Behaviour & Information Technology, 2014
Vol. 33, No. 8, 829–843, http://dx.doi.org/10.1080/0144929X.2013.810781

Taylor & Francis
Taylor & Francis Group

PalmRC: leveraging the palm surface as an imaginary eyes-free television remote control

Niloofar Dezfuli*, Mohammadreza Khalilbeigi, Jochen Huber, Murat Özkorkmaz and Max Mühlhäuser

Computer Science, TU-Darmstadt, Hochschul str. 10, Darmstadt 64289, Germany

(*Received 8 November 2012; accepted 28 May 2013*)

User input on television (TV) typically requires a mediator device such as a handheld remote control. While this is 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 such as 3D mid-air gestures using novel depth sensors (e.g. Microsoft Kinect), aim at overcoming these limitations, but are known to be tiring. In this article, we propose to leverage the palm as an interactive surface for TV remote control. We present three user studies which set the base for our four contributions: We (1) *qualitatively* explore the conceptual design space of the proposed imaginary palm-based remote control in an explorative study, (2) *quantitatively* investigate the effectiveness and accuracy of such an interface in a controlled experiment, (3) identified *user acceptance* in a controlled laboratory evaluation comparing PalmRC concept with two most typical existing input modalities, here conventional remote control and touch-based remote control interfaces on smart phones for their user experience, task load, as well as overall preference, and (4) contribute PalmRC, an eyes-free, palm-surface-based TV remote control. Our results show that the palm has the potential to be leveraged for device-less eyes-free TV remote interaction without any third-party mediator device.

Keywords: alternative remote control; TV; device-less; eyes-free; omnipresent; input; direct touch; non-visual; imaginary

Subject Classification Code: H5.2.

Information interfaces and presentation: user interface, input devices and strategies

1. Introduction

With the evolution of interactive television (iTV) and the ever increasing number of TV services and applications, enhancing TV interactions in today's living rooms has become the focus of previous research. Despite this research on novel interaction modalities and techniques (Zimmermann *et al.* 2003, Bernhaupt *et al.* 2008), TV interaction is predominantly supported through remote controls. Common examples are button-based conventional remotes or touch-based interfaces on smart phones. Thus, users are always required to utilise a particular *mediator device* to interact with the TV. While admittedly a well-established interaction paradigm, it has various drawbacks. On the one hand, the device itself can be out of reach or misplaced or even lost (Freeman and Weissman 1995). In addition, users typically have to deal with several remote controls for different home entertainment devices each with an excessive number of functions assigned to various physical-buttons. This makes remote controls even more complicated and confusing than before (Bernhaupt *et al.* 2008). On the other hand, touch-based interfaces on mobile devices (Cesar *et al.* 2008) require a lot of attention and users have to constantly switch their attention between the device and the content on the TV (Weisz *et al*. 2007). This increases a user's effort for controlling the TV and therefore diminishes the user experience while watching.

Research on device-less TV interaction (Freeman and Weissman 1995, Igrashi and Hughes 2001, Brutti *et al.* 2008) strives to overcome the aforementioned limitations. Prominent examples are speech input (Igrashi and Hughes 2001, Brutti *et al.* 2008) or three-dimensional (3D) mid-air gestures (Freeman and Weissman 1995, Mäntyjärvi *et al.* 2004). These input modalities prompt viewers to engage with TV systems in human-like conversation. However, they still suffer from severe drawbacks: the use of speech input is not always socially appropriate and the technology may fail to recognise commands in noisy environments (Brutti *et al.* 2008). In addition, it is not well suited for common continuous interactions such as scrolling a channel list or adjusting the TV volume (Igrashi and Hughes 2001). 3D gestures aim at overcoming these limitations, but are known to be, e.g. tiring (Freeman and Weissman 1995). These drawbacks might explain why speech and 3D gestures as input modalities are still limited to laboratory environments and not yet widely deployed and included in home TV environments.

*Corresponding author. Email: niloo@tk.informatik.tu-darmstadt.de

In this work, we propose PalmRC, a novel imaginary interface to operate TV systems. PalmRC transforms the palmar of the non-dominant hand into an interactive input surface. Users can then operate the TV through touching the palmar with the other hand's index finger. PalmRC builds on the sense of proprioception (Sherrington 1907): humans are unconsciously aware of the relative position and orientation of their own hands. In particular, the palm can be appropriated for eyes-free TV interaction. PalmRC, therefore, neither demands a user's visual attention nor requires an additional mediator device.

We explore the concept of PalmRC in a series of user 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 moreover conducted a controlled experiment to investigate the human capability of touching owns palm without paying any visual attention to it. More precisely we aimed to quantitatively answer the following 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 its surface without any visual attention?

We then conducted a third study in which we compared PalmRC to conventional button-based and touch screen-based remote controls. In this study, we particularly focused on identifying respective advantages and disadvantages of each input modality.

The results of all three studies provided deep and broad insights into the conceptual design space and clarified concrete design questions like precision, effectiveness and user experience pertaining to the concept of PalmRC. As a proof of concept, we designed and implemented a functional prototype using depth sensing technology (Microsoft Kinect 2013). Our prototype supports common tasks such as zapping through channels, menu navigation or social interaction between remote viewers.

The remainder of this paper is organised as follows. We first review related work. We then present the three studies and discuss their results. Finally, we present a prototypical implementation of PalmRC and a set of interaction techniques, followed by a conclusion and an outlook upon future work.

2. Related work

We categorised user interfaces for TV input based on the required degree of visual attention while interacting with the TV: attentive vs. eyes-free interfaces. In this section, we first review prior work related to eyes-free user interfaces such as voice and 3D mid-air gestures. We then discuss related work in the field of attentive input modalities such as device-based remote controls and body-based wearable systems.

2.1. Eyes-free interfaces

There is large body of research investigating speech and 3D mid-air gestures for TV interaction. Brutti *et al.* (2008) presented a distant-talking interface for the interactive control of a TV set with multichannel acoustic data collection. Igrashi and Hughes (2001) focused on direct control of iTV by using non-verbal low-level features of voice such as pitch and volume. Although speech is a natural input modality, its usage is not always socially appropriate. Furthermore, technology may fail to recognise commands in noisy and unpredictable acoustic environments. Besides being inefficient and not well-scalable, it is also not well-suited for common continuous interactions such as scrolling a channel list or adjusting the TV volume.

Many studies aimed at overcoming these limitations and investigated how viewers can control TV using 3D hand gestures. Freeman and Weismann (1995) have investigated how viewers can remotely control a TV set by hand gestures without extensive user training and memorisation. 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. Mäntyjärvi *et al.* (2004) explored a possible set of gestures suitable for controlling home appliances such as a TV. They showed that 3D hand gestures lack an easy memorisable and universal vocabulary. They reported that mid-air hand gestures are not appropriately recognisable for unpredictable scenes and suffer from scalability issues in group-watching experiences. In addition, their study showed that people find mid-air gestures somewhat uncomfortable and tiring (*fatigue problem*) as well as that they critiqued the lack of haptic feedback. Overall, these drawbacks might explain why speech and 3D gestures as input modalities are still limited to laboratory environments and not yet widely deployed and included in home TV environments.

Recently, researchers have started to investigate the usage of proxemic and spatial information of users to mediate interaction with digital devices. Ballendat *et al.* (2010) proposed proxemic interactions in ubiquitous environments based on interpreting spatial relationships of persons, objects and digital devices. They particularly investigated how fine-grained proxemic knowledge (such as user's presence, position, posture and orientation) can be exploited to design interaction techniques with surrounding digital devices. Similarly, Dezfuli *et al.* (2012) further investigated the concept of proxemic interaction for living room settings. Based on the results of their study, they designed various implicit interaction and awareness techniques to facilitate (re-)engagement for watching activities.

In this work, however, we focus on using the human's sensory system as an omnipresent and eyes-free input channel to facilitate more *fine-grained* TV interactions.

2.2. *Attentive input modalities*

Given the wide adoption of high-tech smart phones, most of modern TV manufacturers enable users to install and use applications featuring a virtual remote control to operate TVs (Samsung Smart TV 2012). They, however, require significant visual attention, as they lack tactile cues. As a result, most of the previous work has focused on appropriating everyday objects such as tables, pillows or paper interface to foster customised and personalised ways for TV interaction (Berglund *et al.* 2006, Hess *et al.* 2008, Brookstone 2013). Although inspiring and novel, these input modalities require user's visual attention similar to touch-based interfaces on mobile devices. This may distract the user since she needs to constantly switch her attention between the device and the content on the TV.

There are a number of wearable and mobile (Kohli and Whitton 2005, Kuester *et al.* 2005, Mistry *et al.* 2009, Gustafson *et al.* 2011, Harrison *et al.* 2011, 2012) systems that leverage the surface of the hand and arm as an always-available input system.

KITTY (Kuester *et al.* 2005) 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.

SixthSense (Wilson and Benko 2010) is a wearable camera-projector unit supporting gestural manipulation of digital artefacts. It augments physical surfaces with digital information and enables users to interact with projected information in mobile contexts. While the system is superior to existing systems in terms of weight and size, the system uses colour markers as artificial features which are put on a user's fingertips to recognise hand gestures.

Skinput (Harrison *et al.* 2010) presents a novel approach to recognise finger tap on arms and hands by analysing mechanical vibrations that propagate through the body. The system uses arrays of bio-acoustic sensors which need to be worn as an armband. Brainy Hand (Tamaki *et al.* 2009) is another example of a wearable interaction device. It is equipped with a colour camera, which captures an image of the user's hand to recognise 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 (Harrison *et al.* 2011), a wearable projection-based prototype, enabling multi-touch applications on everyday surfaces including the body. They used depth sensing technology to track a hand and recognise whether a finger has hovered over or touched the hand surface. This work and the proposed touch recognition algorithm inspired the design of PalmRC prototype. Similar in nature, Armura (Harrison *et al.* 2012) is an interactive on-body projection system that supports both input and graphical output on a user's arms and hands. The authors explored the design space of arm-driven user interfaces by proposing various synergistic arm gestures and atop of that, developed several interaction techniques and applications.

The aforementioned research requires either a mediator device or visual attention. In this article, we focus on leveraging a 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. In contrast, we propose a device-less approach in which the visual attention remains focused on the TV screen. To do this, we draw upon the concept of imaginary user interfaces in which the actual interface elements are not visually projected onto the interactive surface (in our case the palm). They are just imagined by the user. Imaginary interfaces, introduced as a new device-less interaction approach in (Gustafson *et al.* 2010), 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 utilised to recognise on-surface interactions. 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 (Gustafson *et al.* 2011), where users can interact with their cell phone by recalling, mapping and touching different application icons on their hand attentively.

These prior works highly motivated our research. Our work goes beyond what has been proposed in this line of research by investigating interaction with palm-based imaginary user interfaces in *eyes-free* manner. In the following, we outline the design space and show how the hand surface can be leveraged for eyes-free interaction with TV 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. Figure 1). Initially, we asked about (1) how they would hold their hand and which side

Figure 1. Example user interface screens of a Samsung TV used in the first study (http://www.samsung.com/us/article/apps-built-for-your-tv).

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 user interface (UI) elements while mimicking their proposed interactions on their hand surface. To foster the discussion, we utilised 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. Figure 1).

We recruited 10 volunteer participants (3f, 7 m). They were between 22 and 42 years old. All participants spent 2–3 h in average per day watching TV. Each single-user session lasted about 1 h. As data gathering methodologies, we videotaped the sessions and asked participants to think aloud. We then selected salient quotes and analysed both quotes and videos using an iterative open, axial and selective coding approach (Strauss and Corbin 2008). For inter-coder independence, two coders coded the data separately.

4. Results

4.1. 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 (Gustafson *et al.* 2011). 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

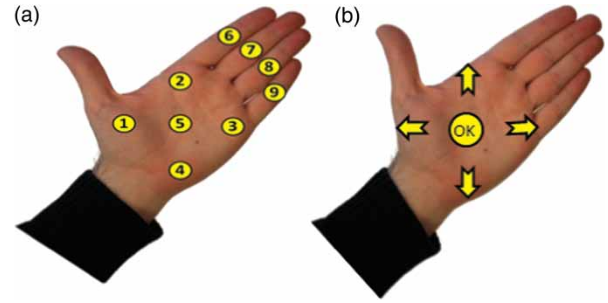

Figure 2. (a) The easily touchable landmarks on the hand. (b) Participants suggested linking the directional keys to the landmarks of the palm while holding the hand diagonally.

revealed nine landmarks on the palm surface, which they believed to be easily touchable without any visual demand based on the proprioceptive sense (Ballendat *et al.* 2010) (cf. Figure 2(a)).

4.2. 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. Figure 2(b)).

In contrast, recalling and transferring digits (typical mapping of 3×3 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.

4.3. Interacting with on-screen UI content

Participants not only suggested two-dimensional (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. Figure 3).

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

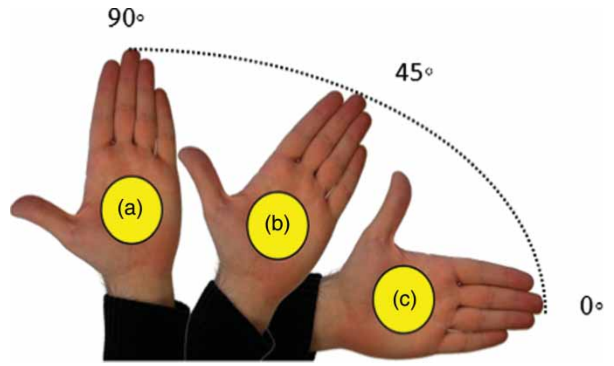

Figure 3. (a) Portrait: pointing towards TV, (b) diagonal: 45° to user's body and (c) landscape: parallel to body.

as 2D-touch gestures, were mainly performed in diagonal orientation.

4.4. 2D-touch gesture interaction on the 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'.

4.5. Pointing on the palm surface

Participants suggested to transfer one-dimensional (1D) 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.

5. Discussion

The results of this study elicit implications for designing a palm-based remote control, which preserves a user's attention to the TV screen during interaction. We found nine 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 visualised 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?

6. 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 mapped UI elements on her palm.

6.1. Experiment set-up

We have conducted the experiment using an optical tracking system (OptiTrack[1] as shown in Figure 4 left) to minimise any noise. We have designed a trackable paper carton apparatus, which the participants wore on the back of their non-dominant hand (cf. Figure 4 right). We have attached three retro-reflective markers as antennas to the paper carton. These markers are then tracked by the OptiTrack system

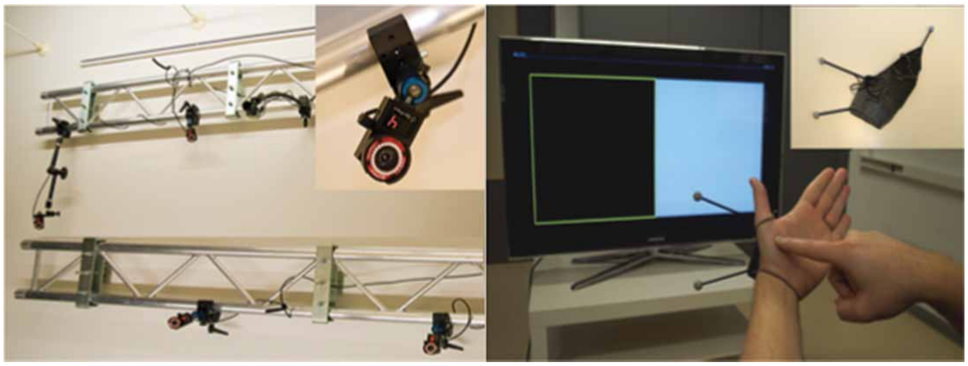

Figure 4. Left: OptiTrack system. Right: the paper carton apparatus used in the controlled experiment.

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, 10 m; 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 min.

6.2. Methodology

The experiment was subdivided into two parts according to our hypotheses. Each part was again subdivided into two tasks (cf. Figure 5). 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 chose a within-subject design. Participants were asked to not look at their hands and only concentrate on the interface shown on the TV screen. We repeated the trials in which the experimenters determined that a 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 comprises two sub-tasks.

- Task 1.1 required participants to map directional keys to their palm (see Figure 2(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 nine different landmarks.
- Task 1.2 required participants to map non-regular grids (see Figure 5) to their palm surface and touch the highlighted position on their palm. Here, participants had to touch eight different landmarks.

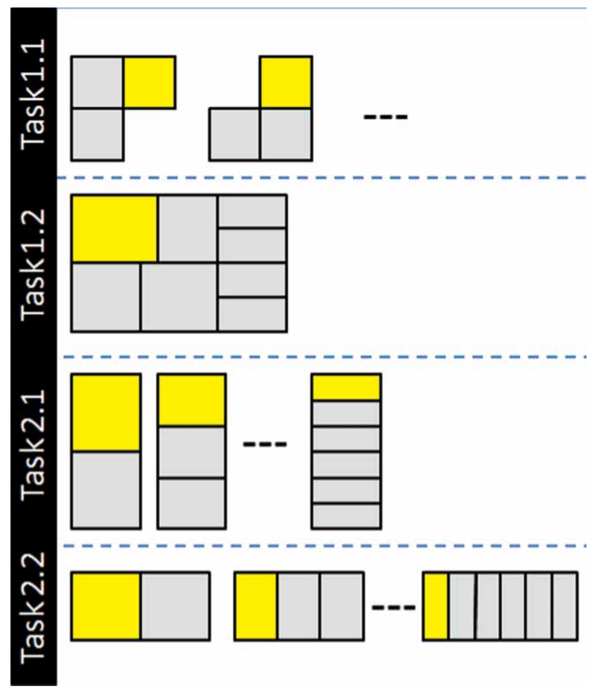

Figure 5. On-screen user interfaces of each task during the experiment.

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.

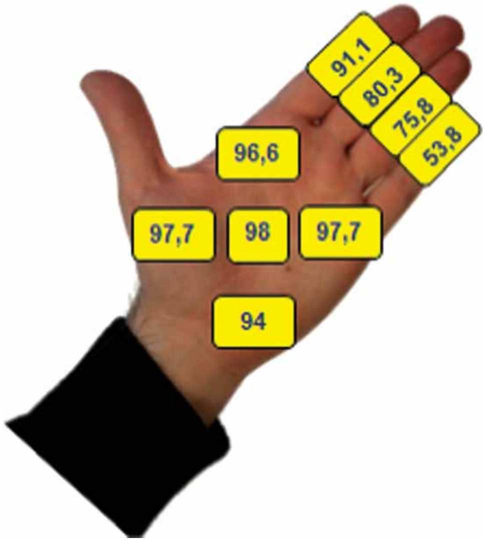

Figure 6. Average effectiveness percentage of targeting each landmark without visual demand.

In order to determine boundaries for the number of targets in this task, we conducted a pilot study. We asked participants to target elements in various density levels starting from two 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 six locations at most. Therefore, the task started with two adjacent targets and step-wise became denser until six targets as depicted in Figure 5, task 2.2.

7. 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 three times the standard deviation away from the centroid. We normalised all hand sizes with the average index finger (7.31 cm).

Part 1: Figure 6 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 analyse targeting, we measured one overall systematic error (offset). On average, the diameter necessary to encompass 90% of all touches is 28 mm ($SD = 0.85$).

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.

Analysis of variance (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 8. The effectiveness decreased

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

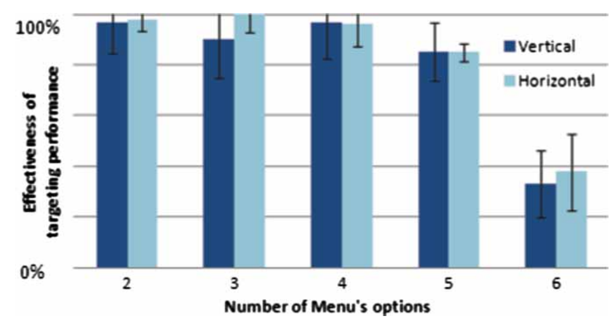

Figure 8. The average effectiveness of targeting vertical and horizontal grids with different equal-sized options.

monotonically for more than three menu options. The average effectiveness is below 90% for more than four options and decreases below 50% for more than five 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.

8. Discussion

Based on the results of the studies, we showed that touching the five landmarks on the palm surface without any visual demand is highly effective. Moreover, it is precise enough to operate interfaces with target sizes of 28 mm 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 four 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 with their palm landmarks, they effectively targeted their index finger. This indicates that also the index finger could be used as an effective input source.

9. Summarising discussion of the 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 recognised 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 utilised 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 28 mm (SD = 0.85) in diameter on the palm surface to encompass 90% of all touches.
- Users can reliably and effectively (>90%) map 1D grid-layout menus with up to four 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 the TV without visual attention in two main ways: first, as a remote control with several functions (e.g. virtual buttons) that are linked to the landmarks. Second, as a unique input surface that the TV user interface is mapped to the entire surface of the palm. Our findings show that under certain circumstances (28 mm button size and four target options) the palm-based remote control is viable. Thus, frequently used functions can be ready at the palm, virtually any time without the need for an additional mediator device.

10. Comparative user study

Our two prior studies presented above examined the general concept of a palm-based remote control. As a consecutive step, we investigate how the concept actually performs in real-world settings. We, therefore, conducted a third study, which we report in the following.

Conventional user input on today's TVs is supported through either button-based remote controls or applications running on Smartphones. Our aim of this user study is to compare these two well-established input modalities with PalmRC in terms of both performance and user experience. We focus on a set of basic tasks such as channel navigation and interaction with common TV applications.

10.1. Study design and methodology

10.1.1. Apparatus

We used a Samsung Smart TV and selected three different input conditions for our comparative study: (1) the *default button-based remote control* for the utilised Samsung Smart TV used in our previous studies. We restricted the interaction to the directional keys and covered the rest of the keypad from the users to focus on basic TV interactions. We further used (2) an *original Samsung Remote application* (ver. 2.2.5) running on an Android-based Smartphone and (3) *PalmRC* with the same setup as in the previous study. We connected PalmRC programmatically to the TV so that users were able to operate the original Samsung TV user interface. The study environment resembled a typical living room.

10.1.2. Tasks

We used the original Samsung user interface for the tasks in all input conditions. The study consisted of three task sets. In tasks 1 and 2, participants were allowed to use four directional keys plus the *OK* button, as well as flick gestures for fast navigation. In the PalmRC condition, the remote control mode with navigational keys was the only active interaction mode – no direct mapping of on-screen user interface was enabled. In task 3, participants could use directional keys, as well as special buttons such as play or pause in both remote control and Smartphone conditions. Direct mapping mode of onscreen interface elements to the palm as region-based UIs was enabled in the PalmRC condition.

The participants had to complete the following tasks:

- *Task 1:* the original list of TV programmes (with a total of 43 programmes – cf. Figure 9(a)) was first shown on the TV screen. Participants had to navigate and find two specific programmes. The programmes were located at two positions in the list, one with a relative short navigational distance to the start position of the task, and the other with a larger navigational distance.
- *Task 2:* In the second set of tasks, participants had to find and watch a movie trailer in the video-on-demand portal of Samsung's Smart TV interface (cf. Figure 9(b) and 9(c)).
- *Task 3:* In this task, we compared the direct mapping of on-screen items offered in the PalmRC condition with the common navigation techniques using directional keys of both button and Smartphone-based remote controls. This task helped us to compare direct selecting a target of TV UI elements with the common way of navigating to the target and selecting it. To do so, we used Shralp (cf. Figure 9(d) and 9(f)), a snowboarding video podcast application due to the simplicity of its user interface. Participants had to first select a video from a menu with four options. Once

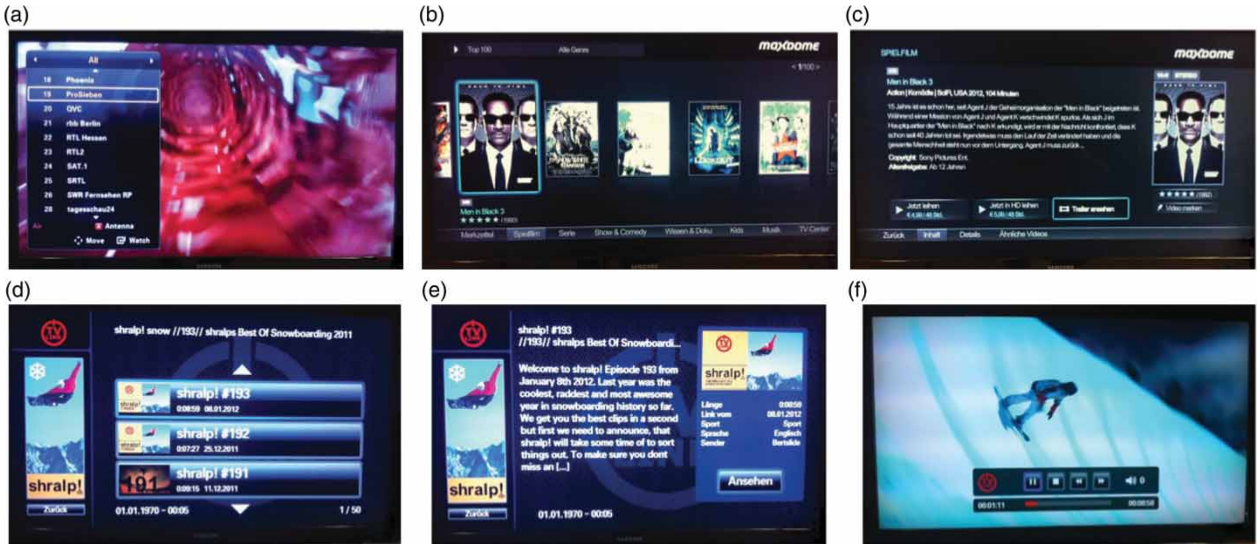

Figure 9. User interfaces and applications used in the comparative user study.

the video was played back, they had to seek for two positions in the video.

10.1.3. Method

We recruited 20 participants (three left-handed, seven female, 35 years of age in average) from various backgrounds such as household, school students, PhD and master students and administrative staff (like secretary). None of the participants took part in any of the previous studies. The participants were introduced to PalmRC concept and the prototype upfront. Each single-user session lasted about 1 h.

The order of input conditions was fully counterbalanced. After each input condition (i.e. standard remote control, Smartphone application and PalmRC), participants were asked to complete an AttrakDiff (Hassenzahl *et al.* 2003) questionnaire, allowing us to measure the user experience, and the (2) NASA task load index (TLI) (Hart and Staveland 1988) which estimates the cognitive demand. Post-study interviews were carried out to collect general comments on the participants' experience during the experiment. All sessions were videotaped. We analysed the interactions, video recordings, salient quotes of the transcripts of the interviews and observational notes using an open coding approach (Strauss and Corbin 2008).

11. Results

In this section, we first present the results of AttrakDiff and the TLI scales followed by behavioural patterns which we derived from our observations. In the end, we discuss some common concerns raised by participants.

11.1. AttrakDiff

Figure 10 shows a user experience portfolio, which situates the three input conditions alongside two quality dimensions: *hedonic* (pleasure) and *pragmatic* (usability) qualities. The portfolio shows that PalmRC excels in terms of hedonic qualities. Its pragmatic qualities are comparable to those of the standard remote control. Overall, the portfolio shows a tendency for PalmRC towards being 'desired'. Among the three input conditions, it was perceived as the most attractive interface with a score of 4.5 on a 7-point Likert scale. The Smartphone interface was perceived as mediocre in terms of both qualities and achieved the lowest score in terms of pragmatic qualities, hence its usability. In terms of attractiveness, it scored 3.5 points. The standard remote control was perceived as the interface with the lowest hedonic qualities, with a slight tendency towards being 'too task-oriented'. It was also evaluated as the least attractive interface with a score of 3. According to the participants, it was further perceived as 'ordinary' and 'unpleasant'.

11.2. Task load index

We collected the perceived workload data using a scale of 1–20 (1 means the least effort) for various types of workloads; mental effort, physical effort, temporal demand, performance, overall effort and frustration. A one-way repeated measure ANOVA found the interface (conditions) to have a main effect on physical effort, temporal demand and frustration (see Figure 11). Post hoc pair-wise comparison with Bonferroni correction revealed that the Smartphone interface caused significantly more temporal demand ($M = 10.9$, $SD = 4.1$) and frustration ($M = 9.6$, $SD = 3.5$) than the other two interfaces (both with $p < .001$). In comparison to the Smartphone and PalmRC conditions, the standard remote control condition resulted in significantly less physical effort being reported ($M = 6.1$, $SD = 3.4$, $p < .001$). We believe that this is because of the two-handed nature of PalmRC (and partially Smartphone) that may require more coordination of both hands.

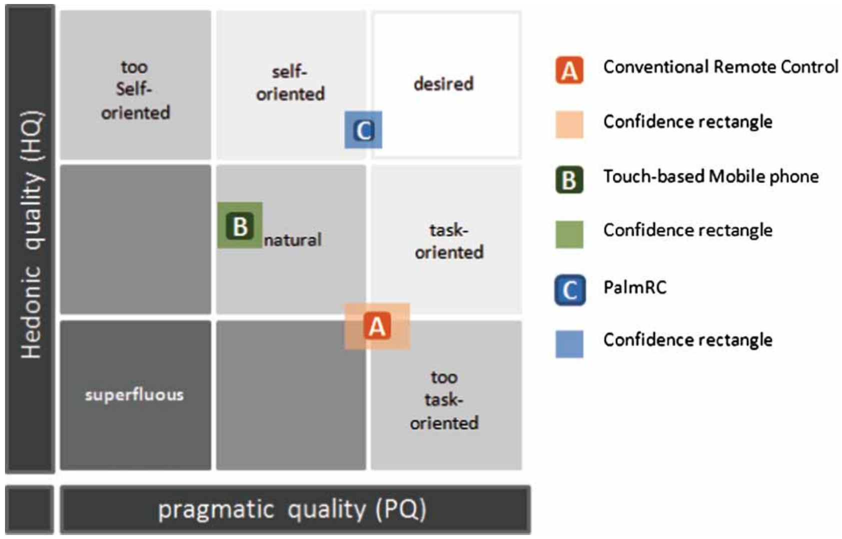

Figure 10. Portfolio with average values of the PQ and HQ dimensions and the respective confidence rectangle of each input conditions.

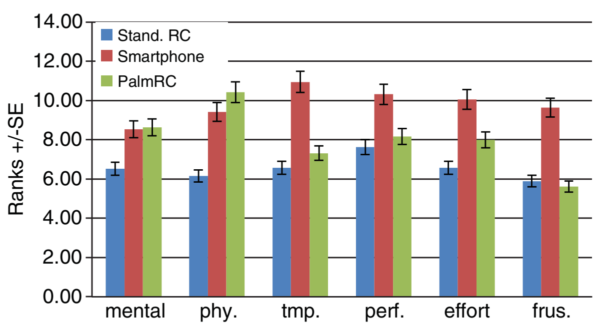

Figure 11. Perceived task load of each interface.

11.3. Observations

11.3.1. General observations

We observed that almost all participants quickly picked up the interaction style of PalmRC. Participants found PalmRC intuitive, fun and entertaining. They appreciated the lack of visual feedback on the palm: 'I just want to see the consequence of my action on the TV and that is enough'. All participants agreed that PalmRC provides shortcuts and immediate interaction. Eight participants stated that PalmRC is very practical in situations where grabbing a conventional remote control or cell phone is difficult, e.g. when either of them are out of reach, or if the hands are soiled while eating something.

All participants found the standard remote control to be practical; for both simple (such as navigation) as well as advanced functions. However, due to its inadequate design it received minimal hedonic qualities. P4 stated 'they [remote controls] are bad designed but well-used'. Nine participants mentioned that they usually have to deal with more than one remote control because of different devices connected to the TV (such as satellite or audio equipments). In contrast, PalmRC was perceived by almost all participants as a *universal* and *personalised* input modality for different devices. Nevertheless, for more non-common interactions participants envisaged using standard remote control instead of PalmRC.

With respect to the Smartphone interface, we observed that participants had more difficulties than with the standard remote control. This was mainly because of the high degree of attention required to look for virtual control widgets on the Smartphone. We observed that the participants held the phone in their hands while watching the TV screen. In many cases, this lead to unwanted touch events and accidental interactions. Eleven participants said that Smartphones are more suitable as a secondary device where one can obtain additional information about the programme. P4 said 'I would like to check related tweets or my Facebook page while watching'.

11.3.2. Visual attention

We analysed how participants interacted with the interface in each condition. We were particularly interested to see how their visual attention was directed while using each input modality. All participants used PalmRC almost eyes-freely. They mainly preserved their attention to the TV.

We observed a different behaviour for the standard remote control condition. The interaction was basically performed in two phases. In the first rather short phase, participants grabbed the device and looked at it for a short while to get accustomed to the layout and find the most common keys (mainly directional keys). This behaviour distracted the participants from the TV. In the second phase, participants left their thumb on one of the main keys (mainly on the *OK* button in the middle of directional keys) while their attention was directed to the TV. In this phase, interaction with the TV was performed almost eyes-freely for navigational tasks. Participants, however, mentioned that in order to look for a specific button, they needed to peek at the remote control.

Based on our observations, we found that the Smartphone was the input modality which required the most visual attention. Participants needed to look at the display of the Smartphone for nearly every single interaction. (1) Lack of an overview of all functions and (2) need to switch to different application modes for different functionalities were the main reasons that participants pointed out.

11.3.3. Concerns

The different orientations designed in PalmRC (cf. Figure 3) turned out to be confusing for eight participants. The participants had particular difficulties with mapping four directions to their hands. Discussion with participants revealed that directional keys should always be oriented towards the TV. This means that the upper part of the hand surface should be mapped to the *UP* key, the lower part to the *DOWN* key.

Nine participants were concerned with the required two-handed usage of PalmRC in contrast to the one-handed usage of the standard remote controls. P4 commented: 'With my remote control at home, I can control the TV while I'm holding a glass in my other hand'. We believe that this issue becomes less severe by extending the PalmRC concept to surfaces of other body parts such as thighs, which affords one-handed interaction. Moreover, the Kinect depth sensing technology has opened up new interactive experiences leveraging any un-instrumented physical surface around users such as couch arms or tables as an input surface to operate TVs (Wilson 2010, Wilson and Benko 2010, Harrison *et al.* 2011).

11.4. Discussion and limitations

Overall, we found that PalmRC provides a usable and foremost joyful way for TV remote interaction. Our observations suggest that is mainly due to its touch-based, eyes-free input characteristic, as well as the natural haptic feedback provided through one's own body parts. It is important to note that PalmRC is not meant as an alternative, but a complementary input technique for TV remote interaction. The study revealed that PalmRC provides a shortcut for common TV interactions and therefore can improve the overall user experience while watching TV.

Furthermore, the study findings confirm existing assumptions (Robertson *et al.* 1996, Cesar *et al.* 2008) that Smartphones and other 'secondary interactive screens' in the living room are more suitable as a companion device than a 'first-class device' demanding a user's complete focus.

While most findings show that PalmRC can enrich the overall living room TV experiences, some limitations apply due to the setting of the study. Although we aimed to create a more realistic environment, the study neglects contextual influences of the real life living room. Among the others, various living room arrangements, number of viewers and their postures in front of the TV as well as their age and health abilities are instances that are not considered in the study. On the other hand, the novelty of the PalmRC concept might influenced why participants rated PalmRC as the most desired input mechanism for TV interactions (cf. Figure 10). While this limits us in investigating the natural user experience in living rooms, the study revealed the salient characteristics of each input mechanism and how they can best complement each other.

As another limitation, the current implementation of PalmRC offers a set of limited interactions with TV systems. It basically supports shortcuts for a few simple functions as well as nonlinear navigation and direct selection of UI elements. As a result, the evaluation focused on studying a set of basic and simple interactions compared with the other input mechanisms. While we believe that this is valid and important as the first step, more advanced interactions need to be examined. Therefore, as future work, we intend to extend the interaction design space of PalmRC to support more advanced interaction techniques.

The results of the three studies provide a fundamental basis for the concept of imaginary hand-based remote controls. Building upon these results, we developed a functional prototype as a proof of concept using a Microsoft Kinect camera as a commodity product. We implemented two main interaction techniques along with several applications to show the usefulness of PalmRC presented in the following section.

12. PalmRC Prototype

PalmRC allows users to operate the TV using empty hands while focusing their visual attention on the TV screen. The users interact by pointing and swiping on their non-dominant hand and the system enables the surface of the palm to be capitalised as an input surface. The TV system receives touch positions and returns appropriate visual feedback on its screen. We developed interaction techniques to perform basic TV interactions such as channel navigation in Electronic Programme Guide, volume adjustment and direct interaction with menu options. PalmRC enables users to use their palm as a shortcut for frequently used commands instead of retrieving a TV remote control. PalmRC supports two main interaction techniques (modes) that make use of pointing and 2D-touch gestures on the palm surface:

12.1. 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 nine landmarks

Figure 12. PalmRC system in different application scenarios.

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 linked to the landmarks of the palm, as revealed in our studies. 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 landmarks 2 and 4 (cf. Figure 2(a)), which are mapped to the up or down keys, respectively. As it is depicted in Figure 2(a), the volume can be adjusted similarly by touching landmark 1 to decrease and 3 to increase the volume. To open up the channel list or menu options users can touch the centre of their palm surface. Similar to touch-enabled devices, swiping upwards 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. Figure 12(a) shows that the user increases the volume by tapping on the corresponding landmark.

12.2. Direct interacting with interface elements

PalmRC directly maps the user interface screen to the entire palm or hand surface. So that 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 iTV interface, which incorporates common social features such as live chat (cf. Figure 12(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 (Kohli and Whitton 2005, Luyten *et al.* 2006) by pointing to the appropriate location of their palm (cf. Figure 12(b)). The technique provides quick and immediate interactions with the social TV interface.

As another application example, while watching a movie or a programme, 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. Figure 12(c)).

12.3. Technical overview

Although the OptiTrack motion capture system used in the controlled experiment and the comparative study enabled us to precisely track the palm and recognise the touch position, it is not practical for TV room settings. As discussed in Section 2, there have been other sensing approaches such as gloves (Kuester *et al.* 2005), Skinput (Harrison *et al.* 2010) and depth cameras (Wilson 2010, Wilson and Benko 2010). We chose to use a Microsoft Kinect (Microsoft Kinect 2013) depth camera because it does not need any instrumentation on the hand of the viewer and also enables and supports recognising touch and drag interactions (Gustafson *et al.* 2010, Gustafson *et al.* 2011).

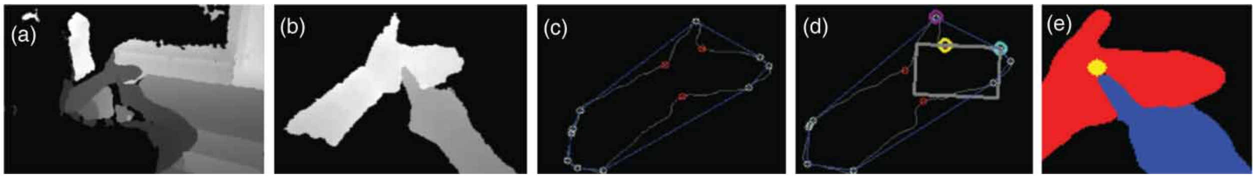

Figure 13. The process of recognising a touch event in PalmRC: (a) original depth image, (b) subtracted background, (c) determining the contour of the reference hand and (d) its palm box, (e) and finally recognised touch on the surface.

In PalmRC, we use the Kinect depth camera to track the non-dominant hand and recognise touch and dragging events with the index finger of the dominant hand. The built-in depth sensors recognise a user's hands in a minimum distance of ~50 cm. Currently, we mount the depth camera on a tripod located at the back of a user's shoulder (cf. Figure 12). We envision 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 recognised in a multistep process similar to (Gustafson *et al.* 2011). The process is depicted in Figure 13. 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 40 cm. We classify the depth values of each hand by calculating the number of peaks in a histogram of all depth values (cf. Figure 13(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 Figure 13(c). The palm box is then calculated based on the prominent defect and the start point (illustrated with yellow and light-blue circles in Figure 13(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 recognised. Due to the local noises and low-resolution of the Kinect depth camera, the precise end of finger tip is not fully recognisable. Similar to Gustafson *et al.* (2011), we determine the touch location by offsetting a small vector in the direction of the finger (yellow circle in Figure 13(e)).

Although the tracking approach requires users to hold their thumb upright while using PalmRC, it robustly recognises different orientations 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.

13. Conclusion

In this article, we explored the concept of leveraging the palm surface as an eyes-free remote control. Through an explorative study, we gained qualitative insights into how people would use their hand as if it were a remote control. Results suggested that users are able to touch several salient 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 direct interaction with on-screen user interface elements. The findings showed that under certain circumstances (e.g. 28 mm button size and four target options) the palm-based remote control is viable.

In the third study, being comparative in nature, we contrasted PalmRC with two common TV input modalities: standard remote controls and Smartphones. The results shed light on advantages and disadvantages of each input modality in terms of both usability and user experience. The results further underline the fact that PalmRC offers an always-available, efficient and effective shortcut for performing frequently used interactions with TV systems without requiring additional mediator devices.

Building upon the results of the three studies, we moreover contributed PalmRC as a prototypical realisation of our concept. We designed two main interaction techniques and showcased them in different application scenarios. The current implementation of the PalmRC prototype support limited number of functions that are mapped to the hand's surface and its landmarks, resulting in a simple and unified user interface design. This is inline with the findings of prior studies (Negroponte and Asher 1995, Pemberton and Griffiths 2003) showing that people want to reduce the overall number of remote controls and the number of keys on each. At the same time, the navigational and direct selection techniques offered in PalmRC can support and cover a wide range of common interactions with TVs (Mirlacher *et al.* 2010, Obrist *et al.* 2010). We, however, believe that the design space of PalmRC has a great potential to support advanced interactions. As an example, the number drawing feature on the palm can be extended in a way that users draw characters to enter text to the TV. Moreover, finger joints of the hand can be also leveraged as landmarks to map additional functions (Gustafson *et al.* 2013).

We conclude that by leveraging different landmarks of the hand, users are able to perform precise interactions, while preserving their visual attention to the TV. At the same time, the palm surface is also appropriate for coarse gestures such as swiping. Based on the initial results, we hypothesise that the interaction style of PalmRC is less tiring than mid-air gestures. Future studies are needed to systematically compare both as device-less input modalities for TV 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. We will further explore the scalability of our concept to handle conflicts when people are watching TV together in co-located settings.

Note

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

References

Ballendat, T., Marquardt, N., and Greenberg, S., 2010. Proxemic interaction: designing for a proximity and orientation-aware environment. *In: ACM International Conference on Interactive Tabletops and Surfaces (ITS'10)*, 7–10 November, Saarbrücken, Germany. New York: ACM, 121–130.

Berglund, A., *et al*., 2006. Paper remote: an augmented television guide and remote control. *Universal Access in the Information Society,* 4 (4), 300–327.

Bernhaupt, R., *et al*., 2008. Trends in the living room and beyond: results from ethnographic studies using creative and playful probing. *Computers in Entertainment (CIE)*, 6 (1), 23 pp. doi:10.1145/1350843.1350848.

Brookstone. 2013. Available from: http://www.brookstone.com/webassets/pdf/650598p_manual.pdf [Accessed 17 September 2013].

Brutti, A., *et al*., 2008. WOZ acoustic data collection for interactive TV. *In*: Nicoletta Calzolari, *et al*., eds. *Proceeding of the sixth international conference on Language and resources and evaluation (LREC'08)*, *European language resources association (ELRA)*, 28–30 May, Marrakech, Morocco. Luxembourg: European Language Resources Association (ELRA).

Cesar, P., Bulterman, D., and Jansen A.J., 2008. *Usages of the secondary screen in an interactive television environment: control, enrich, share, and transfer television content*. Proceedings of the changing television environments, Salzburg, Austria.

Dezfuli, N., *et al*., 2012. *Leveraging the spatial information of viewers for social interactive television systems*. Proceedings of the EuroITV, Berlin, Germany.

Freeman, W.T. and Weissman, C.D., 1995. *Television control by hand gestures*. Proceeding of the international workshop on automatic face and gesture recognition, Santa Barbara, CA, USA. 179–183.

Gustafson, S., Bierwirth, D., and Baudisch, P., 2010. Imaginary interfaces: spatial interaction with empty hands and without visual feedback. *In: Proceedings of the 23rd annual ACM symposium on User interface software and technology (UIST'10)*, New York, NY, USA. New York: ACM, 3–12.

Gustafson, S., Holz, C., and Baudisch, P., 2011. Imaginary phone: learning imaginary interfaces by transferring spatial memory from a familiar device. *In: Proceedings of the 24th annual ACM symposium on User interface software and technology (UIST'11)*, 16–19 October, Santa Barbara, CA, USA. New York: ACM, 283–292.

Gustafson, S., Rabe, B., and Baudisch, P., 2013. Understanding palm-based imaginary interfaces: the role of visual and tactile cues when browsing. *In: Proceedings of the SIGCHI conference on Human factors in computing systems (CHI'13)*, Paris, France. New York: ACM, 889–898.

Harrison, C., Benko, H., and Wilson, A., 2011. OmniTouch: wearable multitouch interaction everywhere. *In: Proceedings of the 24th annual ACM symposium on User interface software and technology (UIST'11)*, Santa Barbara, CA, USA. New York: ACM, 441–450.

Harrison, C., Tan, D., and Morris, D., 2010. Skinput: appropriating the body as an input surface. *In: Proceedings of the SIGCHI conference on Human factors in computing systems (CHI'10)*, Atlanta, GA, USA. New York: ACM, 453–462.

Harrison, C., Ramamurthy, S., and Hudson, S.E., 2012. On-body interaction: armed and dangerous. *In*: Stephen N. Spencer, ed. *Proceeding of the sixth internation conference on Tangible, embedded and embodied interaction (TEI'12)*, Kingston, Canada. New York: ACM, 69–76.

Hart, S.G. and Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. *In*: A. Hancock and N. Meshkati, eds. *Human mental workload.* Amsterdam: North Holland Press, 139–183.

Hassenzahl, M., Burmester, M., and Koller, F., 2003. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. *Proc. Mensch & Computer. Interaktion in Bewegung*, 187–196.

Hess, J., Küstermann, G., and Pipek, V., 2008. pRemote: a user customizable remote control. *In: CHI'08 extended abstracts on Human factors in computing systems (CHI EA '08)*, Florence, Italy. New York: ACM Press, 3279–3284.

Igrashi, T. and Hughes, J.F., 2001. Voice as sound: using nonverbal voice input for interactive control. *In: Proceedings of the 14th annual ACM symposium on User interface software and technology (UIST '01)*, Orlando, FL, USA. New York: ACM, 155–156.

Kohli, L. and Whitton, M., 2005. *The Haptic hand: providing user interface feedback with the non-dominant hand in virtual environments*. Proceedings of the GI 2005, Victoria, BC, Canada. 1–8.

Kuester, F., *et al*., 2005. *Towards keyboard independent touch typing in VR*. Proceedings of the VRST 2005, Monterery, CA, USA. 86–95.

Luyten, K., *et al*., 2006. *Telebuddies: social stitching with interactive television*. Proceedings of the CHI extended abstracts 2006, Montreal, Canada. 1049–154.

Mäntyjärvi, J., *et al*., 2004. *Enabling fast and effortless customization in accelerometer based gesture interaction*. Proceedings of the mobile and ubiquitous multimedia, College Park, MD, USA. 25–31.

Microsoft Kinect. 2013. Available from: http://www.xbox.com/kinect [Accessed 17 September 2013].

Mirlacher, T., *et al*., 2010. *Interactive simplisity for iTV: minimizing keys for navigating content*. Proceedings of the EuroITV 2010, Tampere, Finland. 137–140.

Mistry, P., Maes, P., and Chang, L., 2009. *WUW – wear Ur world: a wearable gestural interface*. Proceedings of the CHI extended abstracts 2009, Boston, MA, USA. 4111–4116.

Negroponte, N. and Asher, M., 1995. *Being digital*. New York, NY: Random House Inc.

Obrist, M., *et al*., 2010. *Field evaluation of a cross platform 6 key navigation model and a unified user interface design*. Proceedings of the EuroITV 2010, Tampere, Finland. 141–144.

Pemberton, L. and Griffiths, R.N., 2003. *Usability evaluation techniques for interactive television*. Proceedings of the HCI international, Crete.

Robertson, S., *et al*., 1996. *Dual device user interface design: PDAs and interactive television*. Proceedings of the CHI 1996, Vancouver BC, Canada. 79–86.

Samsung Smart TV. 2012. Available from: http://www.samsung.com/us/article/apps-built-for-your-tv [Accessed 17 September 2013]

Sherrington, C.S. 1907. On the proprioceptive system, especially in its reflex aspect. *Brain*, 29 (4), 467–482.

Strauss, A. and Corbin, J., 2008. *Basics of qualitative research: techniques and procedures for developing grounded theory*. New York: Sage Publications.

Tamaki, E., Miyaki, T., and Rekimoto, J., 2009. *Brainy Hand: an earworn hand gesture interaction device*. Proceedings of the CHI extended abstracts 2009, Boston, MA, USA. 4255–4260.

Weisz, J.D., *et al*., 2007. *Watching together: integrating text chat with video*. Proceedings of the CHI 2007, San Jose, CA, USA. 877–886.

Wilson, A.D., 2010. *Using a depth camera as a touch sensor*. Proceedings of the ITS 2010, Saarbrücken, Germany. 69–72.

Wilson, A.D. and Benko, H., 2010. *Combining multiple depth cameras and projectors for interactions on, above and between surfaces*. Proceedings of the UIST 2010, New York, USA. 273–282.

Zimmermann, G., *et al*., 2003. *Toward a unified universal remote console standard*. Proceedings of the CHI extended abstracts 2003, Ft. Lauderdale, FL, USA. 874–875.