J. O. (2015). Formalizing agreement analysis for elicitation studies: new measures, significance test, and toolkit. In *Proc. of CHI '15*, pages 1325–1334, New York. ACM.

Vatavu, R.-D. and Wobbrock, J. O. (2016). Between-subjects elicitation studies: Formalization and tool support. In *Proc. of CHI '16*, pages 3390–3402, New York. ACM.

Villarreal-Narvaez, S., Vanderdonckt, J., Vatavu, R., and Wobbrock, J. O. (2020). A systematic review of gesture elicitation studies: What can we learn from 216 studies? In *Proc. of DIS '20*, pages 855–872. ACM.

Wobbrock, J. O., Morris, M. R., and Wilson, A. D. (2009). User-defined gestures for surface computing. In *Proc. of CHI '09*, pages 1083–1092, New York. ACM.