Mid-air Authentication Gestures: An Exploration of Authentication Based on Palm and Finger Motions

Ilhan Aslan+, Andreas Uhl*, Alexander Meschtscherjakov+ and Manfred Tscheligi+
University of Salzburg, Salzburg, Austria
name.lastname@sbg.ac.at+
uhl@cosy.sbg.ac.at*

ABSTRACT

Authentication based on touch-less mid-air gestures would benefit a multitude of ubicomp applications, which are used in clean environments (e.g., medical environments or clean rooms). In order to explore the potential of mid-air gestures for novel authentication approaches, we performed a series of studies and design experiments. First, we collected data from more then 200 users during a three-day science event organised within a shopping mall. This data was used to investigate capabilities of the Leap Motion sensor and to formulate an initial design problem. The design problem, as well as the design of mid-air gestures for authentication purposes, were iterated in subsequent design activities. In a final study with 13 participants, we evaluated two mid-air gestures for authentication purposes in different situations, including different body positions. Our results highlight a need for different mid-air gestures for differing situations and carefully chosen constraints for mid-air gestures.

Keywords

mid-air gestures; authentication; Leap Motion sensor; design;

Categories and Subject Descriptors

H.5.2. [**Information Interfaces and Presentation**]: User Inter- faces - Interaction styles;

1. INTRODUCTION

There is a need for authentication in a world pervaded with more and more digital technologies. We authenticate ourselves by entering a PIN when we want to pay with our credit card. We get access to our mobile phones by connecting points on our smartphones. We type in passwords for authentication on websites. We use iris scanners for authentication in high-security areas. Thus, different interaction modalities are used for authentication in different contexts. The different modalities are often dependent on various factors, such as the degree of security and the availability of technology, as well as environmental factors like background noise or the physical position of interactive elements.

ICMI '14, November 12 - 16 2014, Istanbul, Turkey
Copyright 2014 ACM 978-1-4503-2885-2/14/11 ...$15.00
http://dx.doi.org/10.1145/2663204.2663246 .

Another requirement for the right choice of authentication modalities is the "cleanness" of the interaction context (e.g., aseptic environments). Examples for such environments are well known from the medical domain. In such domains, keyboards or touch screens are not always the best choice for authentication. Users often wear gloves and must not contaminate themselves or the environment. In cooperations with an industrial partner from the semiconductor domain, we recognised that there is also a need for a clean and ideally touch-less way of interaction in the industrial domain, especially in the clean room.

A clean room in the semiconductor context is often a noisy and large environment, where workers need to wear special clothes including rubber gloves and a mask in order to minimise dust, which could cause defects in the production process of integrated circuits (see Fig. 1). Workers in clean rooms are mobile and need to authenticate themselves repeatedly during their daily activities on different workstations and in different situation. This is often realised via text-based passwords. Most workstations in the clean room are desktop computers, which are either used in a standard setting (i.e., seated with monitor, keyboard, and mouse positioned on a desk) or in a setting where the workers have to interact with the system while standing with the monitor and keyboard attached to the wall.

Figure 1: Clean room in a semiconductor factory. Operators are wearing suits and rubber gloves in order to reduce dust circulation.

While keyboards and mice provide an effective and well known way of interaction, they are also considered in-proper for clean rooms, since they attract dust. Therefore, more and more touch interfaces are integrated in clean rooms.

However, even touch interfaces pose an issue in clean rooms, since dust is attracted by touch screens and contaminate operators' gloves when they touch the screens. Following this argument, the interaction in the clean room should be touch-less whenever possible. Apart from using well-known biometric approaches such as face-, iris-, or voice-detection, we propose to use mid-air gestures for a touch-less authentication purposes.

Recently, touch-less interaction has vastly improved that off-the-shelf devices, such as the Leap Motion[1], sensor can be purchased for low costs. These sensors can recognise hand biometry and hand gestures and allow touch-less and gesture based interaction. Moreover, mid-air gestures are a promising modality for interaction in clean rooms and hand biometry information gleaned from mid-air gestures may even provide a clean way for authentication.

In order to explore the potential of mid-air gesture for authentication, we performed a series of studies. Hereby, we were motivated by related work in touch and gesture-based authentication [4, 26] , which has shown that gesture-based authentication can be an alternative to text-based passwords. Therefore, we aim to address the following research question in this paper, can authentication based on mid-air gestures and the Leap Motion controller be used in situation relevant for clean rooms?

In the following section, we first provide some background in gesture-based authentication. Then, we present our exploration of mid-air authentication gestures, which started with a field study and a series of subsequent design experiments based on inspirations gained from visits to actual clean rooms. The design experiments were conducted to specify the design problem and to infer design considerations for authentication gestures. Afterwards, we present a study with 13 participants, which shows authentication accuracy rates for two gestures performed in different situation and an overview of the user experience of the mid-air gestures based authentication method. We conclude that with recognition rates for single gestures around 89% that mid-air based authentication gestures are already an alternative for authentication based on touch gestures. Considering that 3D controllers are compared to touch sensitive screens in their infancy, there is great promise in mid-air gesture based authentication.

2. BACKGROUND

With systems and services becoming increasingly mobile and ubiquitous, novel approaches for authentication need to be easy to use in different situations. For mobile devices, which are used on the go, motion and shape-based approaches (e.g., password patterns used by Android-based mobile phones) have been introduced as alternatives to text-based password schemes (e.g., PIN-authentication). These approaches make use of the pictorial superiority effect [18] and the human motor memory [28] to reduce the cognitive burden of memorizing passwords.

While shape-based authentication approaches are easy to memorise and use, security aspects have been criticised due to the fact that they are easy to spy on (e.g., [3]). In order to improve the security of shape-based authentication approaches on mobile phones related work has investigated the use of an additional implicit authentication layer. Based on a series of user studies, DeLuca et al. [4] were able to show that inter-personal differences in drawing the same shape could be used to improve the security of shape-based authentication on mobile phones.

Inter-personal differences have also been exploited by Sae-Bae et al. [26] in a gesture-based authentication approach for multi-touch devices, which is based on a study with iPads. They developed a comprehensive set of five finger gestures based on hand ergonomics and were able to show that there are behavioural traits in gestures, which can be used to improve shape-based authentication on touch screens. On one hand, these inter-personal differences in human gestures can be disturbing when designing user interfaces due to the introduced inter-personal variability (variances among different users might get interpreted as distinct "commands"). On the other hand, they can be exploited for designing personalised user interfaces (the same gesture is interpreted differently when conducted by different users). In any case, these inter-personal differences are essential for gesture-based authentication.

In recent work the Leap Motion device has been used for recognising handwriting in the air [30] as well as for sign language recognition [21].

However, the capabilities of these sensors for authentication purposes have not yet been investigated in depth, including the influence of context on performing a mid-air authentication gesture.

Human (hand) gesture recognition [17, 14] has already been explored in many applications like sign language recognition [20], device [19] and software control [25], augmented reality application [23] and authentication [1, 7, 10]. We may distinguish *instrumented sytems*, which use trackable markers mounted to the fingers (e.g., infrared targets [23]) and *non-instrumented systems* relying on vision techniques only. Different types of sensors can also be used ranging from classical vision using still-image or video cameras [13, 22, 12, 31, 19, 7] to various depth imaging techniques [2, 32]. Research in mid-air based interaction has increased since sensors, such as Microsoft's Kinect device, have been on the market. Applications involving the Kinect include static hand gesture recognition [11, 24], dynamic hand/arm gesture recognition [33], software control [25], surgeons assistance [15], and authentication [29]. Other sensors employed for (hand) gesture recognition include (smartphone integrated) accelerators [9, 10, 16], virtual reality interfaces like Cyber Glove [20], and specific biometric devices (e.g. a Smart Pen [1]).

With respect to hand gesture-based authentication, few techniques have been proposed. Arm movements are captured for this purpose by (smartphone integrated) accelerators [9, 10, 16]. Mid-air handwriting recognition for authentication is conducted with different sensors (Kinect [29], Bio Smart Pen [1]); sign language recognition for identifying user-specific pass-codes is suggested in [12].

Independent from gesture recognition techniques, the obvious benefit of mid-air hand gestures is that humans already use their hands and fingers in different situations to manipulate real world objects. In the past, interaction design evolved through making use of human skills to improve interaction with the digital world [5]. Therefore, it can be assumed that mid-air gestures have the potential to improve interaction with ubiquitous systems, allowing users to interact more "naturally" with the digital world in different situations. A particular challenge for ubicomp application in

[1]http://www.leapmotion.com

clean rooms is that users need to perform a log-in action in different situations.

Although a multitude of authentication techniques exist, techniques for ubicomp application that use biometric information are favoured. Biometric information is always at hand and available anytime and anywhere. With off the-shelf-3D controllers that can track detailed movements of fingers, the usage of fine-grained mid-air hand gestures for authentication seems to become possible.

3. EXPLORATION OF MID-AIR GESTURES FOR AUTHENTICATION

The Leap Motion controller is a camera-based 3D controller, which uses depth and RGB cameras similar to Microsoft's Kinect device to recognise "body" movement. However, this 3D controller is a light, small sized device with a short range (approximately 1 inch to 2 feet above the device). Its software has been developed for a particular use; i.e., recognising palm and finger movements in detail. Being a small and lightweight 3D controller, it is flexible and can be attached to various surfaces (e.g., walls or desks). In fact, the device is already being embedded in laptops (i.e., the HP envy Leap Motion). Due to the device dimension, it could also be worn on the body (e.g., like a wrist watch).

Being a vision based tracking system, the Leap Motion sensor faces some obvious limitations, such as being able to recognise fingers only if they are in the field of vision. For example, hand rotations along the lower arm are difficult to recognise by the device if the device is used in a seated posture with the device placed on a desk. Based on our initial explorations, we also realised that the hands of some people are recognised better than others. To get an idea about how well the sensor tracks hand movements, we collected data in a field study, where we logged data from more then 200 people.

3.1 Field study

At a science event in November 2014, a few departments of University (blinded) presented their research to the public in a large shopping mall. Everyone could attend this event and engage with demonstrators, prototypes, and design artefacts. At this event, we presented different gesture based interfaces, which made use of the Leap Motion sensor. In one display, visitors could use a capacitive pen in combination with a Leap Motion sensor to create artful drawings. Visitors had to perform a specific mid-air gesture before they could start drawing.

By rotating their palm clockwise by 70 degrees, they could clean the canvas (screen of an iPad device). We logged hand movement data associated with this initial gesture. On a monitor that was positioned in front of the users, this initial gesture was presented as part of a promotional video for the booth. The 3D controller could either be positioned on the desk next to the iPad or worn on the wrist (e.g., as it is done by the participant in Figure 2). To allow users to wear the 3D controller on their wrist, we build an adjustable velcro wristband. Figure 2 presents the final setup, which was used at the field study. As also presented in Figure 2, many of the booth visitors were teenagers and the situation was crowded. We collected data from more then 200 visitors, 195 of those have agreed to fill in a post-hoc questionnaire providing demographic data. Inspection of the data we collected showed that only in 12% of the time 5 fingers were recognised (in 25% of the time 4 fingers, in 37% 3 fingers, in 53% 2 fingers and in 75% 1 finger was recognised) Even though the palm was recognised in 25 %, no fingers were recognised.

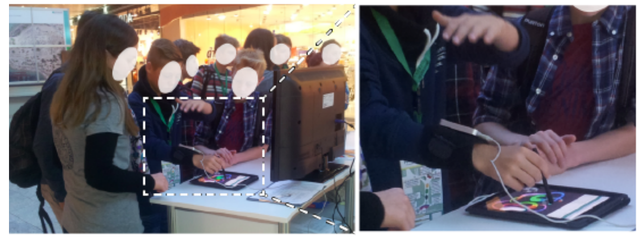

Figure 2: Setup at the science event. A participant has the 3D controller attached to his right wrist and is performing mid-air gestures with his left hand to interact with information presented at the iPad.

Further inspections of the data showed that the Leap Motion SDK only provides finger IDs depending on the order in which fingers are recognised for each frame, thus no semantic information is provided, such as if a recognised finger matches to a thumb or an index finger, etc. Furthermore, the same finger ID is mapped to semantically different fingers of the hand over the period of performing a gesture (see Figure 3). For authentication gestures, it is necessary to map fingers to their respective movement paths otherwise biometric informations is lost. Furthermore, five fingers should be recognised to get maximal data from the hand geometry. The main reason that not all fingers were recognised was that fingers were not clearly separated during gestures. Since gestures are not performed in exactly the same space (e.g., at the same distance to the sensor), it is also necessary to transfer gesture data to a canonical form. This is needed in order to recognise the same gesture even if it is performed with a slightly different orientation and at a different space. Taking these design constraints into account, we decided to design specific authentication gestures.

3.2 Designing authentication gestures

Inspired by the set of categories for five-finger touch gestures, which were reported by Sae-Bae et al. [26], we defined a five finger mid-air gesture for authentication (see Figure 3). We chose a complex gesture including circular finger and palm movements as well as finger closing and finger opening movements (all movements accruing in three dimensional space) to capture maximal biometric information and to study the boundaries of what is feasible.

We presented a video recording of this gesture and asked several people from our own department to replicate this authentication gesture. We used the data for further exploration. Figure 3 also presents the data, which was captured by the Leap Motion device. Even though we asked users to separate their fingers in the beginning of the gesture to ensure that all five fingers were recognised in the beginning, exploration of the captured data showed that for some users the sensor failed to recognise all five fingers in the beginning.

3.2.1 *Normalising hand gesture data*

Recognising five fingers in the beginning is important since we use the initial state of the hand to match movement sequence to individual fingers. Individual fingers are recognised by computing the distance between fingers. If the fingers are spread, the thumb has the highest distance to the other fingers. The index finger is nearest to the thumb, etc.

Once individual fingers are matched, the gesture data can be normalised based on the position of the palm, the index finger and the pinky by performing transformations; i.e., moving the palm to the origin, rotating the index finger twice, so that it's tip is on an the z-axis and on the x-plane. Afterwards, the pinky is rotate also on the x-plane.

In earlier attempts, we used the thumb instead of the palm. However, since the thumb is the most flexible finger it caused higher intra-personal difference (the same persons positioned their thumbs differently in the beginning of a gesture), which is undesirable for an authentication approach.

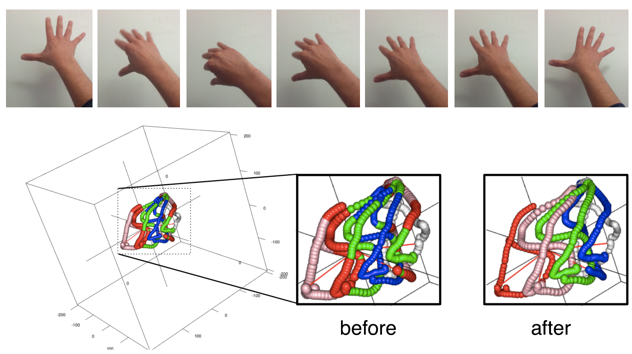

Figure 3: On the upper half a gestures is presented. On the lower half data representing the gesture before and after performing a matching algorithm is presented. The algorithm matches individual fingers to movement paths.

After matching captured data to individual fingers and normalising the gestures using the camera coordinate system, we needed to match individual fingers to the movement paths. This was a challenge since not all fingers were recognised during performing the gesture. Figure 3 visualises captured data before and after performing the matching algorithm.

3.2.2 Design workshop

In order to reflect on results gained and to discuss the design of authentication gestures in general with interdisciplinary professionals, we conducted a design workshop with six participants with backgrounds in computer science, game design, sociology, and psychology. We did this to obtain qualitative and critical feedback to the authentication gesture we designed so far and create new ideas for gestures taking into account restrictions of the Leap Motion sensor and naturalness of performing the gestures for authentication purposes.

In the workshop, we first presented our idea of using mid-air gestures for authentication. Then, we demonstrated the Leap Motion controller, what restrictions it has, and how it could also be worn on the wrist or attached to a wall. Three of the 6 participants have already visited clean rooms and knew the context well. We asked participants to walk around and perform mid-air gestures in situation relevant for the clean room to build empathy for the topic; i.e., seated with imagining the sensor on the desk and standing in front of a whiteboard with a sensor sketched on the whiteboard. Participants could also explore the capabilities of a Leap Motion sensor during the workshop. The workshop took an hour and several ideas were generated on top of our previous results.

The following main issues were raised in the workshop. Without visual feedback, it is difficult to memorise a gestures even if the gesture was chosen by a user themselves. Gestures should be short, simple and easy to repeat. Due to hand and body ergonomics, different contexts might need different gestures which needs to be explored. Shoulder, upper arm, and elbow muscle behaviour could also be used for the authentication.

Based on the insights gained and suggestions for authentication gestures, we refined two gestures and designed visual feedback for those gestures.

- Gesture 1 was an "upward" movement of the hand, with the hand slightly close and and opens. It makes use of finger movement behaviour.
- Gesture 2 is an "upward" movement, with the hand rotating anti-clockwise. It makes use of shoulder, upper arm, and elbow movement and, consequently, associated muscle behaviour.

For both gestures, the start and the end positions have been chosen to be different to allow an improved matching of individual fingers to movement paths. As visual feedback, we created an animation of an hand icon. We chose a very simple and abstract presentation, so that it could be presented on different kinds of small sized displays. Furthermore, presenting an abstract visualisation would allow users to define themselves how exactly the gestures should be performed. The key purpose of visual feedback was twofold: first, to provide feedback on the recognition quality (e.g., hand icon turns green if all five fingers are recognised) and, second, to guide and also remind the user of "what" and "how" to perform the gesture. Once the user is ready and five fingers are recognised, the hand icon was animated slowly to allow the user to "mirror" the gesture. We chose as duration for the animation two seconds based on our own experience.

4. USER STUDY

In order to investigate how the designed gestures would be perceived by users and to explore intra and inter-personal variabilities in performing mid-air authentication gestures, we conducted a user study. Moreover, we investigated their authentication strength in three different contexts

The three contexts were chosen based on prior observations during visits to actual clean rooms. There are two main situation in which workers authenticate themselves on workstations; i.e., seated with monitor, keyboard, and mouse positioned on a desk or standing with the monitor and keyboard attached to the wall. Depending on the physical setting, keyboards are even vertically attached to a wall with the mouse being placed on a small piece which is also attached to the wall. The third context, which we studied, includes a concept for a wearable 3D controller. Authentication on a wearable and mobile interfaces is a future alternative to the currently existing contexts. Using a small sized 3D controller as a wearable on a workers wrist is a realistic alternative for an interface in clean rooms, since workers already have special clothes that they need to wear during their work. In addition, there is a trend for wearable devices outside the clean room context (i.e., smart watches), which might inspire the future interface landscape in clean rooms.

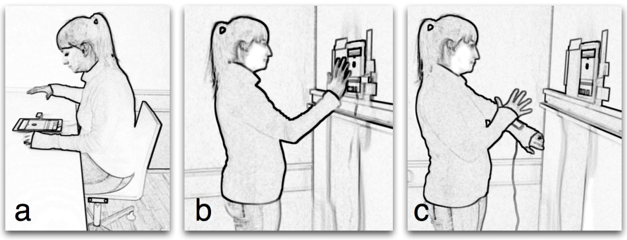

Figure 4: The three conditions in which mid-air gestures for authentication were studied a) 3D controller is positioned on the desk. b) 3D controller is positioned on the wall. c) 3D controller is worn on wrist.

4.1 Study setup

Figure 4 presents the three different physical setups we asked participants to perform mid-air authentication gestures. The visual guidance was presented on a second generation iPad which we positioned in each physical setting in a way that participants could easily perceive the visual feedback. For the study we recruited thirteen right handed participants (7f, 6m) between 26 and 41 years of age. All participants completed the study within 20 minutes.

Participants were asked perform each of the two gestures in 3 different conditions for 10 times in counter balanced order. The rate in which data was collected was between 50 and 55 per second. In sum, we collected data from 260 gestures and, for each gesture, we had between 100 and 110 frames. Participants were also told that the visualisation provides them guidance but that they can define for themselves how to perform the gesture specifically. After completing the tasks, participants were interviewed with the goal to get qualitative feedback on performing the gestures in die different conditions.

4.2 Analysis and results

In order to analyse the hand data, we normalised the data as we described it in the previous section and used dynamic time warping (DTW); i.e., the implementation for R by Toni Giorgino [8]. We chose to use DTW, since it has been already used in related work to compute dissimilarities between gestures for authentication purposes ([4, 26] in a very similar manner with similar sets of data from gestures).

DTW compares two gesture based on time series (e.g., a series of x, y and z coordinates of fingertip positions over the period of a gesture), which can be of different lengths and produces a distance (i.e., Euclidean distance) value between 0 and a positive value. The value is higher for gestures that are less similar to each other. Data from each gesture is formatted in way that it can be used as a parameter for DTW: $Gesture(t) = [palm_1, thumb_1, indexFinger_1, ringFinger_1, pinky_1, ..., palm_n, thumb_n, indexFinger_n, ringFinger_n, pinky_n]$ where n is the number of frames (i.e., a number between 100 and 110). Hereby, $palm_i$ (for example) is the vector of x,y and z co-ordinates of the i^{th} record of the palm position; $palm_i = [x, y, z]$

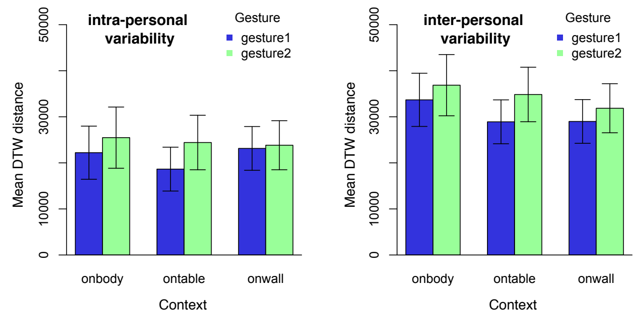

Figure 5: Mean values and intra-personal variability in performing the two gestures are presented on the left (for each condition). Similar values for Inter-personal variability are presented on the right. Y-axis denotes the DTW distance, a value presenting the dissimilarity between two gestures. Error bars denote the standard deviation.

We computed the mean DTW distance and the associated standard deviation (see Figure 5) for each participant, each condition, and each gesture . Based on these values, we computed the intra-personal and inter-personal variability of participants of our study. The difference between intra-personal and inter-personal variability, which is apparent in all conditions, is promising and indicates that mid-air gesture could be used for authentication. Interestingly, the differences seems to be highest for the "onbody" condition. This might be due to the fact that users not only have inter-personal differences in performed gestures but also differences regarding their body posture when having the sensor on their wrist.

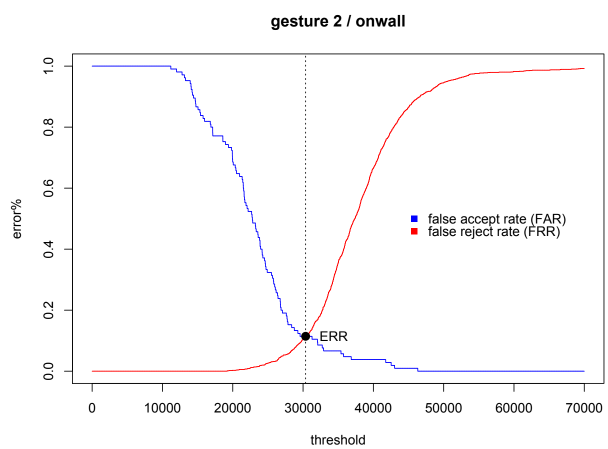

Figure 6: For the condition in which participants were asked to perform gesture 2 with the 3D control attached to the wall FAR and FRR are plotted for a range of thresholds. The Equal Error Rate (EER) is identified and thresholds for dissimilarity score is derived.

4.2.1 Analysis of biometric data

Similar to how Sae-Bae et al. [26] analysed authentication

Table 1: Equal error rates and derived thresholds

Gesture	context	EER	Threshold
1	ontable	0.0973	25926
1	onwall	0.1196	27678
1	onbody	0.1025	29669
2	ontable	0.1322	25926
2	onwall	0.1144	30365
2	onbody	0.1365	32224

gestures based on multi-touch data, we analysed the biometric data in mid-air gestures using Equal Error Rate (EER) to measure accuracy. The EER denotes the rate where False Acceptance Rate (FAR) and False Rejection Rate (FRR) are equal. FAR is compute by dividing the number of incorrectly verified forgery cases by the number of forgery cases. FRR is compute by dividing the number of rejected genuine cases by the number of genuine cases. Since we collected 10 samples (i.e., gestures) from each participant in each context, we used the first five samples to chose a template. The sample with minimum sum distance to the other four samples was taken as the template. The last five samples were used to test for genuine cases.

In order to calculate EER, all thresholds were used to compute FAR and FRR values. FAR and FRR were plotted for all thresholds to identify the ERR (see Figure 6). Table 1 presents the EER values and derived thresholds for both gestures and all three contexts. Overall, the two mid-air gestures achieved EER of 11.71% (as an averaged value between the two gestures). Gesture 2 achieved over all threse context ERR of 12.77 % and gesture 1, EER of 10.65101 %. Figure 7 presents corresponding accuracy rates in percentage.

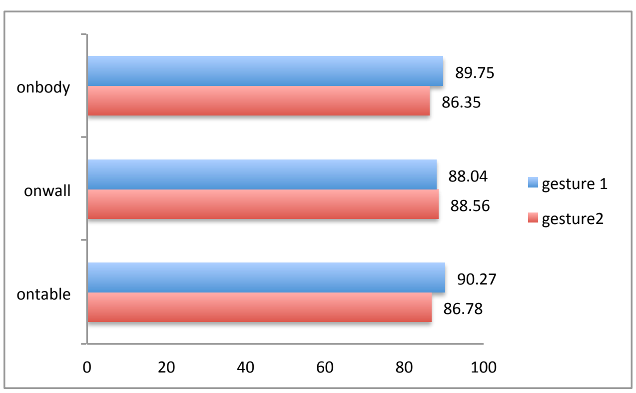

Figure 7: Overview of gesture accuracy (number of authenciatnion gestures that have been correctly accepted or rejected in %) for different contexts.

4.2.2 Analysis of user experience

In addition to measuring the authentication strength of mid-air gestures, we were also interested in exploring how users would perceive mid-air gestures as a modality for authentication. In order to get initial insights we used a semantic differential questionnaire. Participants were asked to answer the questions at the end of the study. Figure 8 provides an overview.

Afterwards, participants were briefly interviewed. Regarding the specific gestures, there was no overall response

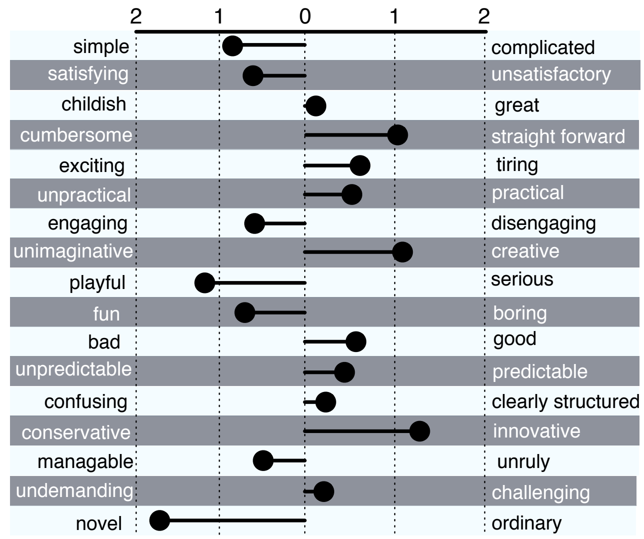

Figure 8: Overview of self-reported items, which participants were asked to provide in order to describe how mid-air gestures for authentication purposes were perceived overall.

on which gesture was easier or better for the purpose of authentication. Half of the participants preferred gesture 2, since rotating their palm did not include fine grade finger movements and was less exhausting. The other half preferred gesture 1, since they did not have to move their elbows and shoulders. However, both gestures were perceived as easy to perform.

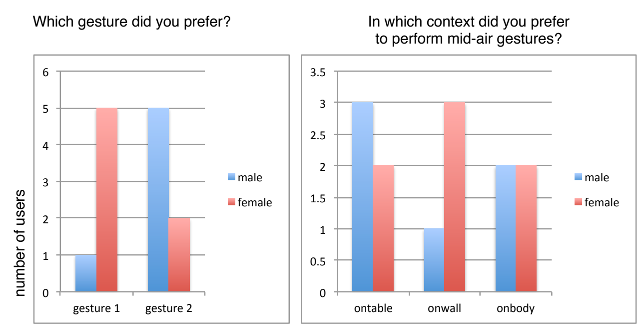

Figure 9: Overview of user preferences regarding mid-air gesture and context in which mid-air gestures were performed.

Interestingly, female participants preferred gesture 1 and male participants gesture 2 (see Figure 9). Participants were also asked to state in which context they felt most comfortable performing mid-air gestures. Our results show no obvious preference for one specific context. However, participants mentioned that they think having the sensor on the arm is, on the one hand practical while, on the other, it is mentally more demanding since they need to coordinate both arms.

The items of the questionnaire (see Figure 8) were inspired by various user experience questionnaires related to the attractiveness of a system or emotions associated with interactive systems, etc. The purpose of the questionnaire was to get an initial impression of how users perceived performing mid-air authentication gestures. Users choose mainly positive attributes to describe their experience, which motivates us to follow this line of research in our future work.

5. DISCUSSION

The accuracy rates, which we have observed in our study are very promising, considering that accuracy rates in multi-touch gestures [26] have been reported to be between 81-93% depending on the individual gestures. We observed accuracy rates between 86-91% for two different kinds of mid-air authentication gestures in different contexts. Sae-Bae et al. [26] suggest that two gestures, which are performed in sequence could improve accuracy rates for touch gestures up to 97%. Furthermore, they report accuracy rates of 97% for gestures, which were user defined. Through allowing users to define their own gestures accuracy rates for mid-air gestures could also be increased, making the technique scalable. However, due to the restrictions of the Leap Motion controller and the fact that users are not yet familiar using mid-air gestures, one would need to provide an easy way to assist in generating authentication gestures. Based on our experience, we would suggest a simple application, which visualises a user's hand gestures while they try out different gestures, providing feedback on how well the gesture would suit for authentication (e.g., how well fingers were recognised over the duration of the gesture).

In comparison to related work, we chose to use only one template (i.e., gesture) as a reference from each user for each context. We did this since we wanted the computation to be fast, so that the algorithm could be implemented and used in a real system in near future. However, it is not unusual to use more then one template; i.e., to use a set of templates and take the best match as the result. Using a set of templates would improve the accuracy rate.

Furthermore, we are aware that DTW produces more false positives due to the fact that the algorithm tries to match two time series (with different lengths). In our future work, we will explore other implementations (e.g., a python based implementation) of the DTW algorithm, which on the one hand might compute dissimilarity scores faster, allowing more than one template to be used as a reference. On the other hand we would investigate implementations [6] to improve the fact that DTW tends to produce false positives.

DTW is not claimed to be a better or worse choice than other possible approaches (e.g., machine learning). However, if there is potential in mid-air authentication gestures, then there is a high chance that DTW will show it [4, 26].

In figure 5 we presented the intra- and inter-personal variability, which seemed to suggest that differences are higher for the "onbody" context. However, the computation of authentication accuracy rates shows no visible difference between contexts. This could be due to possible outliers, and their influence based on the sample size of the user study. Based on our observations during the study, there was no good reason to take out samples from the analyses. However, since participants had to coordinate both hands in the "onbody" context, there might be a greater chance for outliers compared to contexts in which the sensors was physically attached to a still surface. We used the camera coordinate system for the normalization process, which could be an additional explanation for the slightly lower performance of gesture 2 when the sensor is worn on the body.

The rational for using the Leap Motion sensor is that it was a low cost solution, which potentially could be purchased in large numbers and integrated in a clean room environment in near future. Furthermore, it provided an easy way to use palm and finger motions without expertise in vision based recognition systems. Compared to the data provided by the leap motion, sensors using a skeleton-based gesture recognition would include considerably more complexity. Since, we used the Leap Motion controller, our results are specific for this 3D controller. However, with future improved 3D controllers, we expect improvements in usability as well as user experience of mid-air authentication gestures.

In sum, our investigations have shown that authentication based on mid-air gestures and the Leap Motion controller can be used in different contexts relevant for clean rooms.

Interaction based on mid-air gestures has the potential to replace existing interaction modalities in clean rooms, whenever possible. In practice, we expect to see interaction concepts for clean rooms, where mid-air gestures will be used in combination with other modalities. Based on our results, we could imagine to combine touch and touch-less gestures also for authentication purposes. Fusing existing interaction modalities in clean rooms with mid-air gestures would also introduce additional benefits, which are typical for multi-modal interfaces (e.g., naturalness of interaction, or improved performance).

6. CONCLUSION

We argued that clean rooms have particular constraints on interaction modalities and require clean and ideally touch-less ways to interact with workstations, which are distributed in the clean room depending on available space. Throughout the paper we put emphasis on reflective design activities [27] trying to explore mid-air authentication gestures. We used the Leap Motion controller to restrict but also focus our design activities. Based on visits to clean rooms we decided to study mid-air authentication gestures in two situation including body and 3D controller positions (i.e., seated with the 3D controller placed on the desk and standing with the 3D controller attached to a wall). Furthermore, we included a third context, which we thought has future relevance (i.e. using a 3D controller attached on the wrist). We were able to design two authentication gestures, which not only were perceived as easy to perform and suitable for all relevant situations, but also provided hand biometry information which could be used for a clean way of authentication.

7. ACKNOWLEDGMENTS

The financial support by the Austrian Federal Ministry of Science, Research and Economy and the National Foundation for Research, Technology and Development is gratefully acknowledged (Christian Doppler Laboratory for "Contextual Interfaces").

8. REFERENCES

[1] M. Bashir, G. Scharfenberg, and J. Kempf. Person authentication by handwriting in air using a biometric smart pen device. In *Proc. of BIOSIG 2011*, 2011.

[2] F. Coleca1, T. Martinetz, and E. Barth. Gesture interfaces with depth sensors. In M. Grzegorzek et al.,

editors, *Time-of-Flight and Depth Imaging*, volume 8200 of *Springer Lecuture Notes on Computer Science*, pages 207–227. 2013.

[3] A. De Luca, M. Denzel, and H. Hussmann. Look into my eyes!: Can you guess my password? In *Proc. of the 5th Symposium on Usable Privacy and Security*, pages 7:1–7:12, New York, NY, USA, 2009. ACM.

[4] A. De Luca, A. Hang, F. Brudy, C. Lindner, and H. Hussmann. Touch me once and i know it's you!: Implicit authentication based on touch screen patterns. In *Proc. of CHI '12*, pages 987–996, New York, NY, USA, 2012. ACM.

[5] P. Dourish. *Where the Action is: The Foundations of Embodied Interaction*. A Bradford book. Bradford Books, 2004.

[6] H. Feng and C. C. Wah. Online signature verification using a new extreme points warping technique. *Pattern Recogn. Lett.*, 24(16):2943–2951, Dec. 2003.

[7] S. Fong, Y. Zhuang, I. Fister, and I. F. Jr. A biometric authentication model using hand gesture images. *Biomedical Engineering Online*, 12(111), 2013.

[8] T. Giorgino. Computing and visualizing dynamic time warping alignments in r: The dtw package. *Journal of Statistical Software*, 31(7):1–24, 8 2009.

[9] J. Guerra-Casanova, C. Sanchez-Avila, G. Bailador, and A. de Santos Sierra. Authentication in mobile devices through hand gesture recognition. *International Journal on Information Security*, 11(2):65–83, 2012.

[10] J. Guna, E. Stojmenova, A. Lugmayr, I. Humar, and M. Pogacnik. User identification approach based on simple gestures. In *Proc. of SAME*, pages 39–48, 2012.

[11] M. Hamissi and K. Faez. Real-time hand gesture recognition based on the depth map for human robot interaction. *International Journal of Electrical and Computer Engineering (IJECE)*, 3(6):770–778, 2013.

[12] K. Karthik, K. Varalakshmi, and S. Ravi. A file authentication system using hand gesture passcodes. *International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)*, 13(174):394–401, 2013.

[13] W. Kasprzak, A. Wilkowski, and K. Czapnik. Hand gesture recognition based on free£form contours and probabilistic inference. *International Journal of Applied Mathematics and Computational Sciences*, 22(2):437–448, 2012.

[14] S. Kavanagh. Facilitating natural user interfaces through freehand gesture recognition. In *Proc. of CHI '12*, New York, NY, USA, 2012. ACM.

[15] K. Kosic, B. Arzensek, A. Kuhar, and M. Vogrincic. Towards new user interfaces based on gesture and sound identification. In *Proc. of SQAMIA*, pages 45–53, 2013.

[16] J. Liu, Z. Wang, L. Zhong, et al. uWave: Accelerometer-based personalized gesture recognition and its applications. In *Proc. of PerCom'09*, 2009.

[17] S. Mitra and T. Acharya. Gesture recognition: A survey. *IEEE Transactions on Systems, Man, and Cybernetics - PartC: Applications and Reviews*, 37(3):311–324, 2007.

[18] D. L. Nelson, V. S. Reed, and J. R. Walling. Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning and Memory*, 2(5), 1976.

[19] O. Nieto and D. Sasha. Hand gesture recognition in mobile devices: Enhancing the musical experience. In *Proc. of CMMR'13*, 2013.

[20] F. Parvini and C. Shahabi. An algorithmic approach for static and dynamic gesture recognition utilising mechanical and biomechanical characteristics. *International Journal on Bioinformatics Research and Applications*, 3(1):4–23, 2012.

[21] L. Potter, J. Araullo, and L. Carter. The leap motion controller: a view on sign language. In *Proc. of OzHCI'13*, pages 175–178, 2013.

[22] A. Ramamoorthy, N. Vaswani, S. Chaudhury, and S. Banerjee. Recognition of dynamic hand gestures. *Pattern Recognition*, 36:2069–2081, 2003.

[23] S. Reiflinger, F. Wallhoff, M. Ablassmeier, T. Poitschke, and G. Rigoll. Static and dynamic hand-gesture recognition for augmented reality applications. In J. Jacko, editor, *Proc. of HCII 2007*, volume 4552 of *Springer Lecture Notes on Computer Science*, pages 728–737, 2007.

[24] Z. Ren, J. Yuan, and Z. Zong. Robust hand gesture recognition based on finger- earth mover£s distance with a commodity depth camera. In *Proc. of ACM Multimedia 2011*, pages 1093–1096, 2011.

[25] K. Sabir, C. Stolte, B. Tabor, and S. O'Donoghue. The molecular control toolkit: Controlling 3d molecular graphics via gesture and voice. In *Proc. of the IEEE Symposium on Biological Data Visualisation 2013*, pages 49–56, 2013.

[26] N. Sae-Bae, K. Ahmed, K. Isbister, and N. Memon. Biometric-rich gestures: A novel approach to authentication on multi-touch devices. In *Proc. of CHI '12*, pages 977–986, New York, NY, USA, 2012. ACM.

[27] D. A. Schön. Designing as reflective conversation with the materials of a design situation. *Knowledge-Based Systems*, 5(1):3–14, 1992.

[28] R. Shadmehr and T. Brashers-Krug. Functional stages in the formation of human long term motor memory. *Journal of Neuroscience*, 17(1):409–419, Jan. 1997.

[29] J. Tian, C. Qu, W. Xu, and S. Wang. Kinwrite: Handwriting-based authentication using kinect. In *Proc. of the 20th Annual Network & Distributed System Security Symposium*, 2013.

[30] S. Vikram, L. Li, and S. Russel. Writing and sketching in the air, recognizing and controlling on the fly. In *Extended Abstracts on Human Factors in Computing Systems CHI EA '13*, pages 1179–1184, 2013.

[31] Y. Yao and C.-T. Li. Real-time hand gesture recognition for uncontrolled environments using adaptive SURF tracking and hidden conditional random fields. In *Proc. of ISVC'13*, Springer, pages 542–551, 2013.

[32] M. Ye et al. A survey on human motion analysis from depth data. In *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications*, volume 8200 of *Springer*, pages 149–187. 2013.

[33] Y. Yin and R. Davis. Gesture Spotting and Recognition Using Salience Detection and Concatenated Hidden Markov Models. In *Proc. of ICMI'13*, pages 489–494, 2013.