Leveraging the Palm Surface as an Eyes-free TV Remote Control

Niloofar Dezfuli
Technische Universität Darmstadt
Hochschulstraße 10
64289 Darmstadt, Germany
niloo@tk.informatik.tu-darmstadt.de

Florian Müller
Technische Universität Darmstadt
Hochschulstraße 10
64289 Darmstadt, Germany
f_b.mueller@stud.tu-darmstadt.de

Mohammadreza Khalilbeigi
Technische Universität Darmstadt
Hochschulstraße 10
64289 Darmstadt, Germany
khalilbeigi@tk.informatik.tu-darmstadt.de

Max Mühlhäuser
Technische Universität Darmstadt
Hochschulstraße 10
64289 Darmstadt, Germany
max@tk.informatik.tu-darmstadt.de

Jochen Huber
Technische Universität Darmstadt
Hochschulstraße 10
64289 Darmstadt, Germany
jhuber@tk.informatik.tu-darmstadt.de

CHI 2012, May 5–10, 2012, Austin, TX, USA.
ACM 978-1-4503-1016-1/12/05.

Abstract
User input on television typically requires a mediator device such as a handheld remote control. While being a well-established interaction paradigm, a handheld device has serious drawbacks: it can be easily misplaced due to its mobility and in case of a touch screen interface, it also requires additional visual attention. Emerging interaction paradigms like 3D mid-air gestures using novel depth sensors such as Microsoft's Kinect aim at overcoming these limitations, but are known for instance to be tiring. In this paper, we propose to leverage the palm as an interactive surface for TV remote control. Our contribution is two-fold: (1) we have explored the conceptual design space in an exploratory study. (2) Based upon these results, we investigated the accuracy and effectiveness of such an interface in a controlled experiment. Our results show that the palm has the potential to be leveraged for device-less and eyes-free TV interactions—without any third-party mediator device.

Keywords
Omnipresent, input, TV, device-less, eyes-free, imaginary interface, controlled experiment

ACM Classification Keywords
H.5.2 [Information Interfaces And Presentation]: User Interfaces - Input devices and strategies.

General Term
Design, Human Factors, Experimentations

Figure 1. A user touches his non-dominant hand to operate the TV.

Figure 2. Sample on-screen user interface elements of a Samsung TV.

Introduction

User input on television (TV) is typically supported through remote controls. Common examples are conventional remotes with physical buttons or touch-based interfaces on smart phones. In effect, users are always required to utilize a particular *mediator device* to interact with the TVs. While this is a well-established interaction paradigm, it has various drawbacks. The device itself can be easily out of reach or misplaced. Moreover, touch-based interfaces on mobile devices require a lot of attention as they provide visual feedback. This distracts the user, since she needs to switch her attention between the device and the content on the TV. In summary, this increases the actual effort for controlling the TV and therefore, might diminish the user experience while watching.

Researchers have been investigating other input modalities to control TV systems. Speech [2] and mid-air gestures [3] are two well-known approaches for device-less and eyes-free TV interaction. However, these modalities still suffer from important drawbacks such as limitation in precision in unpredictable environment, privacy, efficiency and scalability. In addition, many studies showed that people find mid-air gestures somewhat uncomfortable and tiring [3].

Imaginary interfaces, as a novel device-less interaction approach [4], are conceptually based on the ability of people to attentively transfer their spatial memory on imaginary surfaces. Building on this work, Gustafson et al. designed an always-available imaginary phone [5], in which users could interact with their cell phone by recalling, mapping and touching different application icons on their hand. Although we focus on palm-based interactions without any visual attention to it, this prior work highly motivated our research agenda.

We propose a palm-based approach to operate TVs that is device-less, omnipresent and it provides eyes-free interactions with TV systems. In contrast to touch-based devices, we do not provide any visual feedback on the palm, but similar to the interaction style of conventional remote control, users receive feedback on the TV screen.

We explore our concept through two studies. In the first study, we exploratively investigate how users would perform basic interactions with TVs using their hand. The results of this study empirically elicit implications on how to design palm-based remote controls and lead to two main questions: (1) How *precisely* can people touch their palm landmarks and fingers without looking at them? (2) How *effectively* can they map the on-screen UI elements to their palm and touch them without visual attention? To answer these questions we conducted a controlled experiment in which we quantitatively examine how precise and effective users can interact with their palm landmarks without looking at them.

Explorative Study

Methodology

The first part of the study had a brainstorming character in which participants were asked to discuss high-level aspects of using the hand as a remote control (cf. Fig. 1). Particularly, we asked about how they would hold and which side and parts of their hand are suitable for interacting with the TV. Then, they were asked to envision and mimic the basic TV interactions on their hand such as navigating through lists. At the end, we utilized user interface screens of a Samsung Smart Internet TV and asked participants to directly interact with the elements displayed on the

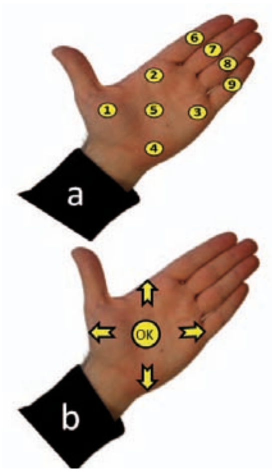

Figure 3. (a) The easily touchable landmarks on hand in eyes-free manner. (b) Participants transfer the mapping of directional keys using hand in diagonal orientation.

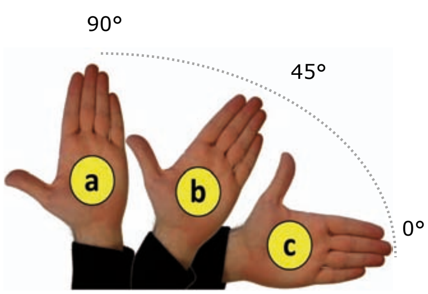

Figure4. (a) Portrait: pointing toward TV. (b) Diagonal: 45◦ to user's body. (c) Landscape: parallel to body.

screens by pointing to their hand surface (cf. Fig. 2). We recruited 10 volunteers (3f, 7m) in single user sessions (each 1h, think-aloud protocol, videotaped). Two coders coded the videos separately and iteratively.

Results

Leveraging the hands as TV remote control:

Unlike the one-hand usage of typical remote control all participants used their palmar (inner side) of the non-dominant hand as an input surface and interacted with the other hand's index finger similar to [5]. They said interacting with the palmar is not only more intuitive, but it also offers several salient regions (landmarks) to easily interact without any visual demand. P3 said *"I am able to* properly *touch any of my fingers as easy as* moving them.*"* and P8 added *"I can touch four curved areas (convex) on my palm surface even in the darkness".* Participants revealed nine landmarks on the hand surface, which they believed to be easily touchable without any visual demand based on the proprioceptive sense [1] (cf. Fig. 3a).

Mapping basic remote controls functionalities:

Participants mentioned that they would only map the functions on their hand, which are frequently used to operate the TV such as volume adjustment. In addition, they offered to properly map these functionalities to the location of landmarks of the palm as they can be easily hit without visual attention. For example, participants stated that mapping of directional keys could exactly match the four convex and one concave landmarks of the palm (cf. Fig. 3b). They also commented, since no digital information is projected on the palm surface, the simplicity of the design of the palm-based remote control is crucial.

Interacting with on-screen UI content:

Participants suggested direct interaction with the UI elements (e.g. media player controls with three buttons including backward, pause and forward). While looking at the TV, participants first mapped the whole screen of the UI to the non-dominant palm surface and then selected/triggered UI elements by pointing to the corresponding location on the palm surface.

They used three different hand orientations including diagonal, landscape and portrait to transfer the on-screen UIs to their palm (cf. Fig.4). The diagonal orientation was stated as the most comfortable form of holding the palm as an interactive surface. The 2D-touch gesture interactions and also interactions in which participants required to map remote control functions to their palm (such as directional keys) were mainly performed in diagonal orientation. However, participants transfer the grid-based vertical and horizontal UI screens to their hand while holding it in portrait or landscape orientations respectively.

Generally, participants appreciated the idea of being able to use the palm surface to operate the TV. The participants' comments highlighted the fact that designing TV UIs with the location of the palm landmarks in mind may be worthwhile. Palm-tailored TV UIs may decrease the cognitive effort of mapping these elements to the palm surface. P4 stated: "*If a menu could have four options, I could easily touch my middle finger to select the second option*".

Discussion

The results empirically ground the requirements and elicit implications for designing a palm-based remote control, which preserve user attention to the TV screen

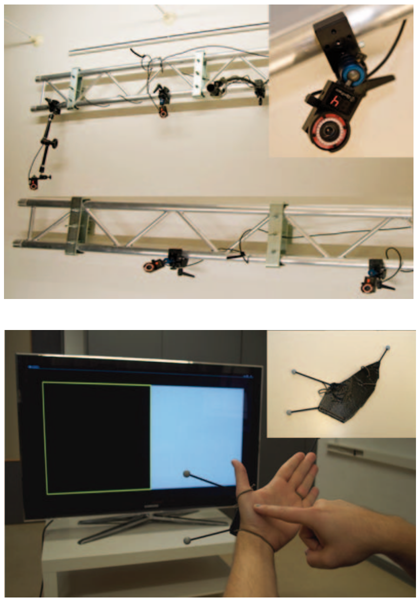

Figure 5. Top: OptiTrack system. Bottom: the paper carton apparatus used in the controlled experiment. We have attached three retro-reflective markers as antennas to the paper carton. These markers are then tracked by the OptiTrack system with 6 IR-cameras and define a 3D plane that corresponds to the hand surface. This allowed us to reliably track the hand without covering the hand completely, e.g. using a glove.

while interaction. We found 9 distinct landmarks on the palm surface, which can be easily touched without visual attention. Participants appreciated the way they can interact with the palm surface without any restriction. Based on the observations, we believe that due to the similar form factor, participants could easily transfer the TV UI screens to the palm surface. Considering the different orientations of the hand, the visualized interface elements on the TV screen can be tailored to the hand orientation. This enables user to easily switch between different menus based on the hand's orientation.

As a first step toward exploring the effectiveness and accuracy of palm-based interfaces for TV remote control, we focus on direct touch. The result discussed above, left us with two unexplored questions: (1) How *precisely* can people touch their palm landmarks and fingers without looking at them? (2) How *effectively* can they map the on-screen UI elements to their palm and touch them without visual attention?

Controlled Experiment

We have formulated the aforementioned questions as hypotheses and verified them in a controlled experiment. The two questions map to the following two hypotheses:

H1: People can touch their palm landmarks and fingers precisely without looking at them (0.90 confidence level).

H2: When mapping on-screen UI elements to palm,

H2.1: the effectiveness will decrease, the denser the UI elements are placed.

H2.2: the effectiveness is independent of the UI elements' alignment; i.e. whether they are horizontally or vertically aligned.

Effectiveness here means, whether a participant successfully touches a mapped UI element on her palm.

Experiment Setup

We have conducted the experiment using an optical tracking system (OptiTrack as shown in Fig. 5 top) to minimize any noise. We have designed a trackable paper carton apparatus, which the participants wore on the back of their non-dominant hand (cf. Fig 5 bottom). To allow for accurate touch input on the non-dominant hand, we placed another marker on the index finger of the dominant hand. A touch then is calculated by projecting the marker position on the palm plane and measuring the distance. We recruited 15 participants (5f, 10m; 32 years of age in average, with near-to-perfect sight). The participants were introduced to the system upfront. Each single-user session lasted about 45 minutes.

Methodology

We chose a within-subject design. The experiment was subdivided into two parts according to our hypotheses. Each part was again subdivided into two tasks (see Fig. 6). The order of the presented targets within each task was completely counterbalanced. The system advanced to the next target after each touch, regardless of whether the participant had successfully touched the target. We only repeated the trials in which the experimenter determined that the participant looked at her palm.

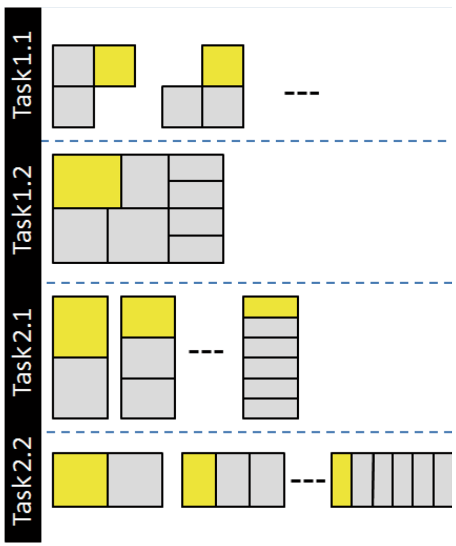

Figure 6. On-screen user interfaces of each task during the experiment.

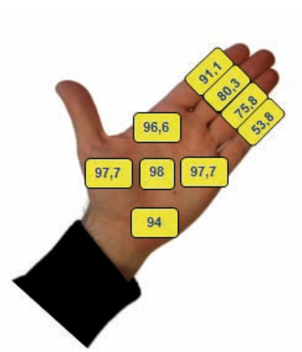

Figure 7. Average effectiveness percentage of targeting each landmark without visual attention.

Part 1: In the first part, participants were asked to touch landmarks without visual attention. Independent variable was the landmark location. Dependent variable was the success rate of a user touching the landmark on her hand. Task 1 was comprised of two sub-tasks.

- Task 1.1 required participants to map directional keys to their palm (see fig. 3.b), and navigate through a path of target items starting from the highlighted one (yellow box). For example, the first layout of task 1.1 in figure 5 required the participant to first touch left, then down. Participants had to touch 9 different landmarks.
- Task 1.2 required participants to map non-regular grids (see Fig. 6) to their palm surface and touch the highlighted position on it. Here, participants had to touch 8 different landmarks.

Part 2: In the second part, participants had to map and touch UI elements on their palm surface. Independent variable was the on-screen layout. Again, dependent variable was the success rate of a user touching the landmark corresponding to the UI element on her palm.

- Task 2.1 required participants to map vertical 1D regular grids to their palm surface and touch the highlighted position on their palm. Each user had to touch 20 different targets.
- Task 2.2 required participants to do the same with horizontal 1D regular grids, again for 20 different targets.

In a pilot study, we determined participants were able to divide and eyes-freely touch the palm surface up to 6 locations at most. Therefore, the task started with 2 adjacent targets and increased stepwise until 6.

Results

Each target was repeated 3 times, leading to a total of 2565 data points over all 15 participants: 15 * 3 * [9 (T1.1) + 8 (T1.2) + 20 (T2.1) + 20 (T2.2)]. We discarded 21 trials as outliers, since they were farther than 3 times the standard deviation away from the centroid. We normalized all palm sizes with the average index finger (7,31cm).

Part 1: Figure 8 shows the distribution of the raw data for tasks 1.1 and 1.2 by 90% confidence ellipses. This illustrates the spatial precision of the touches with respect to the centroid of each landmark. To analyze targeting, we measured one overall systematic error (offset). On average, the diameter necessary to encompass 90% of all touches is 28mm (SD= 0.85).

The average effectiveness for each landmark is shown in figure 7. All of the palm landmarks were effectively touched with at least 94%. The finger landmarks were less effectively touched with as little as 53% for the pinky. ANOVA tests revealed that the difference between palm and finger landmarks is statistically significant ($p<0.001$). Bonferroni post-hoc tests confirmed that this holds for all comparisons ($p<0.001$).

Part 2: The average effectiveness for the target elements is shown in figure 9. The effectiveness decreased monotonically for more than 3 menu options. The average effectiveness is below 90% for more than 4 menu options and decreases below 50% for more than 5 options. ANOVA with Bonferroni post-hoc tests revealed that these effects are statistically significant ($p<0.05$). The differences between horizontal and vertical alignments were not significant.

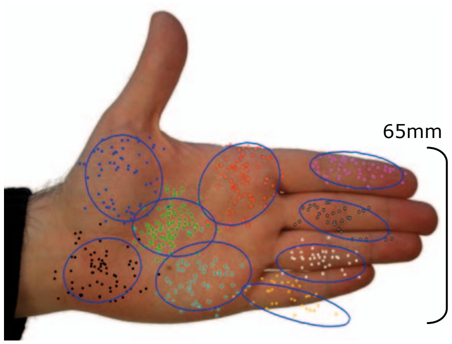

Figure 8. Distribution of the raw data of all participants by 90% confidence ellipses.

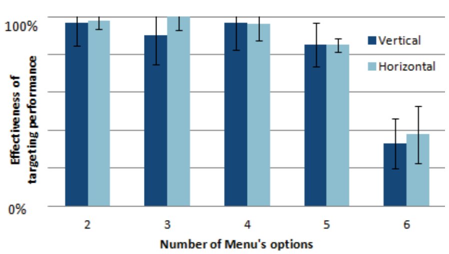

Figure 9. The average effectiveness of targeting vertical and horizontal grids with different equal-sized options.

Discussion and Conclusion

Using their proprioceptive sense [1], people can touch different landmarks of their palm surface and fingers while preserving their visual attention to other targets. This work focused on leveraging this by using the palm as an interactive surface for an eyes-free TV remote control. Concluding the results of two studies, we found that touching the 5 landmarks on the palm surface without any visual demand is highly effective. Moreover, it is precise enough to operate interfaces with target sizes of 28mm in diameter (H1). This implies that future palm-based TV interfaces should not map functions to regions with a smaller diameter. Moreover, this shows that users can effectively map common functions of traditional remote controls such as navigational keys to the landmarks of a palm and touch them to operate a TV.

Our results provide evidence that people can reliably and effectively (>90%) map 1D grid-layout menus with up to 4 options to their palm surface (H2.1), independent of whether the menu is horizontally or vertically aligned (H2.2). For future palm-based TV interfaces, we envision this to be leveraged as region-based shortcuts. While the participants were not as effective when touching their fingers compared to their palm landmarks, they effectively targeted their index finger. This indicates that also the index finger could be used as an effective input source.

In conclusion, our results show the palm has the potential to be leveraged as an interactive surface for TV interaction without any third-party mediator device.

Acknowledgements

This research was supported by the KCC (Korea Communications Commission), under the R&D program supervised by the KCA (Korea Communications Agency) "(KCA-2011-10035587/10921-03001)". We are grateful to Murat Özkorkmaz and Johann Raspasovic for their assistance.

References

[1] Sherrington CS. On the proprioceptive system, especially in its reflex aspect. In *Brain* 1906.

[2] Brutti, A. and Cristoforetti, L. and Kellermann, W. and Marquardt, L. and Omologo, M. WOZ Acoustic Data Collection for Interactive TV. In *Language Resources and Evolutions* 2010, 205-219.

[3] Freeman W.T., and Weissman, C.D. Television Control by Hand Gestures. In *Proc. Int. Workshop on Automatic Face and Gesture Recognition* 1995, 179-183.

[4] Gustafson, S., Bierwirth, D., and Baudisch, P. Imaginary Interfaces: Spatial Interaction with Empty Hands and Without Visual Feedback. In *Proc. UIST 2010*, 3-12.

[5] Gustafson, S., Holz, C., and Baudisch, P. Imaginary Phone: Learning Imaginary Interfaces by Transferring Spatial Memory from a Familiar Device. In *Proc. UIST* 2011, 283-292.

[6] Harrison, C., Benko, H., and Wilson, A. OmniTouch: wearable multitouch interaction everywhere. In *Proc. UIST* 2011, 441-450