PalmRC: Imaginary Palm-based Remote Control for Eyes-free Television Interaction

Niloofar Dezfuli Mohammadreza Khalilbeigi Jochen Huber Florian Benjamin Müller
Max Mühlhäuser
Technische Universität Darmstadt
Hochschulstraße 10
64289 Darmstadt, Germany
{niloo, khalilbeigi, jhuber, müller, max} @tk.informatik.tu-darmstadt.de

ABSTRACT
User input on television (TV) 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 to be e.g. tiring.

In this paper, we propose to leverage the palm as an interactive surface for TV remote control. Our contribution is three-fold: (1) we explore the conceptual design space in an exploratory study. (2) Based upon these results, we investigate the effectiveness and accuracy of such an interface in a controlled experiment. And (3), we contribute PalmRC: an eyes-free, palm-surface-based TV remote control, which in turn is evaluated in an early user feedback session. Our results show that the palm has the potential to be leveraged for device-less and eyes-free TV remote interaction without any third-party mediator device.

Categories and Subject Descriptors
H5.2. [Information interfaces and presentation]: User Interface-Input devices and strategies.

General Terms
Design, Human Factors, Experimentations.

Keywords
Alternative Remote Control, TV, Device-less, Eyes-free, Omnipresent, Input, Direct Touch, Non-visual, Memory.

1. INTRODUCTION
User input on television 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 [12]. Moreover, touch-based interfaces on mobile devices [20] 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 [14]. In summary, this increases the actual effort for controlling the TV and therefore may diminish the user experience while watching.

Figure 1. User taps on his non-dominant hand to operate the TV, while preserving his visual attention to the TV.

Researchers have been investigating other input modalities to control TV systems. Speech input [2, 6] and 3D mid-air gestures [12, 13] are two well-known approaches for device-less and eyes-free TV interaction [21]. However, these modalities still suffer from important drawbacks. Despite advantages of using natural language, the use of speech input is not always socially appropriate and the technology may fail to recognize commands in noisy environments. In addition, it is not well suited for common continuous interactions such as scrolling a channel list or adjusting the TV volume. As another approach, 3D gestures aim at overcoming these limitations, but are known to be e.g. tiring [12]. These drawbacks might explain why speech and 3D gestures as input modalities are still limited to lab environments and not yet widely deployed and included in home television environments.

In this work, we propose a novel approach to operate TVs. We leverage the non-dominant hand as an interactive input surface. Users can then operate the TV by interacting with the other hand's index finger (cf. Fig. 1). Therefore, the means to control a TV is naturally always available to viewers, without requiring any mediator device. Since humans are unconsciously aware of the relative position and orientation of one's own hands through proprioception [1], the palm can be appropriate for eyes-free interactions to operate the TV.

In this paper, we explore our concept in two studies. In the first study, being exploratory in nature, we aim at gaining insights into the conceptual space of palm-based remote controls. We particularly investigate different interaction styles and elicit implications on how to design such remote controls. Based on the results of this study, we designed a prototypical interface to conduct a controlled experiment. The major goal was to get a better understanding of the human capability of touching one's own palm without any visual attention. More precisely, the experiment aims at answering the following questions:

(1) How *precisely* can users touch their palm's salient regions (landmarks) without looking at them?

(2) How *effectively* can users select the target element of transferred on-screen user interface elements on their palm by pointing to the corresponding region on its surface without any visual attention?

The results of these studies gave us broad and deep insights into the conceptual design space. Building upon this, we designed and implemented PalmRC: an eyes-free, palm-surface-based TV remote control. Our prototype supports common tasks such as zapping through channels, menu navigation or social interaction between remote viewers. We evaluated PalmRC in early user feedback sessions.

The remainder of this paper is organized as follows. After reviewing related literature, we present the aforementioned studies and discuss the results. Then, we introduce PalmRC, interaction techniques and give a short technical overview. Finally, we report the results of our early user feedback sessions with PalmRC and conclude with an outlook upon future work.

2. RELATED WORK

While standard remote controls are still the most common interface for controlling TV systems at home, today's research is moving away from device-based interactions by observing and interpreting people's vocalization, movements and gestures with respect to the TV. In this section, we review research on device-less interaction with TVs and hand-based user interfaces. Our work is also inspired by prior work in the area of imaginary user interfaces, which we discuss at the end of this section.

2.1 Eyes-free and Device-less Interactions for Television Systems

Many studies investigated speech and 3D mid-air gestures input modalities to facilitate interaction with TV. Brutti et al. [2] presented a distant-talking interface for the interactive control of a TV set with multichannel acoustic data collection. Igarashi and Hughes [6] focused on direct control of interactive television by using non-verbal lower-level features of voice such as pith and volume. Although speech is a natural input modality, its usage is not always socially appropriate and the technology may fail to recognize commands in noisy environments.

Mäntyjärvi et al. [13] explored and generalized a possible set of gestures suitable for controlling home appliances such as a TV. They showed that 3D hand gestures lack an easy memorizable and universal vocabulary. Freeman and Weismann [12] have investigated how viewers can remotely control a television set by hand gestures without extensive user training and memorization. To do so, they provided visual feedback on the TV screen. This enabled users to move an on-screen pointer coupled to their hand to adjust various graphical controls. They reported that mid-air hand gestures are not appropriately recognizable for unpredictable scenes and suffer from scalability issues in group-watching experiences. In addition, their studies showed that people find mid-air gestures somewhat uncomfortable and tiring.

2.2 Hand-based Interactions

There are a number of wearable and mobile [7, 5, 9, 10] systems that leverage the surface of the hand and arm as always-available input systems.

KITTY [9] is a glove-type input device which covers parts of the hand with electronic contacts to enable touch event detection. An electric circuit is closed and a signal is generated upon closing of one finger-contact with one thumb-contact. This offers both speed and accuracy with a discrete signal input that is continuously ready and provides an ultra-portable solution for data input into portable computer systems. However, we focus on leveraging viewer's hand without any instrumentation such as gloves as this is not practical for TV rooms and also can mar the user experience while watching TV.

SixthSense [11] is a wearable camera-projector unit, which supports gestural manipulation of digital artifacts and augments physical surfaces with digital information. Moreover, it enables users to interact with the information projected on the surfaces in mobile contexts. While the system is superior to existing systems in terms of weight and size, the system uses color markers as artificial features that are placed on a user's fingertips to recognize hand gestures.

Skinput [8] presents a novel approach to recognize finger taps on the arm and hand by analyzing mechanical vibrations that propagates through the body. This is done using arrays of bio-acoustic sensors worn as an armband.

Brainy Hand [10] is another example of a wearable interaction device. It is equipped with a color camera, which captures an image of the user's hand to recognize its movements as input gestures. Since the digital data corresponding to each input gesture is projected as a picture onto the user's palm, it requires a lot of visual attention.

Recently, Harrison et al. presented OmniTouch [5], a wearable depth-sensing and projection prototype enabling multi-touch interaction by projecting on everyday surfaces such as a hand or an arm. They used depth sensing technology to track the hand and recognize whether a finger is hovering over or touched the hand surface. This work and the proposed algorithms inspired us when implementing the touch recognition on the hand surface.

The aforementioned research requires either a mediator instrument such as glove or visual attention to the input data. In contrast, we want to propose a device-less approach in which the visual attention can remain focused on the TV screen. To do this, we draw upon the concept of imaginary user interfaces. In this case, the actual interface elements are not projected onto the interactive surface (in our case the palm). They are just imagined by the user.

2.3 Imaginary User Interfaces

Imaginary interfaces, introduced as a new device-less interaction approach in [3], are based on a human's ability to map the spatial memory to physical surfaces. Here, no user interface is displayed on the surface but various sensing approaches are utilized to recognize on-surface interaction. Although, no information is projected on imaginary interfaces, the original concept requires users to look at their hands to define the origin of an imaginary space and attentively point and draw in the resulting physical space. Building on this work, Gustafson et al. designed an always-available imaginary phone [4], in which users can interact with their cell phone by recalling, mapping and touching different application icons on their hand attentively.

Figure 2. Example user interface screens of a Samsung TV used in the first study.

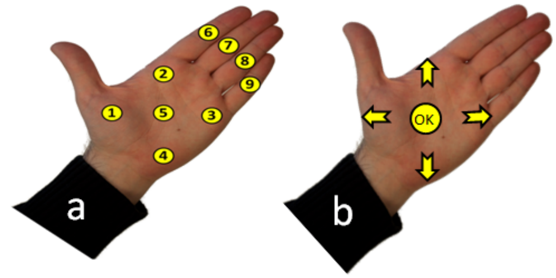

Figure 3. (a) The easily touchable landmarks on the hand. (b) Participants suggested linking the directional keys to the landmarks of the palm while holding the hand diagonally.

Although we aimed at designing interactions not requiring any visual attention to the hand, this prior work highly motivated our research agenda. We therefore combined the concept of proprioception with an imaginary user interface on a user's hand. In the following, we outline the design space and show how the hand surface can be leveraged for eyes-free interaction with television systems.

3. EXPLORATIVE STUDY

We conducted an exploratory study to empirically ground the requirements for designing an eyes-free, palm-based TV remote control. We were particularly interested to see how users would interact with their hand to perform a set of common interactions with TVs, while preserving their attention to the TV screen.

3.1 Design and Methodology

The study had a brainstorming character in which participants were asked to discuss high-level aspects of using the palm as a remote control (cf. Fig. 1). Initially, we asked about (1) how they would hold their hand and which side and parts of their hand would be suitable for interacting with the TV. Then, they were asked to particularly elaborate on (2) how they would transfer the remote control functions on their hand and (3) how they would interact with on-screen UI elements while mimicking their proposed interactions on their hand surface. To foster the discussion, we utilized and displayed some typical user interfaces of a Samsung Smart Internet TV on its screen and asked participants to show how they would interact with these elements using their hand. The user interface screens can be classified into three vertical, horizontal and whole screen grid-based menus (cf. Fig. 2).

We recruited 10 volunteer participants (3f, 7m). They were between 22-42 years old. All participants spent 2-3 hours in average per day watching TV. Each single-user session lasted about 1 hour. As data gathering methodologies, we videotaped the sessions and asked participants to think aloud. We then selected salient quotes and analyzed both quotes and videos using an iterative open, axial and selective coding approach [18]. For inter-coder independence, two coders coded the data separately.

3.2 Results

3.2.1 Using the Palm surface as a TV remote control

Generally, participants appreciated the idea of being able to use the palm surface for operating the TV. Unlike the one-hand usage of typical remote controls, 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 palm surface, which they believed to be easily touchable without any visual demand based on the proprioceptive sense [1] (cf. Fig. 3a).

3.2.2 Mapping basic remote controls functionalities

Participants mentioned that they would only map frequently used functions to their palm such as navigation, selection, digits for direct switching between channels, volume adjustment, or play and pause. In addition, they offered to properly map these functionalities to the location of landmarks of the palm, since they can be easily hit without any visual attention. For example, participants stated that the mapping of directional keys could exactly match the four convex and one concave landmarks of the palm (cf. Fig. 3b).

In contrast, recalling and transferring digits (typical mapping of 3x3 buttons of digits from 1 to 9) to the palm was found to be very complex. P5 said *"Digits may have a conventional mapping but still they lack having a natural mapping and I would prefer to draw digits on my palm to change the channels"*. P7 added: *"Even if I could recall each digit position, I would not know where to map it on the palm surface as no landmarks afford their mapping"*.

Participants also commented, since no digital information is projected on the palm surface, the simplicity of the design of a palm-based remote control is crucial.

3.2.3 Interacting with on-screen UI content

Participants not only suggested 2D-touch gestures (e.g. swipe, scroll, and draw) on the palm, but they also proposed mapping UI elements displayed on the TV screen to the palm's surface. They then imagined triggering the target elements by pointing (tap, click) to the corresponding location on the palm surface. For this purpose, participants used three different hand orientations including diagonal, landscape and portrait (cf. Fig.4).

The diagonal orientation was stated as the most comfortable form of holding the hand as an interactive surface. The interactions requiring participants to map remote control functions to their palm (such as directional keys), as well as 2D-touch gestures, were mainly performed in diagonal orientation.

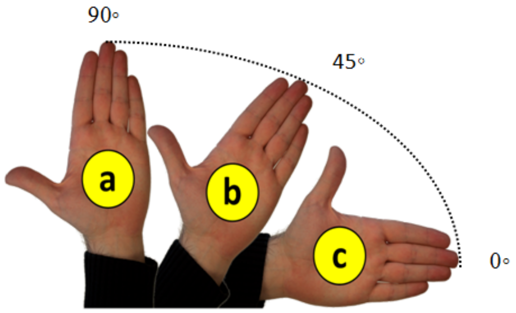

Figure 4. (a) Portrait: pointing toward TV. (b) Diagonal: 45◦ to user's body. (c) Landscape: parallel to body.

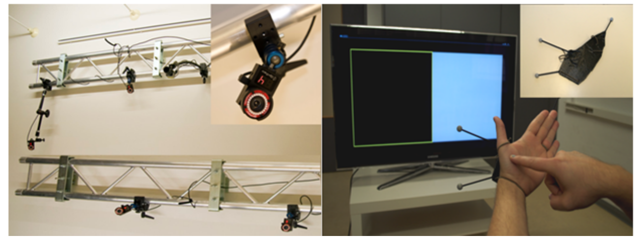

Figure 5. Left: OptiTrack system. Right: the paper carton apparatus used in the controlled experiment.

3.2.4 2D-touch gesture interaction on palm surface

Although a palm is not a flat planar surface, participants considered it as a concrete surface and proposed using 2D touch gestures on it while holding it in diagonal mode. This interaction technique was typically proposed for either efficiently browsing menus with a plethora of options, or mimicking digits on the palm surface for channel navigation, or even nonverbal communication between remote viewers; as P3 stated: *"I could for instance draw a smiley on my palm surface and send it to my online friends who are watching the same program"*.

3.2.5 Pointing on the palm surface

Participants suggested to transfer one-dimensional grid-based UI elements (e.g. list of applications or media player controls with three buttons including backward, pause and forward) onto the palm surface. While looking at the TV, participants first mapped the whole screen of the UI to the non-dominant hand surface and then selected/triggered UI elements by pointing to the corresponding location on the hand surface using the index finger of their dominant hand. Participants transferred the grid-based vertical and horizontal UI screens to their palm while holding it in portrait or landscape orientations respectively.

Participant's comments highlighted the fact that the design of TV UIs elements based on the location of the palm landmarks may improve the mapping. P4 stated: "*If a menu could have four options, I could easily touch my middle finger to select the second option*". Discussion with participants revealed that hand-tailored TV UIs may decrease the cognitive effort of mapping these elements to the surface of palm and eventually results in more secured feeling of hitting appropriate location on the palm while looking at the TV.

3.3 Discussion

The results of this study empirically elicit implications for designing a palm-based remote control, which preserves a user's attention to the TV screen during interaction. We found 9 distinct landmarks on the palm surface which can be easily touched without visual attention. The main benefit of this is that it allows TV viewers to link the common functions of a remote control (e.g. directional keys) to these landmarks for eyes-free TV interaction.

Since no digital information is to be projected on the palm, participants also appreciated the way they can interact with the palm surface. Based on our observations, we believe that due to the similar form factor of the palm surface and the TV screen, participants could easily imagine the UI elements on TV screen to their palm surface and touch the corresponding location of the target elements. Considering the different orientations of the hand, the visualized interface elements on the TV screen can be tailored to the hand orientation. This enables users to easily switch between different menus based on the orientation of the hand.

The results discussed above, left us with two unexplored questions: (1) How *precisely* can users touch their palm's salient regions (landmarks) without looking at them? (2) How *effectively* can they select the target element of transferred on-screen user interface elements on their palm by pointing to the corresponding region on the palm surface without any visual attention?

4. 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 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.

4.1 Experiment Setup

We have conducted the experiment using an optical tracking system (OptiTrack[1] as shown in Fig. 5 left) 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 right). 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 palm surface. This allowed us to reliably track the palm without covering the palm completely, e.g. using a glove. To allow for accurate touch input on the non-dominant hand, we have augmented the index finger of the dominant hand with another marker. A touch then is calculated by projecting the marker position on the hand 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.

4.2 Methodology

The experiment was subdivided into two parts according to our hypotheses. Each part was again subdivided into two tasks (cf. Fig. 6). The order of the presented targets within each task was completely counterbalanced. The system advanced to the next

[1] http://www.naturalpoint.com/optitrack/

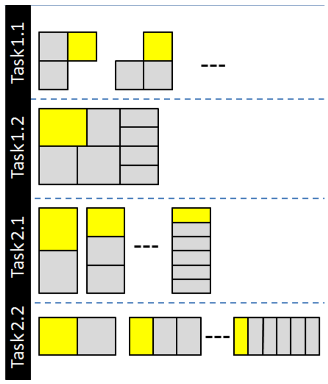

Figure 6. On-screen user interfaces of each task during the experiment.

Figure 7. Distribution of raw data of all participants by 90% confidence ellipses.

target after each touch, regardless of whether the participant had successfully touched the target. We chose a within-subject design. We only repeated the trials in which the experimenter determined that participant looked at her palm.

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 palm. 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 their palm. 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 order to determine boundaries for the number of targets in this task, we conducted a pilot study. We ask participants to target elements in various density levels starting from 2 adjacent targets in both horizontal and vertical orientations. We determined that 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 stepwise became denser until 6 targets as depicted in Fig. 6, task 2.2.

4.3 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 hand sizes with the average index finger (7,31cm).

Part 1: Figure 7 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 8. 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 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.

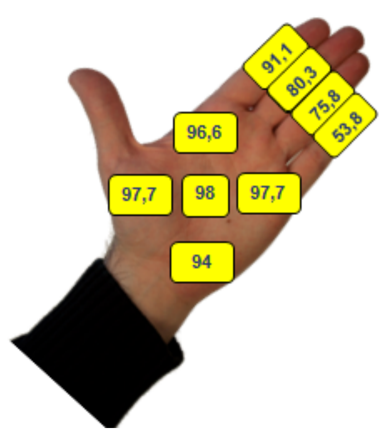

Figure 8. Average effectiveness percentage of targeting each landmark without visual demand.

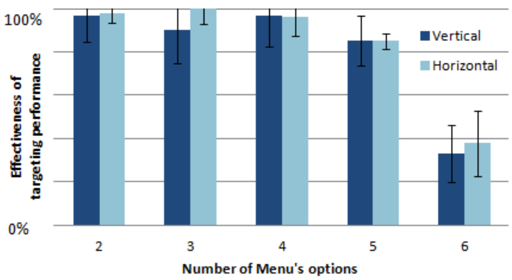

Figure 9. The average effectiveness of targeting vertical and horizontal grids with different equal-sized options.

4.4 Discussion

Based on the results of the studies, we showed 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.

5. SUMMARY OF TWO STUDIES

The results of the studies show that:

- Users preferred to transfer typical remote control functionalities such as directional keys to the palm (inner side) of their non-dominant hands. We also found out that the palm offers nine salient regions (landmarks), which can be easily recognized and touched without requiring any visual attention.
- Users preferred 2D touch gestures such as swiping on the palm surface for efficient browsing of lists with so many options. Our findings also revealed that users utilized the palm surface as a canvas to draw short symbols such as digits or emoticons.
- The landmarks can be touched precisely enough for TV interaction if the size of targets is considered sufficiently large about 28mm (SD= 0.85) in diameter on the palm surface to encompass 90% of all touches.
- Users can reliably and effectively (>90%) map one-dimensional (1D) grid-layout menus with up to 4 options to their palm surface, independent of whether the menu is horizontally or vertically aligned.

The results discussed above show that users are able to use their palm to interact with TV without visual attention in two main ways: first, as a remote control with several functions (virtual buttons) that are linked to the landmarks and second, as an unique input surface which the television user interface is mapped to the entire surface of the palm. Our findings showed that under certain

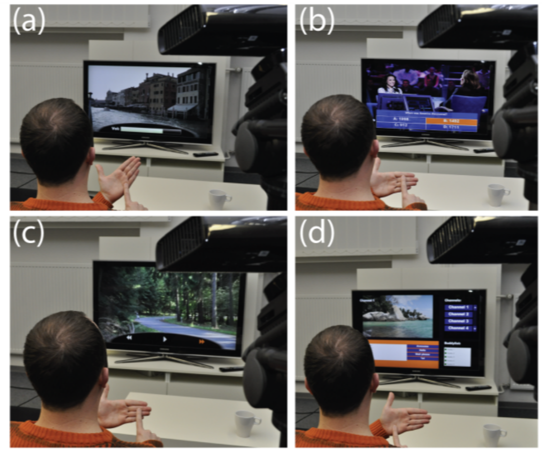

Figure 10. PalmRC system in different application scenarios.

circumstances (28mm button size and 4 target options) the palm-based remote control is viable. Thus, frequently used functions can be ready at palm, virtually any time without need for an additional mediator device.

Building upon these results, we developed a palm-based remote control called PalmRC along with two main interaction techniques. In addition, we implemented several applications to show the usefulness of PalmRC for varying interaction scenarios with TV systems.

6. PalmRC

PalmRC allows users to operate the TV using empty hands while focusing their visual attention on TV screen. The users interact by pointing and swiping on their non-dominant hand and the system enables the palm's surface to be capitalized as an input surface. The TV system receives touch positions and returns appropriate visual feedback on its screen. We developed interaction techniques to perform conventional TV interactions such as channel navigation in Electronic Program Guide (EPG), volume adjustment and direct interaction with menu options. PalmRC enables users to use their palm for various typical commands instead of retrieving a TV remote control.

Based on the orientation of the hand (see Fig. 4), the PalmRC supports two main interaction techniques (modes) that make use of pointing and 2D-touch gesturing on the palm surface:

6.1.1 Linking Functions to the palm's Landmark

Based on the results of the first study, the diagonal orientation of the non-dominant hand was found to be comfortable and resembles the style of holding a remote control in hand. Therefore, in this orientation, PalmRC links the common buttons of the remote control to the 9 landmarks of the palm. Users can trigger buttons by touching the corresponding locations on their palm.

We implemented this mode for directional keys and a confirmation/menu button (as the most frequent used buttons). These are in turn linked to the landmarks of the palm, as revealed in our studies respectively. This technique also allows for a natural spatial mapping between the buttons and the landmarks. Users can zap through the TV channel by tapping on the corresponding landmarks, which are mapped to the up or down keys. The volume can be adjusted similarly by touching the locations of the right and left keys (cf. Fig. 10-a). To open up the channel list or menu options users can touch the center of their palm surface. Similar to touch-enabled devices, swiping upwards

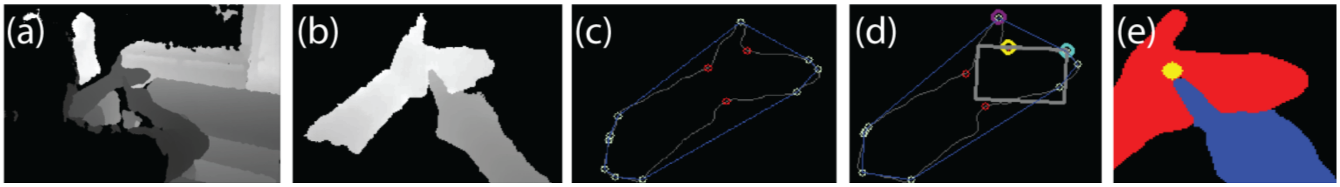

Figure 11. The process of recognizing a touch event in PalmRC: (a) original depth image, (b) subtracted background, (c) determining the contour of the reference hand and (d) its palm box, (e) and finally recognized touch on the surface.

or downwards on the palm surface allows for a fast browsing of the channel list. Users can also directly switch to another channel by drawing its number on the palm surface.

6.1.2 Direct Interacting with Interface Elements

By holding the hand in portrait or landscape mode, PalmRC maps the user interface screen to the entire palm surface. Then, users can touch the corresponding location of a target element on the palm. This interaction mode allows users to directly select a target on the TV screen.

We showcased this technique in a social interactive television interface, which incorporates common social features such as live chat (cf. Fig. 10-d). Once users hold their hands in landscape orientation, the communication mode will be activated and they can directly select and interact with one of the options. We also integrated this interaction technique in an application enabling remote viewers to answer questions of a quiz [15] by pointing to the appropriate location of their palm (cf. Fig. 10-b). The technique provides quick and immediate interactions with the social TV interface.

As another application example, while watching a movie or a program, users can hold their hands in landscape orientation. Thus, the media player menu including three options as backward, pause/play and forward appears on the TV screen. Then users can map it to the palm and touch the corresponding location of the desired option (cf. Fig. 10-c).

6.2 Technical Overview

Although the OptiTrack motion capture system used in the controlled study enabled us to precisely track the palm and recognize the touch position, it is not practical for TV room settings. As we reviewed in the related work section, there have been other sensing approaches such as using gloves [10], Skinput [9], depth cameras [6]. We chose to use a Microsoft Kinect [17] depth camera because it does not need any instrumentation on the hand of the viewers and also enables and supports recognizing touch and drag interactions [6, 5].

In PalmRC, we use the Kinect depth camera to track the non-dominant hand and recognize touch and dragging events with the index finger of the dominant hand. The built-in depth sensors recognize a user's hands in a minimum distance of ~50cm. Currently, we mount the depth camera on a tripod located at the back of a user's shoulder (cf. Fig. 10). We envision the future depth cameras to be small and precise enough to be either unobtrusively worn, or to be integrated into living room furniture. Touch events are recognized in a multistep process similar to [4]. The process is depicted in Fig. 11. In order to subtract the background, we first find the closest pixels in the raw image and remove all other relative depth values greater than 40cm. We classify the depth values of each hand by calculating the number of peaks in a histogram of all depth values (cf. Fig. 11-b). To track the non-dominant hand, we then calculate the contour and the convex hull of the hand including convexity defects (red points) and convexity start points (blue points) depicted in Fig. 11-c. The palm box is then calculated based on the prominent defect and the start point (illustrated with yellow and light-blue circles in Fig. 11 -d accordingly). To determine if and where the touch occurs, we compare the depth values of the finger tip with the surrounding values in the hand box. If the finger tip gets close enough to the reference hand, a touch event will be recognized. Due to the local noises and low-resolution of the Kinect depth camera, the precise end of finger tip is not fully recognizable. Similar to [4], we determine the touch location by offsetting a small vector in the direction of the finger (yellow circle in Fig. 11-e).

Although the tracking approach requires users to hold their thumb upright while using PalmRC, it robustly recognizes different orientation of the non-dominant hand. Future work is needed to improve the hand tracking and touch recognition so that users can arbitrary hold their hands.

7. EARLY USER FEEDBACK

In order to get initial user feedback on the interaction techniques, we conducted a focus group with 5 participants (26 years of age in average, all male and right handed with near-to-perfect sight). None of them had participated in our first studies. We were mainly interested in observing how they would use PalmRC, therefore allowing us to identify strengths and weaknesses of the overall system. Moreover, we wanted to get an initial impression on the usability of PalmRC. After a brief introduction to both, the concept and interaction techniques, the participants explored the interactions while thinking aloud.

The discussion with participants revealed that they found the concept of eyes-free palm-based remote control appealing. It received positive feedback for several reasons:

- Participants were able to quickly understand the interaction techniques. P5 stated that "*I don't need to think about how to interact or look for widgets on my palm*".
- Participants found our concept very practical in terms of providing immediate shortcuts for TV interactions. P3 commented that "*This is great! I can turn the volume down quickly, for instance when my phone is ringing. And I don't need to look for the remote control everywhere, anymore*".
- Three participants, who had used mid-air hand gestures before at home, compared the PalmRC interaction style with 3D mid-air gestures. They stated that they feel PalmRC interactions are less tiring. They added that these types of interactions seem to be more comfortable than 3D hand gestures since they are relaxing or even suitable in other body postures such as lying on the couch in front of the TV.

One participant was concerned with the two-handed usage of the system, in contrast to the one-handed usage of typical remote control. He commented: *"With my remote control at home, I can* *control the TV while I'm holding a glass with the other hand"*. We believe that this issue will become less severe by extending the palm-based remote control to the surface of other parts of the body such as thigh, which affords one-handed interaction. Moreover, the Kinect depth camera has opened up new interactive experiences on any un-instrumented physical surface [5, 18, 19]. Future work should also investigate leveraging physical surfaces around users such as couch arms or tables as an input surface to operate TVs.

The discussion with participants also revealed that they were unsure about the effectiveness of transferring and interacting with more complicated UI elements on the TV such as an EPG. They emphasized that they would like to use PalmRC for sending frequent commands to operate the TV instead of a standard remote control.

8. CONCLUSION AND FUTURE WORK

In this paper, we explored the concept of leveraging the palm surface as an eyes-free remote control. Through an explorative study, we qualitatively gained insights into how people would use their palm as if it were a remote control. Results confirmed that users are able to touch several prominent regions of their palm without looking at them. In a controlled experiment, we quantitatively determined how precisely they could interact with these regions in an eyes-free manner. We also investigated the effectiveness of using the palm as an input surface for directly interacting with on-screen user interface elements. The findings showed that under certain circumstances (28mm button size and 4 target options) the palm-based remote control is viable.

Building upon the results, we introduced PalmRC as a prototypical realization of our concept. We designed two main interaction techniques and showcased them in different application scenarios. The early user feedback underlined the fact that PalmRC offers an always-available shortcut for performing frequently used interactions with TV systems without needing to find and retrieve any mediator.

Based on our results, we conclude that by leveraging different landmarks of the palm, users are able to perform precise interactions, while preserving their visual attention to the TV. Moreover, the palm surface is also appropriate for coarse gestures such as swiping. Initial results suggest that the interaction style of PalmRC is less tiring than mid-air gestures. However, more studies are needed to systematically compare both as device-less input modalities for television systems.

As a next step, we want to investigate more deeply how PalmRC can enrich the user experience while watching TV in real world settings. Finally, we will further explore the scalability of our concept to handle conflicts when people are watching TV together in co-located settings.

9. ACKNOWLEDGEMENT

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.

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

[4] 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.

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

[6] Igrashi, T. Hughes, J.F., Voice as sound: using non-verbal voice input for interactive control . In *Proc.* UIST 2001, 10.1145.

[7] Kohli, L. and Whitton, M. The Haptic Hand: providing user interface feedback with the non-dominant hand in virtual environments. In *Proc.* GI 2005, 1–8.

[8] Harrison, C., Tan, D. and Morris, D. Skinput: appropriating the body as an input surface. In *Proc.* CHI 2010, 453–462.

[9] Kuester, F., Chen, M., Phair, M.E. and Mehring, C. Towards keyboard independent touch typing in VR. In *Proc.* VRST 2005, 86–95.

[10] Tamaki, E., Miyaki, T. and Rekimoto, J. Brainy Hand: an earworn hand gesture interaction device. In *Proc.* CHI Extended Abstracts 2009, 4255–4260.

[11] Mistry, P., Maes, P., and Chang, L. WUW - wear Ur world: a wearable gestural interface. In *Proc*. CHI Extended Abstracts 2009, 4111–4116.

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

[13] Mäntyjärvi, J., Kela, J., Korpipää, P., Kallio, S. Enabling fast and effortless customization in accelerometer based gesture interaction. In *Proc.* Mobile and ubiquitous multimedia 2004, 25-31.

[14] Weisz , J.D., Kiesler, S., Zhang, H., Ren, Y., Kraut R.E, Konstan, J.A. Watching together: integrating text chat with video. In *Proc.* CHI 2007, 877-886.

[15] Luyten, K., Thys, K., Huypens, S., Coninx, K. Telebuddies: social stitching with interactive television. In *Proc.* CHI Extended Abstracts 2006, 1049-154.

[16] Strauss, A., Corbin, J. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. *Sage Publications*, 2008.

[17] Microsoft Kinect, http://www.xbox.com/kinect.

[18] Wilson, A.D. and Benko, H. Combining Multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proc.* UIST 2010, 273-282.

[19] Wilson, A.D. Using a depth camera as a touch sensor. In *Proc.* ITS2010, 69-72.

[20] Robertson, S., C. Wharton, C., Ashworth, C., Franzke, M. Dual device user interface design: PDAs and interactive television. In *Proc.* CHI 1996, 79-86.

[21] Zimmermann, G., Nixon, T., Beard, M., Sitnik, E., LaPlant, B., Trewin, S., Laskowski, Sh. and Vanderheiden, G. Toward a unified universal remote console standard. In *Proc. CHI* Extended Abstracts 2003, 874-875.