A Study on Proximity-based Hand Input for One-handed Mobile Interaction

Florian Müller[1], Mohammadreza Khalilbeigi[1], Niloofar Dezfuli[1],
Alireza Sahami Shirazi[2]*, Sebastian Günther[1], Max Mühlhäuser[1]
[1] Technische Universität Darmstadt, Darmstadt, Germany
[2]Yahoo Labs, 701 First Avenue, Sunnyvale, CA 94089, United States
{florian.mueller, khalilbeigi, niloo, guenther, max}@tk.informatik.tu-darmstadt.de,
alireza@yahoo-inc.com

ABSTRACT

On-body user interfaces utilize the human's skin for both sensing input and displaying graphical output. In this paper, we present how the degree of freedom offered by the elbow joint, i.e., flexion and extension, can be leveraged to extend the input space of projective user interfaces. The user can move his hand towards or away from himself to browse through a multi-layer information space. We conducted a controlled experiment to investigate how accurately and efficiently users can interact in the space. The results revealed that the accuracy and efficiency of proximity-based interactions mainly depend on the traveling distance to the target layer while neither the hand side nor the direction of interaction have a significant influence. Based on our findings, we propose guidelines for designing on-body user interfaces.

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

General Terms: Human Factors, Experimentation

Keywords: Human Factors; Design; Measurement.

1. INTRODUCTION

Technological advances in (depth) sensors and mobile projectors resulted in the emergence of a new class of interfaces that extend interaction to the surface of our body. These so-called on-body interfaces [6, 7] allow ubiquitous and mobile interaction with digital contents by sensing input and projecting graphical output on the skin. The hand and forearm receive particular attention because they are often unclothed and socially acceptable to touch [12]. These advantages resulted in a large body of research for body-based projective [7, 14, 8, 13], augmented [5] or imaginary [3, 4] interfaces.

In most of these systems, the user's non-dominant hand acts as a two-dimensional interactive surface on which the

---

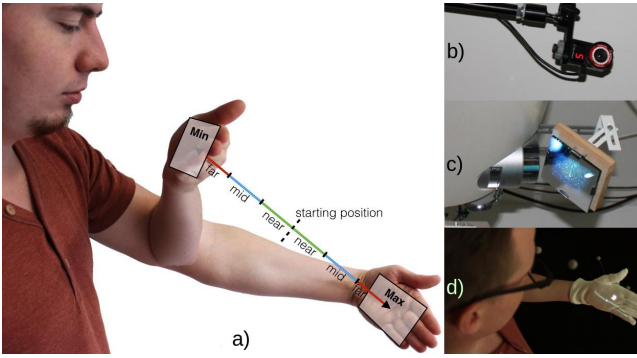*The majority of the work has been conducted while he was a researcher at the University of Stuttgart.

Figure 1: A map application as an example of one-handed (a) proximity-based interaction with a linear layered information space. The user can browse map layers by moving his hand through the space (b).

opposing hand interacts with the content through (multi-)touch gestures. While useful and practical, the interaction space is bound to the two-dimensional surface of the hand. Moreover, this style of interaction requires both hands and therefore hardly supports situations, where users are encumbered. Similar to [2], we believe that the large number of degrees of freedom offered by our hands and arms can support one-handed interaction styles based on proximity. We can rotate and move our hands away or towards our body or we can hold them at a specified position.

We extend the input space of prior on-body user interfaces by focusing on the degree of freedom offered by the elbow joint, i.e., flexion by moving the hand toward and extension by moving the hand away from the body. We propose to use this proximity dimension as an additional input modality for one-handed mobile interaction. The interaction space alongside the user's line of sight can be divided into multiple parallel-planes. Similar to [11], each plane corresponds to a layer with visual content. The user can move his hands to browse through successive layers (cf. Figure 1). Beyond palm-projected interfaces, our approach can also be used as an additional input dimension for devices such as wearables or head-mounted displays with small input spaces for touch interaction. For such devices, our approach allows to expand the interaction space and provide direct manipulation.

In this paper, we investigate the human capabilities for a proximity-based hand input modality in multi-layer information spaces. We contribute the results of a controlled experiment addressing two main questions:

1. How accurate and efficient users can interact with the layered information space in a search task scenario?
2. How to design the interaction space in terms of layer thickness, number of layers, and convenient boundaries of the physical interaction volume?

Figure 2: Traveling distance zones (a) and setup of the study (b-d).

Figure 3: Visual feedback in the study: After reaching the starting position (a), the system showed the direction of interaction (b). The participants task was to browse through a stack of white colored numbers (c) to find the one red colored number (d).

In the remainder of this paper, we present the methodology of the experiment followed by detailed report on our results. Finally, we present the implications and limitations.

2. CONTROLLED EXPERIMENT

The focus of the study was to investigate how efficiently and accurately users can interact in a multi-layer information space in a searching task. Further, we evaluated the influence of the direction of interaction and the side of the hand.

We recruited 14 participants (P1-P14: 4 female, 1 left-handed), aged between 24 and 29 years ($\mu = 26$, $\sigma = 1.6$), using the University's mailing address. The average height was 177cm ($\sigma = 9.5$cm), the average arm length (measured from armpit to carpus) of 59cm ($\sigma = 3.6cm$). We choose a within-subject design. No compensation was provided.

2.1 Design and Task

Similar to [10], we designed a basic multi-layer information space alongside the participants line of sight consisting of randomized integer numbers (each layer displayed one number). We varied the **number of layers** in the available interaction space (which directly correlates with the layers' thickness) as the independent variable with the values of 12, 24, 36, 48, 60 and 72. In addition, we varied the **direction of interaction** between flexion and extension as a second as well as the **side of the hand** (palm or backside) as a third condition. We considered those conditions to assess their influence on the participants performance in terms of accuracy and efficiency.

The participants' first task was to search for the one red colored number in the stack of white colored numbers (cf. Figure 3). Once found, participants confirmed the discovery by pressing a button with their non-interacting hand. Directly after, as the second task, participants had to hold the hand steady at the respective position for 3 seconds to measure the accuracy while trying to hold on a layer.

We defined the maximum boundary of the interaction space with the participant's individual arm-length and the minimum boundary as the near point of the human's eye (not closer than 12.5cm to the user's face). Furthermore, we defined the starting point of all trials as half of the distance between the minimum and the maximum interaction distance, resulting in an elbow joint deflection of around 100 degree. Informal pre-tests showed this to be a natural and relaxed holding position for the hand. To systematically an-

alyze influences of the traveling distance of the users hand, we divided the total available interaction space in each direction into three equal-sized zones: near, medium, and far as shown in Figure 2 a).

2.2 Study Setup and Apparatus

We used an optical tracking system (OptiTrack, cf. Figure 2 b) to precisely measure the linear distance between the participant's hand and his eyes alongside the participant's line of sight. To achieve this, we used two trackable apparatuses: a glasses frame and a glove, each augmented with a number of small retro-reflective markers (cf. Figure 2 d), which participants wore during the study. We further used the real time tracking information to fit the projected feedback to the participant's hand. (cf. Figure 3). For each trial, we measured:

1. the **task completion time (TCT)** as the timespan between starting the trial and confirming the discovery of the target.
2. the **overshooting error** as the maximum deviation in distance (in mm) between the center of the target layer and the participant's hand before confirming the discovery.
3. the **holding error** as the maximum distance (in mm) from the starting point of the holding task.

2.3 Procedure

We used a repeated measure design with 6 levels for numbers of layers, 2 different hand sides, and 2 directions of inter-action with 6 repetitions (two from each zone) resulting in $6 \times 2 \times 2 \times 6 = 144$ trials. The order of the conditions was counterbalanced using a Balanced Latin Square design for the number of layers and the direction of interaction. We excluded the side of the hand condition from the Latin Square design because remounting the trackable marker resulted in also recalibrating the system. However, half of the participants performed all palm-side trials first, while the other half started with the backside trials.

We introduced the participants to the concept and study setup. We mounted the two trackable apparatuses and calibrated the system to adapt it to the respective arm size. Before starting each trial, the system guided the user to the starting position through visual feedback displayed on the users hand. Once in the starting position, the system displayed the direction of the interaction. Each trial started by pressing the button. Once the target was found, the participant confirmed the discovery through another click. After that, the system informed participants to hold their current position for three seconds. Participants did not receive any
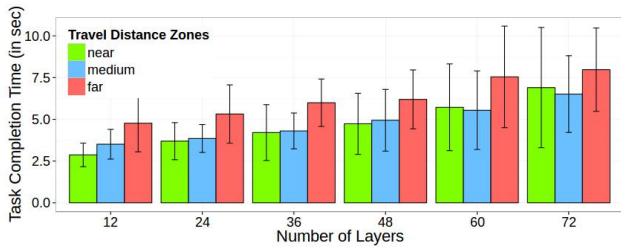
Figure 4: Mean TCT and SD for different numbers of layers.

Figure 5: Error measurements for the three traveling distance zones.

feedback during the holding task and were not informed on the current layer thickness.

After each condition, participants took a 30 seconds break. We concluded the experiment with a semi-structured interview focusing on their overall opinion about the concept, preferred interaction boundaries (minimum/maximum distance), and differences between the tested conditions. The experiment took 60 minutes per participant.

3. RESULTS

We analyzed the data using repeated measures ANOVA. In cases where Mauchly's test indicated a violation of the assumption of sphericity, we corrected the tests using the Greenhouse-Geisser method and report the ϵ. When significant effects were revealed, we applied Bonferroni corrected pairwise t-tests for the post-hoc analysis. For correlation tests, we used Pearson's method.

3.1 Task Completion Time

The analysis unveiled that the traveling distance of the hand had a significant effect on the TCT ($F_{2,62} = 23.27; p < 0.001$). Post-hoc tests confirmed that the TCT for near (μ=4.7s, σ=3.6s) and medium zone (μ=4.8s, σ=2.7s) targets were significantly smaller (p<0.001) than for those in the far zone (μ=6.4s,σ=3.4s). Post-hoc test did not show a significant different TCT between medium and near zone targets.

The number of layers had a significant effect on the TCT (ϵ=0.49;$F_{2.45,31.36}$=45.68; p<0:001). Post-hoc tests revealed a significantly (p< 0.01) larger TCT for higher numbers of layers. The mean TCT increased from 3.7 sec (σ=1.8s) for 12 layers to 7.2 sec (σ=4.7s) for 72 layers. While the mean TCT was faster for extension (μ=5.5s, σ=3.5s) than flexion (μ=5.1s,σ=3.1s), we could not find any significant effects ($F_{1,13}$=2.8, $p = 0.12$). Also, no significant effect of the hand orientation on TCT was found ($F_{1,13}$=0.15, $p = 0.70$, Palm: μ=5.2s, σ=3.2s, Back: μ=5.3s, σ=3.4s). We could not find interaction effects between the conditions. Figure 4 shows the TCT for the explored numbers of layers and target layers.

3.2 Overshooting Error

The traveling distance also had a significant effect on the overshooting error ($\epsilon = 0.63; F_{1.26,16.38} = 39.44; p < 0.001$). Post-hoc tests showed significant differences between all zones (all p<0.05). We observed that participants initially started with fast movements and slowed down towards their physical boundaries in the far zones, resulting in higher overshooting errors in the near (μ=4.4cm, σ=1.7cm) and medium (μ=2.1cm, σ=1.0cm) zones compared to the far
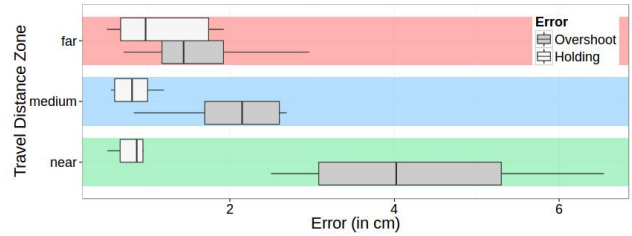
(μ=1.6cm, σ=0.7cm) zone. Figure 5 shows the overshooting error.

The analysis showed neither any significant influence of the direction of interaction on the overshooting error ($F_{1,13} = 0.0008, p = 0.97$; flexion: $\mu = 2.5cm, \sigma = 3.0cm$; extension: $\mu = 2.6cm, \sigma = 3.2cm$) nor the hand orientation ($F_{1,13} = 0.11, p = 0.75$; palm: $\mu = 2.6cm, \sigma = 3.1cm$; back: $\mu = 2.6cm, \sigma = 3.0cm$). Furthermore, we could not find any significant influence ($F_{5,64}$=0.64, p=0.66) of the number of layers (Min: μ=2.2cm, σ=3.1cm for 12 layers, Max: μ=2.8cm, σ=3.4cm for 36 layers). Also, we could not find any significant correlation between the participant's arm-length and their accuracy ($r(166)$= âĹŠ0.8376, p=0.40) in our results.

3.3 Holding Error

We observed significant effects of the distance between the starting point and the holding point on the holding error (ϵ=0.56, $F_{1.12,14.56}$=5.53, p<0.05). Post-hoc tests showed a significant effect (all p<0.05) between targets in the far (μ=1.6cm, σ=1.8cm) and the medium (μ=1.0cm, σ=0.9cm) zone as well as between the far and near (μ=1.1cm,σ=1.1cm) zone. The difference between near and medium zones were however not significant ($p > 0.74$). Figure 5 shows the holding error for the travel distance zones.

We could not find either any significant influences of the direction of interaction ($F_{1,13} = 1.65, p = 0.22$; flexion: $\mu = 0.7cm, \sigma = 0.8cm$; extension: $\mu = 0.8cm, \sigma = 0.8cm$) or the hand orientation ($F_{1,13} = 1.37, p = 0.26$; palm: $\mu = 0.8cm, \sigma = 0.8cm$; back: $\mu = 0.7cm, \sigma = 0.7cm$) on the holding error. Furthermore, we could observe a significant (ϵ=0.29, $F_{1.45,18.85}$=7.21, p<0.001) influence on the number of layers. Post-hoc tests confirmed a significant (p < 0.01) bigger holding error for 12 layers (μ=1.2cm, σ=0.9cm) compared to all higher numbers of layers. The mean hold error further decreased for increasing numbers of layers (min: μ=0.6cm, σ=0.4cm for 72 layers) but was not significant.

3.4 Qualitative Results

In general, all participants appreciated the idea of being able to interact with multi-layer information spaces through movements of their hand. There was a strong consensus among participants (11 out of 14) that this input modality is suitable for immediate and short-term interactions, such as serendipitous discovery of contents, fast peeking into information or executing a shortcut. From the participants' comments, we derived that convenient boundaries for interaction are approximately the near and middle zones in each direction. Far zones turned out to cause more fatigue on arm and upper arm muscles.

4. IMPLICATIONS

Partition the space by layer thickness Our results indicate that the accuracy - measured as an error of absolute distance - of hand movement is not influenced by various sizes of participants' arms. For users with smaller arms, too many and, thus, thin layers would decrease the accuracy. On the other hand, for taller users with greater arm length, insufficient numbers of layers would result in greater traveling distances and, therefore, decreased efficiency. Hence, we propose to design the interaction space based on the layer's thickness. This way, the design results in different numbers of layers for different arm sizes, allowing the user to interact within the borders of their physical abilities.

Use an uneven layer thickness The traveling distance proved to be the most important factor. We observed that the typical overshooting error decreases towards outer regions. Therefore, we propose to use descending layer thicknesses towards outer regions. This allows for smaller layers in outer regions without increasing the interaction time that is introduced due to overshooting the target. Based on our quantitative results, a layer thickness of 7.8 cm for near, 4.2 cm for medium and 3.0 cm for far targets (the respective mean overshoot plus the double standard deviation) would result in >95% accuracy for all traveling distances.

Design with convenient boundaries in mind The qualitative feedback from participants showed that interactions in the far zones are less convenient compared to the closer regions. Therefore, we propose to focus on the near and medium zones for frequent and common interactions. As showed in [1], the slightly uncomfortable hand position in the far zones can be leveraged for important and not reversible actions.

5. APPLICABILITY

We believe that our results help to answer fundamental questions related to the design of proximity-based interaction in front of the user. We imagine a real world system implemented as a bracelet using an attached pico-projector along with an infrared proximity array as presented in [9].

Beyond palm-projected interfaces, proximity interaction can also be used in other domains such as head-mounted displays. For stereoscopic **AR glasses**, our approach can leverage the space in front of the user to present a layered information space (virtually projected on the user's palm). Additionally, touch input on the palm can provide direct manipulation possibilities on each layer. **Head-Up Displays** such as Google Glass present floating interfaces in the view of the user. Proximity interaction can act as a selector for different layers of the UI. The respective layout can be imaginary mapped to the palm's surface. Through the sense of proprioception, users can touch interface elements without visual attention to the hand. Similar to [5], the hand's 3D features can be extracted from a RGB-D attached to the head-mounted display.

6. CONCLUSION

We presented findings derived from a controlled experiment in which we focused on human capabilities for proximity-based hand input. The results confirmed the viability and feasibility of this input modality. The traveling distance to the target layer proved to be the main influence for the accuracy and the efficiency.

8. REFERENCES

[1] S. Benford, C. Greenhalgh, G. Giannachi, B. Walker, J. Marshall, and T. Rodden. Uncomfortable interactions. In *Proc. CHI '12*, page 2005. ACM Press, 2012.

[2] X. A. Chen, J. Schwarz, C. Harrison, J. Mankoff, and S. Hudson. Around-body interaction. In *Proc. MobileHCI '14*, pages 287–290, New York, New York, USA, Sept. 2014. ACM Press.

[3] N. Dezfuli, M. Khalilbeigi, J. Huber, F. Müller, and M. Mühlhäuser. PalmRC: imaginary palm-based remote control for eyes-free television interaction. In *Proc. EuroiTV '12*, pages 27–34. ACM Press, 2012.

[4] S. G. Gustafson, B. Rabe, and P. M. Baudisch. Understanding palm-based imaginary interfaces. In *Pro. CHI '13*, page 889. ACM Press, 2013.

[5] T. Ha, S. Feiner, and W. Woo. WeARHand: Head-worn, RGB-D camera-based, bare-hand user interface with visually enhanced depth perception. In *Proc ISMAR '14)*, pages 219–228. IEEE, Sept. 2014.

[6] C. Harrison and H. Faste. Implications of location and touch for on-body projected interfaces. In *Proc. DIS '14*, pages 543–552. ACM Press, 2014.

[7] C. Harrison, S. Ramamurthy, and S. E. Hudson. On-body interaction. In *Proc. TEI '12*, page 69. ACM Press, 2012.

[8] P. Mistry and P. Maes. SixthSense âĂŞ A Wearable Gestural Interface. In *Sketches. SIGGRAPH ASIA '09*, page 11. ACM Press, 2009.

[9] D. Ryu, D. Um, P. Tanofsky, D. H. Koh, Y. S. Ryu, and S. Kang. T-less: A novel touchless human-machine interface based on infrared proximity sensing. In *Proc. IROS '10*, pages 5220–5225. IEEE, Oct. 2010.

[10] M. Spindler, M. Martsch, and R. Dachselt. Going beyond the surface: studying multi-layer interaction above the tabletop. In *Proc. CHI '12*, pages 1277–1286. ACM Press, 2012.

[11] S. Subramanian, D. Aliakseyeu, and A. Lucero. Multi-layer interaction for digital tables. In *Proc UIST '06*, page 269, New York, New York, USA, Oct. 2006. ACM Press.

[12] J. Wagner, M. Nancel, S. G. Gustafson, S. Huot, and W. E. Mackay. A Body-centric Design Space for Multi-surface Interaction. In *Proc. CHI '13*, page 1299. ACM Press, 2013.

[13] A. D. Wilson and H. Benko. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proc. UIST '10*, page 273. ACM Press, 2010.

[14] C. Winkler, J. Seifert, D. Dobbelstein, and E. Rukzio. Pervasive information through constant personal projection. In *Proc. CHI '14*, pages 4117–4126. ACM Press, 2014.